The Political Geography of Tax H(e)avens and Tax Hells

By NICO S. HANSEN AND ANKE S. KESSLER*

Abstract

The paper explains the existence and the characteristics of tax havens through the interplay of political and geographical factors. In a system of independent democratic jurisdictions among which individuals are mobile, equilibrium tax regimes depend on the relative geographical size of the jurisdictions. If geographical differences are modest or small, jurisdictions independently conduct similar tax policies and average incomes converge. In contrast, if the relative size differentials are substantial, i.e. there are very small and large jurisdictions in the system, tax h(e)avens and tax hells emerge. In equilibrium, small jurisdictions are inhabited by wealthy households and conduct low tax policies (tax heavens) while poor households live in large jurisdictions where taxes are high (tax hells).

Keywords: Tax Havens, Migration, Fiscal Federalism, Redistribution.
JEL–Classification: H71, H73

*Hansen: Apax Partners & Co., Possartstr. 11, D-81679 München, Germany. Kessler: Dept. of Economics, University of Bonn, Adenauerallee 24-42, 53113 Bonn, Germany (kessler@wipol.uni-bonn.de). We would like to thank two anonymous referees, Marcus Berliant, Dennis Epple, Christian Ewerhart, Gerhard Glomm, David Pines, Urs Schweizer, and participants in presentations at the University of Munich, the 1997 SITE meeting (Stanford), the 1997 American Econometric Society Summer Meeting (Pasadena), and the 1996 IIPF Congress (Tel Aviv) for helpful suggestions and discussions. Both authors gratefully acknowledge financial support by the Deutsche Forschungsgemeinschaft, SFB 303 at the University of Bonn. Remaining errors are our own. The views expressed in this paper should not be attributed to Apax Partners & Co.

Worldwide many governments rely on personal income taxation as one of their major sources of tax revenue. Casual empirical evidence suggests that while most developed countries levy substantial taxes, particularly on higher incomes, there also exist a few countries that are characterized by no or very low income taxation. A distinguishing feature of the countries in the latter group is that they are geographically very small as can be seen from Table 1 which presents international income tax policies and geographical dimensions of some selected countries.

[Table 1 about here]

In the present paper, we investigate whether the geography of a country is related to its pattern of taxation. Central to our argument is the ongoing international integration in the last decades. In some cases (e.g. in the EU) the process has advanced to the point where all formal constraints to mobility have been abandoned. This development has also greatly improved the mobility of households across states or national borders. In contrast to the mobility of production factors, however, the effects of household mobility (migration) are not confined to budgetary consequences as tax payers immigrate or emigrate: the in- and outflow of citizens also alters policy objectives by changing the composition of the electorate in a jurisdiction. At the same time migration decisions, especially those of wealthy individuals, are based upon local tax policies. Consequently, the migration of households determines fiscal policies through the interplay of two basic effects: (a) residential choices determine tax rates through a process in which a jurisdiction's inhabitants select their local policies, and (b) tax and welfare policies in each jurisdiction influence residential decisions. As we argue in this paper, this interdependency of residential and political decisions may provide an explanation for the stylized facts illustrated in Table 1.

We consider a simple framework in which households differ in incomes and national tax policies are democratically determined. As a natural implication of their earning characteristics, high income households ceteris paribus prefer to live in countries with low taxation. For ease of exposition, we refer to those countries as *tax h(e)avens* in a slight perturbation of popular nomenclature. Low income households, in contrast,

are more interested in generous public spending than in low income tax rates. Ceteris paribus, they prefer to reside in countries with large welfare programs financed by substantial taxation which we call *tax hells* for obvious reasons. Thus, individual preferences imply a self-selection process which leads to the segregation of households across countries according to income classes.[1] If this segregation is in turn supported by a national vote for low taxes in countries where high-income earners live and high taxes in countries where lower-income earners live, an equilibrium with *tax heavens* - populated by wealthy residents, - and *tax hells* - populated by the less affluent - evolves.

Yet, the geographical size of countries plays a crucial role in this development: first, it affects the number of a country's inhabitants (the population size). Since households sort themselves according to income classes (preferences), small jurisdictions will tend to be inhabited by a more homogeneous population than large jurisdictions. This greater homogeneity of the local population alleviates the political conflict of interest in small jurisdictions. If taxation is redistributive in nature, lower taxes rates will therefore be politically supported by a majority of the inhabitants. Second, to sustain distinct tax rates and segregation there must be a mechanism which prevents middle class households from immigrating into tax heavens in order to escape heavy taxation in the tax hells. The geographical disparities of countries provide such a mechanism: small countries are potential tax heavens because scarcity of land can make property prices prohibitive for some households and a small group of wealthy individuals is able to separate itself from the poor. Conversely, in geographically large jurisdictions land is abundant, and tax enhancing immigration of heterogeneous income groups cannot be prevented through the land market. In particular, we show that there may exist a critical *relative* maximum size of a country which is necessary to support the evolution of a tax heaven and which is rather small. In contrast, if the geographical differences of countries are too modest they will independently

[1]The sorting of individuals by preferences across jurisdictions goes back to the famous contribution of Charles M. Tiebout (1956) on migration as a means to reveal preferences over public goods.

conduct identical policies in equilibrium entailing considerable taxation.[2] Hence, while small countries may be tax heavens medium sized or large countries never are. Beyond the interpretation in the international context, our model could equally be applied to jurisdictional systems within national borders. Specifically, we believe that the described mechanisms may help to explain the demographic and political patterns prevalent in municipalities and metropolitan areas in federal countries as the U.S. where one observes a strong tendency for smaller communities or suburbs to host wealthy populations.[3]

The observation that some countries have lower taxes than others and that this pattern may be related to some notion of 'size' is not new. In Sam Bucovetsky (1991) and John D. Wilson (1991), mobile physical capital is source taxed in countries with immobile labor forces of different size. Due to the equalization of net returns on capital in equilibrium, the large country perceives a smaller elasticity of capital flight on a unilateral tax increase than the small country. As a result, the equilibrium tax rate in the large country is higher and the capital allocation is distorted in favor of the small country. In Kanbur and Keen (1993), countries compete for mobile consumers with sales taxes. Lowering taxes is more attractive for a small than for a large country, because its revenue losses on the home tax base are smaller and its gains through increased cross-border shopping are larger. Again, taxation will be lower in the small country. These contributions certainly describe empirically relevant phenomena, but they rely on source or origin based taxation. Since most countries formally apply residence based individual income taxes (or destination based indirect taxes), these tax competition models thus implicitly involve illegal tax evasion. The present paper, in contrast, focuses on tax havens such as Monaco, Andorra, and the Bahamas. Individuals migrate into those

[2]We will see below that such a symmetric equilibrium in which all countries levy high taxes and have similar population structures always exists, irrespective of relative geographical sizes. Geographical conditions only matter for the existence of asymmetric equilibria. Smallness is therefore a necessary (but not a sufficient) property of tax heavens.

[3]This point is elaborated in Section IV below.

countries to avoid (rather than evade) taxes.[4] Accordingly, our analysis builds on the residence principle of taxation. The emergence of wealthy, low-tax jurisdictions rests on household mobility and geographical (as well as political) arguments, and not on tax competition.[5] In addition, we consider the endogenous change in governments' objective functions due to migration and the political power human capital exerts through voting. Our model is therefore based on - and inspired by - the literature on household mobility and voting on local public services in federal systems. Susan Rose-Ackerman (1979), Dennis Epple et al. (1984, 1993), Epple and Thomas Romer (1991), Epple and Richard Romano (1997) and Raquel Fernandez and Richard Rogerson (1996), for example, also characterize equilibria in which households sort themselves according to preferences (income classes). The main difference between these contributions and the present paper is that we analytically relate political outcomes to a jurisdictions' size.

The outline of the paper is as follows. Section I presents the model. In Section II we prove existence and derive characteristics of equilibria. The last section briefly discusses the results and concludes.

I. The Model

We investigate a system of politically independent but economically integrated jurisdictions among which households are mobile. Political independence means that each jurisdiction determines (some) fiscal policies autonomously. Economic integration does not allow for barriers to trade or migration. The jurisdictions may be interpreted as communities in a metropolitan area, regions, or member states of a federation like the European Union. In order to focus on political motives for government and household behavior, we abstract from allocative reasons for public

[4]In the case of the Bahamas, both tax avoidance and evasion clearly play a role. As in Switzerland, Luxembourg, and the Cayman Islands, comprehensive bank secrecy laws have led to the development of a portfolio capital tax haven based on tax evasion motives.

[5]Note in particular that tax competition is generally driven by the perceived revenue generating effect of lowering taxes. As a consequence, tax competition models cannot account for zero income tax policies as long as some countries still set positive tax rates.

spending and assume that jurisdictions raise income taxes for redistributive purposes only.

Consider an economy populated by a continuum of households with mass normalized to unity. Households have identical preferences over the consumption c of a composite commodity. They are heterogeneous with respect to their exogenously given income y which is distributed across the total population according to the distribution function $F(y)$ with density $f(y) > 0$ on the support $[\underline{y}, \overline{y}] \subset \mathbb{R}^+$. We assume that $f(y)$ is unimodal and that the median income of the overall population y^m is lower than the average income $\overline{Y}$.[6]

Households reside in one and only one of the jurisdictions indexed by $j = 1, ..., J$. For simplicity, we assume that a household occupies exactly one unit of land (called plot) wherever it lives. Each jurisdiction has a given geographical size q_j which represents the land mass (number of plots) available for housing as well as the geographical range of a jurisdiction's political instruments. We impose $\sum_j q_j \geq 1$ and $1 > q_j \geq 0, j = 1, ..., J$, to ensure that the aggregate land area of the economy suffices to accommodate the entire population and that at least two jurisdictions are populated. The plots are owned by competitive absentee landlords from which they have to be rented at the local market price p_j.[7] Without loss of generality, the cost of providing an existing plot of land is zero everywhere.

In each jurisdiction the population democratically implements a local redistributive

[6]While the latter assumption on the skewedness of the income distribution is well in line with empirical observations, the unimodality requirement is more disputable. There are examples of income distributions which are bimodal, e.g., the distribution of household incomes in Great Britain. Most income distributions, however, appear to be unimodal independent of the measurement concept [see e.g. Burkhauser et. al (1997)] Unimodality is also satisfied by the lognormal and the Pareto distribution functions that are often used to approximate real world income distributions.

[7]The absentee landlord assumption allows us to abstract from the effect of land price changes on incomes and the resulting endogeneity of the income distribution. Equivalently, we could assume that each household owns an identical share of plots in each land market. In the literature, this assumption is often imposed to eliminate income effects on demand [see, e.g., Rose-Ackerman (1979) or Epple et al. (1984, 1993)] or the political conflict between property owners and renters [Epple and Romer (1991)]. Neither of these effects plays a role in our model because there is only one consumption good and because of the time structure we impose (see below).

policy scheme that consists of a linear income tax t_j and a per capita grant g_j. Since jurisdictions are financially autonomous, local tax revenues have to match local expenditures for per capita transfers. Denoting the fraction of households with income y who live in jurisdiction j as $f_j(y)$, a jurisdiction's policy bundle (t_j, g_j) is feasible if the following budget constraint holds:

$$g_j \int_{\underline{y}}^{\overline{y}} f_j(y)dy \leq \left(t_j - \frac{1}{2}t_j^2\right) \int_{\underline{y}}^{\overline{y}} y \; f_j(y)dy. \tag{1}$$

The left hand side of (1) represents local spending on redistributive transfers. On the right hand side of the budget are local income tax revenues minus costs of raising public funds which we assume to take the form $t_j^2/2$ for simplicity.[8] If (1) holds with equality, it outlines a concave per-capita 'Laffer-curve'.

There are two explicit decisions each household has to make: it has to choose a jurisdiction to live in and it has to vote together with all other residents in that particular community. We assume the following time structure: in the first stage, each household settles in a jurisdiction. Plot prices endogenously adjust to equate demand and supply for land. In the second stage, the residents of each jurisdiction determine the tax rates through a ballot. In the third stage, they consume.[9] Since residential decisions are made *before* tax policies are determined, this timing eliminates tax competition effects and allows us to focus entirely on the interaction of demographic and geographical conditions in the determination of a jurisdiction's policy. With the model set up as above, an equilibrium is described by a distribution of households over jurisdictions $f_j^*(y)$, fiscal policies (t_j^*, g_j^*), and plot prices

[8] An assumption on such costs is necessary because incomes are exogenous and the absence of negative effects of taxation would allow for arbitrarily high tax rates in the case of pure redistribution, a situation which is clearly unrealistic. In a more general model, deadweight losses of distortive taxation could be induced by endogenous labour supply which would complicate the analysis without providing additional insights (see footnote 17). Introducing costs of public funds is not necessary if one considers local public good provision instead of redistribution (see footnote 10).

[9] See also Epple and Romano (1997) or Fernandez and Rogerson (1996). This sequential structure is equivalent to a setting in which households migrate and vote simultaneously, and voters do not foresee migration responses to their political choices [as, for instance, in Frank Westhoff (1977), Rose-Ackerman (1979), Epple et al. (1984, 1993)].

p_j^* for all $j = 1, ..., J$, such a) given residential choices, the tax/grant combination in each jurisdiction is feasible and preferred to any other feasible combination by a majority of the jurisdiction's inhabitants, b) given plot prices, households decide on their residence optimally, perfectly anticipating subsequent political developments, and c) housing markets clear.

II. Equilibrium Analysis

Let us first consider a household's preferences. Since each household occupies exactly one plot of land independently of the jurisdiction of residence, its utility is fully described by the consumed amount of the composite commodity which is equal to net income minus the rent bill:

$$V(y, t_j, g_j, p_j) = c(y, t_j, g_j, p_j) = (1 - t_j)y + g_j - p_j. \tag{2}$$

Recall that the political equilibrium in j (stage 2) is a redistributive scheme (t_j^*, g_j^*) that is preferred by a majority of the local residents. A voter with income y prefers the redistributive scheme which maximizes his or her indirect utility (2) subject to the budget constraint of the community (1). Note that the housing prices are already determined when the voting occurs and do not influence the political decision. Totally differentiating $V(\cdot)$, we find that the slope of an indifference curve in the plane spanned by t and $g - p$ is positive and increasing in a households income y,

$$\frac{d(g_j - p_j)}{dt_j} \Big|_{V=\overline{V}} = y > 0. \tag{3}$$

Therefore, low-income households prefer regions with high taxes combined with large grant-rent differentials while high-income households prefer regions with low taxes and small grant-rent differentials.[10] Since the budget will always be binding, we can

[10]The analysis here is standard and parallels the general structure of results in the earlier literature [see, e.g., Epple and Romer (1984, 1991) and other papers mentioned in the introduction]. Furthermore, it has to be emphasized that the assumption of purely redistributive policy objectives is made for analytical convenience only and not crucial for our results. Alternatively, we could generalize preference structures by introducing local public services. It is possible to show that the subsequent analysis continues to apply as long as utility functions are homothetic. Note also that in our stylized model, consumption could be substituted for income as the tax base with equivalent implications.

solve the budget constraint for g_j. Substituting this value into (2) and denoting average income in jurisdiction j by $\overline{Y}_j$ yields the preferred tax rate of $1-y/\overline{Y}_j$ for a voter with income y. As preferences for tax rates are single-peaked and monotone in individual income, the feasible redistributive policy which is most preferred by the median income household in jurisdiction j is the unique majority voting equilibrium. If we let y_j^m denote the income of this household, equilibrium tax rates in jurisdiction $j = 1, ..., J$ are thus given by

$$t_j^* = 1 - \frac{y_j^m}{\overline{Y}_j} \tag{4}$$

Tax rates solely depend on the ratio of median to average income in each jurisdiction: the closer median to mean income, the lower the equilibrium tax rates and vice versa. Interpreting the ratio of median to mean as a proxy for income inequality, one may also say that local tax rates are increasing in local income inequality.[11] Employing (4), equilibrium grants are

$$g_j^* = \frac{1}{2}\frac{\overline{Y}_j^2 - (y_j^m)^2}{\overline{Y}_j}, \qquad j = 1, ..., J. \tag{5}$$

Ceteris paribus, the grant will be higher in a region with a higher average income and lower in a region with larger median income.

Having determined fiscal policies, we now turn to the migration decision and ask how the population of the economy locates over jurisdictions. Since households are infinitesimally small, their individual residential choices have no impact on local rents, budgets or political decisions. When deciding where to live, they take plot prices as given and anticipate the political process in the following stage.

Let us first establish the general existence of symmetric perfect foresight equilibria. To see why an equilibrium in which all jurisdictions carry out identical policies always exists, suppose that the population is distributed symmetrically over jurisdictions such that each local income distribution is a smaller copy of the overall

[11] Of course, such an inequality measure is very crude, a drawback it shares with all unidimensional measures.

distribution with $f_j^*(y) = f(y)q_j / \sum_i q_i,\ j = 1, ..., J$. In this case, median and average incomes are the same in all jurisdictions. They are also equal to the overall (federal) median y^m and mean $\overline{Y}$, respectively. Since local tax rates implemented in the second stage depend only on the local median-mean ratio, they will be positive and identical across regions, and so will be the grants. With this particular distribution of the population, there are less people than plots in every jurisdiction, resulting in an economy-wide plot price of zero. This implies that every household is indifferent between jurisdictions, and the presumed distribution, $f_j^*(y)$, $j = 1, ..., J$, is a migration equilibrium in the first stage.[12] Hence, the allocation is an overall equilibrium. The harmonization in this equilibrium not only concerns local fiscal policies but also necessarily average per capita incomes because equal taxes and grants require equal tax bases.[13]

THEOREM 1: *Independent of the jurisdictions' size, a symmetric equilibrium with identical fiscal policies* $(t_j, g_j) = (t^*, g^*) > 0$*, plot prices* $p_j = p^* = 0$*, and per capita average incomes* $\bar{Y}_j = \bar{Y}$ *in all jurisdictions* $j = 1, ...J$ *always exists.*

Next, we investigate the possibility of jurisdictions offering different schemes $(t_j, g_j) \neq (t_k, g_k)$ in equilibrium and to the existence of low tax regions (tax heavens) and high tax regions (tax hells). To simplify the subsequent analysis, it is convenient to focus on the case of two jurisdictions with $(t_1, g_1) \neq (t_2, g_2)$. Note that one jurisdiction must have lower tax rates *and* grant-rent differentials, i.e. $(t_1, g_1 - p_1) < (t_2, g_2 - p_2)$.[14] Otherwise everybody would prefer to live in the region with low taxes and a high grant-rent differential which is impossible due to $q_j < 1$. From (3), if $(t_1, g_1 - p_1) < (t_2, g_2 - p_2)$ the population will sort itself

[12] Population structures across jurisdictions need not be identical: all local distributions which are mean and median preserving spreads of the federation-wide income distribution yield this type of outcome.

[13] Note that this symmetric equilibrium is unstable with respect to various perturbations (see also Fernandez and Rogerson, 1996). This problem is less severe in the more realistic presence of (small) mobility costs which prevent migration if policy differences are insignificant since small shocks to the local population do not cause further migrational responses in this case.

[14] If not indicated otherwise, we assume in what follows without loss of generality that region 1 is the low tax jurisdiction and region 2 the high tax jurisdiction.

across jurisdictions according to income classes - a phenomenon which is already well known from the literature on migration and local public goods and is often referred to as *stratification*. Since both jurisdictions must be populated, there will be a 'boundary' household $\tilde{y}$ which is indifferent between the two jurisdictions: all households $y > \tilde{y}$ will reside in the low tax region 1 whereas all households $y \leq \tilde{y}$ will reside in the high-tax region 2. From (4), we infer that in order to ensure that $t_1 < t_2$ is the outcome of the political process in the second stage, we must have $\frac{y_1^m}{\overline{Y}_1} > \frac{y_2^m}{\overline{Y}_2}$. Thus, a necessary condition for stratification is that income inequality is lower in the wealthy jurisdiction (inhabited by households $y \in (\tilde{y}, \bar{y}]$) than in the poor jurisdiction (inhabited by households $y \in [\underline{y}, \tilde{y}]$.

Now suppose tax rates in a potential stratification equilibrium are non-negative (as we show in the Appendix, any equilibrium satisfies this property under our assumption of a uni-modal distribution function). Since the income of the boundary household $\tilde{y}$ exceeds average income in region 2 and falls short of average income in region 1, this household is then a 'net-contributor' to the redistributive system in region 2 and a 'net-recipient' in region 1. If $p_1 \leq p_2$, i.e., housing is cheaper in the low-tax jurisdiction, the boundary household will therefore strictly prefer to live there, a contradiction. Hence, housing prices in the tax heaven must exceed housing prices in the tax hell.[15] For the land market to clear in this case, however, the demand for available plots in the tax heaven must meet the supply, i.e., p_1 has to increase sufficiently to ensure that exactly the q_1 wealthiest households want to live there. In this case, $p_2 = 0$ and the indifferent boundary household is $\tilde{y} = F^{-1}(1-q_1)$. Substituting $y_1^m = F^{-1}(1 - \frac{1}{2}q_1)$ and $y_2^m = F^{-1}(\frac{1-q_1}{2})$ into (4), the equilibrium tax

[15] Substituting for $g_j = (t_j - \frac{1}{2}t_j^2)\overline{Y}_j$ and using $\overline{Y}_1 > \tilde{y} > \overline{Y}_2$, we have $V(\tilde{y}, t_1, g_1) = (1-t_1)\tilde{y} + (t_1 - \frac{1}{2}t_1^2)\overline{Y}_1 > (1-t_1)\tilde{y} + (t_1 - \frac{1}{2}t_1^2)\tilde{y} - p_1 \geq (1-t_2)\tilde{y} + (t_2 - \frac{1}{2}t_2^2)\tilde{y} - p_2 > (1-t_2)\tilde{y} + (t_2 - \frac{1}{2}t_2^2)\overline{Y}_2 - p_2 = V(\tilde{y}, t_2, g_2, p_2)$ for $0 \leq t_1 \leq t_2$ and $p_1 \leq p_2$. Hence, for $\tilde{y}$ to be indifferent, we must have $p_1 > p_2$.

rates are

$$t_1^* = 1 - q_1 F^{-1}\left(1 - \frac{1}{2}q_1\right) \Big/ \int_{F^{-1}(1-q_1)}^{\bar{y}} y\, f(y)dy, \tag{6}$$

$$\text{and} \quad t_2^* = 1 - (1-q_1)F^{-1}\left(\frac{1-q_1}{2}\right) \Big/ \int_{\underline{y}}^{F^{-1}(1-q_1)} y\, f(y)dy. \tag{7}$$

Using $V(t_1, g_1 - p_1, \tilde{y}) = V(t_2, g_2, \tilde{y})$, the equilibrium plot price in the tax heaven is strictly positive by the argument above and given by

$$p_1^* = (t_2^* - t_1^*)\, F^{-1}(1-q_1) + (g_1^* - g_2^*) > 0 \qquad \text{for} \quad t_2^* \geq t_1^* \geq 0, \tag{8}$$

which is fully determined because the equilibrium policies given by (6), (7) and (5) are functions of q_1 only. Taken together, a stratification equilibrium exists if and only if

$$t_2^*(q_1) > t_1^*(q_1) \tag{9}$$

as given by (6) and (7), i.e., if and only if relative land areas and the overall income distribution are such that there is at least one jurisdiction which can accommodate the most affluent households among whom there is less inequality than among the rest of the population.

The inequality (9) is rather restrictive. At first sight, one may be tempted to argue that the existence of a tax heaven is always supported: if relative land sizes are such that there is no stratification equilibrium where jurisdiction 1 accommodates the rich, there may be an equilibrium with jurisdiction 2 as the low tax region where the wealthy cluster. Yet, this intuition is misleading as can be demonstrated by a simple example.

[Figure 1 about here]

Consider the income distribution $F(y)$ depicted in Figure 1. In Figure 1(a), the wealthiest q_1 individuals of the entire population live in region 1 and are described by the part of the function above $1 - q_1$. The lower part describes the less affluent of mass $1 - q_1$ living in region 2. The respective medians and the boundary household

are indicated on the horizontal axis. One can compare tax rates in this type of figure by comparing the areas A_j and B_j, $j = 1, 2$. If $B_j > A_j$, the median income in jurisdiction j exceeds the average, and vice versa.[16] Due to $A_1 > B_1$ and $A_2 < B_2$, the median to mean ratio is greater in the poor than in the wealthy jurisdiction. Taxation in the latter would be therefore higher (specifically, we would have $t_1 > 0 > t_2$) and this cannot be an equilibrium If, conversely, high income households lived in the smaller region, Figure 2 (b) shows that the poor region would still not levy taxes ($A_1 < B_1$) and hence no stratification equilibrium exists.

The requirement (9) and the above example suggest that the geographical conditions matter qualitatively for political and residential outcomes. To analyze this relationship in more detail note that the tax rate differential $\Delta t^*(q_1) = t_2^*(q_1) - t_1^*(q_1)$ is a continuous function of the geographical dimension q_1. Taking the limits of the tax rates given by (6) and (7) for extreme community sizes yields

$$\begin{aligned} \lim_{q_1 \to 0} t_1^*(q_1) &= \lim_{q_1 \to 0} 1 - y_1^m / \overline{Y}_1 = 1 - \frac{\bar{y}}{\bar{y}} = 0, \\ \lim_{q_1 \to 0} t_2^*(q_1) &= \lim_{q_1 \to 0} 1 - y_2^m / \overline{Y}_2 = 1 - y^m / \overline{Y} > 0. \end{aligned}$$

Thus, $\lim_{q_1 \to 0} \Delta t^*(q_1) > 0$ and a stratification equilibrium exists. Moreover, the small country will be a tax heaven with zero income taxation and the large country will be a tax hell, levying taxes as a centralized political system. By the same line of reasoning, very large countries can never be tax heavens because $\lim_{q_1 \to 1} t_1^*(q_1) = 1 - y^m / \overline{Y} > 0$ and $\lim_{q_1 \to 1} t_2^*(q_1) = 0$. Consequently, a stratification equilibrium with the wealthy living in the large country cannot exist.[17] Continuity of $\Delta t^*(q_1)$ implies

[16] See the Appendix for a formal argument.

[17] These results do not depend on the specific functional form of the costs of taxation we assume. More generally (e.g., with endogenous labor supply), the per capita revenue in j can be expressed as $t_j \bar{Y}_j(t_j)$ where $\bar{Y}$ is a decreasing function of the tax rate t_j. Assuming that the identity of the pivotal voter is unaffected by the tax policy, his preferred tax rate is implicitly determined by the first-order condition [see, e.g., Allan H. Meltzer and Scott F. Richard (1981)]

$$\bar{Y}_j(t_j^*) - y_j^m(t_j^*) + t_j^* \bar{Y}_j'(t_j^*) = 0.$$

Therefore, as $y_1^m(t_1^*)$ approaches $\bar{Y}_1(t_1^*)$ for $q_1 \to 0$, we have $t_1^* \to 0$. Conversely, for $q_1 \to 1$, $y_1^m(t_1^*)$ approaches $y^m(t_1^*)$ and the tax will be generally be positive if overall average income exceeds overall median income at $t = 0$.

that there will be one critical (relative) jurisdiction size q_1^* below which stratification is always supported. These results are summarized in

THEOREM 2: *There is a critical jurisdiction size $q_1^* > 0$ such that a stratification equilibrium with different tax rates always exists if $q_1 \leq q_1^*$. In particular, as $q_1 \to 0$, an equilibrium exists in which a) the very small jurisdiction does not levy taxes (tax heaven) and b) the large jurisdiction levies taxes as a centralized system (tax hell).*

The theorem states that if one jurisdiction is sufficiently small, an equilibrium with the wealthy living in the tax heaven and the rest of the population living in the tax hell always exists. Observe that the condition on q_1 given in the theorem is only a sufficient condition. Unfortunately, it is not possible to derive stronger analytical results since the value of q_1^* crucially depends on the shape of the distribution function $F(y)$.[18] To investigate how small or large jurisdictions have to be to become tax heavens or tax hells, we conducted numerical computations using the log-normal distribution as an approximation of real world income distributions. For the income standard deviation of 0.6 as estimated by Epple and Romer (1991) for the U.S. population in 1979, a stratification equilibrium is only supported if one jurisdictions controls less the critical relative size of $q_1^* = 0.07$. Similarly, the income standard deviation estimate of 0.455 by Eberhard Klein (1987) for Germany in 1984 yields a unique cut-off size of $q_1^* = 0.05$, indicating that less inequality leads to harder constraints on stratification.[19] The computations also confirm that tax rates in the tax heaven decline with the controlled land area and that tax rates in the tax hell increase in its size. These patterns were qualitatively unaltered in simulations with more than two jurisdictions or more extreme standard deviations. There remains a unique and very small critical size that discriminates between existence and non-existence of stratification equilibria. For the log-normal distribution, therefore, the

[18]For distribution functions with a single mode y^M, it is straightforward to show that $q_1^* \leq 1 - F(y^M)$. If $y^M \leq y^m$ (as is the case for log-normal distributions), however, this result does not imply $q_1^* \leq 1/2$.

[19]Since the lognormal distribution was not truncated in the simulations, the findings also show that there being a finite upper bound on income is not necessary for $y_1^m/\overline{Y}_1$ to approach 1 as $\tilde{y}$ gets large (this does not hold for the Pareto distribution, though).

simulations suggest that $q_1 \leq q_1^*$ is a necessary condition for stratification. In this case, values $q_1 > q_1^*$ imply that the tax differential $\Delta t^*(q_1)$ is negative and only symmetric equilibria exist [see Figure 2 (a)]. In general, however, it is possible that a situation as in Figure 2 (b) arises, where stratification can also emerge for intermediate regional sizes.

[Figure 2 about here]

To understand the intuition behind the theorem recall that the extent of local taxation depends on local income inequality. If there is stratification, the income distribution in a region is equal to the upper (respectively lower) part of the federal income distribution truncated at the income of the boundary household $\tilde{y}$. Since a very large region provides housing for almost the entire population, the local median to mean ratio is smaller than unity: the overall median income is smaller than the economy-wide average. As a result, there will be redistributive taxation in large jurisdictions. If, in contrast, a jurisdiction is very small, it accommodates only few households among which (conditional) income inequality is very low. In the limit, the median income of the truncated distribution approaches the mean and taxes are zero. It is important to bear in mind that while this reasoning holds irrespective of which household types reside in a small jurisdiction, only the wealthiest will live there in equilibrium because they have the highest valuation for low-tax policies: the population structure in the small jurisdiction is endogenously determined by the housing market because high income households are willing to pay higher rents than the poor to avoid taxes. Expressed differently, low income inequality in the wealthy jurisdiction can only be sustained through sufficiently high property prices. If land was not scarce, i.e. housing not sufficiently expensive, the wealthiest household in the tax hell had a strong incentive to move to the tax heaven. The rent market prevents this immigration and solves the problem of 'the poor chasing the rich' [William C. Wheaton (1975)]. While in our model the prevention mechanism is strongest through the inelastic demand and supply of housing in the rich jurisdiction, existence of stratification can be shown along similar lines with less restrictive

assumptions – as long as jurisdiction differ significantly in their geographical size.

As mentioned in the introduction, there is a literature which obtains similar results – low taxes in small countries and high taxes in large countries – by postulating tax competition for capital. Equivalent mechanisms and results are present if voters try to attract a tax base in the form of high-income households as in Epple and Romer (1991), who state that therefore "larger communities ... have greater scope for redistribution than smaller ones do". This effect is a *tax base effect*. In contrast, the driving force in our model lies in the endogenous differences in population structures of countries of different size which translate into political divergence. For this reason, it may be called *politico-geographical effect*. Observe that the former effect has slightly different implications than the latter. While tax competition arguments predict a continuous and monotonically increasing relationship between a country's 'size' and its tax rate, our model also allows for a discontinuous pattern of tax policies. If relative geographical conditions are not sufficiently different, migration leads to the equalization of local tax policies and average incomes (only symmetric, high-tax equilibria exist). Also note that the converse is not true: since a symmetric equilibrium can prevail irrespective of countries' relative geographical size, smallness is only a necessary (but not a sufficient) property of a tax heaven.

A look at the geographical and political map of Europe shows that the implications of our theorem fit the empirical evidence very well. The widely known tax heavens such as Andorra and Monaco are the smallest independent countries, levying no or very low income taxes and providing a shelter for elsewhere heavily taxed individuals. None of the large nations as France, Germany, Italy, or the UK are known to be tending the affluent so much. At the same time, land rents in Monaco, for instance, are extraordinarily high. A moderate two-bedroom apartment can easily cost over one million US-dollars, more than tenfold the price observed in large European cities [see *L'Immobilier - Le Marché immobilier de la Principauté de Monaco*, July 1996]. Worldwide, other tax heavens with similar features exist. The Bahamas and the Bermuda Islands do not levy income taxes and it suffices to buy property to become

a resident. With this policy, the islands attract wealthy people from all over the world and land prices, as in Monaco, are extremely high.

Before closing this section, let us consider how the above results generalize to situations with more than two jurisdictions. First, it is evident that one can generate equilibria with several tax heavens and tax hells by replicating an economy with two jurisdictions and stratification.[20] Second, there may exist stratification equilibria in which all countries implement distinct tax policies. In this case, jurisdictions can be ordered in a way such that $j < k$ implies $(t_j^*, g_j^* - p_j^*) < (t_k^*, g_k^* - p_k^*)$. From our reasoning above, we also know that $0 \leq t_j^* < t_k^*$ implies $p_j^* > p_k^*$, i.e., housing prices are inversely related to tax rates. Furthermore, if one jurisdiction is sufficiently small, it is always a potential tax heaven because it can accommodate only a few households among which (conditional) income inequality is low. Since only the wealthiest part of the total population can afford living there, our results on the characteristics of equilibria continue to apply.

III. Discussion and Concluding Remarks

In this paper, we have shown that the degree of democratically chosen taxation is related to a country's relative size. We have focused on a very simple framework with purely redistributive objectives of local governments, income taxes, and inelastic housing supply and demand functions. However, our model could be extended to allow for local public good provision (such as education) or other forms of taxation (such as property taxes) without changing the main insights. In particular, the intuition that only small jurisdiction can be characterized by high property prices and wealthy populations would continue to apply. As long as land available for

[20] More specifically, suppose that the geographical proportions of two jurisdictions are such that a stratification equilibrium exists in which region 1 is a tax heaven and region 2 a tax hell. Now construct an economy with n_1 and n_2 regions, respectively, replicating jurisdictional sizes q_1 and q_2 and the respective income distributions. In this case, policies and plot prices in each newly created jurisdiction are identical to those of the two jurisdictions in the original economy and we have equilibrium consisting of n_1 tax heavens and n_2 tax hells (with total land mass equal to $n_1 q_1 + n_2 q_2$).

housing is sufficiently scarce in a jurisdiction, it can only accommodate a small number of households who belong to the same income interval in a stratification equilibrium. Consequently, local income (preference) inequality will be low in small jurisdictions. This greater homogeneity of the local population implies that political conflicts of interest are minimized: local policies are - irrespective of their specific nature - 'close' to the preferences of *all* inhabitants. If high-income households gain most form the harmonization of political interests, they alone are willing to pay the high property prices which, in turn, help to sustain the homogeneity in the local population.

There are two assumptions which are important to our analysis: first, households are perfectly mobile between jurisdictions and second, they are treated equally everywhere, i.e. independent of their origin they share rights (voting, receiving redistributive transfers) and duties (paying taxes) according to the residence principle. Clearly, these assumptions do not hold in many real world situations. Migration is not costless and informal barriers to free residential choice remain even in political unions as the EU, where free mobility and equal treatment is guaranteed in the Treaties of Rome. Moreover, immigrants do not have all political rights and are sometimes excluded from financial benefits (though less frequently from the duty as a taxpayer).

Both assumptions, however, are unnecessarily strong. First, since tax heavens are very small, only a small fraction of households has to move - in particular those with high incomes - to support the evolution of tax heavens.[21] Second, the right to take part in elections can be viewed as an approximation for implicit political power which may depend more on the fact of being a resident than on the fact of being a voter. Furthermore, the interpretation of our results is not confined to providing an explanation for the existence and the characteristics of international tax havens. Instead, one could apply our model to jurisdictional systems within national bor-

[21] The higher degree of mobility of wealthy individuals is in accordance with empirical observations [see e.g. George J. Borjas et al. (1992)].

ders (e.g. municipalities) where the above assumptions are more in accordance with reality and one frequently observes that wealthy households cluster in small, low tax - high rent, communities. Katharine L. Bradbury (1988), for example, investigates the degree of property taxation in Massachusetts and reports evidence that taxes decline with the population size of a community.[22] In line with our results, the author finds that "...in 1980, effective property tax rates ranged from 11 percent in Boston to 0.5 percent in the tiny (population 500) Martha's Vineyard..." (p. 38). Naturally, such a relationship would also be predicted by tax competition models. In contrast to these models, the force behind our results is a politico-geographical effect that determines local population structures and the demand for taxation. In the present model, the relationship between relative geographical size and tax rates need not be continuous. As the computations show, only significant geographical differences may allow for the divergence of tax policies at all. This may be the reason why, in practice, we do not observe stratification patterns as frequently as it is often suggested by theoretical work. A case in point is Switzerland where real world conditions are probably closest to our assumptions: many jurisdictions, be it Kantone (approximately equivalent to US states) or municipalities, decide independently on local taxation and the provision of services. The democratic processes are direct, and migration is easy and cheap because distances are short. Still, large tax rate differentials and stratification of income groups are only exceptionally observed and the degree of redistribution in the country is not significantly different from its neighbors such as Germany and France where the redistributive function is more centralized [see Werner Pommerehne et al. (1995)].

APPENDIX

Let $F(y)$ be some twice continuously differentiable distribution function on $[\underline{y}, \bar{y}]$ and assume that a) the associated density function $f(y)$ is unimodal with mode

[22] See Table 1 in Bradbury (p. 39): in 1980, for instance, the effective property tax rate was on average 1.9 percent in towns below 5,000 and 4.7 percent in cities with a population over 50,000 people. We thank Dennis Epple for pointing this data source out to us.

y^M and b) the median for the whole distribution, y^m, is less or equal than the mean, $\overline{Y}$. Also, let $F_1(y) = [F(y) - F(\tilde{y})]/[1 - F(\tilde{y})]$ be the truncated distribution restricted to $y \in [\tilde{y}, \bar{y}]$. We show that the truncated distribution is characterized by $y_1^m \leq \overline{Y}_1 \; \forall \; \tilde{y} \in [\underline{y}, \bar{y}]$.

PROOF: Note first that

$$\overline{Y}_1 = \frac{1}{1 - F(\tilde{y})} \int_{\tilde{y}}^{\bar{y}} y f(y) dy = \frac{1}{1 - F(\tilde{y})} \left\{ \bar{y} - F(\tilde{y})\tilde{y} - \int_{\tilde{y}}^{\bar{y}} F(y) dy \right\}$$

by partial integration. Hence,

$$y_1^m \leq \overline{Y}_1 \quad \Leftrightarrow \quad H(\tilde{y}) \equiv \int_{\tilde{y}}^{\bar{y}} F(y) dy + F(\tilde{y})\tilde{y} - \bar{y} + [1 - F(\tilde{y})] y_1^m \leq 0, \tag{10}$$

so that evaluating $y_1^m \gtreqless \overline{Y}_1$ is equivalent to the comparison $B_1 \gtreqless A_1$ in Figure 1. Clearly, $H(\bar{y}) = 0$. Also, $H(\underline{y}) \leq 0$ since $y^m \leq \overline{Y}$ by assumptiom. Thus, (10) holds at $\tilde{y} = \underline{y}$ and $\tilde{y} = \overline{y}$. Taking derivatives of $H(\tilde{y})$, using $F(y_1^m) - F(\tilde{y}) = \frac{1}{2}[1 - F(\tilde{y})]$ and $\partial y_1^m / \partial \tilde{y} = \frac{1}{2}[f(\tilde{y})/f(y_1^m)]$, one obtains $H'(\tilde{y}) = F(y_1^m) - F(\tilde{y}) - f(y_1^m)(y_1^m - \tilde{y})$. Note that since $F(y)$ is convex (concave) for $y < y^M$ ($y > y^M$), $H'(\tilde{y}) < 0$ for $\tilde{y}$ such that $y_1^m(\tilde{y}) \leq y^M$ and $H'(\tilde{y}) > 0$ for $\tilde{y} \geq y^M$. To complete the proof, it suffices to show that $H''(\tilde{y}) > 0$ whenever $H'(\tilde{y}) = 0$. We have at $H'(\tilde{y}) = 0$,

$$H''(\tilde{y}) = -f'(y_1^m)(y_1^m - \tilde{y}) \frac{\partial y_1^m}{\partial y} + f(y_1^m) - f(\tilde{y}). \tag{11}$$

Now, $H'(\tilde{y}) = 0 \Leftrightarrow F(y_1^m) - F(\tilde{y}) = f(y_1^m)(y_1^m - \tilde{y})$, i.e., the area under the density function between $\tilde{y}$ and y_1^m must be equal to a rectangle with height $f(y_1^m)$ and length $y_1^m - \tilde{y}$. For this to be the case, $f(y)$ cannot be monotonic over $[\tilde{y}, y_1^m]$ and can also not attain a minimum at $y = y_1^m$. $H'(\tilde{y}) = 0$ therefore implies $f'(y_1^m) < 0$ and $f(y_1^m) > f(\tilde{y})$ so that $H''(\tilde{y}) > 0$ by (11) which proves the result. Hence, $y_1^m \leq \overline{Y}_1 \; \forall \; \tilde{y} \in [\underline{y}, \bar{y}]$ which together with (4) establishes $t_1 \geq 0 \; \forall \; \tilde{y} \in [\underline{y}, \bar{y}]$.

REFERENCES

Borjas, George J.; Bronars Stephen G. and Stephen J. Trejo. "Self Selection and Internal Migration in the United States." *Journal of Urban Economics*, September 1992, 32(2), pp. 159-85.

Bradbury, Katharine L. "Shifting Property Tax Burdens in Massachusetts." *New England Economic Review*, September/October Issue 1988, pp. 36-48.

Bucovetsky, Sam. "Asymmetric Tax Competition." *Journal of Urban Economics*, September 1991, 30(2), pp. 167-81.

Bundesministerium der Finanzen (BMF). *Die wichtigsten Steuern im Internationalen Vergleich*, Bonn: Informationsdienst des BMF, 1996.

Burkhauser, Richard V.; Crews, Amy D., Daly, Mary C. and Jenkins, Stephen P. *Income Mobility and the Middle Class – AEI Studies on Understanding Inequality*, American Enterprise Institute Press: Washington D.C.

Chambre Immobilière Monégasque. *L'Immobilier - Le Marché immobilier de la Principauté de Monaco*, July 1996.

Epple, Dennis and Romano, Richard E. "Public School Choice and Finance Policies, Neighborhood Formation, and the Distribution of Educational Benefits." *Mimeo*, Carnegie Mellon University.

Epple, Dennis and Romer, Thomas. "Mobility and Redistribution." *Journal of Political Economy*, August 1991, 99(4), pp. 828-58.

Epple, Dennis; Filimon, Radu and Romer, Thomas. "Equilibrium Among Local Jurisdictions: Toward an Integrated Treatment of Voting and Residential Choice." *Journal of Public Economics*, August 1984, 24(3), pp. 281-308.

——**.** "Existence of Voting and Housing Equilibrium in a System of Communities with Property Taxes." *Regional Science and Urban Economics*, November 1993, 23(5), pp. 585-610.

Fernandez, Raquel and Rogerson, Richard. "Income Distribution, Communi-

ties and the Quality of Public Education." *Quarterly Journal of Economics*, February 1996, 111(1), pp. 135-164.

Klein, Eberhard. "Die Neuschätzung des Lohnsteueraufkommens im makroökonomischen Modell des SFB 3." *SFB 03 Arbeitspapier Nr. 232*, Universities of Frankfurt and Mannheim, Germany, 1987.

Meltzer, Allan H. and Richard, Scott F. "A Rational Theory of the Size of Government." *Journal of Political Economy*, October 1981, 89(5), pp. 914-27.

Pommerehne, Werner W.; Kirchgässner, Gebhard and Feld, Lars P. "Tax Harmonization and Tax Competition at State-Local Levels: Lessons from Switzerland," in Giancarlo Pola, George France and Rosella Levaggi, eds., *Developments in Local Government Finance: Theory and Policy*, Cheltenham, U.K.: Elgar, 1996, pp. 292-330.

Rose-Ackerman, Susan. "Market Models of Local Government: Exit, Voting, and the Land Market." *Journal of Urban Economics*, 1979, 6(2), pp. 319-37.

Tiebout, Charles M. "A Pure Theory of Local Expenditures." *Journal of Political Economy*, October 1956, 64(3), pp. 416-24.

Westhoff, Frank. "Existence of Equilibria in Economies with Local Public Goods." *Journal of Economic Theory*, February 1977, 14(1), pp. 84-112.

Wheaton, William C. "Consumer Mobility and Community Tax Bases." *Journal of Public Economics*, November 1975, 4(4), pp. 377-84.

Wilson, John D. "Tax Competition with Interregional Differences in Factor Endowments." *Regional Science and Urban Economics*, November 1991, 21(3), pp. 423-52.

TABLE 1 – TAX RATES AND GEOGRAPHICAL SIZE OF COUNTRIES IN 1996

Country	Tax Rate [percent]	Area [km^2]
Andorra	0	453
Austria	50	83,856
Bahamas	0	11,396
Bermudas	0	53
Belgium	61	30,518
France	57	543,965
Germany	57	357,042
Luxembourg	51	2,586
Monaco	0	2
Netherlands	60	33,975
Spain	56	504,750
UK	40	244,110
USA	44	9,363,125

Notes: Marginal income tax rates for incomes of $ 100.000 (tax rates include average state and local income taxes, USA: state NY). Data source: information service of the German Ministery of Finance, 1996.

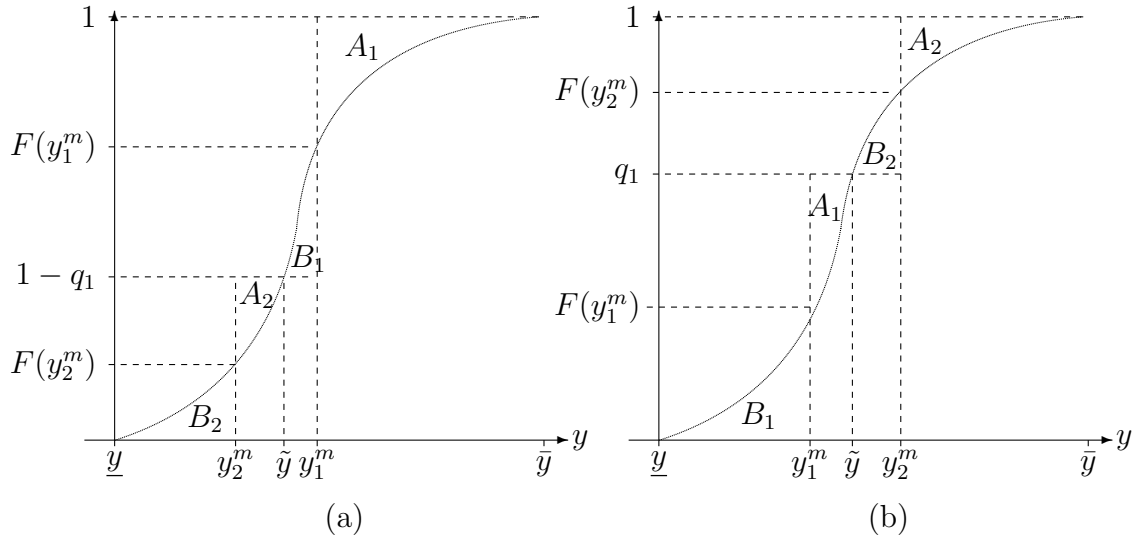

FIGURE 1. A STRATIFICATION EQUILIBRIUM NEED NOT EXIST

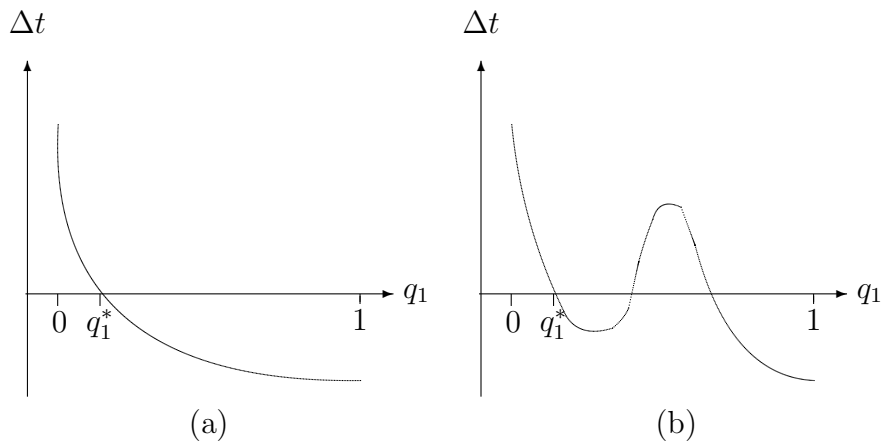

FIGURE 2. STRATIFICATION AND RELATIVE GEOGRAPHICAL SIZE