Anthropogeography and Taxation*

David R. Agrawal[†]
University of Kentucky

Mohammed Mardan[‡]
Norwegian School of Economics

January 2021

Abstract

The common view is that bigger jurisdictions set higher tax rates than smaller jurisdictions. Most tax competition models assume that the cost of tax arbitrage is uniformly distributed and the tax base is only mobile between two jurisdictions. We overturn this conventional result by relaxing these (unrealistic) assumptions, but still assuming that the only source of heterogeneity is due to population differences. Applied to a model of commodity taxes, we show that the more "marginal" people living near borders, the lower will be the jurisdiction tax rate. Then, we derive the optimal distribution of residents and show how jurisdictions can use other fiscal instruments, such as zoning, to mitigate the harmful effects of tax competition. An empirical application that exploits changes over time and space of the population of a jurisdiction. Increases in the number of people living near borders lowers tax rates, and after accounting for this, changes in total population have little effect on tax rates. Our application extends to capital tax competition or profit taxation when moving costs are not uniformly distributed across firms.

Keywords: fiscal competition, asymmetries, jurisdiction size, endogenous elasticities

JEL classification: H2, H7, R5

*The foundations for this project were laid while David Agrawal was a guest researcher at the KOF. He wishes to thank KOF along with the people associated with these institutions for their hospitality and support. We thank David Albouy, Jeffrey Lin, William Hoyt, Joel Slemrod, David Wildasin, and John Wilson who provided comments on a prior version of the paper. Any remaining errors are our own.

[†]University of Kentucky, Department of Economics and Martin School of Public Policy & Administration, 433 Patterson Office Tower, Lexington, KY 40506-0027; email: dragrawal@uky.edu; phone: 001-859-257-8608. Agrawal is also a fellow of CESifo.

[‡]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

We challenge the conventional wisdom in the literature that jurisdiction size is the most important determinant of a country's tax rate. Indeed, most models of asymmetric tax competition focus on interjurisdictional size differences, which leads to the conventional view that larger jurisdictions set higher tax rates than smaller jurisdictions (Kanbur and Keen 1993; Bucovetsky 1991; Haufler and Wooten 2010; Baldwin and Krugman 2004). As one extreme example, it is well known tax havens are often (only) small countries (Dharmapala and Hines 2009; Hines 2010; Hansen and Kessler 2001). Many of the size results are predicated on intuition derived from a standard inverse elasticity rule, where all else equal larger jurisdictions have larger tax bases and thus, lower elasticities that translate into higher tax rates. This intuition, by holding all else equal, imposes that the rate of change of the tax base is the same across large and small jurisdictions. In particular, often as a result of assuming two countries in the world compete for mobile factors, the density of marginal individuals that are indifferent between engaging in tax avoidance and not engaging in tax avoidance are the same in the both regions and does not influence the tax base elasticities. Indeed, we present descriptive stylized facts that indicate high-tax localities have a larger population than their median neighbor in less than 50% of the cases. Moreover, a county that sets a higher tax rate than all of its contigous neighbors is only in less than 15% of the cases the highest population jurisdiction. This initial fact suggests the correlation between population and tax rates need not be as strong as our theoretical models suggest.

Of course, the literature recognizes that other interjurisdictional asymmetries may lead to tax rate differentials across jurisdictions. For example, countries or localities may differ in their preferences for public goods (Haufler 1996), legal systems (Dharmapala and Hines 2009), patterns of agglomeration (Brülhart, Jametti and Schmidheiny 2012), or who they view as competitors (Janeba and Osterloh 2013). Sufficient heterogeneity in these forces may result in larger jurisdictions setting lower tax rates than their neighbors. We place an important qualification on the role of jurisdiction size using the example of cross-border shopping and commodity tax competition. We then extend the intuition of the results to models of capital tax competition and profit taxation.

People are not uniformly distributed across space. Many economic models recognize this phenomenon. For example, models of the monocentric city feature a declining population density away from the central business district (Brueckner 1987). Populations may also cluster near borders or be repelled by borders. Some borders, due to natural or man-made amenities, feature bunching of people on one side with very few on the other side of the border;[1] Moreover, path dependence relating to the presence of certain geographic features (Bleakley and Lin 2012) may make some states have more population near their borders rather than at the interior of the state. This non-uniform distribution of residents has important policy implications for fiscal policy because it influences the elasticity of the tax base. In particular, even if the population of a jurisdiction is held constant, a shift in the population distribution toward the border may lower its tax rate by raising its elasticity. Despite the simplicity of this idea, it has been (almost entirely) ignored by the tax competition literature.

We expand the classic model of Kanbur and Keen (1993) where jurisdictions are all the same area, but they differ in population, which implies one country is more densely populatedthan the other. The population of the country is uniformly distributed across space, but individuals are heterogeneous with respect to their cost of transportation. Jurisdictions set commodity tax rates to compete for cross-border shoppers. In the context of such a model, the larger population jurisdiction sets a higher tax rate than the lower population jurisdiction. Unlike the prior literature, in our model, jurisdictions are potentially homogenous in their total population, but they differ in how that population is distributed across space. Perhaps surprisingly, we show that in the context of *two* competing jurisdictions where the tax base is not mobile to the rest of the world, the standard Kanbur and Keen (1993) 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 within either jurisdiction. Intuitively this is consistent with the inverse elasticity rule. Start from equal tax rates in both jurisdictions. Then, the relative tax base elasticities depend on the size of the tax base, B, and its rate of change, $\triangle B/\triangle\tau$, with respect to the tax rate, τ. Then, because the system of jurisdictions is closed with respect to the rest of the world, and because cross-border shopping must flow from the high-tax to the low-tax jurisdiction, for a small change in either jurisdiction's tax rate, then $\triangle B/\triangle\tau$ must be the same in both jurisdictions. For this reason, starting from equal taxes where no cross-border shopping occurs, all differences in the elasticity of the tax base are driven by population size alone and the Ramsey rule says the large jurisdiction should raise its rate.

However, although much of the tax competition literature focuses on two competing jurisdictions, even for a *locally* mobile base such as shoppers, competition usually occurs between many jurisdictions. In the case of commodity taxes, countries share many borders; in the case of a globally mobile base, our extension will be even more stark. When jurisdictions compete with multiple jurisdictions for mobile factors or shoppers, we show

[1] Sometimes, the sorting of people and firms at borders are due to policy differences themselves (Holmes 1998).

the standard result in the literature that the bigger jurisdiction sets the higher tax rate may no longer hold. In particular, the spatial distribution of economic activity matters. The larger is the concentration of population near borders, the more likely taxes are going to be lower, even if total population is larger than one's neighbors. Critically, we further show that standard comparative statics with respect to jurisdiction population no longer hold: when perturbing a jurisdiction's population, one must account for where the increase in population comes from. Intuitively, returning to the Ramsey rule, $\triangle B/\triangle\tau$ is no longer the same for all jurisdictions, but rather depends upon the density on the high-tax side of borders for all jurisdictions that neighbor it.

Our model has an implication for the role of other fiscal instruments, including zoning. After showing that the standard result concerning size need not hold, we then ask: what is the optimal (revenue maximizing) population distribution within a state and can jurisdictions use other non-tax policies to influence this distribution? In the most extreme, a jurisdiction might relocate a capital city to the center of the state (Campante and Do 2014) in an effort to shift some of its population away from borders. At a less extreme, jurisdictions that are highly reliant on sales taxes might utilize zoning to manipulate the spatial distribution of stores or households such that these types of activity are located far away from borders, instead zoning border areas for industrial activity (Jacob and McMillen 2015) not sensitive to the sales tax base. While zoning has played an important role in Tiebout-style models (Hamilton 1975; Calabrese, Epple and Romano 2012; Barseghyan and Coate 2016) in order to mitigate inefficiencies, the use of zoning in other fiscal competition models remains understudied.[2] We show that zoning can be used to mitigate the harmful effects of commodity tax competition.[3] An immediate implication is that the elasticity of the tax base is not an underlying parameter that the tax authority takes as given. Similar to Slemrod and Kopczuk (2002), through the judicious use of zoning or other fiscal instruments, governments can *choose* the *optimal* elasiticty of the tax base.

The conclusion that small jurisdictions set lower tax rates appears throughout the tax competition literature, even in models that seemingly differ in important ways. Mongrain and Wilson (2018) obtain the same result for a model with two types of firms, domestic and foreign, where there firms have heterogeneous costs of becoming foreign firms by moving from their country of origin and both types of firms pay the same tax. Here, jurisdiction size is measured by the number of domestic firms. The distribution of moving costs plays the same roles as the density of people in commodity tax competition

[2]For models that study competition in the absence of zoning, see Epple and Zelenitz (1981), Wilson (1997), and Hoyt (1993).

[3]This is consistent with Agrawal, Hoyt and Wilson (2020)'s view that restrictions on non-tax instruments such as zoning are one of the main reasons for inefficient fiscal competition.

models. If moving costs are uniform, then similar results arise as in the commodity tax competition literature because the density of marginal firms that are indifferent between moving and not moving are the same in the two regions and does not influence the tax base elasticities. Mongrain and Wilson (2018) relax this assumption and as in our model, density plays a critical role, but these authors do not focus on how the distribution of moving costs affects the effect of jurisdictional size on tax rates; moreover, they assume two jurisdictions which as shown above is critical to overturning the conventional wisdom. In addition, our results generalize to models of profit-shifting (Keen and Konrad 2013; Agrawal and Wildasin 2019), which traditionally impose quadratic costs of shifting for the firms and assumes all profit shifting is from a single high-tax to a single low-tax jurisdiction. In reality, profit shifting occurs between many country pairs, in which case the distribution of shifting costs could, as in our model, overturn the standard role of jurisdiction size.

Although this paper is mainly designed to challenge the conventional theoretical wisdom concerning the role of jurisdictional size, we wish to provide some descriptive empirical evidence that is consistent with our model.[4] To do this, we assemble panel data from the first decade of the 2000s on county sales tax rates in the United States. We combine this with Census data on population and other demographics. Novel to our paper, we construct a metric of the number of "marginal" individuals. For every county border pair in the country, we calculate the number of people on the high-tax side that are within a few minutes of the county border, and those following a tax change likely to change their cross-border shopping decisions. To do this, we take data on Census block points across the United States and calculate the driving time to the nearest major road crossing of a county border. Using this driving time data, and evidence on the spatial reach of cross-border shopping, we can calculate the number of people on the high-tax side of the border and the maximum distance for which cross-border shopping is reasonable. Doing this for 2000 and 2010, then allows us to explain changes in tax rates with changes in the density of people near borders and total population. This approach allows us to exploit variation in counties that grow near the border versus far from the border, perhaps due to idiosyncratic reasons determined long in the past. We find a weak or no correlation of population with tax rates consistent with changes in population having positive or negative effects depending on where that increase happens across space. However, the role of people near borders is strongly negatively correlated with tax changes. We conclude, consistent with a speculative long-lost (or ignored) footnote in Trandel (1994), that the

[4] Most of the empirical literature focuses on estimating reaction function slopes rather than the effect of jurisdiction characteristics on tax rates. For example, see Brueckner (2003), Brueckner and Saavedra (2001), Lyytikäinen (2012), or Parchet (2019). Other important empirical papers on tax competition include Eugster and Parchet (2019) and Mast (2019).

relevant measure of jurisdiction size is not total population, but rather the population "close enough" to the border to make cross-border shopping worthwhile. Our empirical results suggest that empirical researchers should focus not only on measures of jurisdiction size, but rather measures of the "marginal" individuals or factors. This result applies for commodity taxation, or any other tax for which heterogenous agents face differentiated costs of tax avoidance or shifting mechanisms.

2 Comparison to prior literature

In standard tax competition models, if the only difference between jurisdictions is size, measured by number of residents, then the equilibrium tax rates will depend negatively on size. This result reappears among commodity tax competition models (Kanbur and Keen 1993; Trandel 1994; Nielsen 2001; Devereux, Lockwood and Redoano 2007), but as we will see, is common to other seemingly unrelated tax competition models. In these models, two governments set taxes in an open economy context. Consumers face a choice over where to buy goods: at home, paying the home jurisdiction tax rate and no transport costs, or abroad, paying the neighboring jurisdiction tax rates and transport costs. Consumers, located along a line segment with uniform density, incur transport costs equal to the distance needed to travel to the border times a per mile cost of travel. In equilibrium, cross-border shopping must be from the high-tax to low-tax jurisdiction. As individual demand is inelastic and all changes in the tax base are due to changes in cross-border shopping. Then, the rate of change of the tax base is the same in both jurisdictions: $\triangle B_i(\tau_i, \tau_j)/\triangle \tau_i$ is simply the inverse of the cost per unit of distance traveled. A tax increase in one jurisdiction will repel the same number of shoppers as the other jurisdiction will attract: one jurisdiction's loss is another jurisdiction's gain. Suppose that taxes are equal across both jurisdictions, so that no cross-border shopping occurs. Then the equilibrium tax base must be population size. If the rate of change is the same, it is easy to see that the Ramsey rule implies the large jurisdiction should raise its tax rate. However, as we will show, this classic size result is predicated on two assumptions: populations are uniformly distributed and there are two competing jurisdictions only.

Each of these models considers different variants of what size means. 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, population and length, and because uniform density is imposed throughout both jurisdictions, the model does not actually allow for a perturbation 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.

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 country but not within country. Because the Kanbur and Keen (1993) features a discontinuity in the density, unlike the prior model, it allows for specific country perturbations. In particular, an increase in population (also density) increases tax rates. An increase in area but holding population constant (lowering also density) decreases tax rates. However, the model is isomorphic to these two shocks because the ratio of population and area is all that matters. In particular, in both models, once we know area and population, density is irrelevant. Our model is flexible enough and helps separate the confusion over the role of population, which in these models, is confounded with density.

Trandel (1994) realizes this issue and allows density to vary across space, but does so in a linear manner. The assumption of a linearly increasing distribution, combined with a normalization of jurisdiction area, implies that density and population are always positively correlated: the bigger population jurisdiction must also have a higher density near the border.

We break this correlation between density and population and thus the positive correlation between jurisdiction size and tax rates. We do this by allowing an arbitrary distribution of residents such that the bigger jurisdiction may have a lower population near its border than the smaller jurisdiction has near its border. In practice, thinking of distance as proportional to the cost of tax evasion, the distribution of residents can be viewed as the distribution of people/factors/businesses with heterogeneous costs of avoiding taxes. In turn, this will allow us to generalize to capital tax competition models.

Before turning to our model, not the same "size" result applies in models of capital tax competition. The basic reason is that for a larger region, more of a tax increase is capitalized into the after-tax return on capital, resulting in a smaller increase in the required before-tax return on capital. In other words, the large region faces a less elastic supply of capital at a given tax rate. Given these tax rate differences, sufficiently small jurisdictions will import capital from larger jurisdictions, resulting in terms-of-trade effects from a tax increase. But these effects exist only because of the tax rate differences. Wilson (1991) and Bucovetsky (1991) find that in a 2-jurisdiction economy, the small jurisdiction is better off than large jurisdictions under tax competition, because its low tax rate is increasing its tax base at the expense of the large jurisdiction. In fact, Wilson (1991) finds that a sufficiently small jurisdiction will be better off than in the absence of tax competition, where lump-sum taxes on residents finance public good provision. As mentioned in the introduction, if firms were heterogeneous in the cost of moving capital, we can think of this cost being distributed as analogous to residents are distributed over distance. As with much of the literature, with uniform moving costs, the model reduces to variant of Kanbur and Keen (1993). Thus, relaxing this density function in our model also has strong implications for this literature.

3 Stylized Facts

To motivate the importance , or lack of importance, of jurisdictional size, we first present initial correlations between tax rates and population. To do this, we use the local sales tax dataset described in Agrawal (2014). Agrawal (2014) assembles a panel dataset of local sales taxes from 2003 to 2011. We will focus on county governments; counties are more numerous than states and, unlike towns, are sufficiently large in area that the assumptions of our theoretical model to follow are likely to hold.[5] In the United States, sales taxes are set by three levels of government in many states: states, counties, and towns. Our focus in this section are on the tax data at the start and end of this panel. For purposes of this section, we will focus on cross-sectional correlations or long-differences over the entire panel. We combine tax rate data with information on county level population from the decenial Census and the American Community Survey (ACS).[6]

Jurisdiction size, as measured by population, is *the* most important source of heterogeneity found in the theoretical tax competition (Bucovetsky 1991; Kanbur and Keen 1993). It is well-known that such models of tax competition predict that the larger jurisdiction will set a higher tax rate. This heterogeneity leads to inefficient diversity of tax rates and spatial distortions (Fajgelbaum et al. 2019). Jurisdiction size is also the most important predictor of which countries become tax havens (Dharmapala and Hines 2009). While it is well-recognized that other sources of heterogeneity – such as preferences for public goods (Haufler 1996; Nielsen 2002), the number of competitors (Ohsawa 1999), or agglomeration (Agrawal and Wildasin 2019) – result in diversity of tax rates and may overturn the "big jurisdiction sets a higher tax rate" result, size is often refered to as the most important factor contributing to tax rate diversity. Given the theoretical literature has focused almost exclusively on jurisdiction size as a source of asymetry, is this size result the dominant mechanism in the data or are other factors more important?

Table 1 focuses on data in 2011 for counties in states where these jurisdictions

[5] As we will focus on density near the border, given that many towns are very small in area, the assumption that residents sufficiently far away from the border at the town level is unlikely to hold. Counties vary in their size by state, but on average they are sufficiently large for our theory to hold. XX MAX DIST XX

[6] The ACS started in 2005. For panel data regressions, we interpolate data between the decenial census and 2005 to exploit changes from 2003 to 2011. Results are robust to starting in 2005.

Table 1: Number of Counties, Own Tax Relative to Median Neighbor Tax

	Own Tax Relative to Median Neighbor Tax		
	Own > Neighbor	Own < Neighbor	Own = Neighbor
Total Cases	575	530	1269
Own Population > Neighbor	236 [41%]	260 [49%]	566 [45%]
Own Population < Neighbor	339 [59%]	270 [51%]	703 [55%]

This table shows the count of counties in states allowing for county sales taxes for various cases. The columns indicate the county tax rate relative to the median neighbor county tax rate. The rows indicate the county population relative to the median neighbor population. A neighbor is defined as a county that shares a major road crossing. The median is then calculated over all neighboring jurisdictions. Percent of total cases in that column are given in [ ].

have the autonomy to set county-level taxes. For each county, we construct the set of neighboring counties. Following the approach in Agrawal (2015), a neighbor to county i is defined as a county that shares a major road crossing with county i. We use roads to define neighbors rather than the common approach of contiguity because cross-border shopping, which leads to tax competition, requires consumers to be able to cross borders. Indeed for many counties that share a border cross-border shopping is not feasible due to the presence of a large river or mountain range. Furthermore, some counties separated by large bodies of water and thus would not be contigous are connected by roads, such as the Bay Bridge. After calculating all of the neighbors we calculate the median, mean and maximum neighbor tax rate and the median, mean and maximum neighbor population. We then compare these to the tax rate in county i. Table 1 presents the tabulations for the median and table 2 presents the results for the maximum.[7] For the median, only 41% of jurisdictions that set a higher tax rate than their median neighbor have a larger population than the median neighbor. At the same time, when the tax rate is lower than the median neighbor, 50% of jurisdictions are larger in size than the median. This provides some initial descriptive evidence that, as expected, other factors than size matter. In table 2, the results are even more striking. For the 192 counties in the U.S. that set a tax rate strictly larger than all of their neighbors, only in 14% of the cases is this jurisdiction also the largest.

For a more formal test, we look at the correlation of tax rates and jurisdiction size in the cross-section. Given the relationship between tax rates and populations appears to be quadratic, for jurisdictions that are very large, increases in population appear to lower tax rates. Figure 1 shows this relationship.

In the prior figure, one may be worried about unobserved heterogeneity. Thus, we look at the correlation of changes in tax rates and changes in population. To do this, we

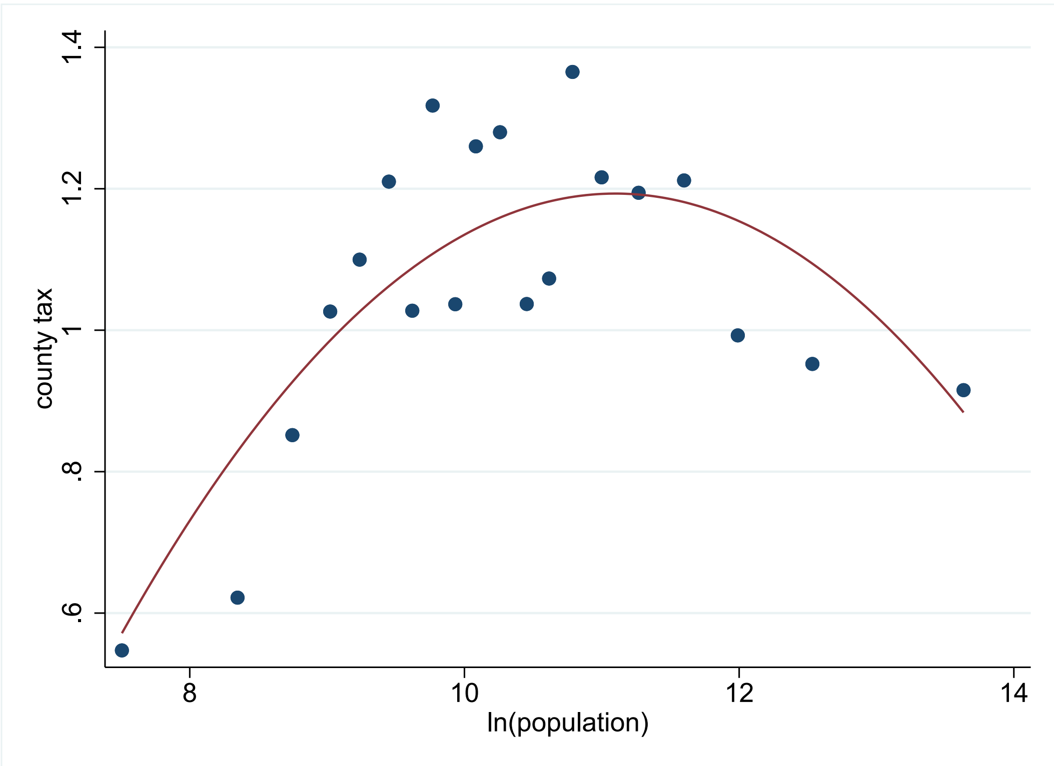
[7] Results using the mean are presented in the appendix as table A.1.

Table 2: Number of Counties, Own Tax Relative to Maximum Neighbor Tax

	Own Tax Relative to Maximum Neighbor Tax		
	Own > Neighbor	Own < Neighbor	Own = Neighbor
Total Cases	192	1062	1120
Own Population > Neighbor	27 [14%]	160 [15%]	169 [15%]
Own Population < Neighbor	165 [86%]	902 [85%]	951 [85%]

This table shows the count of counties in states allowing for county sales taxes for various cases. The columns indicate the county tax rate relative to the maximum neighbor county tax rate. The rows indicate the county population relative to the maximum neighbor population. A neighbor is defined as a county that shares a major road crossing. The maximum is then calculated over all neighboring jurisdictions. Percent of total cases in that column are given in [ ].

Figure 1: Correlation: Taxes and Population



This figure shows the cross-sectional correlation between county tax rates and the natural logarithm of population. Data are binned into twenty equally sized bins along with a quadratc fit.

estimate a regression of the form

$$\triangle\tau_{it} = \alpha + \beta\triangle\frac{P_{it}}{\bar{P}_i} + \triangle\varepsilon_i \tag{1}$$

where $\triangle x_{it}$ denotes the change in a variable in jurisdiction i from $t = 2003$ to $t = 2011$. Then, τ is the county tax rate and p_i is the population. We standardize $\triangle p_i$ across jurisdictions by dividing by the initial population, $\bar{p}_i$ in 2003, so that the independent variable of interest is the "percent change" in population. This equation can be interpreted as a panel data regression that has differenced out county-level fixed effects.[8] Figure 2 shows the results visually in the form of a binned-scatter plot for all jurisdictions, including those that did not change the tax rate and excluding those that did not change the tax rate. Surprisingly a strong negative relationship is present: the largest tax increases were those jurisdictions that grew the smallest amounts. The smallest tax increases were those that grew the most. Formal regression results, clustering standard errros at the state level,[9] indicate these results are statistically significant: the estimated coefficient [p-value] is -0.246** [p = 0.014] when exclusing zeros and is -0.054** [p = 0.014] when including zeros.

The correlations should only be observed as causal if no other unobservable factors are correlated with both population and tax rates. While this is a strong assumption, the evidence above suggests that population may not be as important a factor in explaining tax rates as the theoretical literature would suggest. Moreover, as we explain below, this is because it is the elasticity of the base that matters. This elasticity depends on two things: size and the change in the base due to a tax change. In the context of cross-border shopping, the latter of these is interpreted as density.

4 Model

4.1 The two-jurisdiction case

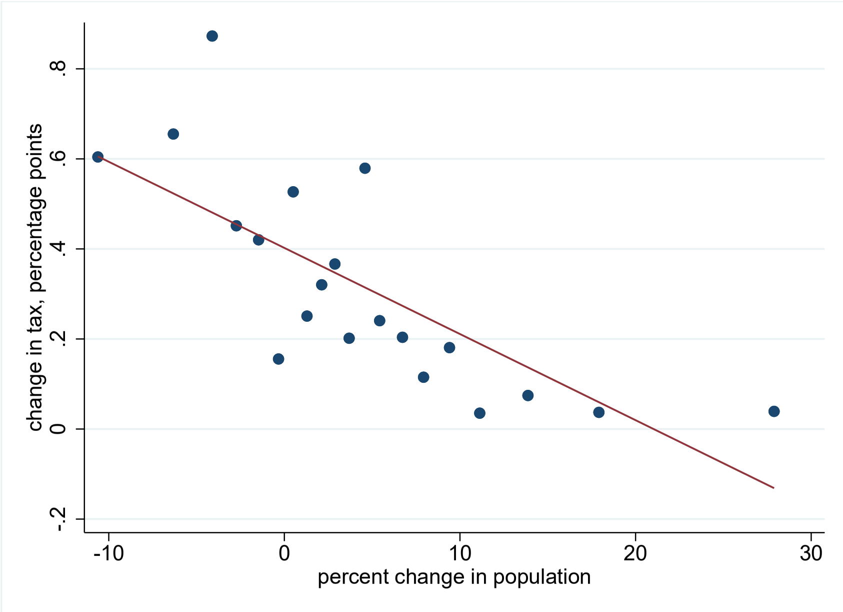
We set up a commodity tax competition model featuring strategic interactions between jurisdictions. Many features in the model share similarities with the classic papers of Kanbur and Keen (1993) and Nielsen (2001). In this section, we consider a model of two jurisdictions on a circle. The total length of the two jurisdictions, and thus the circumference of the circle, is L. We allow jurisdictions to differ in their length, meaning that jurisdiction 1 ranges from $l_{21} = 0$ to l_{12} and has length L_1, while jurisdiction 2

[8]Because we use a single long difference, time fixed effects become the intercept.

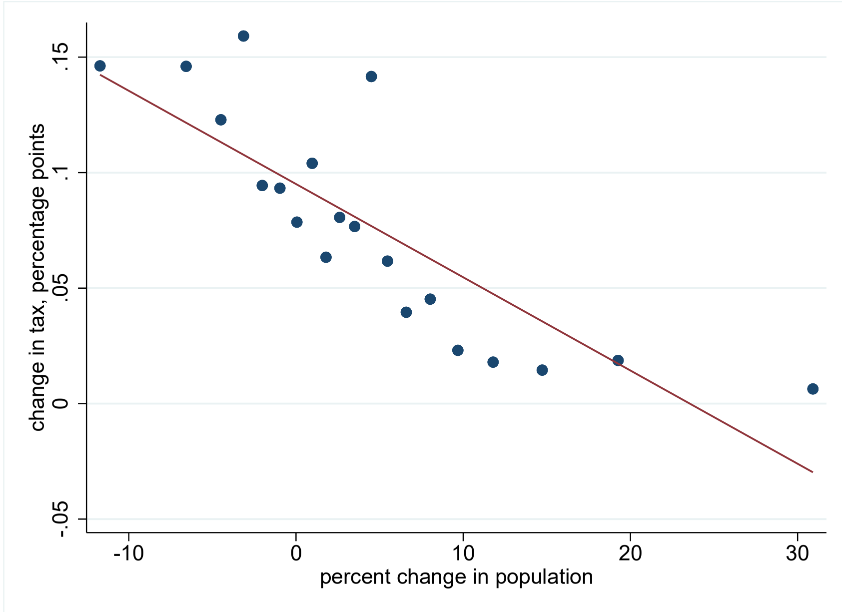
[9]We apply the percentile-t wild cluster bootstrap, imposing the null hypothesis (Cameron and Miller 2015), to obtain corrected p-values of statistical significance (). We do this because we have only 31clusters, the number of states that allow for local taxes.

Figure 2: Change in Taxes and Change in Population

(a) Excluding Counties with No Tax Change



(b) Including Counties with No Tax Change



This figure shows the results of a panel data regression in (long) first differences over a decade long period. The vertical axis represents the change in the county tax rate and the horizontal axis represents the change in the county population. Panel 2a excludes counties that did not change the tax rate, while panel 2b includes counties that did not change the tax rate.

ranges from l_{12} to l_{21} and has length $L_2 = L - L_1$. Each jurisdiction's government levies an origin-based commodity tax to maximize tax revenue, where we denote the tax rate of jurisdiction i by T_i.[10]

Individuals live along the circumference of the circle and wish to purchase one unit of a good from firms, which sell a single composite good. Irrespective of individuals' residence, the maximum willingness to pay is V and large enough to ensure full market coverage. Firms are located potentially anywhere on the circle and offer the good in a perfectly competitive environment resulting in pre-tax prices equaling marginal costs, which we normalize to 1. Although demand is 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 same point on the circle as she resides. Instead, if the individual purchases in the neighboring jurisdiction, she pays the tax-inclusive price in the neighboring jurisdiction, but incurs per unit transportation costs δ for traveling to the border from her home. Define individuals' location on the circle relative to point "0," we can determine the marginal individuals that are inidifferent between shopping at home and in the neighboring jurisdiction as

$$\begin{aligned} x_{12} &\geq l_{12} - \frac{T_1 - T_2}{\delta} \equiv x_{12}^*, \\ x_{21} &\geq \frac{T_1 - T_2}{\delta} \equiv x_{21}^*. \end{aligned} \tag{2}$$

Notice that marginal individuals can be located on one or the other side of the border depending on whether the tax differential $T_i - T_j$ is positive or negative. The setup described above is essentially the same as in Nielsen (2001), except we have two rather than one border. Figure XX summarizes the structure as well as the notation of the model.

Instead of following the convention to assume uniformly distributed individuals, we generalize the standard setup of Nielsen (2001) by allowing individuals to be arbitrarily distributed on the circle.[11] Specifically, we assume residents are distributed on the circle

[10] In the United States, taxes are effectively origin-based. Although cross-border transactions are legally subject to use taxes at the destination rate, these taxes are readily evaded. As a result, most purchases are assessed at the sales tax rate in the place of purchase.

[11] With the exception of Kanbur and Keen (1993) and Trandel (1994), every spatial commodity tax competition has assumed a uniform distribution of residents. Kanbur and Keen (1993) assume the distribution is discontinuous at the border, but is then uniform *within* each jurisdiction. For other examples using the uniform distribution, see Nielsen (2001), Devereux, Lockwood and Redoano (2007), Kessing (2008), and Agrawal (2015). In Trandel (1994) the distribution of residents is a linearly decreasing function. His main interest is to determine how general the result that the larger jurisdiction sets a higher tax rate is to alternative assumptions of welfare maximization and imperfect competition. His non-uniform density is mainly a modification to eliminate the discontinuities in Kanbur and Keen (1993).

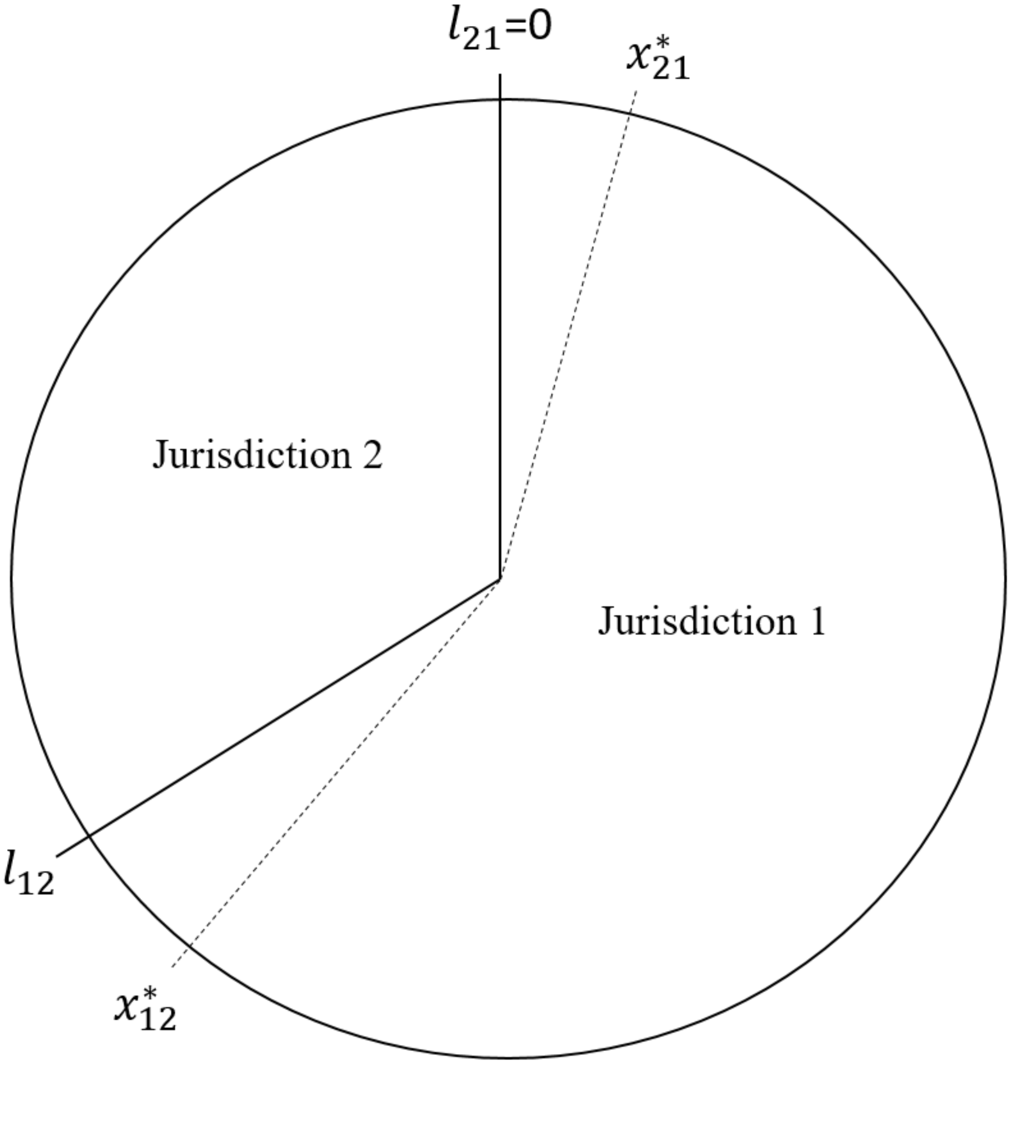


Figure 3: Model structure with two jurisdictions for $P_1 > P_2$

according to a density function $f(x)$ over the space $[-L, L]$. Although $f(x)$ could be any function, including a function containing discontinuities at arbitrary points, we assume that $f(x)$ is continuous and differentiable on $x \in [-L, L]$ as well as periodic with a period of L to avoid discontinuous revenue functions that lead to piecewise best replies.[12] It may appear as if these assumptions limit the number of density functions to which the model applies. However, we note that any function, which does not satisfy these assumptions, can be approximated by a Fourier series $\mathcal{F}(x)$ that is continuous, differentiable and periodic. In Appendix A.3, we show examples of Fourier approximations that allow us to approximate a variety of population distributions.[13] We note that defining the density function over the $x \in [-L, L]$ although the length of the cicle ranges from 0 to L is made for analytical convenience due the possibility that $x_{21}^* < 0$ if $T_2 > T_1$. Specifically, the assumption of a periodic function implies that for any $x < 0$, we have $f(x) = f(x + L)$, which allows us to integrate over a range that contains negative values of x.

Then, the revenue functions of the two jurisdictions are given by

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

where x_{ij}^* are the marginal consumers as defined in (2) and $\bar{P} = \int_0^{l_{21}} f(x)dx$ is the total population of the two jurisdictions. Differentiating the tax revenues with respect to the local tax rate yields the optimal tax rates

$$T_1^N = \frac{\delta \left[F(x_{12}^*) - F(x_{21}^*) \right]}{f(x_{12}^*) + f(x_{21}^*)}, \quad (3)$$

$$T_2^N = \frac{\delta \left[\bar{P} - F(x_{12}^*) + F(x_{21}^*) \right]}{f(x_{12}^*) + f(x_{21}^*)}. \quad (4)$$

Because the denominators are identical, whether the tax rate of jurisdiction 1 or 2 is larger only depends on the relative sizes of the tax bases. Hence, from equations (3) and (4), we can derive the following condition:

$$T_1^N > T_2^N \Longleftrightarrow \left[F(x_{12}^*) - F(x_{21}^*) \right] > \left[\bar{P} - F(x_{12}^*) + F(x_{21}^*) \right]. \quad (5)$$

[12] Then, a periodic function could simply be a function that repeats at regular intervals with disctonuities. For example the piecewise function $f(x) = x$, $0 \leq x < L$ could repeat as $x - L$, $L \leq x < 2L$ and $x + L$, $-2L \leq x < 0$, etc.

[13] Moreover, functions from the class of functions sometimes called "bump functions" or "mollifiers" or "test functions" can be pieced together to obtain arbitrary population distributions.

If $T_1^N > T_2^N$, then it must be that $P_1 > F(x_{12}^*) - F(x_{21}^*)$, where $P_1 = \int_0^{l_{12}} f(x)dx$ is jurisdiction 1's population size, because some individuals residing in jurisdiction 1 decide to cross-border shop in jurisdiction 2. For the same reason $\bar{P} - F(x_{12}^*) + F(x_{21}^*) > P_2$, which using (5) implies that $P_1 > F(x_{12}^*) - F(x_{21}^*) > \bar{P} - F(x_{12}^*) + F(x_{21}^*) > P_2$, with $P_2 = \bar{P} - P_1$ as jurisdiction 2's population size. Thus, it follows that if $T_1^N \geq T_2^N \iff P_1 \geq P_2$, and $T_1^N < T_2^N \iff P_1 < P_2$, which means that in a two jurisdiction setup only the more populated jurisdiction can be the high-tax jurisdiction. We summarize in:

Proposition 1. *When two asymmetric jurisdictions competitively set tax rates, the larger population jurisdiction always sets the higher tax rate regardless of the distribution of residents in the jurisdictions.*

At first glance, it may appear suprising that the large jurisdiction always sets the higher tax rate even in the extreme case when its whole population is located directly at the borders, while the small jurisdictions' population is concentrated far away from the borders. The reason for our stark result in contrast to one's first intuition originates from the fact that the tax base sensitivity is identical for both jurisdictions as they are competing for the same individuals irrespective of how these individuals are distributed. Then, the tax elasticity each jurisdiction's goverment faces only differs because the sizes of the tax bases (which correspond to populations) are different across the jurisdictions. Due to the larger population, the large jurisdiction faces the smaller elasticity, which implies, following a Ramsey rule, that it sets the higher tax rate.

While the distribution of individuals does not have a qualitative effect on the conventional outcome that the larger jurisdiction sets the higher tax rate, the density does indeed quantitatively affect the outcome of tax competition. To get a clear idea of how the density of people at the border affects tax competition, we assume, for simplicity, that jurisdictions are identical in terms of length and population size. Moreover, the distribution of individuals is not only symmetric between the two jurisdictions, but also symmetric within each jurisdiction. Using these simplifications, each jurisdiction's tax base amounts to $\frac{\bar{P}}{2}$, which yields an optimal tax rate for jurisdiction $i = 1, 2$ of

$$T_i = \frac{\delta \bar{P}}{4f(l)}, \tag{6}$$

where $f(l)$ is the density of individuals located directly at any of the two borders due to our simplifying assumption that the distribution of individuals is symmetric between and within jurisdictions.

Equation (6) shows that tax competition will be fiercer the more densely populated location $x^* = l$ is, i.e. the larger $f(l)$. The reason is that the higher density at the borders the more cross-border shoppers can be attracted by a reduction in the tax rate. We summarize in:

Proposition 2. *A larger number of marginal cross-border shoppers increases tax competition and leads to lower equilibirum tax rates.*

4.2 The three-jurisdiction case

In this section, we verify whether the previously derived result extends to a setup with three jurisdictions. While all of the basic assumption remain unaltered, we modify the setup introduced in section 4.1 by only adding a third jurisdiction. Specifically, jurisdiction 1 ranges from $l_{31} = 0$ to l_{12} and has length L_1, jurisdiction 2 ranges from l_{12} to l_{23} and has length L_2, and jurisdiction 3 ranges from l_{23} to $l_{31}$and has length $L_3 = L - L_1 - L_2$. The cut-off rules for the marginal individuals are therefore given by

$$\begin{aligned} x_{12} &\geq l_{12} - \frac{T_1 - T_2}{\delta} \equiv x_{12}^* \\ x_{23} &\geq l_{23} - \frac{T_2 - T_3}{\delta} \equiv x_{23}^* \\ x_{31} &\geq \frac{T_1 - T_3}{\delta} \equiv x_{31}^*. \end{aligned} \tag{7}$$

Based on individuals' shopping behavior, we can express the revenue functions for the three jurisdiction as

$$\begin{aligned} R_1 &= T_1 \int_{x_{31}^*}^{x_{12}^*} f(x)dx = T_1 \left[F(x_{12}^*) - F(x_{31}^*) \right] \\ R_2 &= T_2 \int_{x_{12}^*}^{x_{23}^*} f(x)dx = T_2 \left[F(x_{23}^*) - F(x_{12}^*) \right] \\ R_3 &= T_3 \left[\bar{P} - \int_{x_{31}^*}^{x_{12}^*} f(x)dx - \int_{x_{12}^*}^{x_{23}^*} f(x)dx \right] = T_3 \left[\bar{P} - F(x_{23}^*) + F(x_{31}^*) \right] \end{aligned}$$

Differentiating the revenue function with respect to the local tax rate yields the optimal tax rates, which we can express as follows:

$$T_1^N = \frac{\delta\left[F\left(x_{12}^*\right) - F\left(x_{31}^*\right)\right]}{f\left(x_{31}^*\right) + f\left(x_{12}^*\right)}, \tag{8}$$

$$T_2^N = \frac{\delta\left[F\left(x_{23}^*\right) - F\left(x_{12}^*\right)\right]}{f\left(x_{12}^*\right) + f\left(x_{23}^*\right)}, \tag{9}$$

$$T_3^N = \frac{\delta\left[\bar{P} - F(x_{23}^*) + F(x_{31}^*)\right]}{f\left(x_{23}^*\right) + f\left(x_{31}^*\right)}. \tag{10}$$

Based on our previous analysis, the question arises whether jurisdictions with a larger population will actually levy lower tax rates than jurisdictions with a population. To analyze this question we assume, for convenience, that jurisdictions 2 and 3 have the same length and that the overal population distribution on the circleis axially symmetric with respect to the border between jurisdiction 2 and 3.This implies that jurisdictions 2 and 3 have the same population sizes and distributions . and will therefore levy the same tax rates while jurisdiction 1's tax rate will differ. The optimal tax rates are then given by

$$\begin{aligned} T_1^N &= \frac{\delta\left[F\left(x_{12}^*\right) - F\left(x_{31}^*\right)\right]}{2f\left(x_{12}^*\right)}, \\ T_2^N = T_3^N &= \frac{\delta\left[F\left(x_{23}^*\right) - F\left(x_{12}^*\right)\right]}{f\left(x_{12}^*\right) + f\left(l_{23}\right)}, \end{aligned} \tag{11}$$

where $f\left(l_{23}\right)$ is the density of individuals residing at the border between jurisdiction 2 and 3. Whether the tax rate of jurisdiction 1 or of jurisdictions 2 and 3 are larger does not only depends on the relative sizes of the tax bases, but also on the densities at the respective borders. To show that the larger jurisdiction may set the lower tax rate, we assume without loss of generality that $P_1 > P_2 = P_3$. Then, $T_1^N < T_2^N = T_3^N$ if

$$\frac{F\left(x_{12}^*\right) - F\left(x_{31}^*\right)}{F\left(x_{23}^*\right) - F\left(x_{12}^*\right)} < \frac{2f\left(x_{12}^*\right)}{f\left(x_{12}^*\right) + f\left(l_{23}\right)}. \tag{12}$$

Thus, the question is if and under which conditions it is possible that $T_1^N < T_2^N = T_3^N$. If $T_1^N < T_2^N = T_3^N$, some individuals residing in jurisdictions 2 and 3 decide to cross-border shop in jurisdiction 1. Therefore, it must be that $F(x_{12}^*) - F(x_{31}^*) > P_1$, where we denote $P_1 = \int_0^{l_{12}} f(x)dx$ as jurisdiction 1's population size. For the same reason $P_2 > F\left(x_{23}^*\right) - F\left(x_{12}^*\right)$, where we denote $P_2 = \int_{l_{12}}^{l_{23}} f(x)dx$ as jurisdiction 2's population size, which implies that $F(x_{12}^*) - F(x_{31}^*) > P_1 > P_2 > F\left(x_{23}^*\right) - F\left(x_{12}^*\right)$. Thus, in equation (12), the left-hand side is larger than 1. Consequently, $T_1^N < T_2^N = T_3^N$ cannot be satisfied if $f\left(l_{23}\right) \geq f\left(x_{12}^*\right)$, because the right-hand side of (12) cannot exceed 1. Based on our previous discussion, a necessary condition is that the tax base sensitivity of jurisdictions 2 and 3 has to be lower than the one for jurisdiction 1. This will be the case if $f(l_{23}) < f(x_{12}^*)$, i.e. the number of marginal cross-border shoppers jurisidcions 2 and 3 compete for is smaller than the one jurisdictions 1 and 2 compete for. Moreover, the difference in the two tax bases cannot be too large, which will be the case if the difference in population sizes and the number of cross-border shopper is sufficiently small. To clarify why this has to be the case, we rewrite (12) as

$$T_1^N < T_2^N = T_3^N \quad \Leftrightarrow \quad \frac{P_1 + 2CBS}{P_2 - CBS} < \frac{2f(x_{12}^*)}{f(x_{12}^*) + f(l_{23})}, \tag{13}$$

where we denote as $CBS > 0$ as the number of cross-border shoppers from jurisdiction 2, respectively jurisdiction 3, to jurisdiction 1. In the extreme case when $f(l_{23}) = 0$, jurisdictions 2 and 3 will set a higher tax rate than jurisdiction 1 despite a smaller population size if $CBS < \frac{2P_2 - P_1}{4}$. Thus, apart from that CBS has to be sufficiently small, which will be satisfied if δ is sufficiently large, the aggreagate population of jurisdictions 2 and 3 must be larger than jurisdiction 1's population size for this condition to hold. We summarize in:

Proposition 3. *When two symmetric small jurisdictions and one large jurisdiction competitively set tax rates, the smaller population jurisdictions set the higher tax rate if*

(i) the difference in population sizes between a small and the large jurisdiction is sufficiently small,

(ii) the number of marginal cross-border shoppers the two small jurisdictions compete for is sufficiently small,

(iii) the total number of cross-border shoppers from a small to the large jurisdiction is sufficiently small.

Proposition 3 illustrates that less populated jurisdictions may set higher tax rates than their more populated neighbors despite their smaller tax base. However, it is important to note that such a situation cannot occur when a very large jurisdiction competes with a very small jurisdiction. In this case, the incentive for the large jurisdiction to set a high tax rate prevails because of its large tax base.

5 Examples

In this section, we provide examples to illustrate our main findings. In our first example, we assume symmetry between all jurisdictions to emphasis the role of density for tax competition.

5.1 Step function

Case 1: 2 country KK model, step function. Steps in opposite direction. Redo KK model with step functions. Revenue function for small jurisdiction changes slightly. Maybe do, maybe don't

5.2 Example 1

The purpose of this example is to indicate the relevance of density for tax competition and therefore the level of equilibrium tax rates. Suppose each of the three jurisdictions has a length of 2π, so that the total length is 6π. To illustrates the importance of density for the outcome of tax competition, we analyze two cases with different density functions, one for which $f(x) = cos(x) + 1$ and the other for which $f(x) = cos(x + \pi) + 1$. Both desity functions yield a total population in each jurisdiction of 2π. Because the symmetry condition implies that $CB_{12} = CB_{13} = CB_{23} = 0$, equation (??) simplifies to

$$T = \frac{\delta\pi}{f(l)}, \tag{14}$$

where $f(l)$ indicates the density directly at the border. In the first case, for which density is $f(x) = cos(x) + 1$, the density directly at the border reads $f(0) = f(2\pi) = 2$, resulting in equilibrium tax rates of $T = \frac{\delta\pi}{2}$. In the second case, where density is given by $f(x) = cos(x+\pi)+1$, the density directly at the border is given by $f(0) = f(2\pi) = 0$, resulting in the highest possible equilibrium tax rates, because reducing the tax rate does not provoke an inflow of cross-border shoppers. Overall, this example illustrates again that density determines the strength of tax competition.

5.3 Example 2

The purpose of this example is to show that larger jurisdictions in terms of total population do not necessarily need to set higher tax rates than their smaller counterparts. Suppose jurisdiction 1 has a length of $\frac{7\pi}{10}$, while jurisdictions 2 and 3 have both a length of $\frac{13\pi}{20}$, so that the total length is 2π. Moreover, assume the density function reads $f(x) = cos(x - \frac{7\pi}{20}) + 1$, which implies that density is axially symmetric with respect to the border between jurisdictions 2 and 3. Moreover, because of the symmetry, jurisdiction 2 and 3 will set the same tax rates, which implies that the density directly at the border matters for tax competition between them, which is zero given the density function. Hence, we calculate

Three country case 1: Overturn the classic result that the large jurisditions set higher tax rates

Three country case 2: Asymmetric jurisdictions can set syemmtric tax rates.

6 Verification of the Model

In this section, we verify that the density of people near borders plays a prominent role even after account for (total) jurisdiction size, i.e., population. To do this we use the panel data on local sales tax rates discussed in section 3. We combine this with a novel dataset of density near jurisdiction borders. To the density of "marginal" cross-border shoppers we need to determine how the population of a jurisdiction is allocated across space. To do this, we use data on Census block groups.[14] Although Census block groups are generally small in area, in less dense places they may be relatively large in area. For this reason, we calculate the population weighted centroid of the block group (or the point upon which the group area would balance) using all the Census block centroids and their population counts within the group. Using data on the full road infrastructure of the USA and network data on speeds, we then calculate the driving time from every population-weighted centroid to the nearest intersection of a "major road" crossing and a county border. The use of major roads is designed to capture road usage for cross-border shopping rather than for residential usage. This procedure follows Agrawal (2015), except at a much finer level of geography and for county borders rather than state borders.[15] We repeat this data using Census data from 2000 and 2010.[16] Armed with this data in hand, we know how far (in minutes) each Census block group is from the nearest county border. We thus know the tax differential that residents face at that border. In turn, we will exploit variation in the population of jurisdictions, but also the spatial distribution of that variation.

Figure 4 shows the variation in the dpopulation distribution that we exploit. The green dots represent the the centroids of the Census block groups and the (larger) red dots are major road intersection border crossings. Given these block groups may have different populations, the number of centroids in an area proxies for a density in the figure. In the formal analysis we account for their population. In California, notice how the population clusters in certain areas due to the geographic shape of the valley in the area. Sacremento County has a much higher density near its borders as does Stanislaus County, in comparision to San Joaquin County, which has much of its population clustered

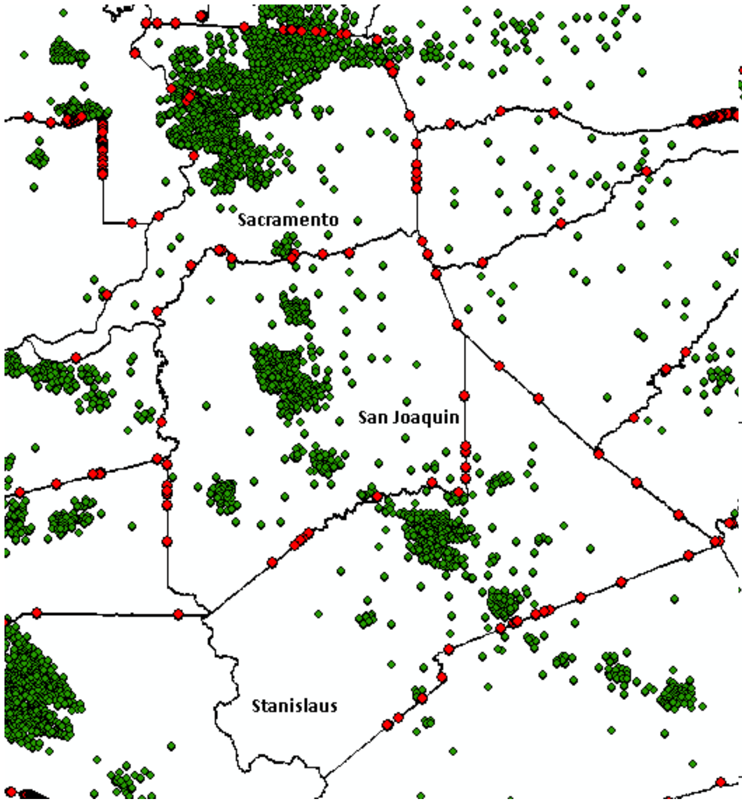
[14]Census block groups are the second smallest Census unit. In 2010, the United States was divided into more than 2010 Census block groups. Typically, Census block groups have a popluation between 600 and 300 people. We use Census block groups because the procedure we implement is computationally intensive and would take several weeks to run on all Census blocks.

[15]Lovenheim (2008) and Harding, Leibtag and Lovenheim (2012) use the crow-fly distance.

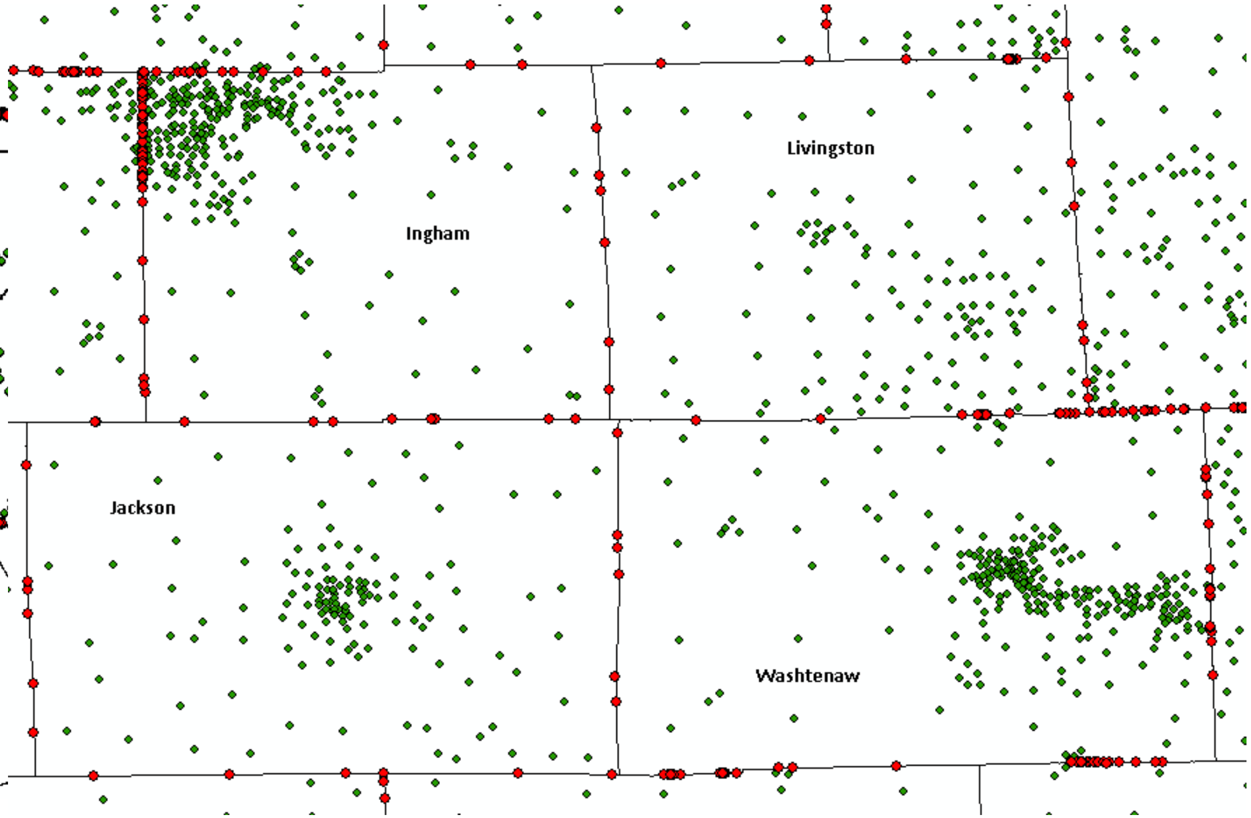
[16]Block data are not available at the ACS. Thus, when we match to tax rates, we will match 2000 Census data to our earliest local tax rates, 2003 and 2010 Census data to our 2011 tax rates.

Figure 4: Examples of Density

(a) California Counties



(b) Michigan Counties



This figure shows examples of the distribution of density for some counties in the United States. The green dots represent the the centroids of the Census block groups. In the formal analysis we account for their population. The red (slightly larger) dots are the intersections between major roads and county borders. We show results for counties in CA and MI, but the latter are not in our empirical analysis as MI does not allow for county tax rates.

in the center. In the Michigan Counties, notice that Ingham and (perhaps) Washtenaw Counties have more of their populations near border crossings. On the other hand, Jackson has most of its population at the center of the county, while Livingston County has its population spread relatively uniformly. Although some of these counties have similar populations, they differ in the spatial arrangement of that population, which is the variation we exploit.

To proceed, we assume the data are in equilibirum and then calculate a density measure for each county. As noted in theory, the relevant density measure is in the neighborhood of the pivotal shoppers, which means in the presence of tax differentials, the relevant density will be inside the higher-tax county; when the tax differential is zero, it will be the density exactly at the border. Consider the three California counties and suppose for simplicity that driving to cross-border shop could only happen on the north-south corridor among these three counties; further assume the tax rate in Sacramento is higher than in San Joaquin, which is higher than in Stanislaus. The density for Sacramento would be the population of people just inside its border because it is higher tax. The density for Stanislaus would be the density just inside San Joaquin's borders because those are the shopers that lower-tax Stanislaus could attract for a marginal change in its rates. For San Joaquin, the density would be the sum of the people just inside Sacramento's plus the sum of people in San Joaquin just inside its own border with Stanislaus. We calculate such a metric of density for each border pair in the country and then aggreage across all border pairs for each county. We do this for both Census years 2000 and 2010 and therefore can also calcualte the change in density near the border.

To do this, for each year, we first determine the driving times for all Census block groups. We then determine whether the nearest neighboring county is high-tax, low-tax or same-tax.[17] Theory then suggests we need to determine the density in the neighborhood of a cutoff-point x^*_{ij} between jurisdiction i and j. Such theoretical guidance is too strong for empirical analysis as consumers at a given distance are likely also heterogeneous with respect to the value of their time and transport costs. Thus, empirically, we believe its likely that there are pivitol consumers at all distances between the border and x^*_{ij}.

To formalize this let the distance to the border range from point 0 (exactly on the border) to point $\bar{x}_{ij}$, the maximal driving time from a census block in county i to the border of county j. Thus, we define x^*_{ij} as the maximum distance in minutes at which a consumer would cross-border shop. This variable depends on the tax differential and we use the empirical literature on cross-border shopping to approximate x^*_{ij} as a linear function of the tax differential $x^*_{ij} = a + b\,|\tau_i - \tau_j|$.[18] Estimates of a and b are taken from

[17]To do this we use the state plus county tax rate.

[18]If a county is very small and the tax differential very large, it could be possible that $x^*_{ij} > \bar{x}_{ij}$. In

the empirical literature and we verify robustness. After determining x^*_{ij}, for each county, we determine the number of pivitol shoppers between the border (point $x_{ij} = 0$) and x^*_{ij}. For places with tax differentials, this is always the number of people in this region on the high-tax side. For same-tax borders, we define this as the number of people between $[0, a]$ on both sides of the border.[19] For subsequent ease of notation, write x^*_{ij} if it is inside county i and $-x^*_{ij}$ if it is inside county j. Formally, the density for county i, is given by

$$D_i = \sum_j 1^j_{\tau_i > \tau_j} \left[\sum_b 1^b_{[0, x^*_{ij}]} \phi_b \right] + \sum_j 1^j_{\tau_i < \tau_j} \left[\sum_b 1^b_{[-x^*_{ij}, 0]} \phi_b \right] + \sum_j 1^j_{\tau_i = \tau_j} \left[\sum_b 1^b_{[-a,a]} \phi_b / 2 \right] \tag{15}$$

where b indexes census blocks and ϕ_b is the population of block b. The indicator function 1^j determine if i is relatively high, low, or same tax and equal one if the subscript holds. The indicator function 1^b determine if Census block b is between the border and the given cutoff point and equals one if the subscript holds. The first term sums the population near the border inside of i near other low-tax counties; the second term sums the population near the border, but inside of j, when j is a high-tax county; the final term is the average density on either side of same-tax county borders.

Unfortunately, density is endogenous to the tax rate for two reasons: first, x^*_{ij} depends on the tax rates and second, because the side of the border that x^*_{ij} is on is determined by the tax differential. Only the last term in (15) is exogenous; it used the density on both sides of the border at within an exogenously given distance of a from the border. Thus, for every border pair, including same-tax and high-tax, we calculate

$$\widetilde{D}_i = \sum_j \left[\sum_b 1^b_{[-a,a]} \phi_b / 2 \right] \tag{16}$$

which is the average density at all county borders (for which there is a road crossing) located between the border and a driving time of a. $\widetilde{D}_i$ then acts as an instrument for D_i. The relevance of this instrument is as follows: in many, but not all cases, if one county has a large population at the border, the other county will also have a large population. The same is true for low-population areas. Note, here we use the instrument to rule out the fact that the postion of the piviotal consumers depends on the tax rates; we make no claim that this instrument would hold if our empirical specification is misspecified because of unobservable variables. In terms of the exlusion restriction, $\widetilde{D}_i$ is unlikely to

this case, we set t $x^*_{ij} = \bar{x}_{ij}$.

[19] As a technical matter, this interval is on both sides of the border. For the reason, we take the average number of people, where we average across both sides of the border. In theory, this should be the mass precisely at the border, but given we know centroids rather than the precise distribution of all people, this approximation is necessary.

Table 3: Cross-sectional Results

	(1)	(2)	(3)	(4)	(5)	(6)
Density	-0.136*	-0.147*	-0.149**	-0.125*	-0.130*	-0.131*
	(0.070)	(0.078)	(0.005)	(0.072)	(0.080)	(0.080)
Population	0.010**	0.010**		0.009**	0.010**	
	(0.004)	(0.005)		(0.004)	(0.005)	
Tax Base			0.011**			0.010**
			(0.005)			(0.005)
State FE	Y	Y	Y	Y	Y	Y
Controls	N	N	N	Y	Y	Y
Number	2374	2374	2374	2374	2374	2374
Kleibergen-Paap rk Wald F statistic	OLS	1251	545	OLS	1103	504

a = 3, b = 1/30

influence taxes directly except via the density expression.

We then proceed exploit cross-sectional and temporal variation. In our first-approach, because the effects of size and density are likely to be equilibirum phenomena, we exploit within-state cross-sectional variation for county i in state s:

$$\tau_{i(s)} = \zeta_s + \beta P_{i(s)} + \gamma D_{i(s)} + \theta X_{i(s)} + \varepsilon_{i(s)}, \tag{17}$$

where τ_i is the county tax rate, P_i is the county population, D_i is border density as calculated above, ζ_s are state fixed effects and X_i are other covariates. The state fixed-effects control for state instittuional features that affect all counties in the state, such as constraints on taxation authority or the size of the tax base. As the reader might worry about unobserved heterogeneity, we then exploit a long-difference that captures tax rate, density, and population changes over a decade. To do this, we estimate for county i in period t:

$$\tau_{i,t} = \zeta_i + \zeta_t + \beta P_{i,t} + \gamma D_{i,t} + \theta X_{i,t} + \varepsilon_{i,t}. \tag{18}$$

In both specifications we expect $\beta > 0$ and $\gamma < 0$. One concern with this specification is that in equilibirum, it is not population and density that matter, but rather the tax base and density that matter. For this reason we replace $P_{i,t}$ with the equilibirum number of shoppers in a jurisdiction. Note that the equlibrium number of shoppers is approximately $P_{i,t}$ plus the first two terms of (15). We can then instrument for the equilibrium tax base with population.

Table 4: Panel Results

	(1)	(2)	(3)	(4)	(5)	(6)
Density	-0.526	-0.616*	-0.615*	-0.507	-0.638**	-0.636**
	(0.440)	(0.346)	(0.345)	(0.420)	(0.300)	(0.298)
Population	-0.020	-0.014		-0.024	-0.015	
	(0.032)	(0.022)		(0.034)	(0.024)	
Tax Base			-0.014			-0.015
			(0.021)			(0.023)
State FE, Year FE	Y	Y	Y	Y	Y	Y
Controls	N	N	N	Y	Y	Y
Number	4744	4744	4744	4744	4744	4744
Kleibergen-Paap rk Wald F statistic	OLS	15	9	OLS	17	10

a = 3, b = 1/30

Table 5: Heterogeneity

	(1) std	(2) big area	(3) lowtax	(5)	(7)	(8
Density	-0.193*	-0.240***	-0.180	-1.039***	-0.208	-1.22
	(0.116)	(0.083)	(0.136)	(0.3811)	(0.335)	(0.3
Tax Base	0.009	0.014***	-0.024	-0.004	-0.013	-0.0
	(0.007)	(0.004)	(0.027)	(0.037)	(0.021)	(0.0
State FE	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y
Number of Obervations	1157	1189	1048	2310	2364	20
Kleibergen-Paap rk Wald F statistic	339	273	234	7	34	7

a = 3, b = 1/30

7 Implications for the empirical commodity tax competition literature

I think it is worthwhile to have a separate section to discuss the empirical implications. Previous studies usually used population size as a measure for tax competition because the Kanbur and Keen as well as the Nielsen model shows that tax rates will be affected by jurisdiction size. If this is what these studies do, population might measure something, presumably the size of the tax base, but I believe it is not tax competition what they measure. Indeed, if you derive the tax externalities in the Nielsen model, a change in b or L will not affect the revenues of the neighboring jurisdiction (remains to be checked). Put differently, suppose there is a line with three jurisdictions and there is a "population shock" or "tax base shock" in the left jurisdiction which increases the tax rate of this jurisdiction. This will, ceteris paribus, indirectly also lead to a higher tax rate in the right jurisdiction although these two jurisdiction do not compete with each other. In our model, the intensity of tax competition is measure by the population *density* at the location of the marginal consumers. In the line example, the left and the right jurisdiction are not competing because they do not have "common marginal consumers". Ideally, it would be nice to run a standard existing tax competition analysis and add the densities and hope that the population coefficient becomes insignificant. Let's wait and see.

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*, 104082.

Agrawal, David R., William H. Hoyt, and John D. Wilson. 2020. "Local Policy Choice: Theory and Empirics." in preparation for the Journal of Economic Literature.

Baldwin, Richard E., and Paul Krugman. 2004. "Agglomeration, Integration and Tax Harmonisation." *European Economic Review*, 48(1): 1–23.

Barseghyan, Levon, and Stephen Coate. 2016. "Property Taxation, Zoning, and Efficiency in a Dynamic Tiebout Model." *American Economic Journal: Economic Policy*, 8(3): 1–38.

Bleakley, Hoyt, and Jeffrey Lin. 2012. "Portage and Path Dependence." *Quarterly Journal of Economics*, 127(2): 587–644.

Brueckner, Jan H. 1987. "The Structure of Urban Equilibria: A Unified Treatment of the Muth-Mills Model." *Handbook of Regional and Urban Economics*, 2: 821–845.

Brueckner, Jan K. 2003. "Strategic Interaction Among Governments: An Overview of Empirical Studies." *International Regional Science Review*, 26(2): 175–188.

Brueckner, Jan K., and Luz A. Saavedra. 2001. "Do Local Governments Engage in Strategic Property-Tax Competition?" *National Tax Journal*, 54: 203–230.

Brülhart, Marius, Mario Jametti, and Kurt Schmidheiny. 2012. "Do Agglomeration Economies Reduce the Sensitivity of Firm Location to Tax Differentials?" *The Economic Journal*, 122(563): 1069–1093.

Bucovetsky, Sam. 1991. "Asymmetric Tax Competition." *Journal of Urban Economics*, 30(2): 167–181.

Calabrese, Stephen M., Dennis N. Epple, and Richard E. Romano. 2012. "Inefficiencies from Metropolitan Political and Fiscal Decentralization: Failures of Tiebout Competition." *Review of Economic Studies*, 79(3): 1081–1111.

Cameron, A. Colin, and Douglas L. Miller. 2015. "A Practitioner's Guide to Cluster-Robust Inference." *Journal of Human Resources*, 50(2): 317–372.

Campante, Filipe R., and Quoc-Anh Do. 2014. "Isolated Capital Cities, Accountability, and Corruption: Evidence from US States." *American Economic Review*, 104(8): 2456–2481.

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.

Dharmapala, Dhammika, and James R. Hines, Jr. 2009. "Which Countries Become Tax Havens?" *Journal of Public Economics*, 93(9-10): 1058–1068.

Epple, Dennis, and Allan Zelenitz. 1981. "The Implications of Competition Among Jurisdictions: Does Tiebout Need Politics?" *Journal of Political Economy*, 89(6): 1197–1217.

Eugster, Beatrix, and Raphaël Parchet. 2019. "Culture and Taxes." *Journal of Political Economy*, 127(1): 296–337.

Fajgelbaum, Pablo D., Eduardo Morales, Juan Carlos Suárez Serrato, and Owen Zidar. 2019. "State Taxes and Spatial Misallocation." *Review of Economic Studies*, 86(1): 333–376.

Hamilton, Bruce W. 1975. "Zoning and Property Taxation in a System of Local Governments." *Urban Studies*, 12(2): 205–211.

Hansen, Nico A., and Anke S. Kessler. 2001. "The Political Geography of Tax H(e)avens and Tax Hells." *American Economic Review*, 91(4): 1103–1115.

Harding, Matthew, Ephraim Leibtag, and Michael F. Lovenheim. 2012. "The Heterogeneous Geographic and Socioeconomic Incidence of Cigarette Taxes: Evidence from Nielsen Homescan Data." *American Economic Journal: Economic Policy*, 4(4): 169–198.

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. 2010. "Competition for Firms in an Oligopolistic Industry: The Impact of Economic Integration." *Journal of International Economics*, 80(2): 239–248.

Hines, Jr., James R. 2010. "Treasure Islands." *The Journal of Economic Perspectives*, 24(4): 103–125.

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. 1993. "Tax Competition, Nash Equilibria, and Residential Mobility." *Journal of Urban Economics*, 34(3): 358–379.

Jacob, Benoy, and Daniel McMillen. 2015. "Border Effects in Suburban Land Use." *National Tax Journal.*

Janeba, Eckhard, and Steffen Osterloh. 2013. "Tax and the City - A Theory of Local Tax Competition." *Journal of Public Economics*, 106: 89–100.

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.

Kessing, Sebastian. 2008. "Discriminatory Indirect Taxation in International Tax Competition." 2008 CESifo Conference Paper.

Lovenheim, Michael F. 2008. "How Far to the Border?: The Extent and Impact of Cross-Border Casual Cigarette Smuggling." *National Tax Journal*, 61(1): 7–33.

Lyytikäinen, Teemu. 2012. "Tax Competition among Local Governments: Evidence from a Property Tax Reform in Finland." *Journal of Public Economics*, 96(7-8): 584–595.

Mast, Evan. 2019. "Race to the Bottom? Local Tax Break Competition and Business Location." *American Economic Journal: Applied Economics.*

Mongrain, Steeve, and John D. Wilson. 2018. "Tax Competition with Heterogeneous Capital Mobility." *Journal of Public Economics*, 167: 177–189.

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. 1999. "Cross-border Shopping and Commodity Tax Competition among Governments." *Regional Science and Urban Economics*, 29(1): 33–51.

Parchet, Raphaël. 2019. "Are Local Tax Rates Strategic Complements or Strategic Substitutes?" *American Economic Journal: Economic Policy*, 11(2): 189–224.

Slemrod, Joel, and Wojciech Kopczuk. 2002. "The Optimal Elasticity of Taxable Income." *Journal of Public Economics*, 84: 91–112.

Trandel, Gregory A. 1994. "Interstate Commodity Tax Differentials and the Distribution of Residents." *Journal of Public Economics*, 53(3): 435–457.

Wilson, John Douglas. 1991. "Tax Competition with Interregional Differences in Factor Endowments." *Regional Science and Urban Economics*, 21(3): 423–451.

Wilson, John Douglas. 1997. "Property Taxation, Congestion and Local Public Goods." *Journal of Public Economics*, 64: 207–217.

Table A.1: Number of Counties, Own Tax Relative to Mean Neighbor Tax

	Own Tax Relative to Mean Neighbor Tax		
	Own > Neighbor	Own < Neighbor	Own = Neighbor
Total Cases	838	831	705
Own Population > Neighbor	232 [28%]	268 [32%]	215 [30%]
Own Population < Neighbor	606 [72%]	563 [68%]	490 [70%]

This table shows the count of counties in states allowing for county sales taxes for various cases. The columns indicate the county tax rate relative to the mean neighbor county tax rate. The rows indicate the county population relative to the mean neighbor population. A neighbor is defined as a county that shares a major road crossing. The mean is then calculated over all neighboring jurisdictions. Percent of total cases in that column are given in [ ].

A Appendices (for online publication only)

A.1 Data Appendix

A.2 Derivation of the optimal tax rates

A.3 Density functions

As noted in the text, the assumption of a periodic, continuous and differentiable function is not restrictive. The reason is that any function $f(x)$ can be approximated using a Fourier approximation $\mathcal{F}(x)$. This Fourier series can then be used in place of $f(x)$ in all formulas. A Fourier approximation of degree T with period L is given by

$$\mathcal{F}(x) = a_0 + \sum_{t=1}^{T} a_t cos(tx) + \sum_{t=1}^{T} b_t sin(tx) \tag{A.1}$$

where

$$a_0 = \frac{1}{L}\int_0^L f(x)\,dx, \tag{A.2}$$

$$a_t = \frac{2}{L}\int_0^L cos(\frac{2\pi tx}{L})\,f(x)\,dx, \tag{A.3}$$

$$b_t = \frac{2}{L}\int_0^L sin(\frac{2\pi tx}{L})\,f(x)\,dx. \tag{A.4}$$

This function then allows for easy integration and differentiability over the entire unit circle. For example, figure A.1 shows the Fourier series over the unit circle $[0, 2\pi)$ assuming the density function is $f(x) = x$ and for the density function with discontinuties given in Kanbur and Keen (1993). To show the Fourier series is periodic, we plot it on $[-2\pi, 2\pi]$.

As an alternative to Fourier approximations, one can obtain a large range of dis-

Figure A.1: Fourier Approximations

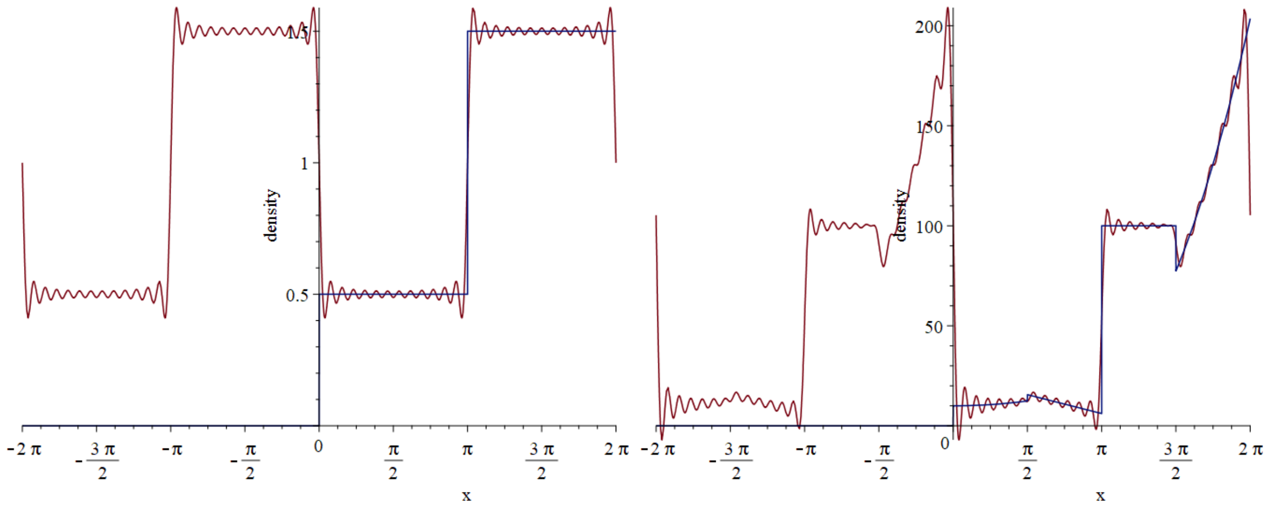
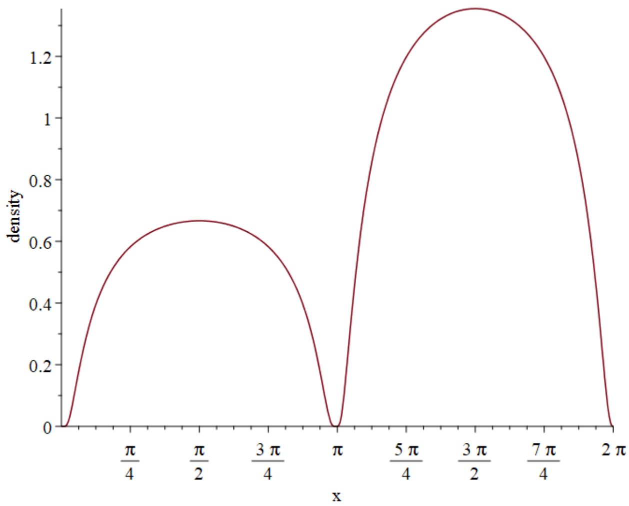


Figure A.2: Bump Functions



tributions using bump functions

$$g(x) = \begin{cases} k\, e^{\frac{1}{(x-a)(x-b)-c}} & a < x < b \\ 0 & otherwise \end{cases} \tag{A.5}$$

where (a, b) is the open interval, and c and k are positive constants. Note that $\frac{1}{(x-a)(x-b)-c}$ is negative and approaches negative infinity near the end points of the interval. Then, $g(x)$ approaches 0 smoothly (exhibits infinite differentiability) as x approaches a from above or b from below. Thus, it is differentiable at a and at b. Then, for various intervals, we can piece together several bump functions to obtain different densities across different jurisdictions. An example is given in figure A.2.

Moreover, now let's focus on jurisdiction 1, which extends (say) from point a to point b on the circle. It has a bump function which we can call g_1. Now identify two points strictly within that jurisdiction's interval, which we can name A_1 and B_1 (i.e., $a_1 < A_1 < B_1 < b_1$). Define the same kind of bump function as above only with the parameters A_1, B_1 instead of a and b. Call this function g_2. Now consider the function $\phi_1 = g_1 + g_2$. (Recall that the bump function is always zero outside of its support interval.) Then we have a new function for jurisdiction 1 that is continuous and infinitely differentiable. It can be connected smoothly at the boundary to the densities for neighboring jurisdictions. However, we have now introduced two new parameters, A_1 and B_1, which define the support of g_2. We can slide this support around while staying strictly with the support of g_1. As we do this, we do not change the integrals of the densities g_1, g_2, so total population remains fixed. However, the distribution of the population within jurisdiction 1 does change.