Sensitivity versus Size: Implications for Tax Competition*

David R. Agrawal† UC Irvine

Adib Bagh‡ U. of Kentucky

Mohammed Mardan§ NHH

December 2024

Abstract

The conventional wisdom is that a big jurisdiction sets a higher tax rate than a small jurisdiction. We show this result arises due to simplifying assumptions that imply tax-base sensitivities are equal across jurisdictions. When more than two jurisdictions compete in commodity taxes, tax-base sensitivities need not be equal across jurisdictions and a small jurisdiction can set a higher tax rate than a big jurisdiction. Our analysis extends to capital and profit taxes, and, more generally, to various types of multi-player asymmetric competition.

Keywords: Ramsey rule, inverse elasticity, fiscal competition, optimal taxation, spatial price competition, sales tax

JEL classification: C7, D4, H2, H7, L1, R5

---

*We thank David Albouy, Masaki Aoyagi, Jan Brueckner, Sam Bucovetsky, Anthony Creane, Aart Gerritsen, Andreas Haufler, Jean Hindriks, William Hoyt, Ravi Kanbur, Jeffrey Lin, Tidiane Ly, Yukihiro Nishimura, Dirk Schindler, Guttorm Schjelderup, Joel Slemrod, David Wildasin, and John Wilson, along with participants at the International Institute of Public Finance Congress and the Symposium in Public Economics for comments on the paper. We are grateful for the extensive comments of the editor, Florian Scheuer, and three anonymous referees. Any remaining errors are our own.

†Department of Economics, 3151 Social Science Plaza, Irvine, CA 92697-5100; email: david.agrawal@uci.edu; phone 001-949-824-3521. Agrawal is also a fellow of CESifo.

‡Departments of Economics and Mathematics, 223F Gatton College of Business and Economics, Lexington, KY 40506-0027; email: adib.bagh@uky.edu; phone: 001-859-257-3057.

§Department of Business and Management Science, Helleveien 30, 5045 Bergen, Norway; e-mail: mohammed.mardan@nhh.no; phone +47 55 95 95 92. Mardan is also an affiliate member of CESifo and NoCeT.

1 Introduction

The inverse elasticity rule, dating to Ramsey (1927), has been widely applied to study taxation, regulatory policy, and the pricing of public utilities. Applied to taxation, the Ramsey rule states that if commodity demands are independent, optimal tax rates should be inversely proportional to the elasticities.[1] A common extension of the inverse-elasticity rule is to tax competition models where jurisdictions set tax rates to attract a mobile tax base. In equilibrium, the revenue-maximizing tax rate T_i in jurisdiction i is characterized by Ramsey pricing:

$$\frac{T_i}{1+T_i} = \frac{1}{\epsilon_i} \quad \text{with} \quad \epsilon_i = -\frac{dB_i(\mathbf{p})/dT_i}{B_i(\mathbf{p})} p_i \ , \tag{1}$$

where after normalizing producer prices to unity, $p_i = 1 + T_i$ denotes the after-tax price and B_i denotes the tax base, which depends on the vector of after-tax prices of all jurisdictions, $\mathbf{p}$. We refer to $\left|dB_i(\mathbf{p})/dT_i\right|$ as the base sensitivity.

A large share of the game theoretic tax competition literature focuses on "duopolity" theory where *two* jurisdictions differ in size (Keen and Konrad 2013). These models conclude that, with mobile factors, a jurisdiction's tax rate is positively correlated with size—bigger jurisdictions set higher tax rates than smaller jurisdictions—because, evaluating (1) at equal tax rates, bigger jurisdictions face a smaller elasticity. We revisit the role of size in the context of spatial competition models concerning the setting of commodity taxes. In the classic models of Kanbur and Keen (1993) and Nielsen (2001), two jurisdictions differ in population, but the population is *uniformly* distributed across space within each jurisdiction. Jurisdictions compete for cross-border shoppers by setting taxes in a Nash game. The more populated jurisdiction always sets a higher tax rate than the smaller.

However, competition does not involve only two jurisdictions nor are people uniformly distributed across space. Most urban economics or trade models recognize that population density differs across space and density has been shown to be critical in the context of cross-border issues (Hindriks and Serse 2019).[2] We therefore extend the classic commodity tax competi-

[1] The assumption that demands are independent is restrictive and makes the inverse elasticity rule less useful for optimal policy (Scheuer and Werning 2016), but it is the basis for much of the intuition concerning the optimal multi-jurisdictional setting of tax rates in open economies.

[2] Friberg et al. (2022) find a non-monotonicity in incentives to cross-border shop.

tion model by allowing jurisdictions to differ in how their populations are *distributed*. One jurisdiction may have many consumers who can readily cross-border shop, while the other may have most of its population far away from borders. Following Caplin and Nalebuff (1991), we show that under mild conditions on the distribution function, a unique Nash equilibrium exists. We first show that if maintaining the assumption of *two* competing jurisdictions, the conventional result remains: a jurisdiction with a larger total population will *always* set a higher tax rate than the smaller jurisdiction *regardless* of how population is distributed.

The intuition for this stark result is that, in the case of two jurisdictions, the marginal benefit of lowering taxes is the same for both the large and small jurisdiction. Thus, the tax-base sensitivity is identical in both jurisdictions, $|dB_1/dT_1| = |dB_2/dT_2|$, and the magnitude of this derivative simultaneously affects the *level*—but not the *pattern*—of tax rates in both jurisdictions. Then, the relative *elasticity*, ϵ_1/ϵ_2, of the tax base evaluated at identical tax rates $T_1 = T_2$, depends *only* on the relative jurisdiction sizes, B_1/B_2. Smaller governments perceive a higher elasticity irrespective of the distribution of individuals.

Our main contribution is then to show that relaxing *both* the assumption of two jurisdictions and a uniform population distribution allows for a much richer pattern of equilibrium tax rates that can differ significantly from the conventional wisdom. Critically, we show that a smaller jurisdiction will set a higher rate than a bigger jurisdiction if there are multiple competitors *and* if the distribution of population is not uniform. This result arises without resorting to any other asymmetries and without adding cross-base interdependencies.[3]

Simplifying to gain intuition, consider the example of Connecticut, Rhode Island, and Massachusetts. Both Connecticut's and Rhode Island's borders with Massachusetts are densely populated due to the Hartford/Springfield and Providence metropolitan areas. But the density at the Rhode Island and Connecticut border is very low, consisting of mainly rural farmland. With multiple borders, the sensitivity of the tax base depends on an average of the responses at both borders of each state. As both borders with Massachusetts are the densest, Massachusetts can attract more cross-border shoppers by lowering its tax rate compared to either of the other states, meaning that the tax-base

[3]Most other models of asymmetric tax competition focus on size differences (Bucovetsky 1991; Haufler and Wooten 1999). The theoretical literature acknowledges that other asymmetries such as preferences for public goods matter (Haufler 1996; Nielsen 2002).

sensitivity in Massachusetts, $|dB_{MA}/dT_{MA}|$, is larger in absolute value than the tax-base sensitivity in either of the other states, $|dB_{RI}/dT_{RI}|, |dB_{CT}/dT_{CT}|$. Thus, even though the tax base of Massachusetts is the largest, if its sensitivity is sufficiently larger in absolute value than the other two sensitivities, Massachusetts' elasticity will be larger. Accordingly, in the presence of multiple jurisdictions, differences in the elasticity of the tax base now depends *both* on population size and the *distribution* of residents across space. There are numerous examples where smaller jurisdictions set higher tax rates than bigger jurisdictions. In the U.S., where counties can set local taxes, we document using data from Agrawal (2014) that only in 16% of cases does the county that sets the highest tax rate compared to its neighbors also have the largest population.

Tax competition remains an important determinant of consumption and excise taxes, even though capital and labor are regarded as more mobile. First, for smaller governments such as states, counties, or towns (or potentially even small countries), cross-border shopping remains important. Second, unlike capital which is globally mobile, consumption tax bases are only locally mobile via cross-border shopping.[4] States, localities and even countries have a small number of neighbors, so that game theoretic interactions become important. Finally, numerous empirical studies document the existence of strategic tax competition in commodity taxes.

Given technological change and globalization have arguably made capital and labor relatively more footloose, we show that our main results are applicable to models of tax competition for corporate profits and capital, which seemingly differ in important ways from the spatial commodity tax model. First, our results generalize to models of tax competition with profit shifting (Keen and Konrad 2013), which traditionally impose restrictions on the cost of shifting profits to another country. In addition, these models generally assume that multinational firms shift profits to a single low-tax country, but, in reality, profit shifting can occur between many country pairs. The curvature of the shifting cost function to the firm has the same qualitative implications on the relationship between tax rates and size as the density function does in the commodity tax setting.

Second, analyzing competition for capital, Mongrain and Wilson (2018)

[4] However, e-commerce and digital services allow households to consume goods from all over the world, potentially making the tax base more mobile.

obtain the standard size result when firms have heterogeneous costs of moving. We show the distribution of moving costs plays the same role as the density of people in commodity tax competition models. If moving costs are uniformly distributed, then the size effect dominates because the number of firms that are indifferent between moving and not moving are the same for the two regions. If moving costs are not uniformly distributed and there are more than two jurisdictions, the results from our commodity tax model extend to capital tax competition.

Our results also have implications for industrial organization and political economy. A strand of industrial organization focuses on spatial price competition with heterogeneous consumers and shows that the distribution of preferences affects firm competition (Neven 1986; Anderson and Goeree 1997; Bloch and Manceau 1999; Calvó-Armengol and Zenou 2002). A similar mechanism is in the spatial voting literature (Bagh 2023) where candidates announce platforms and heterogeneously distributed voters choose the platform closest to their preferences. If candidates view winning as a means to policy, there is a trade-off between the probability of winning and the implementation of the preferred party policy if elected; this trade-off can be influenced by the distribution of voter preferences (Wittman 1983). Both industrial organization and political economy typically restrict attention to *duopolistic* or *symmetric* oligopolistic competition, while we show that the implication of consumer density may *qualitatively* change price-setting behavior with *asymmetric oligopolistic competition*.[5]

2 A General Model of Tax Competition

The result that a bigger jurisdiction sets a higher tax rate than a smaller jurisdiction is common among commodity tax competition models (Kanbur and Keen 1993; Lockwood 1993; Trandel 1994; Nielsen 2001; Wang 1999; Ohsawa and Koshizuka 2003). We expand the classic model of commodity tax compe-

[5]Chen and Riordan (2008) compare monopoly pricing to symmetric duopolistic competition and find that prices under duopoly may be higher. Because a duopolist has a smaller base, duopoly prices tend to be lower. However, the price sensitivity is lower in a duopoly, which leads to comparably higher prices under duopoly. The reason for the latter effect is that a duopolist can only increase the market share by attracting consumers from the competitor, whereas the monopolist can increase the market share by extending market coverage.

tition of Nielsen (2001)[6] and Kanbur and Keen (1993).[7] In the typical model, consumers are uniformly distributed within jurisdictions. Because commodity tax models are closer to a Ramsey framework than capital tax models, we start here.

2.1 Non-uniform Distribution with Two Jurisdictions

Before analyzing the implications of a non-uniform distribution of residents, we adjust the standard linear Hotelling model by locating two jurisdictions on a circle. Transitioning to a circle will allow us to generalize the model to more than two jurisdictions, with each jurisdiction having the same number of borders.[8] We normalize total population to 1 and the circumference of the circle to 1. Jurisdiction 1 ranges from l_{21} to l_{12}; jurisdiction 2 ranges from l_{12} to $l_{21}+1$.

Each jurisdiction's government levies an origin-based commodity tax, where we denote the tax rate of jurisdiction i by T_i. Tax rates are chosen in a Nash game to maximize tax revenue. Firms are potentially located anywhere on the circumference and sell the good in a perfectly competitive environment resulting in producer prices equating to marginal costs, which we normalize to 1. Individuals reside along the circumference and wish to purchase one unit of a composite good from firms. Irrespective of the individuals' residence, the maximum willingness to pay is $\overline{V}$, meaning that individual demand is zero if

[6] Nielsen (2001) normalizes density to be unity across both jurisdictions, allowing him to focus on size differences (area and population) by assuming that one jurisdiction is longer than the other. Given jurisdictions are characterized by two parameters, density and length, and because uniform density is imposed throughout both jurisdictions, the model does not actually allow for a change in one jurisdiction's population *unless* you are willing to change area and population jointly at same time. In particular, increases in population in both jurisdictions (via an increase in density) holding market area constant has no effect on tax rates. However, an increase in market area, holding constant population (i.e., reducing density) increases tax rates.

[7] Kanbur and Keen (1993) normalize area (length) across both jurisdictions allowing them to talk about differences in population even though they have differences in density across countries but not within countries. Because Kanbur and Keen (1993) features a discontinuity in the density at the border, unlike the Nielsen (2001) model, it allows for specific country perturbations. An increase in population (also density) increases tax rates. An increase in area but holding population constant (lowering density also) decreases tax rates. However, these two shocks are isomorphic in this model because the ratio of population and area is all that matters.

[8] The use of a circle follows the industrial organization literature (Salop 1979). In public finance, Trandel (1992) and Agrawal (2015) use a circle with uniformly distributed consumers and no differences in jurisdiction sizes.

the total price exceeds $\overline{V}$. We assume $\overline{V}$ is large enough to ensure full market coverage, that is, larger than the highest gross price inclusive of transportation costs. This upper limit on willingness to pay bounds tax rates from above at a maximum rate $\overline{T}$.

Although demand is perfectly inelastic, individuals have choice over where to buy the good. A purchase at home incurs no transport costs because the individual shops at the firm located at the point where she resides, thus paying the tax rate there. Instead, if the individual purchases the good in the neighboring jurisdiction—doing so at the first store after crossing the border—she pays the tax-inclusive price in the neighboring jurisdiction, but incurs transportation costs δ per unit of distance traveled to the nearest border from her home.

We generalize the standard model by allowing individuals to be non-uniformly distributed.[9] Let x denote the clockwise distance from the border l_{21}, which starts at point 0 on the circle. Residents are distributed on the circle according to a continuous and differentiable probability density function (pdf), $f(x)$ with $f(x) > 0$, on the interval $[0,1]$. It has a cumulative distribution function denoted by $F(x)$. Populations are given by $P_1 = \int_0^{l_{12}} f(x)dx$ and $P_2 = \int_{l_{12}}^1 f(x)dx$.

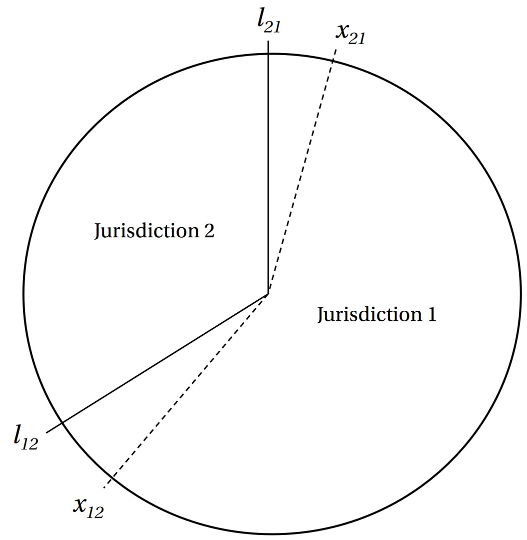
A consumer will purchase the good in the neighboring jurisdiction if the tax savings (the tax differential) are greater than or equal to the cost of travel to the nearest border (δ times distance to the respective border). For each border, the location of the marginal individuals that are indifferent between shopping at home and in the neighboring jurisdiction, are given by:

$$x_{12}(T_1, T_2, l_{12}) = l_{12} - \tfrac{T_1 - T_2}{\delta}, \qquad x_{21}(T_1, T_2, l_{21}) = l_{21} - \tfrac{T_2 - T_1}{\delta}. \tag{2}$$

We normalize l_{21} to be point 0 so that x represents distance from this border. Individuals located within the (clockwise) range enclosed by the points $x_{21}(T_1, T_2, l_{21})$ and $x_{12}(T_1, T_2, l_{12})$ purchase the good in jurisdiction 1, whereas the remaining individuals shop in jurisdiction 2. The model does not presuppose any pattern on tax rates: these cutoff rules encompass both cases where $T_1 > T_2$ and $T_1 \leq T_2$. Figure 1 summarizes the geography of the model for one case.

[9] In Nielsen (2001) and Kanbur and Keen (1993), once area and population are known, density is irrelevant. Trandel (1994) realizes this issue and allows density to vary across space, but does so in a linear manner and with two jurisdictions. A linearly increasing distribution implies that density near the border and population are always positively correlated.

Figure 1: Model structure with two jurisdictions when $T_1 > T_2$



This figure shows the geographic layout. Consumers are located on the circumference of a circle where jurisdiction 1 ranges from l_{21} to l_{12} and jurisdiction 2 ranges from l_{12} to $l_{21}+1$. We normalize $l_{21}=0$. For illustrative purposes, we show the case where $T_1 > T_2$. Individuals located within the (clockwise) range enclosed by the points x_{21} and x_{12} purchase the good in jurisdiction 1, whereas the remaining individuals shop in jurisdiction 2.

For notational convenience, due the possibility that $x_{21} < 0$ if $T_2 > T_1$, we define the density function over the interval $x \in [-1, 1]$, where 1 is the length of the circumference. For the same reason, we assume $f(x)$ is periodic with a period of 1, which implies that for any $x < 0$, we have $f(x+1) = f(x)$. This allows us to integrate over a range containing negative values of x, thus expressing the revenue functions elegantly. These assumptions are not critical to derive any results.

The revenue (payoff) functions are the tax rate times the tax base, $B_i(T_1, T_2)$:

$$\begin{aligned} R_1 &\equiv T_1 B_1(T_1, T_2) = T_1 \left[\int_{x_{21}}^{x_{12}} f(x)\, dx \right] = T_1 \left[F\left(x_{12}\right) - F\left(x_{21}\right) \right], \\ R_2 &\equiv T_2 B_2(T_1, T_2) = T_2 \left[1 - \int_{x_{21}}^{x_{12}} f(x)\, dx \right] = T_2 \left[1 - F\left(x_{12}\right) + F\left(x_{21}\right) \right], \end{aligned}$$

where x_{12} and x_{21} are the locations of the marginal consumers as defined in (2), suppressing the functional notation for convenience.

Differentiating tax revenues to solve the revenue-maximization problem yields the first-order conditions for T_1 and T_2, respectively:

$$\frac{\partial R_1}{\partial T_1} = F\left(x_{12}\right) - F\left(x_{21}\right) - T_1 \frac{f\left(x_{12}\right) + f\left(x_{21}\right)}{\delta} = 0, \tag{3}$$

$$\frac{\partial R_2}{\partial T_2} = 1 - F\left(x_{12}\right) + F\left(x_{21}\right) - T_2 \frac{f\left(x_{12}\right) + f\left(x_{21}\right)}{\delta} = 0. \tag{4}$$

To prove existence and uniqueness of a Nash equilibrium in the tax competition game, we make the following assumption after defining $\beta(x) = f'(x)/f(x)$:

Assumption 1. (log-concavity) *The distribution $f(x)$ is log-concave and, therefore, the ratio $\beta(x) = f'(x)/f(x)$ is non-increasing on $[0,1]$.*

The assumption of log-concavity is frequently used in studies that have a non-linear distribution of consumers (Anderson et al. 1995; Bloch and Manceau 1999). Caplin and Nalebuff (1991) show that log-concavity is a sufficient condition for the existence of equilibrium in a general class of games. As noted in Caplin and Nalebuff (1991), the class of log-concave densities covers many frequently used probability distribution functions such as the normal, exponential, gamma, beta, Weibull, logistic, Laplace, and uniform distributions.

Remark 1. Although many log-concave pdfs have supports larger than $[0,1]$, we can always truncate a pdf with support larger than $[0,1]$ by defining the truncated pdf $\tilde{f} = f(x)/[F(1) - F(0)]$ with support $[0,1]$. As shown in Bagnoli and Bergstrom (2005), when f is log-concave, its truncation $\tilde{f}$ will also be log-concave.

Thus, Assumption 1 allows for a large variety of distribution functions.

Remark 2. As shown in Bagnoli and Bergstrom (2005), the log-concavity $f(x)$ implies the log-concavity of $F(x)$, that is, $f(x)/F(x)$ is also non-increasing on $[0,1]$.

We need an additional technical assumption for existence. Define $\rho = \delta/\overline{T}$.

Assumption 2. (Technical) *The distribution $f(x)$ satisfies $f'(0)/f(0) < \delta/\overline{T} \equiv \rho$.*

Remark 3. This condition is about the growth rate of f at zero. It will hold if—near zero—the graph of f is below the graph of the function $h(x) = f(0)e^{\rho x}$.

As long as $f(x)$ is below $h(x)$ near zero, the slope of f will be less than h and $f'(0) < h'(0)$. Therefore, $f'(0)/f(0) < \rho f(0)e^{\rho \cdot 0}/f(0) = \rho$. Since h grows exponentially, the condition is not very restrictive, allowing for a wide variety of pdfs. While ρ can theoretically take on any finite number, as a matter of practicality, we can think of a reasonable range for ρ. If taxes are expressed in ad valorem form, tax rates are usually bound above by one. With respect to δ, if distance is measured in time (hours), then δ is proportional to the opportunity value of time (wages) plus driving costs (gasoline). Thus, ρ need not be small.

There are numerous examples of log-concave distribution functions—or their truncation over $[0,1]$—for which Assumption 2 holds. This includes: the uniform distribution for any value of ρ; the exponential distribution $f(x) = \lambda e^{-\lambda x}$ truncated over $[0,1]$ satisfies $f'(0)/f(0) = -\lambda < \rho$ for all values of ρ; and the normal distribution's truncation over $[0,1]$ with parameters μ and σ has $f'(0)/f(0) = 2\mu/\sigma$, which will be smaller than any ρ for a small enough μ or large enough σ. With the normal distribution, for $\mu = 0$, the condition will hold for any ρ and any σ.

We can now state:

Proposition 1. (Existence & Uniqueness) *Suppose Assumptions 1 and 2 hold. A Nash equilibrium exists and is unique.*

Proof. See Appendix A.1. □

This Nash equilibrium is characterized by

$$T_1^N = -\frac{B_1(T_1^N, T_2^N)}{\frac{\partial B_1}{\partial T_1}\left(T_1^N, T_2^N\right)} = \frac{F\left(x_{12}^N\right) - F\left(x_{21}^N\right)}{\left[f\left(x_{12}^N\right) + f\left(x_{21}^N\right)\right]/\delta}, \quad (5)$$

$$T_2^N = -\frac{B_2(T_1^N, T_2^N)}{\frac{\partial B_2}{\partial T_2}\left(T_1^N, T_2^N\right)} = \frac{1 - F\left(x_{12}^N\right) + F\left(x_{21}^N\right)}{\left[f\left(x_{12}^N\right) + f\left(x_{21}^N\right)\right]/\delta}, \quad (6)$$

where we define $x_{21}^N \equiv x_{21}(T_1^N, T_2^N, 0)$ and $x_{12}^N \equiv x_{12}(T_1^N, T_2^N, l_{12})$ as the values of the cutoff rules evaluated at the Nash tax rates. Assuming $f(x)$ follows a uniform distribution implies that, after accounting for the second border, the optimal tax rates align with those in Nielsen (2001). Further, dividing both sides of each equation by $1 + T_i^N$ yields the standard inverse elasticity formulation given by (1):

$$\frac{T_i^N}{1 + T_i^N} = \frac{1}{\epsilon_i} \text{ with } \epsilon_i = -\frac{\frac{\partial B_i}{\partial T_i}\left(T_1^N, T_2^N\right)}{B_i(T_1^N, T_2^N)}\left(1 + T_i^N\right). \quad (7)$$

The numerators in the right-hand-side of (5) and (6) are i's equilibrium *tax base*; in the case of jurisdiction 1 this is $F\left(x_{12}^N\right) - F\left(x_{21}^N\right)$. And the (absolute value of) changes in the tax base, $\partial B_i/\partial T_i < 0$, which are the denominators the right-hand side, are i's equilibrium *tax-base sensitivity*; in the case of two jurisdictions, this term is identical: $|\partial B_i/\partial T_i| = \left[f\left(x_{12}^N\right) + f\left(x_{21}^N\right)\right]/\delta$.

Based on equations (5) and (6), we establish the following result:

Proposition 2. (Two Jurisdictions & Tax Rates) *Suppose Assumptions 1 and 2 hold. The unique Nash equilibrium satisfies $T_1^N > T_2^N$ if and only if $P_1 > P_2$.*

Proof. Given the denominators are equal in (5) and (6), the pattern of tax rates depends on the relative sizes of the tax bases in the numerator:

$$T_1^N > T_2^N \Longleftrightarrow \left[F\left(x_{12}^N\right) - F\left(x_{21}^N\right)\right] > \left[1 - F\left(x_{12}^N\right) + F\left(x_{21}^N\right)\right]. \tag{8}$$

If $T_1^N > T_2^N$, then $P_1 > F(x_{12}^N) - F(x_{21}^N)$, because some individuals residing in jurisdiction 1 cross-border shop into jurisdiction 2. For the same reason, it must be that $1 - F(x_{12}^N) + F(x_{21}^N) > P_2$. Then (8) implies that $P_1 > F\left(x_{12}^N\right) - F\left(x_{21}^N\right) > 1 - F\left(x_{12}^N\right) + F\left(x_{21}^N\right) > P_2$. Thus, $T_1^N > T_2^N \Rightarrow P_1 > P_2$.

We show $P_1 > P_2$ implies $T_1^N > T_2^N$ by contradiction. Let $T_2^N > T_1^N$ despite $P_1 > P_2$. Then by (8), $P_2 - CBS^N > P_1 + CBS^N$, where CBS^N denotes the total number of cross-border shoppers evaluated at the Nash tax rates. If $T_2^N > T_1^N$, then it must be that $CBS^N > 0$. Given equality of denominators, for $T_2^N > T_1^N$ to arise, it must be that $P_2 - CBS^N > P_1 + CBS^N$, which is impossible given that $P_1 > P_2$ and $CBS^N > 0$. Thus, $P_1 > P_2 \Rightarrow T_1^N > T_2^N$. □

This classic result—previously derived in more stylized models—has led to the the intuition underlying many tax competition models. Intuitively, starting from equal tax rates, a change in a jurisdiction's own-tax rate will have a smaller percent change on its tax base if the jurisdiction is larger. Thus, the larger jurisdiction perceives a smaller elasticity, which under the inverse-elasticity rule implies its optimal tax rate must rise relative to the jurisdiction with the smaller population.

At first glance, it may appear surprising that the large jurisdiction always sets the higher tax rate even in the extreme case when almost all of its population is located directly at its borders, while the small jurisdiction's population is concentrated at its interior. One might initially think that the jurisdiction with its population at its interior is more inelastic. The reason for this stark result, however, originates from the fact that the tax-base sensitivity is identical for both jurisdictions: they are competing for the same marginal individuals irrespective of how these individuals are distributed, i.e. $|\partial B_1/\partial T_1| = |\partial B_2/\partial T_2| = \left[f\left(x_{12}^N\right) + f\left(x_{21}^N\right)\right]/\delta$. In other words, the density of the marginal consumers at x_{ij}^N is identical for both jurisdictions, implying that any marginal change in the tax differential results in the same tax-base change for each jurisdiction. Accordingly, differences in relative elasticities, and thus tax rates, are solely determined by relative differences in the size of the tax bases across the jurisdictions.

2.2 Non-uniform Distribution with Three Jurisdictions

In this section, we analyze whether the previously derived result extends to a setup with more than two jurisdictions. While all of the basic assumptions remain unaltered, we modify the setup by adding a third jurisdiction. Specifically, jurisdiction 1 ranges from l_{31} to l_{12}, jurisdiction 2 ranges from l_{12} to l_{23}, and jurisdiction 3 ranges from l_{23} to $l_{31}+1$. The cut-off rules for the marginal individuals are:

$$x_{12}(T_1,T_2,l_{12}) = l_{12} - \tfrac{T_1-T_2}{\delta}, \qquad x_{23}(T_2,T_3,l_{23}) = l_{23} - \tfrac{T_2-T_3}{\delta}, \qquad x_{31}(T_1,T_3,l_{31}) = l_{31} - \tfrac{T_3-T_1}{\delta}. \tag{9}$$

Individuals located within the range enclosed by the points x_{31} and x_{12} purchase the good in jurisdiction 1, individuals located between x_{12}^* and x_{23}^* shop in jurisdiction 2, and the remaining individuals buy the good in jurisdiction 3. Again, we normalize l_{31} to be 0. Figure 2 displays the geography of the model.

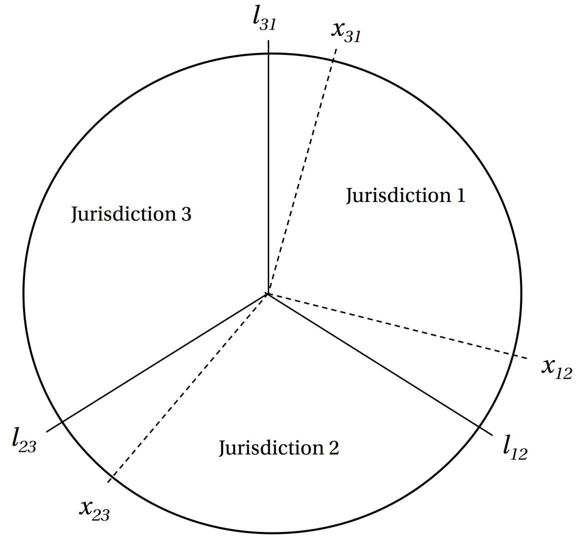
As previously, we can express the revenue functions as:

$$\begin{aligned}
R_1 \equiv T_1 B_1(T_1,T_2,T_3) &= T_1\left[\int_{x_{31}}^{x_{12}} f(x)dx\right] = T_1\left[F\left(x_{12}\right) - F\left(x_{31}\right)\right],\\
R_2 \equiv T_2 B_2(T_1,T_2,T_3) &= T_2\left[\int_{x_{12}}^{x_{23}} f(x)dx\right] = T_2\left[F\left(x_{23}\right) - F\left(x_{12}\right)\right],\\
R_3 \equiv T_3 B_3(T_1,T_2,T_3) &= T_3\left[1 - \int_{x_{31}}^{x_{12}} f(x)dx - \int_{x_{12}}^{x_{23}} f(x)dx\right] = T_3\left[1 - F\left(x_{23}\right) + F\left(x_{31}\right)\right].
\end{aligned}$$

Then, proceeding as previously, we can show:

Remark 4. Suppose Assumptions 1 and 2 hold. A Nash equilibrium with three

Figure 2: Model structure with three jurisdictions when $T_1 > T_2 > T_3$



This figure shows the geographic layout of a three jurisdiction model. Consumers are located on the circumference of a circle where jurisdiction 1 ranges from l_{31} to l_{12}, jurisdiction 2 ranges from l_{12} to l_{23}, and jurisdiction 3 ranges from l_{23} to $l_{31}+1$. We normalize $l_{31}=0$. For illustrative purposes, we show the case where $T_1 > T_2 > T_3$. Individuals located within the (clockwise) range enclosed by the points x_{31} and x_{12} purchase the good in jurisdiction 1, individuals located within the (clockwise) range enclosed by the points x_{12} and x_{23} purchase the good in jurisdiction 2, and the remaining individuals shop in jurisdiction 3.

jurisdictions again exists (and is unique). The proof is a straightforward extension of that for Proposition 1, but is presented in Appendix A.2.

The optimal tax rates, can be expressed as:

$$T_1^N = -\frac{B_1(T_1^N,T_2^N,T_3^N)}{\frac{\partial B_1}{\partial T_1}\left(T_1^N,T_2^N,T_3^N\right)} = \frac{F\left(x_{12}^N\right)-F\left(x_{31}^N\right)}{\left[f\left(x_{31}^N\right)+f\left(x_{12}^N\right)\right]/\delta} = \frac{P_1 - CBS_{12}^N - CBS_{31}^N}{\left[f\left(x_{31}^N\right)+f\left(x_{12}^N\right)\right]/\delta}, \quad (10)$$

$$T_2^N = -\frac{B_2(T_1^N,T_2^N,T_3^N)}{\frac{\partial B_2}{\partial T_2}\left(T_1^N,T_2^N,T_3^N\right)} = \frac{F\left(x_{23}^N\right)-F\left(x_{12}^N\right)}{\left[f\left(x_{12}^N\right)+f\left(x_{23}^N\right)\right]/\delta} = \frac{P_2 + CBS_{12}^N - CBS_{23}^N}{\left[f\left(x_{12}^N\right)+f\left(x_{23}^N\right)\right]/\delta}, \quad (11)$$

$$T_3^N = -\frac{B_3(T_1^N,T_2^N,T_3^N)}{\frac{\partial B_3}{\partial T_3}\left(T_1^N,T_2^N,T_3^N\right)} = \frac{F\left(x_{31}^N\right)-F\left(x_{23}^N\right)}{\left[f\left(x_{23}^N\right)+f\left(x_{31}^N\right)\right]/\delta} = \frac{P_3 + CBS_{23}^N + CBS_{31}^N}{\left[f\left(x_{23}^N\right)+f\left(x_{31}^N\right)\right]/\delta}, \quad (12)$$

where, again, we simplify notation by letting a superscript N on the cut-off rules denote that they are evaluated at the Nash tax rates. On the right-most side of the equations, we rewrite the tax base as the jurisdictions' population adjusted for cross-border shopping. Specifically, population sizes are $P_1 = F(l_{12}) - F(0)$, $P_2 = F(l_{23}) - F(l_{12})$, and $P_3 = F(1) - F(l_{23})$, while the cross-border shoppers evaluated at the Nash equilibrium tax rates are given by $CBS_{12}^N =$

$F(l_{12}) - F\left(x_{12}^N\right)$, $CBS_{31}^N = F\left(x_{31}^N\right) - F(l_{31})$, and $CBS_{23}^N = F(l_{23}) - F\left(x_{23}^N\right)$. As previously, dividing by $1 + T_i^N$ yields the standard inverse-elasticity formulation given by (7).

To determine if, and under what conditions, a smaller jurisdiction will set a higher tax rate, we proceed in two steps. In a first step, we rely on the simplifying assumption of symmetry between two of the three jurisdictions, as this allows us to prove our claim in an easy and elegant way. In a second step, we relax the symmetry assumption and show that a smaller jurisdiction can set a higher tax rate than the larger jurisdiction under a very general population distribution, albeit at the cost of more involved proofs.

First, in a very general, but symmetric setting, we can always find a range of jurisdictional boundaries such that, in equilibrium, at least one smaller jurisdiction posts a higher tax rate than a larger jurisdiction.

Proposition 3. (Overturn Classic Result with Symmetry) *Suppose Assumptions 1 and 2 hold. Assume that $f(x) = f(1-x)$, so that the distribution is symmetric about $x = 1/2$. Let $l_{12} = l$ and $l_{23} = 1 - l$ so that jurisdiction 1 and 3 are symmetric. Then we have values of l where $P_1 < P_2$ and $T_1^N > T_2^N$. That is, there exist distribution functions and jurisdiction lengths where the Nash equilibrium is such that a bigger population jurisdiction sets a strictly lower tax rate than at least one smaller jurisdiction.*

Proof. The value of l that gives us $P_1 = P_2 = P_3$ is obtained by solving $F(l) = 1/3$, which has a unique solution denoted by $\bar{l}$. Using (10)-(12) and imposing symmetry, i.e. $T_1^N = T_3^N$, $f(l) = f(1-l)$, and $F(1-l) = 1 - F(l)$, yields the Nash equilibrium $T_1^N = T_3^N = \delta F(x_{12}^N) / \left[f(0) + f(x_{12}^N)\right]$ and $T_2^N = \delta\left[1 - 2F(x_{12}^N)\right] / \left[f(x_{23}^N) + f(x_{12}^N)\right]$. The value of l that then gives $T_1^N = T_2^N = T_3^N$, denoted by $\underline{l}$, is given by solving

$$M(l) \equiv \frac{2f(l)F(l)}{f(0) + f(l)} = 1 - 2F(l) \equiv H(l), \tag{13}$$

where $M'(l) > 0$ while $H'(l) < 0$. Furthermore, $H(0) = 1$, $H(1/2) = 0$ and $M(0) = 0$. Therefore, (13) has a unique solution $0 < \underline{l} < 1/2$. Therefore, for any $\underline{l} < l < \bar{l}$ or we have $T_1^N > T_2^N$ despite $P_2 > P_1$.

To derive a sufficient condition for $\underline{l} < \bar{l}$, we can rewrite (13) as:

$$F(l) = \frac{f(0) + f(l)}{2f(0) + 4f(l)} \equiv \varphi(l). \tag{14}$$

By definition, $\underline{l}$ is the solution to the equation $F(l) = \varphi(l)$ whereas $\bar{l}$ is the solution to the equation $F(l) = 1/3$. Clearly, $\varphi(0) > F(0)$. Moreover, assuming $f(0) < f(\bar{l})$ implies $\varphi(\bar{l}) < 1/3 = F(\bar{l})$. The continuity of φ and F now implies that F and φ must intersect over $(0, \bar{l})$. Hence, $f(0) < f(\bar{l})$ is a sufficient condition for $\underline{l} < \bar{l}$. □

The symmetry invoked to derive this proposition is a powerful tool to show that there exist parameter constellations of jurisdiction boundaries such that a smaller jurisdiction will set a higher tax rate than a larger jurisdiction. Thus, in a very general, but symmetric setting, we can always find a range of jurisdictional boundaries such at least one jurisdiction posts higher tax rate than a smaller sized jurisdiction. Interestingly, if $f(0) = 0$ the values $\underline{l}$ and $\bar{l}$ satisfy $F^{-1}(1/4) < l < F^{-1}(1/3)$, though in this case, because Assumption 2 does not hold, it would also need to be verified that the revenue functions are quasi-concave for existence. To derive a precise analytical solution, we use a triangular distribution as an example.

Example 1. (Triangular Distribution and Symmetric Jurisdictions) Jurisdictions 1 and 3 are symmetric, and range from $[0; l]$ and $[1 - l; 1]$, respectively. The rest of the circumference encloses jurisdiction 2. The distribution of population is triangular, symmetric around its maximum at $x = 1/2$, and with slopes of m and $-m$, respectively. It satisfies $f(0) = \kappa > 0$ with $4\kappa + m = 4$ in order for the area under f to integrate to one. Thus, the density is $f(x) = mx + 1 - m/4$ for $x \leq 1/2$ and $f(x) = -mx + 1 + 3m/4$ for $x > 1/2$, with $F(x) = (m/2)x^2 + (1 - m/4)x$ for $x \leq 1/2$ and $F(x) = -(m/2)x^2 + (1 + 3m/4)x - m/4$ for $x > 1/2$. Then, focusing on a specific example by letting $m = 2$ and assuming $\rho > 4$, if $\sqrt{73}/16 - 3/16 < l < \sqrt{57}/12 - 1/4$, we have $P_2 > P_3 = P_1$, but $T_2^N < T_3^N = T_1^N$, i.e., the bigger jurisdiction 2 sets the lower tax rate, overturning the classic result.

Proof. We could proceed by explicitly solving for the Nash equilibrium and then comparing relative tax rates and populations. Alternatively, we can apply Proposition 3 by using the the functional form given in Example 1, noting that both Assumption 1[10] and 2 are satisfied when $m = 2$ and $\rho > 4$. Then, $F(l) < 1/3$ becomes $l^2 + l/2 < 1/3$ implying we must have $l < \bar{l} = \sqrt{57}/12 - 1/4 \approx 0.379$. Further, (13) implies $l > \underline{l} = \sqrt{73}/16 - 3/16 \approx 0.347$. Thus, if $l < \underline{l}$, then $P_2 > P_3 = P_1$

[10] Our triangular pdf is log-concave but is not differentiable at 1/2, as it has a sharp peak at 1/2, but we can "smooth out" this peak over an interval that is arbitrarily small around 1/2.

and $T_2 > T_3 = T_1$, i.e., the bigger jurisdiction 2 sets the higher tax rate. If $\bar{l} < l$, then $P_3 = P_1 > P_2$ and $T_3^N = T_1^N > T_2^N$, i.e., the bigger jurisdictions 1 and 3 set the higher tax rate. However, if $\underline{l} < l < \bar{l}$, then $P_2 > P_3 = P_1$, but $T_3^N = T_1^N > T_2^N$, so that the bigger jurisdiction sets the lower tax rate. □

While these symmetric examples are sufficient to make the claim that there exist distribution functions that allow the smaller jurisdiction to set the higher tax rate, one may wish to determine if the result is due to the symmetry assumption or if a more general result is available. Next, we generalize to an asymmetric setting.

To formalize this strategy of proof, we start by embedding $f(\cdot)$, the population pdf, within a family of pdfs—perturbations—$f(\cdot,\varepsilon)$ where $\varepsilon \in [0,\nu]$. In the unperturbed (original) game, $\varepsilon = 0$, and we drop it from our notation, e.g., $f(\cdot,0) = f(\cdot)$. The corresponding family of cdfs will be denoted by $F(\cdot,\varepsilon)$ with $F(\cdot,0) = F(\cdot)$. We will consider games with a perturbed pdf $f(\cdot,\varepsilon)$ and parameters l_{12} and l_{23}. We denote such a game by $G(\varepsilon, l_{12}, l_{23})$. The corresponding population sizes and equilibrium—when it exists—will be respectively denoted by $P_i(\varepsilon, l_{12}, l_{23})$, and $T_i^N(\varepsilon, l_{12}, l_{23})$, for $i = 1,2,3$.

The strategy of our proof relies on small—but very specific— perturbations of the density function that *change the populations* of jurisdictions while leaving *taxes unchanged*. Generally, different distributions of individuals can directly affect tax rates in potentially two ways: through alterations in the size of the tax bases and modifications in the tax-base sensitivity. This means that arbitrary changes in the distribution function will usually have an ambiguous effect on jurisdictions' tax rates. Thus, we will consider a *specific* perturbation. First, the perturbation cannot affect the number of marginal cross-border shoppers at any x_{ij}^N, which ensures that tax-base sensitivities remain unaffected. Second, to ensure that tax bases remain unchanged, changes in a jurisdiction's population that originate from the perturbation need to be appropriately matched by changes in the number of non-marginal cross-border shoppers from another jurisdiction because a jurisdiction's tax base is its population adjusted for cross-border shoppers.

Suppose without loss of generality, l_{12} and l_{23} are such that $P_1 = P_2 > P_3$ and that Assumptions 1 and 2 hold such that a Nash equilibrium exists. In the main text, we focus on the most challenging case to prove our result—where that Nash equilibrium is:[11]

$$T_1^N = T_2^N > T_3^N. \tag{15}$$

We define a population perturbation. First, let (a,b) be an open subinterval of $(0,1)$ with $0 < a < 1$ and $l_{23} < b < 1$; (a,b) must be picked so that it contains x_{12}^N, x_{23}^N and x_{31}^N, that is, the cut-off rules evaluated at the Nash equilibrium before the perturbation. We introduce a population redistribution of size ε from jurisdiction 3 to jurisdiction 1 *around the outside* of the interval (a,b) in the following manner. Consider intervals (a_1,b_1) and (a_2,b_2) such that the first interval satisfies $0 < a_1 < b_1 < a$ and the second interval satisfies $b < a_2 < b_2 < 1$. Define two continuous functions g_1 and g_2 where $g_1 \geq 0$ and it is zero outside (a_1,b_1) whereas $g_2 \leq 0$ and is zero outside (a_2,b_2). We define the perturbed function, $f(x,\varepsilon)$, as:

$$f(x,\varepsilon) = f(x) + g_1(x) + g_2(x).$$

Assume further that g_1 and g_2 are chosen such that i) $f(x,\varepsilon) \geq 0$ on $[0,1]$ and ii) $\int_{a_1}^{b_1} g_1(x)\,dx = \varepsilon$ and $\int_{a_2}^{b_2} g_2(x)\,dx = -\varepsilon$. Our assumptions on g_1 and g_2 imply that $f(\cdot,\varepsilon)$ is a pdf on $[0,1]$. Figure 3 shows the construction of $f(\cdot,\varepsilon)$ graphically. Visually, we can see that the above re-distribution moves a population of size ε from jurisdiction 3 to jurisdiction 1. However, the population we move continues to shop in jurisdiction 3. Therefore, for small enough ε, we expect the above population re-distribution to—very slightly—increases the population in jurisdiction 1 without impacting the tax bases or sensitivities, and thus not changing the equilibrium conditions or the pre-redistribution equilibrium tax rates.

Indeed, Lemma 1 will be used to show that specific population movements to/from particular points will yield the same equilibrium tax rates.

Lemma 1. (Perturbations and Equilibrium) *Let $\phi : \left[0,\overline{T}\right] \to \mathbb{R}$ be a continuous function. Let z^* be the unique maximizer of ϕ over $\left[0,\overline{T}\right]$. We can embed ϕ into a family of perturbations $\phi(\cdot,\varepsilon)$ with $\varepsilon \geq 0$ such that $\phi(z,0) = \phi(z)$ for all $z \in \left[0,\overline{T}\right]$. Assume further that:*

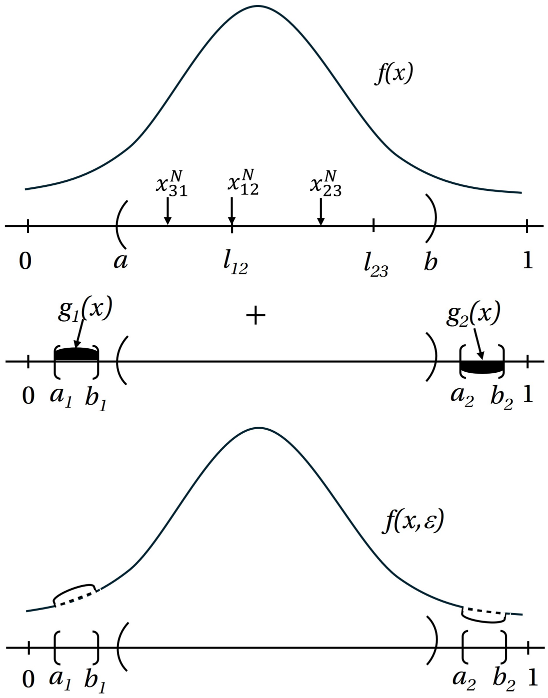
i) there exists an open interval S that is a subset of $[0,\overline{T}]$ and that contains z^* such that for all $z \in S$ and all ε, we have $\phi(z,\varepsilon) = \phi(z)$.

ii) for all $z \in \left[0,\overline{T}\right]$ and all $\varepsilon \in [0,\varepsilon]$, we have $\left|\phi(z,\varepsilon) - \phi(z)\right| < \varepsilon$.

There exists $\tilde{\varepsilon}$ such that $\forall\ \varepsilon < \tilde{\varepsilon}$, z^* is the unique maximizer of $\phi(\cdot,\varepsilon)$ over $\left[0,\overline{T}\right]$.

[11] The other two cases can be shown trivially to lead to our desired result (Appendix A.5).

Figure 3: Graphical Representation of Perturbation



The horizontal x axis shows an interval (a, b) relative to the borders and the cutoff rules for an initial equilibrium corresponding to $T_1^N = T_2^N > T_3^N$. This figure then shows the specific perturbation to the distribution functions given by the continuous functions g_1 and g_2, which equal zero for all x other than the where the functions are nonzero. The shaded area of g_1 is ε while it is $-\varepsilon$ for g_2. The perturbations are amplified graphically. The shape of the distribution function, position of cut-off rules, and jurisdiction borders are not drawn to scale.

Proof. See Appendix A.3. □

Intuitively, given a function with a unique maximizer z^* over some compact interval, we can introduce a "very small" perturbation that occurs "far enough" from z^*. After such perturbation, z^* continues to be a unique maximizer for the perturbed function. We will apply the above lemma to the payoffs functions R_1, R_2, and R_3 and to specific perturbations of the underlying population distribution. This will allow us to show in Appendix A.5 that the equilibrium of the original unperturbed game (with a population distribution satisfying Assumptions 1 and 2) is also an equilibrium to a perturbed game (with a population distribution that may fail to satisfy Assumptions 1 and 2).[12]

However, this perturbation alone does not yield our result, but now

[12]For small enough ε, the solution of the unperturbed problem is an equilibrium for the perturbed problem. However, the perturbed problem may have additional equilibria. This possibility in the perturbed game does not invalidate the subsequent proposition because we are claiming the existence of a game with at least one equilibrium that has specific properties.

must also be combined with changes in jurisdiction boundaries. With this new population distribution, starting from the boundaries such that $T_1^N = T_2^N$, we can then make sufficiently small changes in jurisdictional boundaries such that $T_1^N < T_2^N$ without changing the inequality on the relationship between populations.

Lemma 2. (Comparative Statics) *Suppose Assumptions 1 and 2 hold. The equilibrium of the game $G(l_{12}, l_{23})$ satisfies*

$$\begin{array}{ccc} \frac{\partial(T_1^N - T_2^N)}{\partial l_{12}} > 0, & \frac{\partial(T_1^N - T_3^N)}{\partial l_{12}} > 0, & \frac{\partial(T_2^N - T_3^N)}{\partial l_{12}} < 0, \\ \frac{\partial(T_1^N - T_2^N)}{\partial l_{23}} < 0, & \frac{\partial(T_1^N - T_3^N)}{\partial l_{23}} > 0, & \frac{\partial(T_2^N - T_3^N)}{\partial l_{23}} > 0 \end{array}. \tag{16}$$

Proof. See Appendix A.4. □

Intuitively, focusing on the case of l_{12}, if l_{12} increases clockwise, all else equal, the tax differential $T_1^N - T_2^N$ increases because the direct effect of an increase in length of jurisdiction 1 (simultaneously shrinking 2's size) is to raise the tax base in 1 and decrease tax base in 2. Further, jurisdiction 3 shrinks in size relative to jurisdiction 1 but increases in relative size compared to jurisdiction 2, and tax rates follow these relative patterns. We further prove in the Appendix that this lemma will hold for the perturbed game $G(\varepsilon, l_{12}, l_{23})$ as well.

Under the new perturbed distribution corresponding to the specific perturbation $\tilde{\varepsilon}$, we have that the population of jurisdiction 1 is larger than that of jurisdiction 2, but the Nash tax rates of the two jurisdictions are unchanged, and thus the equality in (15) still holds. Thus, in the final step of the proof, we move l_{12} to a slightly lower level, $\tilde{l}_{12}$. Then by Lemma 2, we obtain $T_1^N(\tilde{\varepsilon}, \tilde{l}_{12}, l_{23}) < T_2^N(\tilde{\varepsilon}, \tilde{l}_{12}, l_{23})$ while $P_1(\tilde{\varepsilon}, \tilde{l}_{12}, l_{23}) > P_2(\tilde{\varepsilon}, \tilde{l}_{12}, l_{23})$, overturning the classic result.

Alternatively, we could have started from $P_1 = P_2 > P_3$, moved some population from jurisdiction 1 to jurisdiction 3 without changing taxes, and then increased l_{12} slightly. This would yield $T_1^N(\tilde{\varepsilon}, \tilde{l}_{12}, l_{23}) > T_2^N(\tilde{\varepsilon}, \tilde{l}_{12}, l_{23})$ while $P_1(\tilde{\varepsilon}, \tilde{l}_{12}, l_{23}) < P_2(\tilde{\varepsilon}, \tilde{l}_{12}, l_{23})$. We can state:

Proposition 4. (Three Jurisdictions & Tax Rates) *Suppose Assumptions 1 and 2 hold. There exist $\tilde{\varepsilon}$, $\tilde{l}_{12}$, and l_{23} and a population distribution $f(x, \tilde{\varepsilon})$ such that in the game $G(\tilde{\varepsilon}, \tilde{l}_{12}, l_{23})$, the populations of the jurisdictions and the Nash equilibrium are such that $P_1(\tilde{\varepsilon}, \tilde{l}_{12}, l_{23}) > P_2(\tilde{\varepsilon}, \tilde{l}_{12}, l_{23})$ and $T_1^N(\tilde{\varepsilon}, \tilde{l}_{12}, l_{23}) < T_2^N(\tilde{\varepsilon}, \tilde{l}_{12}, l_{23})$. Alternatively, we can find a population distribution with $P_1(\tilde{\varepsilon}, \tilde{l}_{12}, l_{23}) < P_2(\tilde{\varepsilon}, \tilde{l}_{12}, l_{23})$*

and $T_1^N(\tilde{\epsilon}, \tilde{l}_{12}, l_{23}) > T_2^N(\tilde{\epsilon}, \tilde{l}_{12}, l_{23})$. *That is, a smaller jurisdiction can set a higher tax rate than the next largest jurisdiction.*

Proof. See Appendix A.5. □

The reason why Proposition 2 can be overturned is based on the fact that with three jurisdictions, a jurisdiction can attract cross-border shoppers from two—instead of one jurisdiction—at different magnitudes and sensitivities. Specifically, for jurisdiction 1 to set a higher tax rate than jurisdiction 2 it must be that:

$$T_1^N > T_2^N \iff \frac{P_1 - CBS_{12}^N - CBS_{31}^N}{\left[f\left(x_{31}^N\right) + f\left(x_{12}^N\right)\right]/\delta} > \frac{P_2 + CBS_{12}^N - CBS_{23}^N}{\left[f\left(x_{12}^N\right) + f\left(x_{23}^N\right)\right]/\delta}.$$

If the tax-base sensitivities evaluated at the Nash equilibrium, $|\partial B_i/\partial T_i|$, are the same for all jurisdictions—as is the case under a uniform distribution—a jurisdiction would set a higher tax rate if and only if it was the larger jurisdiction (for the same reasons as in Proposition 2). To see this, if $T_1^N > T_2^N$ we have $CBS_{12}^N > 0$ and the only way the classic result can be overturned, i.e., when $P_2 > P_1$, is if $CBS_{23}^N - CBS_{31}^N > 0$. In the case of the uniform distribution these magnitudes are simply proportional to the cutoff rule's tax differential, and it can easily be seen that $CBS_{31} < 0$ requires $T_1^N < T_3^N$ while $CBS_{23} > 0$ requires $T_2^N > T_3^N$, which contradicts $T_1^N > T_2^N$. Similarly, if $CBS_{23} > 0$, but $CBS_{31} > 0$, for $T_1^N > T_2^N$ to arise it must be that $CBS_{31} > CBS_{23}$, which contradicts that $CBS_{23}^N - CBS_{31}^N > 0$. When density is non-uniform, the tax rate in jurisdiction 1 can indeed be larger than in jurisdiction 2 if, evaluated at the Nash tax rate, $|\partial B_2/\partial T_2|$ is sufficiently larger than $|\partial B_1/\partial T_1|$ in order to compensate for the fact that $P_1 < P_2$ and thus that the tax base in 2 may be larger. To see this, suppose that the numerators of the prior conditions are approximately equal. Then, the relative relationship between $f\left(x_{31}^N\right)$ and $f\left(x_{23}^N\right)$ determines the relationship above.

The prior results in Proposition 4 focused, without loss of generality, on comparing jurisdictions 1 and 2 (a smaller jurisdiction with a larger jurisdiction). But the results can easily be extended to other pairwise comparisons of tax rates (including comparing the smallest jurisdiction with the largest). We can compare jurisdiction 1 and 2, 1 and 3, and 2 and 3. This can easily be seen by example.

Example 2. (Triangular Distribution and Asymmetric Jurisdictions) In this example, we revisit Example 1. Let $\hat{l}$ be a value that satisfies $\sqrt{73}/16 - 3/16 < l < \sqrt{57}/12 - 1/4$. Then, $l_{12} = \hat{l}$ and $l_{23} = 1 - \hat{l}$. Given the results of Example 1, we already know that at such a value of $\hat{l}$, $P_3 = P_1 < P_2$ and $T_1^N = T_3^N > T_2^N$. In this example, we fix $l_{12} = \hat{l}$ and increase l_{23} from $1 - \hat{l}$ to $1 - \hat{l} + \varepsilon$, with ε very small, thus breaking the symmetry of the jurisdictions. Then, the equalities in Example 1 can be replaced with inequalities such that pairwise comparisons of two jurisdictions, can yield the smallest setting a rate higher than the largest.

We state this result as:

Proposition 5. (Smallest versus Largest Jurisdiction) *Suppose Assumptions 1 and 2 hold. There exists distribution functions such that, in the Nash equilibrium, the smallest jurisdiction can set a tax rate that is higher than the largest jurisdiction, that is, where* $P_3 < P_1 < P_2$ *and* $T_3^N > T_2^N$.

Proof. To show there exist such a distribution, start with the scenario in the example. The ε increase in l_{23} raises the population of jurisdiction 2 at the expense of 3. By Lemma 2 we know the effect of this on relative tax differentials. Thus, when $l_{12} = \hat{l}$ and $l_{23} = 1 - \hat{l} + \varepsilon$, we have:

$$P_3 < P_1 < P_2 \tag{17}$$

and

$$T_1^N > T_3^N > T_2^N. \tag{18}$$

□

Armed with these additional pairwise comparisons, we next pursue the question of whether it is possible to have a full reversal of tax rates where populations and taxes have the opposite rank orderings for all three jurisdictions. Without loss of generality, we focus on the case where $P_1 < P_3 < P_2$. Using a similar strategy of shifting populations with taxes unchanged and changing border lengths, as in Proposition 4, we can show:

Proposition 6. (Three Jurisdictions & Complete Ordering of Tax Rates) *Suppose Assumptions 1 and 2 hold. There exist a distribution function and values of* l_{12} *and* l_{23} *such that* $T_2^N < T_3^N < T_1^N$ *despite* $P_2 > P_3 > P_1$.

Proof. Start with $P_2 > P_3 = P_1$, but $T_3^N = T_1^N > T_2^N$, which we know is possible by Example 1. As in the proof of Proposition 4, we introduce a specific population move ε from jurisdiction 2 to jurisdiction 3, while maintaining the old equilibrium. This will give a new ordering $P_2 > P_3 > P_1$ with the same equilibrium $T_3^N = T_1^N > T_2^N$. Finally, imposing Assumption 2 on our population distribution, we apply the second set of inequalities in Lemma 2 and increase l_{23} slightly to maintain $P_2 > P_3 > P_1$ and obtain the final equilibrium rates $T_2^N < T_3^N < T_1^N$. □

Proposition 6 not only shows that a smaller jurisdiction can set the higher tax rate, but that the full ranking of tax rates follows the reverse order of populations. Proposition 6 follows from the pairwise comparisons of tax rates between jurisdictions 1 and 2 and between jurisdictions 2 and 3 following Example 2 and Proposition 5. We conclude that with multiple competitors and a general population distribution, many different parameter values can yield interjurisdictional tax differentials where larger jurisdictions set lower rates than smaller jurisdictions.

3 Broadening the Model to Other Settings

While the focus of our previous analysis was on commodity taxation, this section highlights that the setup—and main message of our paper—can be easily applied to corporate or capital taxation. Size matters in these contexts as well.[13] Although models of corporate/capital tax competition do not generally rely on Hotelling-style models used in the commodity tax competition literature, the spatial dimension can be reinterpreted in terms of profit shifting or capital mobility.

3.1 Profit Shifting

Keen and Konrad (2013) have already shown that the spatial commodity tax framework can be used to study international profit shifting by redefining the travel costs that individuals incur to cross-border shop as profit shifting costs.[14]

[13] Wilson (1991) and Bucovetsky (1991) find that in a two-jurisdiction economy, the small jurisdiction is better off than the larger jurisdiction under tax competition, because its low tax rate is increasing its tax base at the expense of the large jurisdiction.

[14] Other extensions of the spatial tax competition framework to profit shifting include Agrawal and Wildasin (2019) and Hebous and Keen (2023).

More specifically, they consider a representative multinational enterprise that earns fixed profit $\Pi_i, i = 1,2$ in each jurisdiction in the absence of profit shifting. The multinational enterprise can shift an amount $x_{ij}\left(T_i, T_j\right)$ between countries i and j in order to minimize its overall tax payments. The extent of profit shifting depends on the difference in tax rates between the two jurisdictions. For notational convenience, we drop the arguments in $x_{ij}\left(T_i, T_j\right)$ and use x_{ij}. Note $x_{ij} > 0$ if $T_i > T_j$ and $x_{ij} \leq 0$ if $T_i \leq T_j$. However, like cross-border shopping, profit shifting is costly, which Keen and Konrad (2013) assume to be of the quadratic form $C\left(x_{ij}\right) = \delta x_{ij}^2/2$. The assumption of quadratic shifting costs imply an optimal amount of shifted profits similar to the cut-off rule for cross-border shopping given by (2).[15]

To align our commodity tax framework with the profit shifting model, we have to extend the Keen and Konrad (2013) adaptation to include a third jurisdiction to which the multinational firm can shift profits. This adds an element of realism to the model as firms generally have multiple subsidiaries located across different countries. To ensure that the multinational firm does not shift profits to only one jurisdiction, we assume shifting costs are bilateral, that is, $C_{ij}\left(x_{ij}\right)$.[16] The assumption of bilateral costs maps to placing jurisdictions along a circle.

Next, to map the commodity tax model to profit shifting, we show how the population distribution function relates to the shifting cost function. Like density, the shape of the shifting cost function has important implications for tax competition. Analogous to the assumption of uniform density, the prior literature has assumed the cost function is quadratic. We allow for a more general form where shifting costs are strictly convex i.e., $\text{sign}\left(\partial C_{ij}/\partial x_{ij}\right) = \text{sign}\left(x_{ij}\right)$, $\text{sign}\left(\partial^m C_{ij}/\partial x_{ij}^m\right) = \text{sign}\left(\left|x_{ij}\right|\right)$ and $\partial^n C_{ij}/\partial x_{ij}^n > 0$, $1 < m < n$, where n denotes the highest-order derivative. This says that the sign of the first derivative is either positive or negative depending on the tax differential, the signs of all higher-order derivatives are non-negative, and the sign of the highest-order

[15] The reason is that the marginal benefits/costs of cross-border activity becomes the same in the two models. In the commodity tax model, a consumer living z units away from the border will purchase abroad if the cost of traveling δz is smaller than the tax savings given by the the tax rate differential $T_i - T_j$. In the profit-shifting model, a company will shift additional units of profit as long as the marginal cost of shifting $\partial C(x)/\partial x$ is smaller than the tax savings given by the tax rate differential $T_i - T_j$. If shifting costs are quadratic then $\partial C(x)/\partial x = \delta x$, which resembles the cost of traveling in the commodity-tax framework.

[16] See Huizinga et al. (2008) and van't Riet and Lejour (2018) for empirical evidence showing that multinational firms base their decisions on bilateral costs.

derivative is strictly positive. For simplicity, assume $n = 3$. Relaxing the assumption about the third derivative of the shifting costs is the linking element between the profit shifting model and our commodity tax model. The multinational's after-tax profits are:

$$(1-T_1)(\Pi_1 - x_{12} - x_{13}) + (1-T_2)(\Pi_2 + x_{12} - x_{23}) + (1-T_3)(\Pi_3 + x_{13} + x_{23}) - C_{12} - C_{13} - C_{23},$$

which yields the following optimal levels of profit shifting:

$$\frac{\partial C_{ij}}{\partial x_{ij}} = T_i - T_j, \qquad j = \{2,3\}, i = \{1,2\}, j \neq i. \tag{19}$$

Denoting $x^*_{ij}(T_i, T_j)$ as the optimal levels of shifting implied by (19), we derive the sensitivity of profit shifting:

$$\frac{\partial x^*_{ij}}{\partial T_i} = -\frac{\partial x^*_{ij}}{\partial T_j} = \frac{1}{\frac{\partial^2 C_{ij}}{\partial x^2_{ij}}}, \qquad j = \{2,3\}, i = \{1,2\}, j \neq i.$$

The tax-sensitivity of profit shifting depends on the magnitude of $\partial^2 C_{ij}/\partial x^2_{ij}$ and thus ultimately on the shape of the shifting cost function. Based on the multinational firm's trade-offs, we can formulate the jurisdictions' tax revenues as

$$\begin{aligned} R_1 &= T_1\left[\Pi_1 - x^*_{12} - x^*_{13}\right], \\ R_2 &= T_2\left[\Pi_2 + x^*_{12} - x^*_{23}\right], \\ R_3 &= T_3\left[\Pi_3 + x^*_{13} + x^*_{23}\right]. \end{aligned}$$

Differentiating R_i with respect to T_i implicitly determines the optimal taxes, where we let $x^*_{ij}(T^N_i, T^N_j) \equiv x^N_{ij}$ and $\partial^2 C_{ij}/\partial x^2_{ij} \equiv c''_{ij}$:

$$\begin{aligned} T^N_1 &= -\frac{\Pi_1 - x^N_{12} - x^N_{13}}{-\frac{\partial x_{12}}{\partial T_1}\left(T^N_1, T^N_2\right) - \frac{\partial x_{13}}{\partial T_1}\left(T^N_1, T^N_3\right)} = \frac{\Pi_1 - x^N_{12} - x^N_{13}}{\left(c''_{12}\left(T^N_1, T^N_2\right)\right)^{-1} + \left(c''_{13}\left(T^N_1, T^N_3\right)\right)^{-1}}, \\ T^N_2 &= -\frac{\Pi_2 + x^N_{12} - x^N_{23}}{\frac{\partial x_{12}}{\partial T_2}\left(T^N_1, T^N_2\right) - \frac{\partial x_{23}}{\partial T_2}\left(T^N_2, T^N_3\right)} = \frac{\Pi_2 + x^N_{12} - x^N_{23}}{\left(c''_{12}\left(T^N_1, T^N_2\right)\right)^{-1} + \left(c''_{23}\left(T^N_2, T^N_3\right)\right)^{-1}}, \\ T^N_3 &= -\frac{\Pi_3 + x^N_{13} + x^N_{23}}{\frac{\partial x_{13}}{\partial T_3}\left(T^N_1, T^N_3\right) + \frac{\partial x_{23}}{\partial T_3}\left(T^N_2, T^N_3\right)} = \frac{\Pi_3 + x^N_{13} + x^N_{23}}{\left(c''_{13}\left(T^N_1, T^N_3\right)\right)^{-1} + \left(c''_{23}\left(T^N_2, T^N_3\right)\right)^{-1}}. \end{aligned}$$

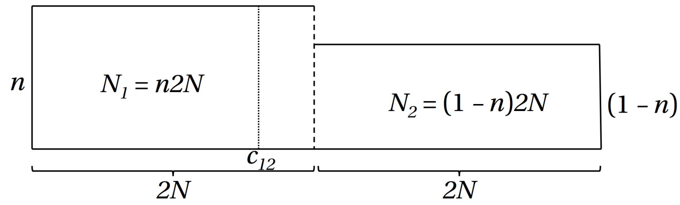
The optimal corporate tax rates have the same structure as the equations (10)-(12), where the exogenous firm profits Π_i play the same role as the exogenous population size P_i. Profit shifting, x_{12}^N, x_{13}^N and x_{23}^N, is analogous to cross-border shopping, CBS_{12}^N, CBS_{31}^N and CBS_{23}^N. The functional form of the shifting costs C_{ij} affects the tax-base sensitivity in a similar fashion as the functional form of the population distribution $f(x)$, where $f\left(x_{12}^N\right)$, $f\left(x_{23}^N\right)$ and $f\left(x_{31}^N\right)$ in the commodity tax model matches $\left(c_{12}''\left(T_1^N, T_2^N\right)\right)^{-1}$, $\left(c_{23}''\left(T_2^N, T_3^N\right)\right)^{-1}$, $\left(c_{13}''\left(T_1^N, T_3^N\right)\right)^{-1}$ in the profit shifting model. If the cost function is quadratic, these terms are all equal, yielding the standard result that bigger jurisdictions set higher rates. For ease of exposition, we focus on illustrating how our result in Proposition 3 can be derived in the profit shifting model. Note that we can define the allocation of real profits across countries as $\Pi_1 = l_{12}\Pi$, $\Pi_2 = (l_{23} - l_{12})\Pi$, and $\Pi_3 = (1 - l_{23})\Pi$, where Π denotes worldwide profits and l_{12}, $(l_{23} - l_{12})$ and $(1 - l_{23})$ are the exogenous shares of multinational's production in each country, with $0 < l_{12} < l_{23} < 1$, where due to symmetry $l_{12} = l$ and $l_{23} = 1 - l$. Then it becomes clear that if $\Pi_1 < \Pi_2$ there exist cost functions such that $T_1^N > T_2^N$ if and only if $\underline{l} < l < \bar{l}$.

3.2 Capital Mobility

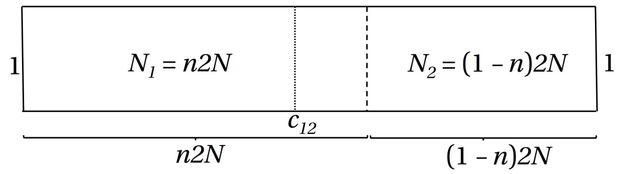
Next, we demonstrate similarities with models of capital mobility by drawing on Mongrain and Wilson (2018), which analyzes tax competition in a setting where firms face heterogeneous moving costs. More specifically, there are two jurisdictions $i = 1, 2$ and a mass of firms of $2N$. Jurisdictions may differ in their size, that is, jurisdiction 1 is assumed to be the larger jurisdiction with initially $N_1 = n2N$ firms, whereas jurisdiction 2 comprises $N_1 = (1 - n)2N$, where $n \in [1/2, 1)$. Firms generate exogenous profits $\gamma > 1$ and can relocate to the other jurisdiction, which results in idiosyncratic moving costs $\hat{c} \in [0, 1]$ distributed according to a cumulative distribution function $F(\hat{c})$, with density $f(\hat{c})$. Tax competition arises because firms are taxed depending on where they are located, that is, each jurisdiction levies a source-based tax T_i.

Figure 4a illustrates the situation of $T_1 > T_2$, which implies that some firms with moving costs $\hat{c} \leq c_{12} \equiv \gamma(T_1 - T_2)$ initially located in jurisdiction 1 relocate to jurisdiction 2, so that the tax bases in jurisdictions 1 and 2 are given by $\gamma N_1 [1 - F(c_{12})]$ and $\gamma N_2 + \gamma N_1 F(c_{12})$. The setup described in Figure 4a is

Figure 4: Eliminating Discontinuities in Capital Tax Models



(a) Model setup in Mongrain and Wilson (2018)



(b) Adjusted Model Setup without Discontinuities

Panel A shows the setup of Mongrain and Wilson (2018), including a discontinuity in the distribution. Panel B shows how we eliminate the discontinuity to simplify the problem, but still preserving all qualitative results of the model. For illustrative purposes, the Figure depicts a cut-off rule if $T_1 > T_2$.

identical to the setup in Kanbur and Keen (1993) and therefore implies that the reaction function of the small jurisdiction, here jurisdiction 2, features a discontinuity. This feature could considerably complicate our analysis.

For this reason, we modify the Mongrain and Wilson (2018) model to eliminate the discontinuity in the best response function without affecting the model's results in a qualitative manner. Our model circumvents the discontinuity by altering the length of the jurisdictions (instead of the height) to make one bigger than the other. After applying this adjustment to the Mongrain and Wilson (2018) model, the model setup can be summarized in the Figure 4b. Figure 4b illustrates the situation of $T_1 > T_2$, which implies that some firms with moving costs $\hat{c} \leq c_{12} \equiv \gamma(T_1 - T_2)$ initially located in jurisdiction 1 relocate to jurisdiction 2, with the difference that there is no discontinuity at the border and, in turn, in the small jurisdiction's tax reaction function. The tax bases in jurisdictions 1 and 2 are now given by $\gamma[N_1 - F(c_{12})]$ and $\gamma[N_2 + F(c_{12})]$.

Based on these adjustments, we extend the Mongrain and Wilson (2018) model to three jurisdictions. To align the capital-mobility model with the commodity-tax model, we need a few additional assumptions. First, following Janeba and Schulz (2023) and Fuest and Sultan (2019), there are three industries but each industry links only two countries, i.e., industry ij links countries i and j. Firms in industry ij can only locate in these two countries because firms cannot change industries. The idea that countries differ in industries is consistent with the Ricardian idea of specialization resulting from regulatory or technological differences across countries. Second, firms draw an industry-specific moving cost $\hat{c}_{ij} \in [0,1]$ from an industry-specific cumulative distribution function $F_{ij}\left(\hat{c}_{ij}\right)$ with density $f_{ij}\left(\hat{c}_{ij}\right)$. Firms located in jurisdiction i will move to jurisdiction j if $\hat{c}_{ij} \leq c_{ij} \equiv \gamma\left(T_i - T_j\right)$, where the cost without the "hat" denotes the optimal cutoff rule.

Based on these assumptions, we can formulate the jurisdictions' revenues:

$$\begin{aligned}
R_1 &= T_1\left[N_1 - F_{12}\left(c_{12}\right) - F_{13}\left(c_{13}\right)\right], \\
R_2 &= T_2\left[N_2 + F_{12}\left(c_{12}\right) - F_{23}\left(c_{23}\right)\right], \\
R_3 &= T_3\left[N_3 + F_{13}\left(c_{13}\right) + F_{23}\left(c_{23}\right)\right],
\end{aligned}$$

where $N_1 = l_{12}N$, $N_2 = \left(l_{23} - l_{12}\right)N$ and $N_3 = \left(1 - l_{23}\right)N$ with N denoting the total number of firms and l_{12}, $\left(l_{23} - l_{12}\right)$ and $\left(1 - l_{23}\right)$ denoting the exogenous shares of initial firms in each country $\left(0 < l_{12} < l_{23} < 1\right)$. Differentiating the revenue functions R_i implicitly determines the optimal taxes

$$\begin{aligned}
T_1^N &= \frac{N_1 - F_{12}\left(c_{12}^N\right) - F_{13}\left(c_{13}^N\right)}{\gamma\left[f_{12}\left(c_{12}^N\right) + f_{13}\left(c_{13}^N\right)\right]}, \\
T_2^N &= \frac{N_2 + F_{12}\left(c_{12}^N\right) - F_{23}\left(c_{23}^N\right)}{\gamma\left[f_{12}\left(c_{12}^N\right) + f_{23}\left(c_{23}^N\right)\right]}, \\
T_3^N &= \frac{N_3 + F_{13}\left(c_{13}^N\right) + F_{23}\left(c_{23}^N\right)}{\gamma\left[f_{13}\left(c_{13}^N\right) + f_{23}\left(c_{23}^N\right)\right]},
\end{aligned}$$

where c_{ij}^N are the optimal cutoff rules evaluated at Nash tax rates.

Again, the structure of the optimal tax rates is qualitatively the same as (10)-(12), where the initial number of firms located in a jurisdiction, N_i, plays the same role as the exogenous population size P_i. The number of firms moving, $F_{12}\left(c_{12}^N\right)$, $F_{13}\left(c_{13}^N\right)$ and $F_{23}\left(c_{23}^N\right)$ corresponds to the number of cross-border shoppers CBS_{12}^N, CBS_{31}^N and CBS_{23}^N, and the distribution of moving costs, $f_{ij}\left(c_{ij}^N\right)$, plays the same role as the distribution of population, $f\left(x_{ij}^N\right)$. If we assume symmetry in the same way as under Proposition 3, then it is clear that if $N_1 < N_2$ there exists moving-cost distributions such that $T_1^N > T_2^N$ if and only if $\underline{l} < l < \bar{l}$.

4 Conclusions

Declining mobility costs, technological change, and reductions in border controls pose substantial challenges to the design of tax policies in an open economy. Many standard *strategic* tax competition models assume duopolistic competition leading to the conventional view that bigger jurisdictions set higher tax rates. However, in reality, competition for mobile tax bases is usually not just a bilateral, but a multilateral matter. We show that allowing for oligopolistic competition can lead to fundamentally different outcomes in the tax competition game irrespective of whether jurisdictions compete for cross-border shoppers, capital, or profits. In the commodity tax setting, the shape of the distribution of residents is critical. In companion work, using data on the distribution of households within jurisdictions, we empirically show that increases in the density of marginal households near the border is negatively correlated with a jurisdiction's sales tax rate and negatively correlated with it being higher tax than its neighbors.

Our analysis focuses on a single-tax policy. However, an important feature of a tax system is that jurisdictions decide on multiple tax policies. In the case of commodity taxes, for example, governments may set different excise tax rates on products that may be complements or substitutes (Hoyt 2017).[17] We could extend our model to multiple excise taxes on different products (e.g., beer, wine, and spirits). To do so, we would need to relax the assumption of inelastic demand following Devereux et al. (2007), and then add a second commodity that influences demand for the first. We have focused on the case in which tax-base interdependencies play a subordinate role. In a quasi-linear model with two goods and two tax rates, if consumption of one commodity is independent of consumption of the other commodity—consumer utility is separable in the commodities—cross-price elasticities of demand are zero and our results would carry through. However, Scheuer and Werning (2016)

[17] More generally, as noted in Slemrod (2019) and Keen and Slemrod (2017), a tax system consists of more than tax bases and rates, with remittance rules, enforcement policies, and information exchange potentially influencing the elasticity of the tax base

note that the inverse elasticity rule may provide little guidance for policy when cross-price elasticities are not zero. Concerning the results of our model, this implies that the conventional view that bigger jurisdictions set higher tax rates may no longer hold even in the two-jurisdiction case. Future research might explore the role of such tax-base interdependencies more thoroughly in order to think about tax competition as it relates to the tax system—not just to a specific tax instrument in isolation.

Although our focus is on competition between governments, our framework shares important commonalities with industrial organization models that consider price competition with more than two firms (Aoyagi and Okabe 1991; Caplin and Nalebuff 1991 Chen and Riordan 2007; Zhou 2017; Tarbush 2018) or price competition in networks (Bloch and Querou 2013; Mossay and Picard 2011; Ushchev and Zenou 2018), "spatial" voting models, where voters differ in preferences (Wittman 1983), and the role of border-effects in trade (Anderson and van Wincoop 2003; Evans 2003) and in urban economics (Holmes 1998).

With reference to border-effects in urban economics, our model implies that the population distribution is critical for the elasticity of the tax base. While it is reasonable to believe that sales-tax differentials are not a major determinant of housholds' (residential) migration decisions, jurisdictions have alternative instruments that influence where people live. For example, land use and zoning regulations may allow jurisdictions to influence the distribution of firms and individuals and thus *choose* the elasticity of the tax base, as in Slemrod and Kopczuk (2002), to maximize tax revenues. Indeed, Jacob and McMillen (2015) document that commercial and industrial parcels are significantly more likely to be located near municipal boundaries, which reduces the likelihood of own-residents shopping in the neighboring jurisdiction and, at the same time, increases the likelihood of attracting neighboring cross-shoppers due to reduced travel times. Whether such a policy is desirable from a welfare perspective needs to be evaluated in a general equilibrium model that takes into account repercussions on, inter alia, the housing market—again highlighting the role of cross-price elasticities.

A Appendix: Proofs

A.1 Proof of Existence and Uniqueness (2 Jurisdictions)

The game is supermodular if the strategy set is compact and the payoff functions display strategic complementarity in the taxes (Rota-Graziosi 2019). The strategy set is the compact set $\left[0, \overline{T}\right]$.

The FOCs are given by (3) and (4). We can rewrite (3) as:

$$\frac{\partial R_1}{\partial T_1} = \underbrace{F(x_{12})\left[1 - \frac{T_1}{\delta}\frac{f(x_{12})}{F(x_{12})}\right]}_{A} - \underbrace{F(x_{21})\left[1 + \frac{T_1}{\delta}\frac{f(x_{21})}{F(x_{21})}\right]}_{B}. \tag{A.1}$$

Taking the derivative of term A with respect to T_2 yields

$$A' = \frac{f(x_{12})}{\delta}\left[1 - \frac{T_1}{\delta}\frac{f'(x_{12})}{f(x_{12})}\right]. \tag{A.2}$$

Then using our assumptions, we have:

$$\frac{f'(z)}{f(z)} \leq \frac{f'(0)}{f(0)} < \frac{\delta}{\overline{T}}, \ \forall\, z \in [0,1], \tag{A.3}$$

where the first inequality follows from the log-concavity of f and the second inequality follows from Assumption 2. Thus, A' in positive for all T_2. By the log-concavity of f and the definitions in (2), term A is strictly increasing in T_2, where the strict condition follows from the from the strict inequality in (A.3). As shown in Bagnoli and Bergstrom (2005), the log-concavity of f implies that F is also log-concave and therefore $f(x_{21})/F(x_{21})$ is weakly decreasing in x. By (2), x_{21} decreases in T_2 and therefore term B is a product of two positive weakly decreasing functions in T_2. Thus, combined with the fact that A strictly increases in T_2, we can conclude that the right hand side of (A.1) is strictly increasing in T_2. Therefore, $\partial^2 R_1/\partial T_2 \partial T_1 > 0$. Applying the same argument to R_2, we conclude that $\partial^2 R_2/\partial T_1 \partial T_2 > 0$. Hence, the game is supermodular and has an equilibrium (Topkis 1979).

Furthermore, we can make use of the dominant diagonal argument to prove uniqueness of the equilibrium (Vives 1999, page 47). We can compute

$$\Phi_1 \equiv \frac{\partial^2 R_1}{\partial T_1^2} + \frac{\partial^2 R_1}{\partial T_2 \partial T_1} = -\frac{f\left(x_{12}\right) + f(x_{21})}{\delta} < 0.$$

Similarly, we have

$$\Phi_2 \equiv \frac{\partial^2 R_2}{\partial T_2^2} + \frac{\partial^2 R_2}{\partial T_1 \partial T_2} < 0.$$

Thus, the equilibrium in unique.[18]

A.2 Proof of Existence and Uniqueness (3 jurisdictions)

Focusing on jurisdiction 1, differentiating the revenue functions yields:

$$\frac{\partial R_1}{\partial T_1} = \underbrace{F(x_{12})\left[1 - \frac{T_1}{\delta}\frac{f(x_{12})}{F(x_{12})}\right]}_{A} - \underbrace{F(x_{31})\left[1 + \frac{T_1}{\delta}\frac{f(x_{31})}{F(x_{31})}\right]}_{C}. \tag{A.4}$$

Term A is identical to that in (A.1) and thus we can repeat the argument in Appendix A.1 to show it is strictly increasing in T_2. This proves the claim for T_2 as x_{31} is unaffected by T_2. Similarly, we know that x_{31} *decreases* in T_3,which means that term C decreases in T_3 because the log-concavity of f in Assumption 1 implies that F is also log-concave. Hence, the negative second term of jurisdiction 1's first-order condition becomes less negative as T_3 increases, which proves the claim for T_3 as x_{12} is unaffected by T_3. Thus $\partial^2 R_1/\partial T_1 \partial T_2 > 0$ and $\partial^2 R_1/\partial T_1 \partial T_3 \geq 0$. Applying this logic to all jurisdictions implies that the game is supermodular under Assumption 1.

With respect to uniqueness, for $i \in \{1,2,3\}$, let

$$\Phi_i \equiv \frac{\partial^2 R_i}{\partial T_i^2} + \sum_{j \neq i} \frac{\partial^2 R_i}{\partial T_j \partial T_i}.$$

[18] The definition of Φ_1, the fact that $\Phi_1 < 0$, and the supermodularity of the game imply that the second partial of R_1 with respect to T_1 is strictly negative and R_1 is strictly concave in T_1. Therefore, the best response is single-valued. The same is true for the second jurisdiction.

We compute the elements of the Hessian matrix:

$$
\begin{aligned}
\gamma_1 = \frac{\partial^2 R_1}{\partial T_1^2} &= -\frac{1}{\delta}\left[f(x_{12}) + f(x_{31})\right] - (\gamma_2 + \gamma_3) < 0, \\
\gamma_2 = \frac{\partial^2 R_1}{\partial T_2 \partial T_1} &= \frac{1}{\delta}\left[f(x_{12}) - \frac{T_1}{\delta} f'(x_{12})\right] > 0, \\
\gamma_3 = \frac{\partial^2 R_1}{\partial T_3 \partial T_1} &= \frac{1}{\delta}\left[f(x_{31}) + \frac{T_1}{\delta} f'(x_{31})\right] \geq 0, \\
\gamma_4 = \frac{\partial^2 R_2}{\partial T_1 \partial T_2} &= \frac{1}{\delta}\left[f(x_{12}) + \frac{T_2}{\delta} f'(x_{12})\right] \geq 0, \\
\gamma_5 = \frac{\partial^2 R_2}{\partial T_2^2} &= -\frac{1}{\delta}\left[f(x_{12}) + f(x_{23})\right] - (\gamma_4 + \gamma_6) < 0, \\
\gamma_6 = \frac{\partial^2 R_2}{\partial T_3 \partial T_2} &= \frac{1}{\delta}\left[f(x_{23}) - \frac{T_2}{\delta} f'(x_{23})\right] > 0, \\
\gamma_7 = \frac{\partial^2 R_3}{\partial T_1 \partial T_3} &= \frac{1}{\delta}\left[f(x_{31}) - \frac{T_3}{\delta} f'(x_{31})\right] > 0, \\
\gamma_8 = \frac{\partial^2 R_3}{\partial T_2 \partial T_3} &= \frac{1}{\delta}\left[f(x_{23}) + \frac{T_3}{\delta} f'(x_{23})\right] \geq 0, \\
\gamma_9 = \frac{\partial^2 R_3}{\partial T_3^2} &= -\frac{1}{\delta}\left[f(x_{23}) + f(x_{31})\right] - (\gamma_7 + \gamma_8) < 0.
\end{aligned}
$$

The indicated signs of these γ's follows from the supermodularity of the payoffs functions. It can then be immediately seen that under assumption 1 and 2, $\Phi_i < 0$ for all jurisdictions, which proves uniqueness (Vives 1999).

A.3 Proof of Lemma 1

Let $S = (\alpha, \beta)$. Let ζ be the maximum value of ϕ on $[0, \alpha]$, and let ξ be the maximum value of ϕ on $\left[\beta, \overline{T}\right]$. Since z^* is the unique maximizer of ϕ, we can find $\tilde{\varepsilon}$ to be such that $\phi(z^*) > max\{\zeta + \tilde{\varepsilon}, \xi + \tilde{\varepsilon}\}$. Assumption (ii) and the definition of $\tilde{\varepsilon}$ imply that, for all $\varepsilon < \tilde{\varepsilon}$, the maximum of $\phi(\cdot, \varepsilon)$ over $[0, \alpha] \bigcup [\beta, \overline{z}]$ is strictly less than $\phi(z^*)$, which is equal to $\phi(z^*, \varepsilon)$ by assumption (i). Moreover, assumption (i) implies that, for all ε, $\phi(z^*)$ is the maximizer of $\phi(\cdot, \varepsilon)$ over (α, β). Therefore, for all $\varepsilon < \tilde{\varepsilon}$, we have $\phi(z^*, \varepsilon) > \phi(z, \varepsilon)$, for all $z \in \left[0, \overline{T}\right]$ that is different fromz^*. Hence, z^* is the unique maximizer of $\phi(\cdot, \varepsilon)$ over $\left[0, \overline{T}\right]$.

A.4 Proof of Lemma 2

Multiplying the first-order conditions of the game by delta and totally differentiating yields the following system of equations:

$$\begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \alpha_4 & \alpha_5 & \alpha_6 \\ \alpha_7 & \alpha_8 & \alpha_9 \end{bmatrix} \times \begin{bmatrix} dT_1 \\ dT_2 \\ dT_3 \end{bmatrix} = \begin{bmatrix} -\delta\alpha_2 \\ \delta\alpha_4 \\ 0 \end{bmatrix} dl_{12} + \begin{bmatrix} 0 \\ -\delta\alpha_6 \\ \delta\alpha_8 \end{bmatrix} dl_{23} + \begin{bmatrix} \delta\alpha_3 \\ 0 \\ -\delta\alpha_7 \end{bmatrix} dl_{31},$$

where $\alpha_i = \delta\gamma_i$. Using Cramer's rule, we can derive the effect of a change in l_{12} on the equilibrium tax rates as follows:

$$\begin{aligned} \frac{dT_1^N}{dl_{12}} &= -\frac{\alpha_2(\alpha_5\alpha_9 - \alpha_6\alpha_8) + \alpha_4(\alpha_2\alpha_9 - \alpha_3\alpha_8)}{\delta^2|\Gamma|}, \\ \frac{dT_2^N}{dl_{12}} &= \frac{\alpha_4(\alpha_1\alpha_9 - \alpha_3\alpha_7) + \alpha_2(\alpha_4\alpha_9 - \alpha_6\alpha_7)}{\delta^2|\Gamma|}, \\ \frac{dT_3^N}{dl_{12}} &= -\frac{\alpha_4(\alpha_1\alpha_8 - \alpha_2\alpha_7) + \alpha_2(\alpha_4\alpha_8 - \alpha_5\alpha_7)}{\delta^2|\Gamma|}, \end{aligned}$$

and the effect of l_{23} as

$$\begin{aligned} \frac{dT_1^N}{dl_{23}} &= \frac{\alpha_6(\alpha_2\alpha_9 - \alpha_3\alpha_8) + \alpha_8(\alpha_2\alpha_6 - \alpha_3\alpha_5)}{\delta^2|\Gamma|}, \\ \frac{dT_2^N}{dl_{23}} &= -\frac{\alpha_6(\alpha_1\alpha_9 - \alpha_3\alpha_7) + \alpha_8(\alpha_1\alpha_6 - \alpha_3\alpha_4)}{\delta^2|\Gamma|}, \\ \frac{dT_3^N}{dl_{23}} &= \frac{\alpha_6(\alpha_1\alpha_8 - \alpha_2\alpha_7) + \alpha_8(\alpha_1\alpha_5 - \alpha_2\alpha_4)}{\delta^2|\Gamma|}, \end{aligned}$$

where $|\Gamma| < 0$ to obtain a maximum. By Assumption 2, we have $\gamma_2, \gamma_6, \gamma_7 > 0$, that is, a strict inequality. Therefore, we also have $\alpha_2, \alpha_6, \alpha_7 > 0$. Moreover, we can derive the following relationships:

$$\begin{aligned} \frac{dT_1^N}{dl_{12}} - \frac{dT_2^N}{dl_{12}} = &\frac{[f(x_{12}) + f(x_{23})]\alpha_2\alpha_9 + [f(x_{12}) + f(x_{31})]\alpha_4\alpha_9}{\delta^2|\Gamma|} \\ &- \frac{[f(x_{23}) + f(x_{31})](\alpha_2\alpha_6 + \alpha_3\alpha_4)}{\delta^2|\Gamma|} \equiv \frac{\Omega_{l_{12}}^{12}}{\delta^2|\Gamma|} > 0, \end{aligned}$$

$$\frac{dT_1^N}{dl_{12}} - \frac{dT_3^N}{dl_{12}} = \frac{\left[f\left(x_{23}\right)+f\left(x_{31}\right)\right]\alpha_2\left(\alpha_4+\alpha_5\right)}{\delta^2\left|\Gamma\right|} - \frac{\left[f\left(x_{12}\right)+f\left(x_{31}\right)\right]\alpha_4\alpha_8+\left[f\left(x_{12}\right)+f\left(x_{23}\right)\right]\alpha_2\alpha_8}{\delta^2\left|\Gamma\right|} \equiv \frac{\Omega_{l_{12}}^{13}}{\delta^2\left|\Gamma\right|} > 0,$$

$$\frac{dT_2^N}{dl_{12}} - \frac{dT_3^N}{dl_{12}} = \frac{\left[f\left(x_{12}\right)+f\left(x_{23}\right)\right]\alpha_2\alpha_7+\left[f\left(x_{23}\right)+f\left(x_{31}\right)\right]\alpha_3\alpha_4}{\delta^2\left|\Gamma\right|} + \frac{\left[f\left(x_{12}\right)+f\left(x_{31}\right)\right]\left[f\left(x_{23}\right)+f\left(x_{31}\right)\right]\alpha_4+\left[f\left(x_{12}\right)+f\left(x_{31}\right)\right]\alpha_4\alpha_7}{\delta^2\left|\Gamma\right|} \equiv \frac{\Omega_{l_{12}}^{23}}{\delta^2\left|\Gamma\right|} < 0, \tag{A.5}$$

and

$$\begin{aligned}
\frac{dT_1^N}{dl_{23}} - \frac{dT_2^N}{dl_{23}} &= \frac{\alpha_6\left[\left[f\left(x_{12}\right)+f\left(x_{31}\right)\right]\alpha_7-\left(\alpha_1+\alpha_2\right)\left[f\left(x_{23}\right)+f\left(x_{31}\right)\right]\right]}{\delta^2\left|\Gamma\right|} \\
&+ \frac{\alpha_3\alpha_8\left[f\left(x_{12}\right)+f\left(x_{23}\right)\right]}{\delta^2\left|\Gamma\right|} \equiv \frac{\Omega_{l_{23}}^{12}}{\delta^2\left|\Gamma\right|} < 0, \\
\frac{dT_1^N}{dl_{23}} - \frac{dT_3^N}{dl_{23}} &= \frac{\left[f\left(x_{12}\right)+f\left(x_{31}\right)\right]\left(\alpha_5+\alpha_6\right)\alpha_8-\alpha_2\alpha_8\left[f\left(x_{12}\right)+f\left(x_{23}\right)\right]}{\delta^2\left|\Gamma\right|} \\
&- \frac{\alpha_2\alpha_6\left[f\left(x_{23}\right)+f\left(x_{31}\right)\right]}{\delta^2\left|\Gamma\right|} \equiv \frac{\Omega_{l_{23}}^{13}}{\delta^2\left|\Gamma\right|} > 0, \\
\frac{dT_2^N}{dl_{23}} - \frac{dT_3^N}{dl_{23}} &= \frac{\alpha_1\alpha_6\left[f\left(x_{23}\right)+f\left(x_{31}\right)\right]-\left[f\left(x_{12}\right)+f\left(x_{31}\right)\right]\left(\alpha_6\alpha_7+\alpha_4\alpha_8\right)}{\delta^2\left|\Gamma\right|} \\
&+ \frac{\alpha_1\alpha_8\left[f\left(x_{12}\right)+f\left(x_{23}\right)\right]}{\delta^2\left|\Gamma\right|} \equiv \frac{\Omega_{l_{23}}^{23}}{\delta^2\left|\Gamma\right|} > 0.
\end{aligned} \tag{A.6}$$

A.5 Proof of Proposition 4

A.5.1 Formalities and Outline

Given any vectors $\mathbf{T} = (T_1, T_2, T_3)$ and $\mathbf{l} = (l_{12}, l_{23})$, we define the following quantities representing the marginal shoppers by $x_{12}(T_1, T_2, l_{12}) = l_{12} - (T_1 - T_2)/\delta$, $x_{23}(T_2, T_3, l_{23}) = l_{23} - (T_2 - T_3)/\delta$, and $x_{31}(T_3, T_1) = l_{31} - (T_3 - T_1)/\delta$, where l_{31} is normalized to zero. Under perturbation ε, we also denote the (parameterized) payoff functions:

$$\begin{aligned}
R_1(T_1, \mathbf{T_{-1}}, \varepsilon, l_{12}) &= T_1\left[F(x_{12}(T_1, T_2, l_{12}), \varepsilon) - F(x_{31}(T_3, T_1), \varepsilon)\right], \\
R_2(T_2, \mathbf{T_{-2}}, \varepsilon, l_{12}, l_{23}) &= T_2\left[F(x_{23}(T_2, T_3, l_{23}), \varepsilon) - F(x_{12}(T_1, T_2, l_{12}), \varepsilon)\right], \\
R_3(T_3, \mathbf{T_{-3}}, \varepsilon, l_{23}) &= T_3\left[1 - F(x_{23}(T_2, T_3, l_{23}), \varepsilon) + F(x_{31}(T_3, T_1), \varepsilon)\right],
\end{aligned}$$

where $\mathbf{T_{-i}}$ denote the actions of players other than player i (e.g. $\mathbf{T_{-1}} = (T_2, T_3)$).

When $\varepsilon = 0$ (i.e. when we are considering a game with the original—unperturbed—pdf and the only parameters that can take different values are l_{12} and l_{23}), we simply drop ε from the notation. We write $G(l_{12}, l_{23})$ to denote the game. We write $R_1(T_1, \mathbf{T_{-1}}, l_{12})$, $R_2(T_2, \mathbf{T_{-2}}, l_{12}, l_{23})$, $R_3(T_3, \mathbf{T_{-3}}, l_{23})$ for the payoffs in $G(l_{12}, l_{23})$. We also write $P_1(l_{12})$, $P_2(l_{12}, l_{23})$, and $P_3(l_{23})$; $T_1^N(l_{12}, l_{23})$, $T_2^N(l_{12}, l_{23})$, and $T_3^N(l_{12}, l_{23})$ to respectively denote the population sizes and equilibrium of $G(l_{12}, l_{23})$. We proceed as follows:

Step 1: Given the population distribution f, we can find values l_{12}, l_{23} and a perturbation size $\tilde{\varepsilon}$ such that, in the perturbed game $G(\tilde{\varepsilon}, l_{12}, l_{23})$, we have $P_1(\tilde{\varepsilon}, l_{12}) > P_2(\tilde{\varepsilon}, l_{12}, l_{23})$ and $T_1^N(\tilde{\varepsilon}, l_{12}, l_{23}) = T_2^N(\tilde{\varepsilon}, l_{12}, l_{23})$.

Step 2: We move the border l_{12} to a slightly lower level $\tilde{l}_{12}$ so that in the new game $G(\tilde{\varepsilon}, \tilde{l}_{12}, l_{23})$, we have $P_1(\tilde{\varepsilon}, \tilde{l}_{12}) > P_2(\tilde{\varepsilon}, \tilde{l}_{12}, l_{23})$ and $T_1^N(\tilde{\varepsilon}, \tilde{l}_{12}, l_{23}) < T_2^N(\tilde{\varepsilon}, \tilde{l}_{12}, l_{23})$.

A.5.2 Step 1

We start by assuming, without loss of generality, that the parameters l_{12} and l_{23} are such that $P_1(l_{12}) = P_2(l_{12}, l_{23}) > P_3(l_{23})$.[19] As shown in Appendix A.2, a unique equilibrium exists for the game $G(l_{12}, l_{23})$.

Case 1: If $T_3^N(l_{12}, l_{23}) > \min\{T_1^N(l_{12}, l_{23}), T_2^N(l_{12}, l_{23})\}$, there is nothing to prove.

Case 2: Similarly, if $T_1^N(l_{12}, l_{23}) < T_2^N(l_{12}, l_{23})$, we can slightly increase l_{12} to a new value l_{12} so that we have $P_1(l_{12}) > P_2(l_{12}, l_{23})$ while the continuity of T_1^N and T_2^N with respect to l_{12} implies that we still have $T_1^N(l_{12}, l_{23}) < T_2^N(l_{12}, l_{23})$. A similar argument can be used if $T_1^N(l_{12}, l_{23}) > T_2^N(l_{12}, l_{23})$.

Case 3: Therefore, the only case we need to consider in our proof is when the equilibrium of the game $G(l_{12}, l_{23})$ is

$$T_3^N(l_{12}, l_{23}) < T_1^N(l_{12}, l_{23}) = T_2^N(l_{12}, l_{23}). \tag{A.7}$$

To to deal with this case, we can find a very specific (small) population redistribution of size ε from jurisdiction 3 to jurisdiction 1. This will correspond to a perturbation $f(\cdot, \varepsilon)$ of f that will increase the population of jurisdiction 1 by $\varepsilon > 0$ while keeping the population of jurisdiction 2 the same and maintaining the equilibrium tax rates at their pre-perturbation levels. In other words, in the perturbed game $G(\varepsilon, l_{12}, l_{23})$, we have that $P_1(\varepsilon, l_{12}) > P_2(\varepsilon, l_{12}, l_{23})$ while

[19] See Lemma 3 in Appendix A.5.4 for details.

$T_1^N(\varepsilon, l_{12}, l_{23}) = T_2^N(\varepsilon, l_{12}, l_{23})$ holds.

First, we simplify our notation with $(T_1^N, T_2^N, T_3^N) = (T_1^N(l_{12}, l_{23}), T_2^N(l_{12}, l_{23}), T_3^N(l_{12}, l_{23}))$, and defining $x_{12}^N = x_{12}^N(T_1^N, T_2^N, l_{12})$, $x_{23}^N = x_{23}^N(T_2^N, T_3^N, l_{23})$, $x_{31}^N = x_{12}^N(T_3^N, T_1^N)$.

Let (a, b) be an open subinterval of $(0, 1)$ with $0 < a < 1$ and $l_{23} < b < 1$. Given (A.7), (a, b) must be picked so that it contains the three points x_{12}^N, x_{23}^N and x_{31}^N. The location of the above points relative to (a, b) is shown on Figure 3.

We introduce a population re-distribution of size ε from jurisdiction 3 to jurisdiction 1 *around the outside* of the interval (a, b) in the following precise manner. Consider intervals (a_1, b_1) and (a_2, b_2) such that the first interval is to the left of (a, b), i.e. $0 < a_1 < b_1 < a$, and the second interval is to the right of (a, b), i.e. $b < a_2 < b_2 < 1$. Consider two continuous functions g_1 and g_2 where $g_1 \geq 0$ and it is zero outside (a_1, b_1) whereas $g_2 \leq 0$ and is zero outside (a_2, b_2) as depicted in the second panel of Figure 3. Define

$$f(x, \varepsilon) = f(x) + g_1(x) + g_2(x).$$

Assume further that g_1 and g_2 are chosen such that i) $f(x, \varepsilon) \geq 0$ on $[0, 1]$ and ii) $\int_{a_1}^{b_1} g_1(x)\, dx = \varepsilon$ and $\int_{a_2}^{b_2} g_2(x)\, dx = -\varepsilon$. The final panel of Figure 3 shows the construction of $f(\cdot, \varepsilon)$. graphically.

Our assumptions on g_1 and g_2 imply that $f(\cdot, \varepsilon)$ is a pdf on $[0, 1]$, and we have the following observations:

Observation 1: $f(x, \varepsilon) = f(x)$ on (a, b).

Observation 2: $F(x, \varepsilon) - F(x', \varepsilon) = F(x) - F(x')$ for any x, x' in (a, b).

Observation 3: $|F(x, \varepsilon) - F(x)| \leq \varepsilon$ on $[0, 1]$.

Observation 3 implies the difference between the perturbed $F(\cdot, \varepsilon)$ and the original $F(\cdot)$ can be arbitrarily small—over all of $[0, 1]$—by choosing ε small enough.

Moreover, the above re-distribution moves a population of size ε from jurisdiction 3 to jurisdiction 1. However, the population we move continues to shop in jurisdiction 3. Therefore, for small enough ε, we expect the above population re-distribution to—very slightly—increases the size of the population in jurisdiction 1 without impacting the pre-redistribution equilibrium tax rates. To see this, note that we can find small enough open intervals S_1, S_2, S_3 respectively containing T_1^N, T_2^N, T_3^N such that for any $T_1' \in S_1$, $T_2' \in S_2$, $T_3' \in S_3$, we have $x_{12}(T_1', T_2', l_{12})$, $x_{23}(T_2', T_3', l_{23})$ and $x_{31}(T_3', T_1')$ arbitrarily close to x_{12}^N,x_{23}^N, and

x_{31}^N.

Observation 4: Points $x_{12}(T_1', T_2^N, l_{12})$, $x_{23}(T_2^N, T_3^N, l_{23})$, $x_{31}(T_3^N, T_1')$ are in (a, b).

Observation 2 and 4 together imply that, for all T_1' in S_1, we have

$$R_1(T_1', \mathbf{T^N_{-1}}, \varepsilon, l_{12}) = R_1(T_1', \mathbf{T^N_{-1}}, l_{12}). \tag{A.8}$$

Furthermore, Observations 3 implies that we have, for all T_1 on $\left[0, \overline{T}\right]$:

$$\left|F(x_{12}(T_1, T_2^N), \varepsilon) - F(x_{12}(T_1, T_2^N, l_{12}))\right| < \varepsilon \tag{A.9}$$

$$\left|F(x_{31}(T_3^N, T_1), \varepsilon) - F(x_{31}(T_3^N, T_1))\right| < \varepsilon \tag{A.10}$$

Since $R_1 = T_1\left[F(x_{12}) - F(x_{31})\right]$, this implies

$$\begin{aligned} &\left|R_1(T_1, \mathbf{T^N_{-1}}, \varepsilon, l_{12}) - R_1(T_1, \mathbf{T^N_{-1}}, l_{12})\right| \\ &< \overline{T}\left|F(x_{12}(T_1, T_2^N, l_{12}), \varepsilon) - F(x_{12}(T_1, T_2^N, l_{12}))\right| \\ &+ \overline{T}\left|F(x_{31}(T_3^N, T_1)) - F(x_{31}(T_3^N, T_1), \varepsilon)\right| < 2\overline{T}\varepsilon \end{aligned} \tag{A.11}$$

for all T_1 on $\left[0, \overline{T}\right]$.

In other words, (A.8) implies that the perturbed R_1 and the unperturbed R_1 agree on S_1 and (A.11) implies that the difference between the values of the perturbed R_1 and the unperturbed R_1 can be made arbitrarily small over $\left[0, \overline{T}\right]$. Hence, by Lemma 1, there exists ε_1 such that for all $\varepsilon < \varepsilon_1$, we have

$$T_1^N = \arg\max R_1(\cdot, \mathbf{T^N_{-1}}, \varepsilon_1, l_{12}).$$

Using a similar argument, we show there exists ε_2 such that for all $\varepsilon < \varepsilon_2$, we have

$$T_2^N = \arg\max R_2(\cdot, \mathbf{T^N_{-2}}, \varepsilon_2, l_{12}, l_{23}).$$

and there exists ε_3 such that for all $\varepsilon < \varepsilon_3$, we have

$$T_3^N = \arg\max R_3(\cdot, \mathbf{T^N_{-3}}, \varepsilon_3, l_{23}).$$

Let $\tilde{\varepsilon} = \min\{\varepsilon_1, \varepsilon_2, \varepsilon_3\}$. Then, (T_1^N, T_2^N, T_3^N) is an equilibrium for the game $G(\tilde{\varepsilon}, l_{12}, l_{23})$ for all $\varepsilon \leq \tilde{\varepsilon}$. In particular, in the game $G(\tilde{\varepsilon}, l_{12}, l_{23})$, the equilibrium tax rates of jurisdictions 1 and 2 are equal and they are the same as the the equilibrium tax rates in the unperturbed game $G(l_{12}, l_{23})$. Therefore, we have established that, for $\varepsilon \leq \tilde{\varepsilon}$,

$$T_1^N(\tilde{\varepsilon}, l_{12}, l_{23}) = T_1^N(l_{12}, l_{23}) = T_2^N(l_{12}, l_{23}) = T_2^N(\tilde{\varepsilon}, l_{12}, l_{23}),$$

when

$$P_1(\tilde{\varepsilon}, l_{12}) > P_2(\tilde{\varepsilon}, l_{12}, l_{23}).$$

A.5.3 Step 2

We can show that in the game $G(\varepsilon, l_{12}, l_{23})$, and for small enough ε, we have

$$\frac{\partial \left[T_1^N(l_{12}, l_{23}, \varepsilon) - T_2^N(l_{12}, l_{23}, \varepsilon)\right]}{\partial l_{12}} > 0, \tag{A.12}$$

in the same manner we used to establish the corresponding inequality for the game $G(l_{12}, l_{23})$ in Lemma 1.[20] Therefore, we perturb l_{12} to a slightly lower level $\tilde{l}_{12}$ so that in the new game $G(\tilde{\varepsilon}, \tilde{l}_{12}, l_{23})$, we continue to have[21]

$$P_1(\tilde{\varepsilon}, \tilde{l}_{12}) > P_2(\tilde{\varepsilon}, \tilde{l}_{12}, l_{23}),$$

and

$$T_1^N(\tilde{\varepsilon}, \tilde{l}_{12}, l_{23}) < T_2^N(\tilde{\varepsilon}, \tilde{l}_{12}, l_{23}),$$

which completes the proof of Proposition 4.

Alternatively, in Step 1, and starting with $P_1 = P_2$ and $T_1^N = T_2^N$, we could have moved an $\tilde{\varepsilon}$ amount of population from jurisdiction 1 to jurisdiction 3 such that the pre-perturbation equilibrium rates are maintained. Therefore, after perturbation, we continue to have $T_1^N = T_2^N$ while we now have $P_1 < P_2$. Then in Step 2, we increase l_{12} to a slightly higher value $\tilde{l}_{12}$ such that while $P_1(\tilde{\varepsilon}, \tilde{l}_{12}) < P_2(\tilde{\varepsilon}, \tilde{l}_{12}, l_{23})$, we have $T_1^N(\tilde{\varepsilon}, \tilde{l}_{12}, l_{23}) > T_2^N(\tilde{\varepsilon}, \tilde{l}_{12}, l_{23})$.

A.5.4 Technical Lemmas Used in Proof

Lemma 3. *Let f be a pdf with $f > 0$ on $[0,1]$. There exist $\bar{l}_{12}$ and $\bar{l}_{23}$ such that $P_1(\bar{l}_{12}, \bar{l}_{23}) = P_2(\bar{l}_{12}, \bar{l}_{23}) > P_3(\bar{l}_{12}, \bar{l}_{23})$.*

[20] See Lemma 4 in Appendix A.5.4 for details.

[21] Suppose $P_1(\tilde{\varepsilon}, \tilde{l}_{12}) > \psi > P_2(\tilde{\varepsilon}, \tilde{l}_{12}, l_{23})$. Since populations vary continuously with l_{12}, we can change l_{12} by a small amount such that we continue to have $P_1(\tilde{\varepsilon}, \tilde{l}_{12}) > \psi > P_2(\tilde{\varepsilon}, \tilde{l}_{12}, l_{23})$.

Proof. Since $0 \le F \le 1$ is strictly increasing, we can choose $\bar{l}_{12}$ and $\bar{l}_{23}$ in $(0,1)$ such that $F(\bar{l}_{23}) > 2/3$ and $F(\bar{l}_{12}) = F(\bar{l}_{23})/2$. Then, we have $P_1(\bar{l}_{12}, \bar{l}_{23}) = F(\bar{l}_{12}) = F(\bar{l}_{23})/2 > 1/3$, $P_2(\bar{l}_{12}, \bar{l}_{23}) = F(\bar{l}_{23}) - F(\bar{l}_{12}) = F(\bar{l}_{23})/2 > 1/3$, and $P_3(\bar{l}_{12}, \bar{l}_{23}) = 1 - F(\bar{l}_{23}) < 1/3$. Therefore, $P_1(\bar{l}_{12}, \bar{l}_{23}) = P_2(\bar{l}_{12}, \bar{l}_{23}) > P_3(\bar{l}_{12}, \bar{l}_{23})$. □

Lemma 4. *In game $G(\varepsilon, l_{12}, l_{23})$, for small enough ε, $\partial\left[T_1^N(l_{12}, l_{23}, \varepsilon) - T_2^N(l_{12}, l_{23}, \varepsilon)\right]/\partial l_{12} > 0$.*

Let $f(\cdot, \varepsilon)$ be obtained as in Step 1 above. The first-order conditions for the optimal tax rates for the resulting perturbed game are:

$$F\left(x_{12}, \varepsilon\right) - F\left(x_{31}, \varepsilon\right) - \frac{T_1\left[f\left(x_{31}, \varepsilon\right) + f\left(x_{12}, \varepsilon\right)\right]}{\delta} = 0, \tag{A.13}$$

$$F\left(x_{23}, \varepsilon\right) - F\left(x_{12}, \varepsilon\right) - \frac{T_2\left[f\left(x_{12}, \varepsilon\right) + f\left(x_{23}, \varepsilon\right)\right]}{\delta} = 0, \tag{A.14}$$

$$1 - F\left(x_{23}, \varepsilon\right) + F\left(x_{31}, \varepsilon\right) - \frac{T_3\left[f\left(x_{23}, \varepsilon\right) + f\left(x_{31}, \varepsilon\right)\right]}{\delta} = 0. \tag{A.15}$$

For $i = 1, \ldots 9$, following A.2, we compute $\gamma_i(\varepsilon)$ and $\alpha_i(\varepsilon) = \delta\gamma_i(\varepsilon)$ by taking taking the total derivative of the above FOCs and following the definitions at the end of Section A.2. For example, for the game $G(l_{12}, l_{12})$,

$$\gamma_2 = \frac{1}{\delta}\left[f\left(x_{12}\right) - \frac{T_1}{\delta}f'\left(x_{12}\right)\right] \ge 0,$$

whereas for the game $G(l_{12}, l_{23}, \varepsilon)$,

$$\gamma_2(\varepsilon) = \frac{1}{\delta}\left[f\left(x_{12}, \varepsilon\right) - \frac{T_1}{\delta}f'\left(x_{12}, \varepsilon\right)\right].$$

Letting $\mathbf{T^N} = (T_1^N, T_2^N, T_3^N)$, we have established that for $\varepsilon < \tilde{\varepsilon}$, an equilibrium $\mathbf{T^N}(l_{12}, l_{23})$ of the game $G(l_{12}, l_{23})$ is also an equilibrium $\mathbf{T^N}(l_{12}, l_{23}, \varepsilon)$ of $G(l_{12}, l_{23}, \varepsilon)$. Since this equilibrium is in the interval (a, b), and using Observation 1, we conclude that γ_2 must be equal to $\gamma_2(\varepsilon)$ when evaluated at vectors $\mathbf{T^N}(l_{12}, l_{23})$ and $\mathbf{T^N}(l_{12}, l_{23}, \varepsilon)$, respectively. Similarly, we can show for $i = 1, \ldots, 9$, $\gamma_i = \gamma_i(\varepsilon)$ and $\alpha_i = \alpha_i(\varepsilon)$ when evaluated at $\mathbf{T^N}(l_{12}, l_{23})$ and $\mathbf{T^N}(l_{12}, l_{23}, \varepsilon)$, respectively.

Therefore, for small enough ε, we have

$$\text{sign}\left[\frac{\partial(T_1^N(l_{12}, l_{23}, \varepsilon) - T_2^N(l_{12}, l_{23}, \varepsilon))}{\partial l_{12}}\right] = \text{sign}\left[\frac{\partial(T_1^N(l_{12}, l_{23}) - T_2^N(l_{12}, l_{23}))}{\partial l_{12}}\right] > 0,$$

where the inequality involving the second term was established in the proof of Lemma 2, and the first term is computed the way the second term is computed in section A.4 but using $\alpha_i(\varepsilon)$ instead of α_i.

References

Agrawal, David R. (2014), "LOST in America: Evidence on local sales taxes from national panel data." *Regional Science and Urban Economics*, 49, 147–163.

Agrawal, David R. (2015), "The tax gradient: Spatial aspects of fiscal competition." *American Economic Journal: Economic Policy*, 7(2), 1–30.

Agrawal, David R. and David E. Wildasin (2019), "Technology and tax systems." *Journal of Public Economics*, 185, 104082.

Anderson, James E. and Eric van Wincoop (2003), "Gravity with gravitas: A solution to the border puzzle." *American Economic Review*, 93(1), 170–192.

Anderson, Simon P., André de Palma, and Yurii Nesterov (1995), "Oligopolistic competition and the optimal provision of products." *Econometrica*, 63(6), 1281–1301.

Anderson, Simon P. and Jacob K. Goeree (1997), "Location, location, location." *Journal of Economic Theory*, 77, 102–127.

Aoyagi, Masaki and Atsuyuki Okabe (1991), "Existence of equilibrium configurations of competitive firms on an infinite two-dimensional space." *Journal of Urban Economics*, 29(3), 349–370.

Bagh, Adib (2023), "Existence and statistical estimation of equilibria in stochastic electoral competitions." *Journal of Convex Analysis*, 30(2), 481–498.

Bagnoli, Mark and Ted Bergstrom (2005), "Log-concave probability and its applications." *Economic Theory*, 26(2), 445–469.

Bloch, Francis and Delphine Manceau (1999), "Persuasive advertising in hotelling's model of product differentiation." *International Journal of Industrial Organization*, 17(4), 557–574.

Bloch, Francis and Nicolas Querou (2013), "Pricing in social networks." *Games and Economic Behavior*, 80, 243–261.

Bucovetsky, Sam (1991), "Asymmetric tax competition." *Journal of Urban Economics*, 30(2), 167–181.

Calvó-Armengol, Antoni and Yves Zenou (2002), "The importance of the distribution of consumers in horizontal product differentiation." *Journal of Regional Science*, 42(4), 793–803.

Caplin, Andrew and Barry Nalebuff (1991), "Aggregation and imperfect competition: On the existence of equilibirum." *Econometrica*, 59(1), 25–59.

Chen, Yongmin and Michael H. Riordan (2007), "Price and variety in the spokes model." *The Economic Journal*.

Chen, Yongmin and Michael H. Riordan (2008), "Price-increasing competition." *RAND Journal of Economics*, 39(4), 1042–1058.

Devereux, Michael P., Ben Lockwood, and Michela Redoano (2007), "Horizontal and vertical indirect tax competition: Theory and some evidence from the USA." *Journal of Public Economics*, 91, 451–479.

Evans, Carolyn L. (2003), "The economic significance of national border effects." *American Economic Review*, 93(4), 1291–1312.

Friberg, Richard, Frode Steen, and Simen A. Ulsaker (2022), "Hump-shaped cross-price effects and the extensive margin in cross-border shopping." *American Economic Journal: Microeconomics*, 14(2), 408–438.

Fuest, Clemens and Samina Sultan (2019), "How will Brexit affect tax competition and tax harmonization? the role of discriminatory taxation." *National Tax Journal*, 72(1), 111–138.

Haufler, Andreas (1996), "Tax coordination with different preferences for public goods: Conflict or harmony of interest." *International Tax and Public Finance*, 3(1), 5–28.

Haufler, Andreas and Ian Wooten (1999), "Country size and tax competition for foreign direct investment." *Journal of Public Economics*, 71(1), 121–139.

Hebous, Shafik and Michael Keen (2023), "Pareto-improving minimum corporate taxation." *Journal of Public Economics*, 225, 104952.

Hindriks, Jean and Valerio Serse (2019), "Heterogeneity in the tax pass-through to spirit retail prices: Evidence from Belgium." *Journal of Public Economics*, 176, 142–160.

Holmes, Thomas J. (1998), "The effect of state policy on the location of manufacturing: Evidence from state borders." *The Journal of Political Economy*, 106(4), 667–705.

Hoyt, William H. (2017), "The assignment and division of the tax base in a system of hierarchical governments." *International Tax and Public Finance*, 24(4), 678–704.

Huizinga, Harry, Luc Laeven, and Gaätan Nicodème (2008), "Capital structure and international debt shifting." *Journal of Financial Economics*, 88(1), 80–118.

Jacob, Benoy and Daniel McMillen (2015), "Border effects in suburban land use." *National Tax Journal*.

Janeba, Eckhard and Karl Schulz (2023), "A theory of economic disintegration." *International Economic Review*.

Kanbur, Ravi and Michael Keen (1993), "Jeux Sans Frontières: Tax competition and tax coordination when countries differ in size." *American Economic Review*, 83(4), 877–892.

Keen, Michael and Kai A. Konrad (2013), "The theory of international tax competition and coordination." *Handbook of Public Economics*, 5, 257–328.

Keen, Michael and Joel Slemrod (2017), "Optimal tax administration." *Journal of Public Economics*, 152, 133–142.

Lockwood, Ben (1993), "Commodity tax competition under destination and origin principles." *Journal of Public Economics*, 52(2), 141–162.

Mongrain, Steeve and John D. Wilson (2018), "Tax competition with heterogeneous capital mobility." *Journal of Public Economics*, 167, 177–189.

Mossay, Pascal and Pierre M. Picard (2011), "On spatial equilibria in a social interaction model." *Journal of Economic Theory*, 146(6), 2455–2477.
Neven, Damien J. (1986), "On hotelling's competition with non-uniform customer distributions." *Economics Letters*, 21(2), 121–126.
Nielsen, Søren Bo (2001), "A simple model of commodity taxation and cross-border shopping." *The Scandinavian Journal of Economics*, 103(4), 599–623.
Nielsen, Søren Bo (2002), "Cross-border shopping from small to large countries." *Economics Letters*, 77(3), 309–313.
Ohsawa, Yoshiaki and Takeshi Koshizuka (2003), "Two-dimensional fiscal competition." *Journal of Economic Geography*, 3(3), 275–287.
Ramsey, Frank P. (1927), "A contribution to the theory of taxation." *The Economic Journal*, 37(145), 47–61.
Rota-Graziosi, Grégoire (2019), "The supermodularity of the tax competition game." *Journal of Mathematical Economics*, 83, 25–35.
Salop, Steven C. (1979), "Monopolistic competition with outside goods." *The Bell Journal of Economics*, 10(1), 141–156.
Scheuer, Florian and Ivaán Werning (2016), "Mirrlees meets Diamond-Mirrlees: Simplifying nonlinear income taxation." NBER Working Paper 22076.
Slemrod, Joel (2019), "Tax compliance and enforcement." *Journal of Economic Literature*, 57(4), 904–954.
Slemrod, Joel and Wojciech Kopczuk (2002), "The optimal elasticity of taxable income." *Journal of Public Economics*, 84, 91–112.
Tarbush, Bassel (2018), "Hotelling competition and the gamma distribution." *Games and Economic Behavior*, 111, 222–240.
Topkis, Donald M. (1979), "Equilibrium points in nonzero-sum n-person submodular games." *SIAM Journal on Control and Optimization*, 17, 773–787.
Trandel, Gregory A. (1992), "Evading the use tax on cross-border sales: Pricing and welfare effects." *Journal of Public Economics*, 49(3), 313–331.
Trandel, Gregory A. (1994), "Interstate commodity tax differentials and the distribution of residents." *Journal of Public Economics*, 53(3), 435–457.
Ushchev, Philip and Yves Zenou (2018), "Price competition in product variety networks." *Games and Economic Behavior*, 110, 226–247.
van't Riet, Maarten and Arjan Lejour (2018), "Optimal tax routing: Network analysis of FDI diversion." *International Tax and Public Finance*, 25(5), 1321–1371.
Vives, Xavier (1999), *Oligopoly Pricing: Old Ideas and New Tools*. MIT Press.
Wang, You-Qiang (1999), "Commodity taxes under fiscal competition: Stackelberg equilibrium and optimality." *American Economic Review*, 89(4), 974–981.
Wilson, John Douglas (1991), "Tax competition with interregional differences in factor endowments." *Regional Science and Urban Economics*, 21, 423–451.
Wittman, Donald (1983), "Candidate motivation: A synthesis of alternative theories." *American Polical Science Review*, 77(1), 142–157.
Zhou, Jidong (2017), "Competitive bundling." *Econometrica*, 85(1), 145–172.