# Inter-Federation Competition:

## Sales Taxation with Multiple Federations

David R. Agrawal*
University of Georgia

This Version: August 2013

**Abstract**

Existing models of sales tax competition focus on competition within a federation. This paper analyzes how introducing multiple competing federal governments affects the strategic nature of sub-federal tax competition. In the context of a Nash equilibrium, the paper shows that lower levels of government will react heterogeneously to the tax rate of the higher level of government depending on the local government's proximity to the federation borders. Inter-federation competition will also introduce "diagonal externalities," which are fiscal externalities between neighboring jurisdictions that are of a different level of government. Diagonal tax competition will have a similar – but a distinct and smaller – effect on local tax rates compared to competition between neighboring jurisdictions of the same level. The paper uses two unique data sets to test for the effect of local fiscal competition: a cross-section of all local tax rates in the United States and spatial proximity data. The empirical specifications allow for vertical and horizontal externalities to have interaction effects and allow for strategic reactions that vary based on proximity to the nearest neighboring federation. Accounting for federal competition increases in absolute value the vertical strategic reaction by approximately 75%. A one percentage point increase in the county tax rate implies that municipal tax rates in that county will be approximately 0.40 percentage points lower. The results also indicate the sign and magnitude of the horizontal and vertical interactions are heterogeneous across localities. Diagonal interactions are found to have the same sign as horizontal interactions, but are smaller in magnitude.

JEL: H25, H71, H73, H77, R50

Keywords: Commodity Taxation, Tax Competition, Fiscal Federalism

*Contact information: University of Georgia Department of Economics, 527 Brooks Hall, Athens, GA 30602. Email: dagrawal@uga.edu. Thank you to the International Institute of Public Finance for awarding this paper the IIPF Young Economists Award. I am especially grateful to Joel Slemrod along with David Albouy, Robert Franzese, James R. Hines Jr. I also wish to thank Charles Brown, Raj Chetty, Paul Courant, Michael Devereux, Michael Gideon, Makoto Hasegawa, Ben Lockwood, Olga Malkova, Lovisa Persson, Jim Saunoris, Jeffrey Smith, Caroline Weber, David Wildasin, and Christobal Young for helpful suggestions and discussions. Suggestions from participants at Clemson University, the Michigan Public Finance Lunch Seminar, the 68th Annual Congress of the International Institute of Public Finance, the National Tax Association Annual Conference on Taxation, and the APPAM Fall Research Conference greatly improved the paper. Thanks to Pro Sales Tax for providing me access to the sales tax data; Nicole Scholtz provided excellent ArcGIS assistance. Any remaining errors are my own.

# 1 Introduction

Existing models of sales tax competition study multiple levels of government in the context of a unitary federation.[1] Most systems of government, however, are often characterized by multiple federations (i.e., fifty states in the United States with many sub-federal governments likes towns and counties). In such a context, how does introducing inter-federation competition into a model of sales taxation change the strategic interaction between governments of different levels? Inter-federation competition is traditionally used to mean competition among nations. However, I use the term inter-federation competition to highlight competition among any higher level of government within the federalist structure. As such, counties are federal to towns just as the national government is federal to the state governments. Throughout this paper, inter-federation competition will refer to competition among multiple counties that are composed of multiple towns, but the theoretical models in the paper could similarly apply to the national-state relationship as well. Any federalist system with three or more levels of government is multiple in the sense suggested by this paper. The model presented in this paper is the first analysis of sales tax competition allowing for both inter-federation competition (counties competing with other counties), intra-federation competition (competition of towns within the federation), along with competition across federal borders (competition with towns in neighboring federations and with neighboring counties).[2] The inclusion of inter-federation competition results in a model of spatial tax competition that is highly applicable to the local context.

Horizontal externalities occur between neighboring jurisdictions with separate tax bases. Vertical externalities arise between different levels of government that share the same tax base. Fiscal externalities arise because jurisdictions do not account for the effect that a tax rate change will have on a another jurisdiction's tax base. Inter-federation competition is essential to fully understand horizontal and vertical externalities, especially at the local level. Modeling and estimating vertical interactions with only one "federal" government ignores horizontal competition among neighboring federal governments. Inter-federation competition

[1] Models of sales tax competition giving authority in its own right to the federal government are Keen (1998) and Devereux, Lockwood and Redoano (2007). Keen (1998) assumes that no cross-border shopping can occur, while Devereux, Lockwood and Redoano (2007) relaxes this assumption with an emphasis on analyzing intra-federation competition. Hoyt (2001) allows for a dual purpose to the federal government – to maximize welfare and to make transfers to correct horizontal externalities. An extensive literature on sales tax competition with one level of government: Mintz and Tulkens (1986), Kanbur and Keen (1993), Trandel (1994), Nielsen (2001).

[2] Multiple federations are studied in the context of capital tax competition in Wilson and Janeba (2005), Kelders and Koethenbuerger (2010), and Becker and Büttner (2012). In addition to these articles, other papers – for example, Keen and Kotsogiannis (2002) and Keen and Kotsogiannis (2003) – focus on unitary federal relationships. Unlike competition for capital, the sales tax base is only locally mobile.

inevitably constrains the federal governments by inducing horizontal externalities at the federal level, while also triggering additional competition at the sub-federal level across federation borders. These differences have strong implications for estimating the strategic reaction functions in a federation. An assumption of a unitary federation may be valid when the federal government is the national government – as any leakage out of the United States boundaries is relatively small. However, when the "federal" governments are state or county governments, such an assumption is no longer valid. The more decentralized the "higher" level of government, the more likely it will have significant horizontal competitors of its own.[3] Empirically analyzing the reaction functions of governments has been mostly restricted to how states respond to national taxes. However, there is no reason to believe that the nature of the strategic reaction function of states is the same as that of localities – especially given that more decentralized levels of government have more horizontal competitors – not to mention institutional differences.

Allowing for multiple competing federations will result in "spatial economics" taking on two unique dimensions – "spatial interdependence" and "spatial location." Spatial interdependence is the process by which one jurisdiction has a contagion effect on another (perhaps neighboring) jurisdiction's tax rate. For example, when a jurisdiction sets a tax rate, it maximizes an objective function that aggregates the welfare of residents within the jurisdiction, but does so while competing with neighboring jurisdictions for a mobile tax base. This competitive process will influence the tax setting behavior of other geographically close and possibly overlapping jurisdictions. Spatial location is the process by which distance from or proximity to a particular geographic feature influences tax setting behavior.[4] For example, tax rates may be a function of proximity to a border or to an amenity. In the presence of multiple federations, the strength of spatial interdependence can be *heterogeneous* with respect to a jurisdiction's spatial location.

Spatial location effects arise in a model with multiple federations because governments have multiple borders. In traditional single federation models, there are usually only two-sub federal governments. With only two sub-federal governments, both governments have only

[3]The literature on tax competition has several interesting parallels with the industrial organization literature. Mathewson and Winter (1984) and Rey and Tirole (1986) starts by analyzing the problem of a single manufacturer with several retailers; the existing tax competition literature has focused on a single federation with multiple sub-federal governments. Bonanno and Vickers (1988) expanded the literature on vertical integration by considering the case of two manufacturers each with one retailer. Saggi and Vettas (2002) allow for markets with multiple upstream and downstream firms, which results in both intrabrand and interbrand competition. Such a setup is analogous to this paper, which will allow for both inter-federation and intra-federation competition.

[4]The concept of "spatial location" is similar to the characteristics-based approach of Gorman (1980) and Lancaster (1966). The location of sale within a county can be thought of as a differentiating characteristic to goods that are identical in all other respects.

one border (one competitor). In such a model, all members of the federation are peripheral (they are located at the federation's borders). In reality, not all members of a federation are located at the border. This paper shows that vertical externalities make it so that the strategic reaction of a locality to its county tax rate is more likely to be negative for jurisdictions located at the county's periphery. In addition to this heterogeneity, the presence of fluid federal borders will give rise to a new type of externality not yet discovered in the literature: "diagonal externalities." A diagonal externality is the effect of a county's tax rate on a municipality in the *neighboring* county. The diagonal externality fails to have the main feature of the vertical externalities – the shared tax base – and lacks one feature of a horizontal externality – the same level of government. For example, if the state of California increases its tax rate, it increases the tax base in Clark County in Nevada. However, the channel by which this externality occurs is very different than the horizontal externality imposed on Clark County in Nevada if San Bernadino County in California had increased its tax rate.[5] Although what matters for cross-border shoppers is the final price of a good, irrespective of which level of government taxes it, a diagonal externality plays a different role than a horizontal externality. The strategic reaction to a neighboring county tax rate will always be less intense than a similar increase in the average neighboring town rates. These differences arise because county borders do not surround the entire municipality and because some towns are relatively far away from nearby counties.

Because of these theoretical considerations, the estimation strategy for determining the strategic response to vertical and horizontal externalities becomes more complex in the presence of inter-federation competition.[6] The empirical methodology presented in this paper indicates, even if the federal government has no horizontal competitors of its own, it is essential to consider the interaction of horizontal and vertical competition. Additionally, when the federal government has horizontal competitors, externalities induced by the federal government on its competitor must also be considered. The empirical results suggest that, with multiple federations, in order to obtain accurate estimates the researcher must empirically allow for strategic reactions to vary based on distance to the federation's borders. The paper also uses decentralized data within federations to determine if the strategic reaction at lower levels of government differs from existing estimates of the state-national interactions mainly estimated in the literature. These modifications to the methods complicate the analysis, but they also provide the researcher with the ability to identify heterogeneous strategic inter-

[5] Geys and Osterloh (forthcoming) show that government officials near borders do not view tax competition as being constrained to neighbors within a federation.

[6] Besley and Rosen (1998), Esteller-Moré and Solé-Ollé (2001) , Hayashi and Boadway (2001), Fredriksson and Mamun (2008), Burge and Rogers (2011) and Brülhart and Jametti (2006) are examples of papers that estimate both horizontal and vertical interactions.

actions. If these heterogeneities match the theoretical predictions, the researcher can more convincingly argue that the identified effects are indeed the result of tax competition rather than the result of common shocks.

The baseline empirical results show that a one percentage point increase in county sales tax rates reduces municipal sales tax rates by .40 percentage points. For the horizontal interaction, a one percentage point increase in the average of a town's neighbors' tax rates will increase a municipality's local sales tax by .59 percentage points. Accounting for horizontal and vertical interaction effects, diagonal externalities, and distance based effects resulting from inter-federation competition increases (in absolute value) the vertical strategic reaction by approximately 75% relative to the baseline specification estimated in the literature. The bias in previous estimates results, in a large part, because a good sold in one town is not identical to a good sold in another town. Although the same good sold in two towns may look identical, they differ in their characteristics on the basis of the location of sale within the county. Because the goods can be viewed as distinct along one characteristic (spatial location within a county), the researcher must account for this difference and its influence on the cross-price elasticities across jurisdictions.

# 2 Model

The model expands Devereux, Lockwood and Redoano (2007). The geographic setup of the model features sub-federal jurisdictions (towns) located on a possibly infinite length horizontal line segment. Federations (counties) are indexed $j = 1, 2, ....$ and each county is composed of $m$ towns. Towns are indexed $i = 1, 2, ....$ and the index does not reset across counties. The federations in the model compete with each other and with the towns in the model. The towns compete with other towns and with the federations. All of the federations are within a common union (state). Each town has $n_i$ residents and is $l_i$ units long on the line segment. The relationship between length and population is that $n_i = \phi_i l_i$ so that $\phi_i$ denotes the population density. Consumers and producers are located everywhere.

Consumers have preferences over a consumption good and another untaxed good (i.e., dollars or leisure). I assume that the producer price of the consumption good is fixed across all towns and normalize it to one. Demand for the taxed good is denoted $x$. The untaxed good is assumed to be the numéraire. Every consumer has a utility function $u(x, \bullet)$ that is strictly increasing and concave with respect to $x$. Utility is linear in the untaxed numéraire.

Every level of government can set a specific commodity tax on the consumption good. Taxes based on the location of the transaction – not the consumer's residence.[7] Towns

[7]In the United States, use taxes are levied under the destination principle, but the use tax is notoriously under-enforced and evaded. The model is analogous to levying use taxes with no enforcement.

(sub-federal governments) set a local tax rate $t_{ij} \equiv t_i$, counties (federal governments) set a county tax rate $\tau_j$ that applies to all towns within the county, and the state government (the union government) sets an exogenously given state tax rate $T$. In any specific town, the after tax price $q_{ij} \equiv q_i$ is equal to $1 + t_i + \tau_j + T$. Governments compete in a Nash game. Counties and towns are considered simultaneous movers in the game.[8] The objective function of governments is to maximize revenue.

When purchasing the consumption good in the home town of residence, no transportation cost is incurred and the resident pays $q_i x_i$. Alternatively, the shopper can purchase the good in a neighboring town. If the resident of town $i$ shops in jurisdiction $k \neq i$, the individual will pay $q_k x_k$ plus any transportation cost of traveling to the border. The transportation cost function is assumed to be linear in distance to the border, $d$, such that the cost is $c_i d$. All cross-border shoppers will purchase the good from the first store in the neighboring jurisdiction and are constrained from shopping multiple towns over.[9]

Individuals will cross-border shop if the utility benefit from shopping abroad exceeds the utility received from shopping at home. Denote $v(q) = \max\left[u(x) - qx\right]$ as the indirect utility from the taxed good. Denote $x(q) = \operatorname{argmax}\left[u(x) - qx\right]$ as the demand for the taxed good for a resident of town $i$ when the price of the good is $q$. Comparing the indirect utility from cross-border shopping with the transportation cost function, a consumer living in $i$ will only shop in $k \neq i$ if $q_i > q_k$ and if she lives at a distance from the border of town $k$ of

$$d \leq \frac{v(q_k) - v(q_i)}{c_i}. \tag{1}$$

The tax base is defined as the sum of residents who shop at home plus the individuals that cross-border shop, which are multiplied by the demand function $x(q)$ to account for elastic demand. In order to define the tax base, the direction of cross-border shopping needs to be specified. Defining $\rho_i = \frac{\phi_i}{c_i}$ and noting that because each jurisdiction has two neighbors, there are four possible cases, tax revenue for towns can now be written as:

$$R_i = \begin{cases} t_i x(q_i)[n_i + \rho_{i+1}(v(q_i) - v(q_{i+1})) + \rho_i(v(q_i) - v(q_{i-1}))] & \text{if } q_i \leq q_{i+1} \,\&\, q_i \geq q_{i-1} \\ t_i x(q_i)[n_i + \rho_i(v(q_i) - v(q_{i+1})) + \rho_{i+1}(v(q_i) - v(q_{i-1}))] & \text{if } q_i \geq q_{i+1} \,\&\, q_i \leq q_{i-1} \\ t_i x(q_i)[n_i + \rho_i(v(q_i) - v(q_{i+1})) + \rho_i(v(q_i) - v(q_{i-1}))] & \text{if } q_i \geq q_{i+1} \,\&\, q_i \geq q_{i-1} \\ t_i x(q_i)[n_i + \rho_{i+1}(v(q_i) - v(q_{i+1})) + \rho_{i+1}(v(q_i) - v(q_{i-1}))] & \text{if } q_i \leq q_{i+1} \,\&\, q_i \leq q_{i-1} \end{cases}, \tag{2}$$

[8]While a leader-follower assumption is realistic for higher levels of government (national), it seems plausible that towns and counties are simultaneous movers.

[9]Relaxing the assumption of shopping in only one town over adds significant complexity. The assumption of shopping one town over will create stark results for contiguous neighbors.

where the term in [ ] of the town tax base is defined as $s_i$ and $\rho_i$ is interpreted as the intensity of the horizontal competition.

I assume that a Nash equilibrium exist.[10] Specifically, in order to reduce the possible number of cases from the revenue function, assume that in equilibrium $q_1 \leq q_2 \leq q_3$....[11]

The local tax rates in this equilibrium are implicitly defined by the reaction function:

$$\frac{\partial R_i}{\partial t_i} = B_i + t_i \frac{\partial B_i}{\partial q_i} = x_i s_i + t_i s_i x_i' - t_i x_i^2(\rho_i + \rho_{i+1}) = 0, \tag{3}$$

where $x' = \frac{\partial x(q_i)}{\partial q_i}$. The reaction function depends on the responsiveness of cross-border shoppers out of $i$ via $\rho_i$ and into $i$ via $\rho_{i+1}$.

The expression can be rewritten an inverse elasticity rule for town tax rates:

$$\frac{t_i}{q_i} = \frac{1}{-\frac{q_i}{B_i}\frac{\partial B_i}{\partial q_i}} = \frac{1}{\varepsilon_i + \theta_i}, \tag{4}$$

where

$$\varepsilon_i = -\frac{q_i x_i'}{x_i} \tag{5}$$

is the elasticity of demand for the consumption good and

$$\theta_i = \frac{q_i x_i(\rho_i + \rho_{i+1})}{s_i} \tag{6}$$

is the elasticity of the number of shoppers in town $i$. It differs from Devereux, Lockwood and Redoano (2007) by accounting for both the in- and out-flows in two directions. Both $\varepsilon_i$ and $\theta_i$ are defined to be positive numbers under the assumption that demand curves slope downward. Let $\eta_i = \frac{q_i x_i''}{x_i'}$ denote the curvature of the demand function. Similarly for counties, an inverse elasticity rule can also be defined as in equation 4.

In order to characterize the strategic interactions that follow, define the following cross-price elasticity, which is scaled by the price ratios:

$$\theta_{i,k} = \frac{q_k}{B_i}\frac{\partial B_i}{\partial q_k}\frac{q_i}{q_k} = \begin{cases} \frac{\rho_k x_k q_k}{s_i}\frac{q_i}{q_k} & \text{for } i < k = i+1 \\ \frac{\rho_i x_k q_k}{s_i}\frac{q_i}{q_k} & \text{for } i > k = i-1 \end{cases} \tag{7}$$

Cross-price elasticities arise if we think of this problem in a Gorman (1980) and Lancaster

[10] This paper is not focused on characterizing the existence and uniqueness of an equilibrium, so, as in other papers in this literature, I assume the existence of a Nash equilibrium.

[11] In the subsequent sections, I will characterize the a symmetric equilibrium where $q_1 = q_2 = q_3$.... For the moment, I characterize a more general solution because the appendix details the solution to the asymmetric case where taxes are monotonically increasing.

(1966) framework where towns sell goods that are identical in all respects except that the goods differ on the basis of one characteristic: the location of sale. The interpretation of $\theta_{i,k}$ is the elasticity of the number of shoppers purchasing goods in town $i$ with respect to a change in the neighboring price. Note that if in a symmetric equilibrium, $\theta_i$ is two times $\theta_{i,k}$ because the town's own price influences two borders.

Define $D_i = -\frac{\partial^2 R_i}{dt_i^2} = 2\varepsilon^2 + 2\theta_i^2 + \varepsilon\eta + \theta_i\varepsilon$ as the negative of the second derivative of the town's revenue function. It is guaranteed that $D_i$ is positive.

## 2.1 Strategic Interaction

The goal of the subsequent exercise is to derive the slope of the reaction functions with respect to the tax rates of a higher level of government in the neighborhood of the equilibrium. Figure 1 presents two possible reaction functions. In the left panel, the reaction function is upward sloping – implying that increases in the county or state rate will raise the town tax rate. In the right panel, the reaction function slopes down – implying that increases in the county rate result in lower town rates. The dotted lines are examples of reaction functions that respond most aggressively to the tax rate. Reaction functions with larger slopes (in absolute values) will be the most responsive to changes in federal rates. An increase in the slope of the reaction function for the right graph implies that the reaction function becomes flatter – and may even change the sign of the slope.

The equations above implicitly determine tax rates as a function of the county and state tax rates. Using the implicit function theorem and Roy's identity, the slopes of the reaction functions can be calculated. It is useful to state the conditions found in the literature for upward and downward sloping reaction functions. In a unitary federation with a symmetric equilibrium, the horizontal interaction is always positive and if $\theta_i - \varepsilon_i - \eta_i > 0$, the reaction function to the federal tax is upward sloping. If less than zero, it is downward sloping. Thus, $\theta_i > \varepsilon_i + \eta_i$ implies that the elasticity of cross-border shopping is large relative to the demand elasticity and demand characteristics. $\theta_i$ represents the strength of horizontal tax competition and can increase if transportation costs fall or the population density increases near the border. In Devereux, Lockwood and Redoano (2007), the assumption of iso-elastic demand guarantees the vertical reaction function is positive, but this will not be the case in the presence of inter-federation competition. I make three simplifying assumptions.

**Assumption 1:** All towns are identical with respect to the model's parameters.

**Assumption 2**: Demand is iso-elastic. This assumption simplifies the intuition by holding $x_i''$ and $x_i'$ constant along the demand curve such that the demand elasticity $\varepsilon_i = \varepsilon$ and curvature $\eta_i = \eta = -(1 + \varepsilon)$ are constant. The assumption will be important for the asymmetric case in the appendix.

**Assumption 3**: The Nash equilibrium is symmetric. In all of the subsequent reaction functions, I derive the reaction function for the case $q_1 \le q_2 \le q_3 ....$ and then evaluate it at the symmetric equilibrium $q_1 = q_2 = q_3 ....$

### 2.1.1 Vertical Reactions

A town is called "interior" if it borders two towns within the same county. A town is called "peripheral" if it borders one town in another county. I am interested in showing whether the strategic reaction by a town to its own county's tax rate ($\frac{\partial t_i}{\partial \tau_j}$ for town $i$ in county $j$) varies by whether a town is internal or peripheral to the county borders.

The slope of the reaction function for towns with respect to its own county's tax rate is calculated by differentiating equation 4. For town $i$ located in county $j$, the slope of the vertical reaction function for the town with respect to its own county's tax rate (evaluated at the symmetric equilibrium) is given by:

$$\frac{\partial t_{ij}}{\partial \tau_j} = \begin{cases} [\varepsilon(\theta_i - \varepsilon - \eta) + \theta_i(\theta_{i,k} - \theta_i)]/D_i = [\varepsilon(\theta - \varepsilon - \eta) - \theta^2/2]/D \gtreqless 0 & \text{for peripheral towns} \\ [\varepsilon(\theta_i - \varepsilon - \eta) + \theta_i(\theta_{i,i+1} + \theta_{i,i-1} - \theta_i)]/D_i = [\varepsilon(\theta - \varepsilon - \eta)]/D \gtreqless 0 & \text{for interior towns} \end{cases} \tag{8}$$

where $k = i+1$ if $i$ is the left-most town in a county or where $k = i-1$ if $i$ is the right-most town in a county and $[\varepsilon(\theta - \varepsilon - \eta) - \theta^2/2]/D < [\varepsilon(\theta - \varepsilon - \eta)]/D$. The derivation of the reaction function is in the appendix. In the appendix, I also demonstrate how an asymmetric equilibrium affects the underlying results.

**Proposition 1.** *In the neighborhood of a symmetric Nash Equilibrium, the slope of a towns' reaction functions with respect to the county tax rate is larger for interior towns than for periphery towns.*

*(1) If the slope of a town's reaction function with respect to its county tax rate is upward sloping, then the reaction function will be steeper for towns at the interior of the county than for towns at the periphery.*

*(2) If the slope of a town's reaction function with respect to its county tax rate is downward sloping, then the reaction function will be less steep (more likely to be positive) for towns at the interior of the county than for towns at the periphery.*

Consider a town at the interior of the county. This town neighbors two other towns that fall under the jurisdiction of the same county rate. Changes in the county tax rate will directly affect the tax base through demand: $x$ is a function of $q$. However, because the price rises by the same amount in both neighboring locations, cross-border shopping remains similar to the symmetric equilibrium. On the other hand, for a town at the county border, the change in the county tax rate will directly distort the demand function for individuals and the number of individuals cross-border shopping in one direction. The post-tax price remains unchanged at one of the town's borders. For a border town, an increase in the county rate will substantially increase outflows. Either way, the tax base becomes much smaller for a periphery town relative to an interior town. The peripheral town has more incentive to lower its rate in response to a high county rate relative to the interior neighbor.

The intuition of proposition one is that periphery towns want to capture additional cross-border shoppers from the neighboring county if their county lowers its rate or they want to discourage their residents from leaving the county if their county raises its rate. For interior towns, the pressures from cross-border shopping are muted.

#### 2.1.2 Horizontal Reactions

The slope of the reaction function for town $i$'s tax rate with respect to neighboring town $k$'s tax rate is given by

$$\frac{\partial t_i}{\partial t_k} = [\theta_i \theta_{i,k}]/D_i = \frac{\theta^2}{2}/D > 0, \tag{9}$$

where $k = i + 1$ or $k = i - 1$. But, the above expression tells the researcher what happens if one neighboring tax rate increases by one unit. But, the town has two neighbors and empirically, the researcher can estimate strategic interactions based on a one unit increase in the *average* neighbor's tax rate. As such, the effect of an increase in the average of the neighboring jurisdiction's tax rates by one unit is equal to

$$\frac{\partial t_i}{\partial t_{i-1}} + \frac{\partial t_i}{\partial t_{i+1}} = \theta^2/D > 0. \tag{10}$$

#### 2.1.3 Diagonal Reactions

Because there are multiple federations, towns at the eastern border of county $j$ directly compete with the border town in county $j + 1$ as well as county $j + 1$. This occurs because the peripheral town shares a border with the neighboring county. When the neighboring county changes its tax rate, towns bordering it will react to this change and adjust their tax rates accordingly. Such an relationship is horizontal because it involves a neighbor's tax rate, but it is vertical because it is with respect to another level of government. Call the resulting externality a diagonal externality. Thus, a diagonal externality on jurisdiction $i$ is the externality imposed by a neighboring jurisdiction that does not cohabit the same tax base as jurisdiction $i$, but that is not at the same level of governance. One example of a diagonal externality is the relationship between Clark County, Nevada and the state of

California. Clark County borders California but is not in California. However, if the state of California raises its tax rate, Clark County will see an increase its tax base. This effect is at a different level of government and California's externality on Clark County should not be mistaken for a horizontal relationship. Formally:

**Definition.** Consider two governments at *different* levels of the federalist hierarchy that do *not share* the same tax base. A diagonal externality results when a tax increase [decrease] levied by one level of government *increases [decreases]* the size of the tax base of the other level of government.

In contrast, a vertical externality arises when *different* levels of government *share* the same tax base and an increase in taxes levied by one level *reduces* the tax base of the other level.[12] It is also different from a horizontal fiscal externality. A horizontal externality arises when the governments are of the *same* level and therefore do *not share* the same tax base (rather they compete for it), which implies that an increase in taxes levied by one government *increases* the the tax base of the other government. Unlike Wilson and Janeba (2005) – where the total tax rate is what matters – the following proposition shows diagonal externalities are different than horizontal externalities in magnitude.

**Proposition 2.** *In the neighborhood of a symmetric Nash equilibrium, the slope of a peripheral town's reaction function with respect to the neighboring county tax rate is identical in sign but smaller in magnitude to the slope of the reaction function with respect to a similar change in the average of the neighboring towns' tax rates. For an interior town, the slope of the reaction function is smaller than the peripheral town.*

Proposition 2 can be shown by deriving the reaction function for an internal town $i$ in county $j$ with respect to county $j+1$'s tax rate. When town $i$ is located at the eastern periphery of its own county, this strategic reaction can be found by differentiating equation 4 with respect to $\tau_{j+1}$. Letting the neighboring county's tax rate be defined as $\tau_r$, the slope of the reaction function of a border town $i$ with respect to neighboring county's tax rate is

$$\frac{\partial t_{ij}}{\partial \tau_r} = \begin{cases} [\theta_i \theta_{i,k}]/D_i = \frac{\theta^2}{2}/D > 0 & \text{for peripheral towns} \\ 0 & \text{for interior towns} \end{cases} \tag{11}$$

where $k = i+1$ and $r = j+1$ if the neighboring county is at the town's eastern border or $k = i-1$ and $r = j-1$ if the neighboring county is at the town's western border. Notice that as a town is further from the border, the diagonal reaction shrinks; under the assumption that shopping is only one town over, the magnitude shrinks to zero.

[12] Note that a diagonal externality lacks the main feature of a vertical externality: the shared base aspect. The diagonal externality is only vertical in the sense that it occurs between governments of different levels.

This may seem intuitive and simple, but the theoretical and empirical literature have ignored this form of tax competition. Failing to account for this form of tax competition in any regression will yield biased estimates of the horizontal interactions if the researcher specifically desires to estimate the same-level-of-government horizontal effect. Notice that the slope of the reaction function with respect to the neighboring county tax rate is unambiguously positive – which is true of standard horizontal tax competition models. This implies that town tax rates are strategic complements with respect to the neighboring county rate. However, the effect of the diagonal interaction is smaller than the effect of an equivalent change in the average neighboring town tax rate because (1) neighboring towns entirely surround town $i$ and (2) the diagonal interaction depends on proximity to the border.

Intuitively, consider a town that has only two neighbors and one neighboring town is in a separate county. If the neighboring town raises its tax rate, this gives rise to a classical horizontal externality. When the neighboring county raises its rate, it is as if the neighboring town were raising its tax rate. For the consumer deciding where to shop, she does not care if the rate in the neighboring jurisdiction rose because of the town raising its rate or a higher level-of government (which has no jurisdiction over her home rate) raising its rate. However, in the data, if the average of the two neighboring town tax rates increase one percentage point, the town should react more intensely compared to if the neighboring county (there is only one) raises its tax rate one percentage point.

# 3 Implications for Empirical Analysis

The model above highlights several implications for the empirical analysis. The traditional empirical analysis is given by:

$$t_{ij} = \alpha_0 + \alpha_1 t_{-i} + \alpha_2 \tau_j + \delta X_{ij} + \epsilon_{ij} \qquad (12)$$

where $X$ are local controls. The variable on the left ($t_{ij}$) is the local tax rate of town $i$ in county $j$. Define $t_{-i}$ as the weighted average of other neighboring tax rates such that

$$t_{-i} = \sum_{k \neq i} w_{ik} t_k. \qquad (13)$$

Denote $w_{ik}$ as exogenous weights normalized such that $\Sigma_{k \neq i} w_{ik} = 1$.[13] Tax rates on the right-hand side are usually instrumented for with the weighted average of the neighbors' $X$. The coefficient on $t_{-i}$ identifies the effect of a jurisdiction's neighboring tax rates. The

[13]See Brueckner (2003) for a survey of weighting schemes used in the literature.

coefficient on $\tau_j$ estimates the effect of the federal tax rate.

The theoretical results above show the following. One, distance to the county border should be included to account for the fact that the slope of the reaction function is different for towns near a border. Two, the right side of the regression equation must include the neighboring county rate in addition to the neighboring town rate. This variable may be interacted with a distance variable for towns that are close to the neighboring county border. Three, the right side must include interactions of the county rate with respect to other neighboring town tax rates on the right side.

## 3.1 Why Interaction Effects Are Essential

To understand the final point, note that from the slopes of the reaction functions, it is evident that $\theta_i$ captures (in part) the strength of horizontal tax competition. Because the slopes of the reaction functions depend on $\theta_i$, the reaction to the higher level of government becomes more intense as the horizontal interaction is increased. Devereux, Lockwood and Redoano (2007) recognize "...there is an interaction between vertical and horizontal tax competition. ...an increase in horizontal tax competition makes it more likely that the vertical slope is positive" but do not include an interaction term.

A specification omitting this interaction may suffer an omitted variable bias. To see this empirically, it is useful to consider a multi-level model of tax competition.[14] Letting $i$ to continue to index the local government and $j$ to index the county level of government, consider the following multi-level model. For simplicity, consider the following univariate regression of local tax rates on neighboring tax rates

$$t_{ij} = \alpha_{0j} + \alpha_{1j} t_{-i} + \epsilon_{ij}, \tag{14}$$

but where it is also known that each $i$ jurisdiction is within a $j$ jurisdiction. As a result of having multiple levels of government, it is known that county tax rates $\tau_j$, which only vary across the $j$ level and not within the $j$ level of the model, affect $t_{ij}$ with some error. The following equation demonstrates this effect:

$$\alpha_{0j} = \gamma_{00} + \gamma_{01}\tau_j + u_{0j}. \tag{15}$$

It is clear that substituting (15) into (14) will yield (12) without controls. This is where the literature on tax competition within federations stops. However, the theoretical results

[14]For a summary of multi-level modeling, see Franzese (2005). It is useful to think of this empirical problem in the context of a multi-level model. However, the empirical strategy that will follow this section will use a reduced form setup to the problem. The reason for this is that estimating the model using hierarchical linear modeling will place stricter assumptions on the problem.

in this paper indicate that the empirical specification is further complicated by an interaction effect which also determines $t_{ij}$. From the theory, the effect of $t_{-i}$ depends on $\tau_j$ and vice versa. Assuming that this interaction also occurs with error, this implies

$$\alpha_{1j} = \gamma_{10} + \gamma_{11}\tau_j + u_{1j}, \tag{16}$$

and substituting (15) and (16) into (14) yields

$$t_{ij} = \gamma_{00} + \gamma_{01}\tau_j + (\gamma_{10} + \gamma_{11}\tau_j)t_{-i} + u_{1j}t_{-i} + u_{0j} + \epsilon_{ij}, \tag{17}$$

which is not the same as equation (12). The implication is that the existing literature has estimated the effect of the federal tax rate as $\frac{\partial t}{\partial \tau} = \gamma_{01}$ despite the true effect being $\frac{\partial t}{\partial \tau} = \gamma_{01} + \gamma_{11}t_{-i}$ and where the estimates of $\gamma_{01}$ are different across the two specifications because they are derived from different models.

# 4 Empirical Methodology

## 4.1 Data

I have unique cross-sectional data from April 2010 that includes local, county, and state sales tax rates for all jurisdictions in the United States.[15] Local tax rates share the same tax base as state tax rates. Municipal, county, and state rates are added together to obtain the total tax rate in a jurisdiction; the county and town sales tax rates are determined through separate political processes. In unincorporated areas only the county and state tax rate apply to a sale; but in incorporated areas the sum of the town, county, and state rate matter. I only use states where both towns and counties have tax setting authority in the analysis, however, when calculating the average neighboring town tax rates, all states are used. Twenty-eight states have sales taxation at three levels of government.

In addition to the tax data, I have also generated several comprehensive data sets concerning the spatial proximity of jurisdictions. I have generated the following variables: the driving time from the population weighted centroid of each Census Place[16] to the nearest intersection of a major road crossing at state and county borders – denoted $d$; measures of the proximity of Census Places to other places; and each jurisdiction's perimeter and area.

Driving times are measured from the population weighted centroid of each Census Place to the nearest intersection of a major road and a state or county border crossing. The driving

[15] The data are proprietary. For a complete description of the data see http://www.prosalestax.com/

[16] A Census Place is generally an incorporated place with an active government and definite geographic boundaries such as a city, town, or village. The reason I do not use the "town" as the level of jurisdiction is that the Census Place is the closest level of statistical analysis for which control variables are available.

time calculated is the time that minimizes the time to drive to the closest border. For a detailed description of this calculation and the tax data, see Agrawal (2011).

Contiguity is often not a satisfactory measure of horizontal competition because many places are small and may have zero contiguous neighbors. To calculate a broader measures of neighborliness, I define a jurisdiction as neighbors if they are within a fifty miles of each other. To measure diagonal tax competition, I define the neighboring county as the county that is closest to the town based on the driving time criteria.

Control variables are from the 2010 United States Census plus geographic and political controls. Table 1 lists all of the control variables used in the regression equation, along with summary statistics. In addition, Agrawal (2011) shows that distance to the nearest state border is an essential determinant of local tax rates. Therefore, as controls, I include the tax differential at the nearest state border, a dummy for whether the town is in a high- or same-tax state relative to the nearest neighbor, the log of driving time to the nearest state border and a complete set of interactions of these variables.[17] Including these terms will help to control for diagonal externalities across state lines.

## 4.2 Estimation in a Cross-Section

I will focus on how municipalities react to the county level rates, neighboring county rates, and neighboring town rates.[18] After defining all terms in the following paragraphs, I estimate

$$\begin{aligned} t_{ij} = \quad & \alpha_0 \underbrace{+\alpha_1 \tau_{i,j}}_{Vertical} \underbrace{+\alpha_2 t_{-i}}_{Horizontal} \underbrace{+\alpha_3 t_{-i}\tau_{i,j}}_{Interaction} \underbrace{+\alpha_4 \tau_{i,-j}}_{Diagonal} \\ & \underbrace{+\alpha_5 \tau_{i,j} d_i}_{Vertical-Distance} \underbrace{+\alpha_6 \tau_{i,-j} d_i}_{Diagonal-Distance} \\ & +X_{ij}\beta + \zeta + \epsilon_{ij} \end{aligned} \tag{18}$$

where $X_{ij}$ are the local controls listed in table 1 and $\zeta$ are state fixed effects – that control for the level of the state tax rate in a state along with other within state policies.[19] The tax

[17]Imposing this log functional form is consistent with Agrawal (2011) where the marginal effects of distance are steepest near the border and decreasing to zero in absolute value. State tax reaction functions are not estimated directly because state fixed effects prohibit the inclusion of state tax rates directly.

[18]One could also estimate separate reaction functions for county tax rates with respect to neighboring counties, town tax rates within the county, and town tax rates outside of the county. Such an estimation strategy is beyond the scope of the paper given that much of the existing literature focuses on estimating vertical reaction functions for the lower levels of government. Estimating these reaction functions for the higher level of government is possible in this case given that counties are likely quite sensitive to town decisions, but I believe this is a separate paper.

[19]Elhorst (2010) compares spatial lag models with more general models such as the spatial Durbin model. I use a spatial lag model because as noted in Elhorst (2010) [p. 15], single equation maximum likelihood procedures cannot easily handle multiple endogenous regressors and the ML estimators of models such as the spatial Durbin model would be "difficult, if not impossible, to derive"

rate $t_{ij}$ is the municipal (inclusive of special district) local option tax and $t_{-i}$ is defined in equation 13. Specifically, if town $k$ is within fifty miles of town $i$, the weights in the main specification are equal to one divided by the number of jurisdictions within fifty miles of town $i$ and zero otherwise. Thus, the interpretation of $t_{-i}$ is the average tax rate of town $i$'s neighbors. Defining $N_i$ as the set of towns within a fifty mile region of town $i$, then

$$w_{ik} = \begin{cases} \frac{1}{n_i} & \text{if } k \in N_i \\ 0 & \text{if } k \notin N_i \end{cases} \tag{19}$$

where $n_i$ is the number of towns in $N_i$.

Define $\tau_{i,j}$ as the county tax rate that town $i$ is located in and $\tau_{i,-j}$ is the nearest neighboring county's tax rate to town $i$. Because $\tau_{i,-j}$ is not a weighted average of all the neighboring counties, I implicitly assume that diagonal tax competition only manifests itself for the nearest county neighbor. One reason for this assumption is that I would like to test how the diagonal effect varies with distance, which would not be feasible if multiple counties are considered as neighbors. Making this assumption allows me to reduce what would be a multidimensional problem into a single dimension. Finally, in a similar spirit, $d_i$ is a measure of distance from the town centroid to the nearest county border and is linear in the driving time to the county border, $d_i$. The assumption that the effects decline in a linear manner from the border is realistic given that the average town is only 12 minutes away from the closest border. I show the results are robust to more non-parametric assumptions. The inclusion of this distance function is driven by Proposition 1.

Note that the effect of the federal, neighbor's and neighboring federation tax rate are now given by the mean analytic derivatives:

$$\begin{aligned} E[\tfrac{\partial t}{\partial \tau_j}] &= \frac{1}{M}\sum_{i=1}^{M}(\alpha_1 + \alpha_3 t_{-i} + \alpha_5 d_i) \\ E[\tfrac{\partial t}{\partial t_{-i}}] &= \frac{1}{M}\sum_{i=1}^{M}(\alpha_2 + \alpha_3 \tau_{i,j}) \\ E[\tfrac{\partial t}{\partial \tau_{-j}}] &= \frac{1}{M}\sum_{i=1}^{M}\alpha_4 + \alpha_6 d_i, \end{aligned} \tag{20}$$

where $M$ is the total number of observations in the estimating sample.

Standard ordinary least squares results will be biased because neighboring town and county tax rates are endogenous. Two possible solutions exist – maximum likelihood estimation (Case, Hines and Rosen 1993) and instrumental variables estimation (Figlio, Koplin and Reid 1999). Instrumental variables via generalized method of moments has the advantage of generating a consistent estimate, even in the presence of the spatial error dependence (Kelejian and Prucha 1998).

For neighboring tax rates, the standard instruments in the literature for $t_{-i}$ are the weighted average of several control variables; the method of Kelejian and Prucha (1998) requires using a subset of the neighbors' control variables. Simply put, the standard instrument for neighboring tax rates is $\Sigma_{k \in N_i} w_{ik} x_k$, where $x_k$ is a variable in $X$. Although this instrument – using the weighted average of several $x$'s – is used mostly for state level tax competition, it has also been applied to local level tax competition. Instead of using the entire subset of the $X$'s as instruments, I will only use geographic variables – predetermined variables – as instruments. Specifically, I will use area and perimeter of the town as instruments. For the county tax rate, I will use the county area and perimeter as instruments. For the neighboring county, I will use its county perimeter and area as instruments. To instrument for $t_{-i} = \Sigma_{k \in N_i} w_{ik} t_k$, I use $\text{area}_{-i} = \Sigma_{k \in N_i} w_{ik} \text{area}_k$ and $\text{perimeter}_{-i} = \Sigma_{k \in N_i} w_{ik} \text{perimeter}_k$ as instruments. Of course, the regression specifications above also include interaction terms, in which case they are instrumented for with the interactions of the respective terms.

In order to justify the instruments, recall that the regression equation controls for town area and town perimeter. Then, the exclusion restriction requires that the instruments should have no partial effect on local taxes *after* controlling for these variables – which include the town's own area and perimeter. Absent any non-linear relationships between county variables and local variables, this is likely to be the case. The direct impact of county area and county perimeter on local taxes is likely to be zero. County area and perimeter affect the county's tax rates, but will have no direct impact on the locality's tax rate so long as there are multiple jurisdictions within a county and so long as counties are sufficiently large in size. The theoretical tax competition literature implies that area and perimeter are important determinants of a jurisdiction's own tax rates. Further, county borders were likely to be historically drawn on latitudes and longitudes or broader geographic features. The area and perimeter of a county depend on a county's characteristics such as whether along a body of water, broader geographic features, and how counties were divided historically. Because area and perimeter are historically drawn, the evolution of time with these variables strengthens the case for their exogeneity. Similarly, the town's area and perimeter often depend on how municipalities were historically formed within the county and the characteristics within the county when the town borders were historically drawn – which in most cases were not at the same time county lines were delineated.[20]

[20] Although the area and perimeter are pre-determined variables, a worry with the identification strategy is that town or county unobservables may be correlated with these variables. If these unobservables are also correlated with tax rates, estimation in a cross-section could be limiting. However, estimation in a panel is arguably even more problematic, where the presence of time varying predetermined variables as valid

## 4.3 Hypotheses to Test

Before presenting the results, recall that the theoretical model provides the following testable hypotheses regarding the sign on the coefficients from equation 18. The coefficient representing the horizontal interaction, $t_{-i}$, is expected to be positive on average based on equation 10. Diagonal tax competition represented by $\tau_{i,-j}$ should be positive on average because the diagonal interaction is similar in sign to a horizontal interaction in the local region of the border from equation 11. The vertical interaction to $\tau_j$, is ambiguous in equation 8. The interaction effect $t_{-i}\tau_j$ should be significant if the intensity of horizontal competition influences vertical competition. Vertical interactions will be affected by distance $d_i\tau_j$ in a positive manner as the lower branch of equation 8 is greater than the upper branch. Finally, the diagonal reaction will be affected by distance through $d_i\tau_{i,-j}$ and the effect is expected to be negative as the lower branch of equation 11 is less than the upper branch.

# 5 Empirical Results

## 5.1 Main Results

Before presenting the instrumental variable results, column 1 of table 2 estimates the baseline specification currently estimated in the literature using ordinary least squares (OLS). The second column estimates the complete specification suggested by a theory of inter- and intra-federation competition. The OLS results for the full specification make both the horizontal and vertical interactions closer to zero in absolute value.

Table 3 presents the baseline results using spatial GMM-IV estimation. Although previous papers have focused on the coefficients directly, these estimates are likely to be misspecified if the interactions are not included. The mean derivative is the most important measure of the effect of tax interactions and, therefore, will be the focus of the following discussion. Each of the columns from 1 to 7 build sequentially on what the literature estimates, with column 7 being the specification that the theory suggests is accurately specified.

Column 1 presents the results where the vertical effect is estimated alone as in the baseline specification of Besley and Rosen (1998). Although the theory predicts the sign of this reaction may be ambiguous, Besley and Rosen (1998) find that the coefficient is positive for state cigarette and gasoline taxes with respect to the level of the federal tax. I find a significant and large negative result that is consistent with the regression discontinuity results in Agrawal (2011). Several explanations exist for the opposite finding. Towns may react in

instruments would be nearly impossible to identify at the local level. Further, I would argue that the use of area and perimeter as instruments is much less likely to suffer from this threat than the standard spatial lags of the demographic variables of the neighboring jurisdictions.

a different manner to county rates than states will react to the federal government as the institutional structure of lower level governments is different. Alternatively, municipalities face a much more mobile tax base relative to states. The increased mobility of the tax base implies that the elasticity of cross-border shopping, $\theta$, is more likely to be larger for local governments than for state governments – and the theoretical model implies the larger $\theta$ is relative to $\varepsilon$, the more likely that county and local rates will be strategic substitutes. Column 2 estimates the horizontal reaction function and finds a significant positive relationship with neighboring jurisdictions' tax rates; this result is standard in the literature – but as the results below will indicate, need not be true for all towns.

Column 3 presents a similar specification to what is used in the literature, where the vertical and horizontal reactions are estimated jointly, but without any interaction effects. In general, the literature finds positive coefficients on the sign of the horizontal tax variable and positive but insignificant results for the vertical variable. Again, the results in Column 3 indicate large and positive effects for neighboring local tax rates and a large negative effect with respect to the county rate. The estimates in this equation suggest that a 1 percentage point increase in county tax rates lowers municipal tax rates by .224 percentage points. Contrarily, a 1 percentage point increase in the average of the neighbors tax rate, increases a municipality's local tax rate by .507 percentage points. It is useful to compare the future results to these benchmark numbers. Comparing this to the OLS coefficients shows that the bias from the OLS estimations exists but does not change the sign of the estimates.

Before proceeding, whether the instruments are valid and are strong is important for identifying consistent estimates of the coefficients. The first stage $R^2$, the magnitude and the precision of the instrumental variables – area of the county, perimeter of the county, and the averages of neighboring areas and perimeters – indicate that the instruments are able to explain variation in the endogenous regressors. Instrument weakness does not appear to be a concern. In the table, I report the robust Kleibergen-Paap Wald rk F statistic and the Stock and Yogo (2005) critical values for tests of 10 percent maximal bias induced by weak instruments. When the critical value falls below the test statistic, the bias from weak instruments is less than 10%. In most every specification where critical values are tabulated, the bias is less than 5%. Unfortunately, Stock and Yogo (2005) do not report critical values for some specifications, so for these cases, I will conduct extra robustness checks. Rejecting instrument weakness using this test is comparable to rejecting the instrument validity with an F statistic less than 10 in cases of a single endogenous regressor. The tables also report the p-values for a Hansen J test of over-identification. Failure to reject the null hypothesis suggests that if one instrument is valid, the other instrument is also valid.

Specification 4 adds the interaction of the vertical and horizontal tax rates and specification 5 accounts for diagonal tax competition. Although effect of the neighboring county is of the expected sign, it is not significant. Specification 6 and 7 (the preferred specification) add in the interaction of the vertical and diagonal tax variables with a linear distance function. Notice that the effect of the county tax rate increases in absolute value to -.400 and the slope of the neighbor's tax rate fall to .466. For the horizontal interaction, the estimate in column 7 is 8% smaller in comparison to the more restrictive specification in column 3. The effect of the county tax rate is 75% larger in absolute value than the result estimated using the more restricted specification in the literature. This suggests that failing to account for inter-federation competition attenuates the vertical reaction closer to zero; as predicted, the vertical effect is most sensitive to inter-federation competition.

The results from columns 6 and 7 suggest that the spatial location effects are extremely important to calculating the mean derivatives. Columns 8 and 9 include only a sub-set of the endogenous regressors in column 7 in order to demonstrate what the reaction functions would look like if only the vertical and diagonal elements were included. Notice that in column 9, which adds only spatial location effects on the vertical reaction function, increases the intensity of the strategic reaction in absolute value relative to column 3. This suggests that much of the bias is a result from leaving out the distance function and not the interaction of the two tax rates. From an economic perspective this bias is a result of the fact that although goods sold within a county are identical in most every respect, they differ in their characteristics (Gorman 1980; Lancaster 1966) where the defining characteristic of each town's good is the location of sale within the county. Thus, although towns compete over tax rates they are not doing so over goods that are entirely identical in their characteristics; each town sells distinct goods – where the distinction is driven entirely on the basis of their sale location. Leaving out the spatial location effects omits this heterogeneity in the characteristics of the goods and thus is akin to an omission of a cross-price elasticity across these goods.

Finally, column 10 provides an important verification that the instruments in column 7 are not weak because the Stock and Yogo (2005) critical values are not tabulated for the case of that many endogenous regressors. With many variables in need of instrumenting, the concern of weak instruments become more worrisome. As an alternative, I estimate the equation by limited information maximum likelihood (LIML). If the estimating equation is correctly specified and the distributional assumptions hold, LIML will produce unbiased coefficient estimates even if weak instruments are present. Because the point estimates in column 10 are similar to column 7 and because the mean derivatives change in the same manner, it suggests that weak instruments in the GMM estimates should not be a concern. The results comparing the full specification with the baseline specification in the literature suggest not allowing for inter-federation competition provide powerful evidence that having a more general estimating equation is important.

#### 5.1.1 Heterogeneity in Reaction Functions

Up until now, I have focused on how correctly specifying the reduced form equation impacts the mean derivatives without discussing the heterogeneous responses that arise. The effect of the neighboring jurisdictions' tax rate on the tax rate in town $i$ will depend on the tax rate of the county it is in. Figure 2 depicts this heterogeneity – keeping in mind that the instrumental variable strategy limits me to imposing that the reaction function interacts linearly with the county tax rate. Notice that neighboring tax rates are strategic complements for jurisdictions in counties with less than a 4% tax rate. However, neighboring tax rates become strategic substitutes after the county tax rate become sufficiently high. Approximately 7% of towns are in counties with a sufficiently high county taxes to imply that town $i$'s tax rate is a strategic substitute with its neighbor's tax rates. Such a result of strategic substitutes is not traditionally found in the literature, but no studies allow for the interaction effect that drives this heterogeneity.[21] I find that for a subset of towns in the sample, neighboring jurisdictions' rates may be strategic substitutes.

The empirical specification also allows town $i$ to realize heterogeneous effects from the tax rate in county $j$ along two dimensions. It will depend on distance to the county border and on the average of the neighboring town tax rates. Figure 3 shows these effects under the limitation that the IV approach requires I impose that heterogeneous responses are linear in each of these effects. The left panel indicates that the effect of a change in a town's own-county tax rate is most negative when the town has high-tax neighbors and when that town is near the county border, which is consistent with the theory. The contour map illustrates that holding fixed your neighbors' tax rate, towns that are internal to the county have reaction functions that are mildly closer to zero. Furthermore, holding distance to the county border constant, each increase in your average neighbors' tax rate more dramatically intensifies the negative effect of the county rate by making it more negative. When neighboring tax rates are low, tax competition is likely the most intense. The results in these figures validate the theoretical prediction that an increase in horizontal tax competition makes the vertical reaction slope more likely to be positive.

[21] One study that empirically finds downward sloping reaction functions is Parchet (2012). However, Wildasin (1988) and Vrijburg and de Mooij (2012) provide theoretical justification for taxes as strategic substitutes.

## 5.2 How Robust Are The Results?

In the following sections, I will discuss four sets of robustness checks: redefining what constitutes a neighboring jurisdiction, specifying a more non-parametric distance function by using a dummy variable approach, ample restrictions, and focusing on various types of borders.

### 5.2.1 Alternative Distance Functions

In all of the previous specifications, I have assumed that the distance function to county borders is linear. Such a parametrization imposes that as driving times become very large, the effect of distance on the strategic reaction remains constant. Although some counties are very large in size, most counties are small in size. Because distances to county borders are relatively small, the effect of distance on the strategic interactions may actually be linear in distance for most observations in the sample. However, I also want to treat distance in a non-parametric manner. I do this by defining a dummy variable that is equal to one for towns that are in the 90th percentile of driving time to the county border. Thus, the dummy variable equals one for towns that are more than 22 minutes from the nearest county border (only twice the distance of the average town to a border) and zero otherwise. In Table 4, I report the results when the distance function is a dummy variable. Columns 1, 2, and 3 correspond to the results in Columns 6, 7 and 10 from Table 3.

In Column 2, $\tau_{i,j} d_i$ has a positive coefficient of .141. Notice this estimate is much larger than the effect from a linear distance interaction, which implied that a one minute increase in distance from the border increased the effect by a very small gradient. The result of the dummy variable approach suggests that for towns that are extremely far from borders, the town government cares much less about leakages across the county border, which creates upward pressure on the strategic reaction. Such a result is consistent with the theory – interior jurisdictions are more likely to have upward sloping reaction functions.

### 5.2.2 Various Model Specifications and Sample Restrictions

Table 5 reports robustness checks. Column 1 and 2 drop jurisdictions for which the nearest county border is also a state border. This specification reduces the possibility that the neighboring state is also producing diagonal tax competition for towns proximate to its borders. The estimates of the effect of higher levels of government become smaller suggesting that the intensity of interactions may be stronger in the region of salient state borders.

In columns 3 and 4, I restrict the sample to towns within five miles of a county border. This specification would be most applicable for identifying the effect of the neighboring county tax rate. The theory predicts that the diagonal interaction is most salient and positive for a local region of county borders. In this specification, where the effect is expected to be salient, the diagonal interaction is positive and much larger than previous specifications. In the more localized region of the border the more intense the vertical reaction is as well. This suggests that for towns relatively proximate to the county border, the diagonal interaction is highly positive and the vertical interaction is more extremely negative relative to internal towns. However, the sample shrinks such that I cannot determine this with any degree of statistical significance and I would note that the preferred way to do such an exercise is by controlling for distance as the previous tables do.

Neighbor weights may also be varied. Table 6 specifies different weighting scheme to calculate the weighted average of the neighboring jurisdictions' rates. Thus the only difference in this table is that each set of columns alters the exogenous weights that determine $t_{-i}$. Define $N_i$ as the set of towns within a fifty mile radius of town $i$, $\ell_{ik}$ as the distance between town $i$ and town $k$, and $\phi_k$ as the population of town $k$. Recall that $t_{-i} = \Sigma_{k \neq i} w_{ik} t_k$. Then columns 1-2 use exogenous weights that are normalized to sum to 1 given by Equation 21:

$$w_{ik} = \begin{cases} \dfrac{1/\ell_{ik}}{\sum_{k \in N_i} (1/\ell_{ik})} & \text{if } k \in N_i \\ 0 & \text{if } k \notin N_i \end{cases}, \tag{21}$$

which can be interpreted as inverse distance weights. In this specification towns closer to town $i$ are given more weight than towns far away. Column 3-4 estimate the equation using the weights from Equation 22:

$$w_{ik} = \begin{cases} \dfrac{\phi_k}{\sum_{k \in N_i} \phi_k} & \text{if } k \in N_i \\ 0 & \text{if } k \notin N_i \end{cases}, \tag{22}$$

such that neighboring towns within the fifty mile radius of town $i$ are given more weight if they have a larger population. Equation 23 specifies the weights used in Columns 5-6:

$$w_{ik} = \begin{cases} \dfrac{\phi_k/\ell_{ik}}{\sum_{k \in N_i} \phi_k/\ell_{ik}} & \text{if } k \in N_i \\ 0 & \text{if } k \notin N_i \end{cases}, \tag{23}$$

which gives the most weight to highly populated towns that are closer to town $i$ and the least weight to towns with small populations that are far from town $i$. I also restrict neighbors to be defined within a twenty-file mile region.

When using inverse distance based weights, the effect of the county tax rate remains approximately the same magnitude as the baseline specification, although both interactions shrink in absolute value. This is consistent with the results where the size of the buffer region was shrunk. The diagonal effect also shrinks but remains positive. The population-distance weights shrink the vertical effect but increase the horizontal effect relative to the baseline specification. The population weights and population inverse distance weights result in smaller effects relative to the binary weights. The sizes of the horizontal interactions shrink slightly. When using population-distance weighting schemes, the diagonal effect intensifies in magnitude and becomes highly significant. A recent theoretical model and survey results of local governments (Janeba and Osterloh forthcoming) provides evidence that cities compete both locally and with other large population centers, but that small municipal governments are much more likely to only compete within a particular region. Such evidence is inconsistent with weighting neighbors by population because municipalities would not account for large jurisdictions that are far away. Most jurisdictions in America are relatively small. Distance based weights would be more reasonable for small municipalities, but less reasonable for large cities who may compete with other cities. However, small cities may compete with larger city centers if they view their consumers as deciding between purchasing a good in a small hometown store or in a big city where many more prominent and perhaps luxurious shopping opportunities may exist. For these reasons and because of its ease in interpretation, I prefer using the unweighted average of tax rates.

#### 5.2.3 By Type of Border

As a final robustness check, Table 7 breaks down the results by whether the town is in a county that is a high-tax county, low-tax county, or a county that sets the same tax rate as the neighboring county. When doing this exercise, the effect of the town's own county tax rate becomes closer to zero to zero for towns that are in high-tax counties. This suggests that for a town on the high-tax side of the border, if its county raises its tax rate, the town only partially offsets the county's tax increase. On the other hand, for a town on the low-tax side of the border, when its county raises its tax rate, the town will offset the county's tax increase more intensely. Diagonal effects from the neighboring county are large and significant among the sub-set of towns in low-tax counties.

### 5.3 Discussion

This paper develops a theoretical model that shows that sales tax competition is not constrained to occur within a federation. When multiple federations exist, horizontal tax competition will cross federation boundaries and diagonal tax competition will arise between lower levels of government and neighboring federations. In addition to tax competition across boundaries, I show that vertical interactions between a lower level of government and its own federal government will depend on the lower government's spatial location within the federation. I proceed empirically by gradually adding these factors into the standard spatial IV approach to estimating fiscal interactions. Doing so requires that I instrument for a number of variables. While having multiple endogenous regressors in an estimating equation is not ideal and may work to obscure the channels at work, I demonstrate (in the context of my local data set) that not accounting for these effects will result in the researcher finding effects that are biased toward zero with regard to the vertical strategic interaction. Further, not accounting for these effects will result in the researcher over-estimating the horizontal interaction. It appears that omitting spatial location effects are the biggest driver of the omitted variable bias. Although the effect of spatial location is relatively small, it induces large biases on the other variables of interest in the regression.

While spatial location effects explain much of the bias, allowing vertical and horizontal reactions to interact results in large heterogeneities in the strategic reaction functions. This interaction effect is important for jurisdictions in unitary federations as well. Towns that are in high-tax counties are much more likely to perceive their neighboring tax rates as strategic substitutes than a town in a low-tax county. These interaction effects also make the vertical interactions extremely heterogeneous across the population. These heterogeneities are important in their own right.

Given that the tax competition literature often has difficulty distinguishing between strategic interaction and common spatial shocks, exploiting spatial location attributes is important to helping the researcher identify tax competition. These spatial location effects are analogous to the "ripple effects" in Benjamin and Dougan (1997) where cigarette taxes are reduced in jurisdictions away from a low-tax point of production and the "tax gradient" effects of Agrawal (2011). This paper, unlike those papers, provides a direct test of the strategic reaction function by exploiting policy discontinuities at borders to provide a more convincing way of identifying strategic interaction beyond a simple spatial lag model.

The paper fails to find a significant effect on the interaction term between neighboring jurisdiction tax rates and the federal government's tax rate. In a large data set such as the one used in this paper, such a negative result may be important in its own right. Although the term is not significant in its own right, the mean derivative can change by relatively large magnitudes and failure to exclude this term masks seemingly important heterogeneity as noted in Figure 2 and 3. The implication seems to be that for researchers with a large data set of tax rates, the term should always be included to mitigate this bias. However, for researcher's working with smaller data sets, a bias-efficiency trade-off exists. Because the interaction effect is not significant in its own right, a simpler specification may be appropriate in some circumstances. Much of the bias in the estimates is driven by the inclusion of spatial location effects rather than the interaction effects. Accounting for inter-federation competition can increase vertical strategic interactions by approximately 75% of the baseline estimated in the literature without the effects of inter-federation competition.

I demonstrate both theoretically an empirically that diagonal interaction are important in their own right – and different from horizontal interactions – even though studies such as Wilson and Janeba (2005) suggest that it is the total (local plus federal) tax rate that matters. This difference arises for two reasons: (1) diagonal interaction are not salient for towns that are located relatively far from the border and (2) even for towns at the border, the effect of a neighboring federation's tax rate will only influence cross-border shopping in one direction, while changes in the average neighboring town tax rate has the potential to effect shopping in multiple directions. Other possible theoretical reasons exist for this difference, but I do not explore them in this paper. For example, consider a model of yardstick competition where voter's compare own-jurisdiction taxes with taxes in "similar" jurisdictions when deciding to re-elect local officials. In this case, distance may still matter as a determinant of the jurisdictions voter's use as a comparison. However, the tax rate in a neighboring county may be less important than the tax rate in neighboring towns because voters may be only comparing tax rates across "similar" jurisdictions and not across governments of a higher level. Empirically, I verify that the effect of diagonal interactions is less intense than horizontal interactions. Such a distinction seems to justify the diagonal externality as a separate concept distinct from a horizontal externality by showing that the town and county tax rates of neighbors matter separately as opposed to their sum.

Finally, the magnitudes of the strategic interactions in this paper are different and, for the vertical externalities, opposite in sign to the traditional literature using state level data. This suggests that using comprehensive local data will produce different strategic interactions than using state level data. The intuition is that the elasticity of cross-border shopping is much larger for smaller jurisdictions, which from the theory suggests that the vertical reaction is more likely to be negative. The evidence in this paper suggests that it would be incorrect to assume that the slope of the reaction function for a state government looks anything like the reaction function for a local government.

# 6 Conclusion

Introducing inter-federation competition into a model of sales tax competition that combines the vertical elements of Keen (1998) and the horizontal elements of Kanbur and Keen (1993) and Nielsen (2001) indicates that the spatial composition of towns within a federation is essential to determine the strategic nature of the tax competition. First, this paper argues that the geo-spatial nearness to borders of sub-federal governments in a federation – particularly the spatial proximity to discontinuous changes in the tax rate resulting from the federation's borders – makes it less likely that a peripheral local government will mimic the federal government. Second, inter-federation competition results in diagonal tax competition – competition induced by a different level of government that does not share the same tax base – which has similar but smaller consequences as a horizontal competition.

The theoretical predictions of the model shed light on the appropriate estimation strategy. In this paper, I define the "federal" government as the county government and the sub-federal government as a municipal government. Using a comprehensive data set on a cross-section of local sales taxes in the United States and using constructed spatial data, I test how local governments strategically interact with county governments. The empirical results validated the theoretical implications and stress the importance of having an accurately specified estimating equation. Whenever estimating horizontal and vertical reaction functions, the interaction of the two must be included in the regression. Further, when considering municipal governments (and even the national government if it competes with other nations), the researcher must consider how the tax rates in neighboring federations and a municipality's proximity to the neighboring federation affect the municipality's tax rate.

With respect the nature of the strategic interaction, local sales taxes are strategic complements with neighboring sales taxes – a finding consistently found in the literature. However, for a subset of towns in high-tax counties, local tax rates may be strategic substitutes with their neighbors. With respect to vertical interactions, the theoretical literature suggests the interaction may be positive or negative. I find that a one percentage point increase in county tax rates lowers municipal tax rates by about .40 percentage points; this suggests that county and municipal taxes are strategic substitutes. The result is opposite to the positive and small effects found for state level governments in response to the federal government. The results in this paper suggest it would be inappropriate to generalize results from state level studies to municipal level interactions. But, as relatively few studies of tax competition have exploited comprehensive local tax data, the use of local data will provide a continued avenue for future research within federations with multiple levels of government.

# References

**Agrawal, David R.** 2011. "The Tax Gradient: Do Local Sales Taxes Reduce Tax Differentials at State Borders?" Working Paper.

**Becker, Johannes, and Theiss Büttner.** 2012. "Tax Competition among Federations." Working Paper Presented at the 2012 IIPF Congress.

**Benjamin, Daniel K., and William R. Dougan.** 1997. "Efficient Excise Taxation: The Evidence from Cigarettes." *Journal of Law and Economics*, 40(1): 113–136.

**Besley, Timothy J., and Harvey S. Rosen.** 1998. "Vertical Externalities in Tax Setting: Evidence from Gasoline and Cigarettes." *Journal of Public Economics*, 70: 383–398.

**Bonanno, Giacomo, and John Vickers.** 1988. "Vertical Separation." *The Journal of Industrial Economics*, 36(3): 257–265.

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

**Brülhart, Marius, and Mario Jametti.** 2006. "Vertical Versus Horizontal Tax Externalities: An Empirical Test." *Journal of Public Economics*, 90: 2027–2062.

**Burge, Gregory, and Cynthia Rogers.** 2011. "Local Option Sales Taxes and Consumer Spending Patterns: Fiscal Interdependence under Multi-tiered Local Taxation." *Regional Science and Urban Economics*, 41: 46–58.

**Case, Anne C., James R. Hines, and Harvey S. Rosen.** 1993. "Budget Spillovers and Fiscal Policy Interdependence." *Journal of Public Economics*, 52(3): 285–307.

**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.

**Esteller-Moré, Alex, and Albert Solé-Ollé.** 2001. "Vertical Income Tax Externalities and Fiscal Interdependence." *Regional Science and Urban Economics*, 31(2-3): 247–272.

**Figlio, D., V. Koplin, and W. Reid.** 1999. "Do States Play Welfare Games?" *Journal of Urban Economics*, 46: 437–454.

**Franzese, Robert J.** 2005. "Empirical Strategies for Various Manifestations of Multilevel Data." *Political Analysis*, 13: 430–446.

**Fredriksson, Per G., and Khawaja A. Mamun.** 2008. "Vertical Tax Externalities in Cigarette Taxation: Do Tax Revenues Go Up in Smoke?" *Journal of Urban Economics*, 64: 35–48.

**Geys, Benny, and Steffen Osterloh.** forthcoming. "Borders as Boundaries to Fiscal Policy Interactions? An Empirical Analysis of Politicians' Opinions on Rivals in the Competition for Firms." *Journal of Regional Science*.

**Gorman, W. M.** 1980. "A Possible Procedure for Analysing Quality Differentials in the Egg Market." *Review of Economic Studies*, 47: 843–856.

**Hayashi, Masayoshi, and Robin Boadway.** 2001. "An Emperical Analysis of Intergovernmental Tax Interaction: The Case of Business Income Taxes in Canada." *Canadian Journal of Economics*, 34(2): 481–503.

**Hoyt, William.** 2001. "Tax Policy Coordination, Vertical Externalities, and Optimal Taxation in a System of Hierarchical Governments." *Journal of Urban Economics*, 50: 491–516.

**Janeba, Eckhard, and Steffen Osterloh.** forthcoming. "Tax and the City - A Theory of Local Tax Competition." *Journal of Public Economics*.

**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.** 1998. "Vertical Tax Externalities in the Theory of Fiscal Federalism." *IMF Staff Papers*, 45(3): 454–485.

**Keen, Michael, and Christos Kotsogiannis.** 2002. "Does Federalism Lead to Excessively High Taxes?" *American Economic Review*, 92(1): 363–370.

**Keen, Michael, and Christos Kotsogiannis.** 2003. "Leviathan and Capital Tax Competition in Federations." *Journal of Public Economic Theory*, 5(2): 177–199.

**Kelders, Christian, and Marko Koethenbuerger.** 2010. "Tax Incentives in Fiscal Federalism: An Integrated Perspective." *Canadian Journal of Economics*, 43(2): 683–703.

**Kelejian, Harry H., and Ingmar R. Prucha.** 1998. "A Generlized Spatial Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances." *Journal of Real Estate Finance and Economics*, 17: 99–121.

**Lancaster, Kevin J.** 1966. "A New Approach to Consumer Theory." *The Journal of Political Economy*, 74(2): 132–157.

**Mathewson, G. F., and R. A. Winter.** 1984. "An Economic Theory of of Vertical Restraints." *RAND Journal of Economics*, 15: 27–38.

**Mintz, Jack, and Henry Tulkens.** 1986. "Commodity Tax Competition Between Member States of a Federation: Equilibrium and Efficiency." *Journal of Public Economics*, 29(2): 132–172.

**Nielsen, Søren Bo.** 2001. "A Simple Model of Commodity Taxation and Cross-Border Shopping." *The Scandinavian Journal of Economics*, 103(4): 599–623.

**Parchet, Raphaël.** 2012. "Are Local Tax Rates Strategic Complements or Strategic Substitutes?" Working Paper.

**Rey, Patrick, and Jean Tirole.** 1986. "The Logic of Vertical restraints." *American Economic Review*, 76: 923–939.

**Saggi, Kamal, and Nikolaos Vettas.** 2002. "On Intrabrand and Interbrand Competition: The Strategic Role of Fees and Royalties." *European Economic Review*, 46: 189–200.

**Stock, James H., and Motohiro Yogo.** 2005. "Testing for Weak Instruments in Linear IV Regression." In *Identification and Inference for Econometric Models.* , ed. Donald W.K. Andrews, 80–108. New York, NY:Cambridge University Press.

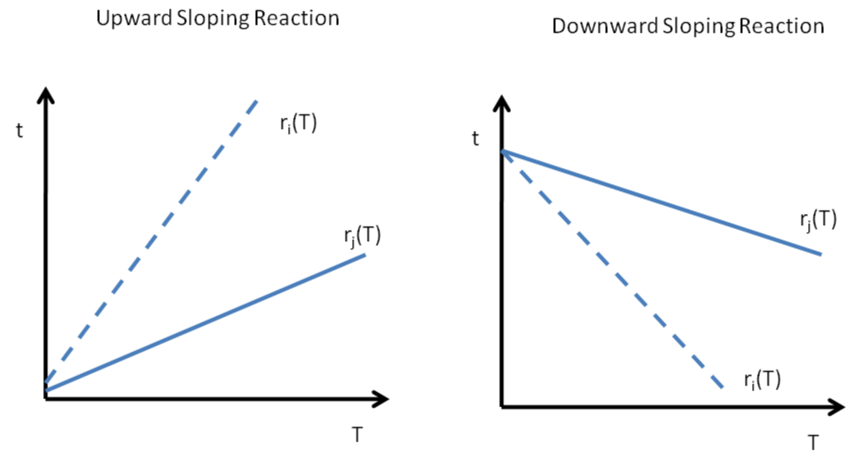
**Trandel, Gregory.** 1994. "Interstate Commodity Tax Differentials and the Distribution of Residents." *Journal of Public Economics*, 53(3): 435–457.

**Vrijburg, Hendrik, and Ruud A. de Mooij.** 2012. "Tax Rates as Strategic Substitutes." Tinbergen Institute Discussion Paper.

**Wildasin, David.** 1988. "Nash Equilibria in Models of Fiscal Competition." *Journal of Public Economics*, 32(2): 229–240.

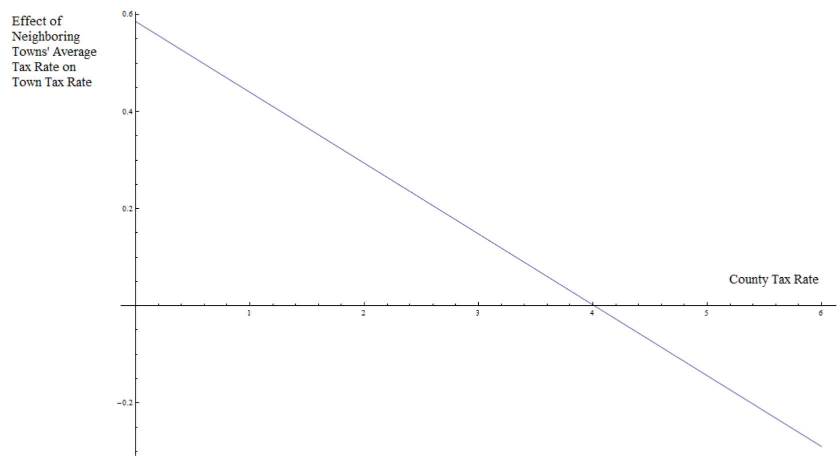
**Wilson, John Douglas, and Eckhard Janeba.** 2005. "Decentralization and International Tax Competition." *Journal of Public Economics*, 89: 1211–1229.

Figure 1: Reaction Functions



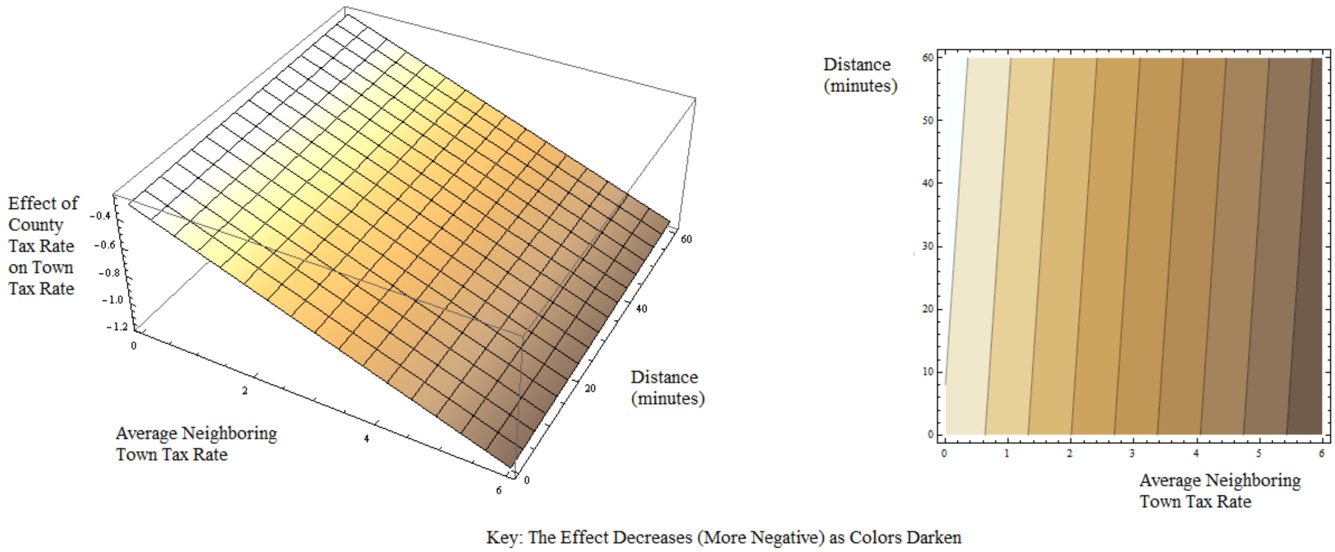

The left panel displays strategic reaction functions where tax rates at the lower level of government ($t$) and tax rates at the higher level of government ($T$) are strategic complements. The right panel illustrates strategic substitutes. The dotted lines represent the most intense strategic reaction in each panel.

Figure 2: Heterogeneity in Horizontal Reaction




The vertical axis represents the effect of the average neighboring town's tax rate on the tax rate of town $i$. The graph is derived from column (7) in table 3. The effect of the horizontal interaction is a function of the county tax rate because the regression specification allows for horizontal and vertical reactions to interact. Notice that towns in relatively low tax counties perceive their neighbors' tax rates as strategic complements. However, towns in very high tax counties view their neighbors' tax rates as strategic substitutes.

Figure 3: Heterogeneity in Vertical Reaction Functions




In the left panel, the vertical axis represents the effect of the federation's tax rate on a town's tax rate. The graph is derived from column (7) in table 3. The effect the vertical reaction interaction is a function of the neighboring town tax rates because the regression specification allows for horizontal and vertical reactions to interact. The regression specification also allows the strategic reaction to differ on the basis of proximity to the nearest county border. For simplicity I only show the reaction functions for towns within an hour of the nearest county border and for neighboring tax rates that average between zero and five percent. The left panel illustrates the heterogeneity in the reaction functions using a three dimensional figure. The right panel condenses the left figure into a contour plot where the darker colors correspond to strategic reactions that are more intense – which for downward sloped reaction functions evidenced above implies that darker colors imply the reaction function slopes are further from zero.

Table 1: Summary Statistics
Averages of Variables by Type
Standard Deviations in ( )

| Type | Variable | Mean |
|---|---|---|
| Dependent | Local Tax Rate | .768 |
| | | (1.164) |
| Other | Distance to County Border (minutes) | 12.099 |
| | | (19.859) |
| Endogenous Regressors | County Tax Rate | 1.037 |
| | | (1.188) |
| | Unweighted Average Neighbors' Rate | .746 |
| | (50 Mile Radius) | (.864) |
| | Unweighted Average Neighbors' Rate | .750 |
| | (25 Mile Radius) | (.943) |
| | Neighboring County Tax Rate | .991 |
| | | (1.173) |
| Control Variables | Town Perimeter | 14.311 |
| | | (23.799) |
| | Town Area | 5.550 |
| | | (17.751) |
| | Number of Neighbors | 1.761 |
| | | (1.768) |
| | Population | 9032 |
| | | (58,169) |
| | Senior (%) | 16.035 |
| | | (8.160) |
| | Less than College (%) | 81.802 |
| | | (14.113) |
| | Income | 56,665 |
| | | (30,817) |
| | Male | 49.129 |
| | | (5.447) |
| | Percent on Public Assistance | 2.434 |
| | | (3.520) |
| | White (%) | 84.442 |
| | | (19.986) |
| | Age | 39.352 |
| | | (8.023) |
| | Ratio of Private School : Public School | .131 |
| | Students | (.487) |
| | Number of Rooms | 5.564 |
| | | (.733) |
| | Age of Structure | 44.245 |
| | | (15.712) |
| | Non-citizen (%) | 3.149 |
| | | (5.817) |
| | Obama Vote Share | 43.747 |
| | | (13.773) |
| | Sample Size | 12,996 |

The log of driving time to the state border, a dummy for the relatively high-tax side of the border and same-tax side of a border, the tax differential at the border and a complete set of interactions is also included in every regression specification.

Table 2: Slopes of Reaction Functions – OLS

| Variable | (1) | (2) |
|---|---|---|
| $\tau_{i,j}$ | -.440*** | -.175*** |
| | (.021) | (.024) |
| $t_{-i}$ | .498*** | .635*** |
| | (.028) | (.031) |
| $t_{-i}\tau_j$ | | -.171*** |
| | | (.015) |
| $\tau_{i,-j}$ | | .004 |
| | | (.014) |
| $\tau_{i,j}d_i$ | | .0008** |
| | | (.0004) |
| $\tau_{i,-j}d_i$ | | .0003 |
| | | (.0004) |
| $E[\frac{\partial t}{\partial \tau_j}]$ | -.440*** | -.292*** |
| | (.021) | (.019) |
| $E[\frac{\partial t}{\partial t_{-i}}]$ | .498*** | .495*** |
| | (.028) | (.028) |
| $E[\frac{\partial t}{\partial \tau_{-j}}]$ | | .007 |
| | | (.014) |
| Buffer | - | 50 |
| Distance | - | linear |
| Method | OLS | OLS |
| $R^2$ | .689 | .697 |
| N | 12,996 | 12,996 |

***99%, **95%, *90%, standard errors are robust, standard errors for mean derivatives are calculated using the Delta Method

All of the specifications above include state fixed effects and the local control variables outlined in the text. The above results are estimated using OLS – thus ignoring endogeneity issues. Column (1) estimates the standard equation used in the literature. Column (2) accounts for the effects of inter-federation competition as suggested by the theoretical model in the paper. The mean derivatives represent the slope of the reaction functions.

Table 3: Slopes of Reaction Functions – Baseline Specification

| Variable | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|---|---|---|---|
| $\tau_{i,j}$ | -.280** | | -.224*** | -.240* | -.213 | -.325** | -.308** | -.277*** | -.409*** | -.342** |
| | (.118) | | (.096) | (.143) | (.131) | (.137) | (.135) | (.098) | (.1143) | (.151) |
| $t_{-i}$ | | .521*** | .507*** | .485*** | .558*** | .563*** | .586*** | .507*** | .486*** | .580*** |
| | | (.066) | (.064) | (.091) | (.132) | (.130) | (.129) | (.063) | (.0625) | (.166) |
| $t_{-i}\tau_j$ | | | | .017 | -.094 | -.128 | -.146 | | | -.075 |
| | | | | (.0834) | (.163) | (.160) | (.160) | | | (.349) |
| $\tau_{i,-j}$ | | | | | .114 | .103 | .135 | .060 | | .151 |
| | | | | | (.146) | (.138) | (.139) | (.081) | | (.199) |
| $\tau_{i,j}d_i$ | | | | | | .001** | .001** | | .001** | .001** |
| | | | | | | (.0006) | (.0006) | | (.0006) | (.0007) |
| $\tau_{i,-j}d_i$ | | | | | | | -.0002 | | | -.0003 |
| | | | | | | | (.0005) | | | (.0004) |
| $E[\frac{\partial t}{\partial \tau_j}]$ | -.280** | | -.224*** | -.230** | -.264*** | -.406*** | -.400*** | -.277*** | -.393*** | -.437** |
| | (.118) | | (.096) | (.113) | (.097) | (.108) | (.107) | (.098) | (.110) | (.140) |
| $E[\frac{\partial t}{\partial t_{-i}}]$ | | .521*** | .507*** | .500*** | .481*** | .458*** | .466*** | .507*** | .486*** | .457*** |
| | | (.066) | (.064) | (.061) | (.061) | (.061) | (.061) | (.063) | (.063) | (.069) |
| $E[\frac{\partial t}{\partial \tau_{-j}}]$ | | | | | .114 | .103 | .131 | .060 | | .147 |
| | | | | | (.146) | (.138) | (.138) | (.081) | | (.197) |
| Buffer | - | 50 | 50 | 50 | 50 | 50 | 50 | 50 | 50 | 50 |
| Distance | - | - | - | - | - | linear | linear | - | linear | linear |
| Method | GMM | GMM | GMM | GMM | GMM | GMM | GMM | GMM | GMM | LIML |
| Over-id† | .193 | .104 | .221 | .378 | .382 | .656 | .589 | .244 | .670 | .593 |
| Weak Inst.‡ | 48.661 | 564.63 | 26.618 | 25.610 | 3.804 | 3.221 | 2.639 | 14.776 | 10.643 | 2.639 |
| Critical‡ | 19.93 | 19.93 | 7.56 | 7.77 | - | - | - | 7.77 | 9.01 | - |
| $R^2$ | .659 | .651 | .680 | .677 | .690 | .690 | .691 | .682 | .689 | .689 |
| N | 12,996 | 12,996 | 12,996 | 12,996 | 12,996 | 12,996 | 12,996 | 12,996 | 12,996 | 12,996 |

***99%, **95%, *90%, standard errors are robust, standard errors for mean derivatives are calculated using the Delta Method

All of the specifications above include state fixed effects and the local control variables outlined in the text. The instruments are always the area and perimeter of the respective jurisdictions. In the case of horizontal externalities, the instruments are the average area and perimeter in the buffer zone of fifty miles around each town. The instrument for interactions are the interaction of the instruments. When the regression includes a tax term interacted with distance, the distance variable also enters the regression equation as a stand-alone variable. The measure of distance is the time to the nearest county border and it enters all specifications in a linear manner. Column (1) allows for vertical competition only. Column (2) allows for only horizontal competition. Column (3) is the benchmark specification estimated in the literature. Column (4) adds interaction effects between the horizontal and vertical expressions. Column (5) allows for interaction effects and diagonal interactions. Column (6) allows the vertical effect to vary based on proximity to the county border. Column (7) allows the diagonal effect to also vary with respect to proximity to the border. Columns (8) and (9) demonstrate the effect of only adding the neighboring federation's tax rate or proximity effects on the federation's tax rate to the baseline specification. Column (10) is the same as column (7) except that it is estimated using limited information maximum likelihood. $E[\frac{\partial t}{\partial \tau_j}]$ represents the effect of the federation's tax rate. $E[\frac{\partial t}{\partial t_{-i}}]$ represents the effect of the average neighboring tax rate. $E[\frac{\partial t}{\partial \tau_{-j}}]$ represents the effect of the neighboring town tax rate.

†The test of over-identification reports the p-value of the Hansen J test.

‡The test for weak instruments is the the robust Kleibergen-Paap Wald rk F statistic. I report the critical values (when available) for which the bias induced by weak instruments is less than 10% if the test statistic is greater.

Table 4: Slopes of Reaction Functions – When Distance Is a Dummy Variable

| Variable | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| $\tau_{i,j}$ | -.197 | -.223* | -.275* | -.212* |
| | (.121) | (.131) | (.158) | (.125) |
| $t_{-i,j}$ | .617*** | .632*** | .650*** | .511*** |
| | (.123) | (.122) | (.176) | (.063) |
| $t_{-ij}\tau_j$ | -.177 | -.205 | -.246 | |
| | (.157) | (.156) | (.259) | |
| $\tau_{i,-j}$ | .167 | .216 | .268 | |
| | (.129) | (.137) | (.236) | |
| $\tau_{i,j}d_i$ | .026 | .141 | .199 | .013 |
| | (.027) | (.144) | (.190) | (.0223) |
| $\tau_{i,-j}d_i$ | | -.116 | -.172 | |
| | | (.144) | (.181) | |
| $E[\frac{\partial t}{\partial \tau_j}]$ | -.280*** | -.270*** | -.319* | -.207* |
| | (.108) | (.109) | (.171) | (.118) |
| $E[\frac{\partial t}{\partial t_{-i}}]$ | .472*** | .463*** | .449** | .511*** |
| | (.062) | (.063) | (.081) | (.063) |
| $E[\frac{\partial t}{\partial \tau_{-j}}]$ | .167 | .163 | .191 | |
| | (.129) | (.127) | (.199) | |
| Buffer | 50 | 50 | 50 | 50 |
| Distance | dummy | dummy | dummy | dummy |
| Method | GMM | GMM | LIML | GMM |
| Over-id† | .521 | .568 | .586 | .235 |
| Weak Inst.‡ | 3.419 | 2.648 | 2.648 | 10.974 |
| Critical‡ | - | - | - | 7.77 |
| $R^2$ | .689 | .682 | .670 | .679 |
| N | 12,996 | 12,996 | 12,996 | 12,996 |

***99%, **95%, *90%, standard errors are robust, standard errors for mean derivatives are calculated using the Delta Method

The estimating equations in this table are identical to Tables 3 except the interaction with the distance function uses a dummy variable interaction, rather than a linear distance term. All instruments and control variables remain the same. The dummy variable takes on the value of one if the driving time to the county border is more than 22 minutes and a value of zero otherwise. The cutoff of 22 minutes corresponds 90th percentile of towns with respect to time from the border. The expected sign on $\tau_{i,j}d_i$ remains positive while the expected sign on$\tau_{i,-j}d_i$ remains negative. Columns (1)-(3) corresponds to columns (6), (7) and (10) in table 3. Column (4) corresponds to column (9) in Table 3.$E[\frac{\partial t}{\partial \tau_j}]$ represents the effect of the federation's tax rate. $E[\frac{\partial t}{\partial t_{-i}}]$ represents the effect of the average neighboring tax rate. $E[\frac{\partial t}{\partial \tau_{-j}}]$ represents the effect of the neighboring town tax rate.

†The test of over-identification reports the p-value of the Hansen J test.

‡The test for weak instruments is the the robust Kleibergen-Paap Wald rk F statistic. I report the critical values (when available) for which the bias induced by weak instruments is less than 10% if the test statistic is greater.

Table 5: Slopes of Reaction Functions – Robustness Checks: Restrictions and Weighting

| Variable | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| $\tau_{i,j}$ | -.238** | -.284** | -.391 | -1.54 | -.308*** | -.378** |
| | (.101) | (.138) | (.275) | (1.21) | (.117) | (.180) |
| $t_{-i}$ | 516*** | .521*** | .635*** | .996** | .458*** | .451*** |
| | (.073) | (.104) | (.139) | (.404) | (.092) | (.191) |
| $t_{-i}\tau_j$ | | -.022 | | -.757 | | .006 |
| | | (.089) | | (.579) | | (.235) |
| $\tau_{i,-j}$ | | .018 | | 1.01 | | .299 |
| | | (.151) | | (1.03) | | (.201) |
| $\tau_{i,j}d_i$ | | .0009 | | .268 | | .0002 |
| | | (.0005) | | (.284) | | (.001) |
| $\tau_{i,-j}d_i$ | | -.0004 | | -.282 | | -.0004 |
| | | (.0005) | | (.309) | | (.0007) |
| $E[\frac{\partial t}{\partial \tau_j}]$ | -.238** | -.289*** | -.391 | -1.28** | -.308*** | -.370** |
| | (.101) | (.110) | (.275) | (.666) | (.117) | (.161) |
| $E[\frac{\partial t}{\partial t_{-i}}]$ | 516*** | .498*** | .635*** | .179 | .458*** | .455*** |
| | (.073) | (.072) | (.139) | (.358) | (.092) | (.081) |
| $E[\frac{\partial t}{\partial \tau_{-j}}]$ | | .014 | | .197 | | .293 |
| | | (.150) | | (.351) | | (.198) |
| Buffer | 50 | 50 | 50 | 50 | 50 | 50 |
| Distance | | linear | | linear | | linear |
| Restriction | Border | Border | $D < 5$ | $D < 5$ | Weights By State | Weights By State |
| Method | GMM | GMM | GMM | GMM | GMM | GMM |
| Over-id† | .186 | .434 | .308 | .910 | .1674 | .551 |
| Weak Inst.‡ | 22.353 | 7.568 | 2.544 | .303 | 4.280 | .150 |
| Critical‡ | 7.56 | - | 7.56 | - | 7.56 | - |
| $R^2$ | .689 | .696 | .737 | .221 | .657 | .631 |
| N | 11,807 | 11,807 | 3147 | 3147 | 12,996 | 12,996 |

***99%, **95%, *90%, standard errors are robust, standard errors for mean derivatives are calculated using the Delta Method

All the specifications include the same set of controls and fixed effects as the previous tables. The instruments also remain the same. (1)-(2) drops towns where the nearest county border is also a state border in order to eliminate diagonal externalities between the towns and state governments. (3)-(4) restrict the estimating sample to towns within five miles of a county border. Columns (5)-(6) weight the observations in the sample so that each state receives equal weight and such that states with more towns do not give undue weight to the results. $E[\frac{\partial t}{\partial \tau_j}]$ represents the effect of the federation's tax rate. $E[\frac{\partial t}{\partial t_{-i}}]$ represents the effect of the average neighboring tax rate. $E[\frac{\partial t}{\partial \tau_{-j}}]$ represents the effect of the neighboring town tax rate.

†The test of over-identification reports the p-value of the Hansen J test.

‡The test for weak instruments is the the robust Kleibergen-Paap Wald rk F statistic. I report the critical values (when available) for which the bias induced by weak instruments is less than 10% if the test statistic is greater.

Table 6: Slopes of Reaction Functions – Robustness Checks: Various Neighbor Weights

| Variable | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| $\tau_{i,j}$ | -.195* | -.376*** | -.300*** | -.184 | -.281** | -.052 | -.186* | -.387*** |
| | (.100) | (.130) | (.105) | (.172) | (.110) | (.177) | (.100) | (.146) |
| $t_{-i}$ | .440*** | .416*** | .050 | .384* | .058 | .596*** | .400*** | .286** |
| | (.071) | (.103) | (.040) | (.197) | (.055) | (.227) | (.072) | (.112) |
| $t_{-i}\tau_j$ | | .007 | | -.098* | | -.174** | | .169 |
| | | (.102) | | (.053) | | (.069) | | (.110) |
| $\tau_{i,-j}$ | | .053 | | .212 | | .331** | | -.121 |
| | | (.112) | | (.141) | | (.162) | | (.147) |
| $\tau_{i,j}d_i$ | | .001** | | .001 | | .0003 | | -.0002 |
| | | (.0006) | | (.001) | | (.0007) | | (.006) |
| $\tau_{i,-j}d_i$ | | -.0003 | | -.001 | | -.001** | | -.001 |
| | | (.0006) | | (.001) | | (.0004) | | (.0009) |
| $E[\frac{\partial t}{\partial \tau_j}]$ | -.195* | -.354*** | -.300*** | -.288** | -.281** | -.245* | -.186* | -.246** |
| | (.100) | (.104) | (.105) | (.134) | (.110) | (.130) | (.100) | (.114) |
| $E[\frac{\partial t}{\partial t_{-i}}]$ | .440*** | .422*** | .050 | .303** | .058 | .453*** | .400*** | .425*** |
| | (.071) | (.069) | (.040) | (.153) | (.055) | (.172) | (.072) | (.075) |
| $E[\frac{\partial t}{\partial \tau_{-j}}]$ | | .049 | | .204 | | .320** | | -.124 |
| | | (.111) | | (.139) | | (.158) | | (.145) |
| Buffer | 50 | 50 | 50 | 50 | 50 | 50 | 25 | 25 |
| Distance | - | linear | - | linear | - | linear | - | linear |
| Restriction | $1/D$ | $1/D$ | Population | Population | $Pop/D$ | $Pop/D$ | - | - |
| Method | GMM | GMM | GMM | GMM | GMM | GMM | GMM | GMM |
| Over-id† | .590 | .580 | .132 | .292 | .053 | .279 | .190 | .386 |
| Weak Inst.‡ | 28.443 | 2.946 | 35.204 | 3.504 | 37.877 | 2.504 | 29.783 | 2.236 |
| Critical‡ | 7.56 | - | 7.56 | - | 7.56 | - | 7.56 | - |
| $R^2$ | .690 | .697 | .664 | .665 | .664 | .647 | .691 | .659 |
| N | 12,996 | 12,996 | 12,996 | 12,996 | 12,996 | 12,996 | 12,996 | 12,996 |

***99%, **95%, *90%, standard errors are robust, standard errors for mean derivatives are calculated using the Delta Method All the specifications include the same set of controls and fixed effects as the previous tables. The instruments also remain the same. Towns receive non-zero weight in the neighboring weight matrix if they are within the fifty mile buffer zone. (1)-(2) uses inverse distance weights to calculate the weighted average of neighboring tax rates such that it weights each town in $t_{-i}$ by the inverse distance to the neighbor. (3)-(4) weights each town in $t_{-i}$ by the population of the neighbor such that the most weight is given to the largest jurisdictions in the 50 mile radius. (5)-(6) weights each town in $t_{-i}$ by the population of the neighbor neighbor and then by the inverse of distance to the neighbor such that the most weight is given to large towns that are also close. Columns (7) and (8) town receive non-zero weight if they are outside a 25 mile buffer of the jurisdiction. $E[\frac{\partial t}{\partial \tau_j}]$ represents the effect of the federation's tax rate. $E[\frac{\partial t}{\partial t_{-i}}]$ represents the effect of the average neighboring tax rate. $E[\frac{\partial t}{\partial \tau_{-j}}]$ represents the effect of the neighboring town tax rate.

†The test of over-identification reports the p-value of the Hansen J test.

‡The test for weak instruments is the the robust Kleibergen-Paap Wald rk F statistic. I report the critical values (when available) for which the bias induced by weak instruments is less than 10% if the test statistic is greater.

Table 7: Slopes of Reaction Functions – Robustness Checks: Type of Border

| Variable | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| $\tau_{i,j}$ | -.152 | -.349* | -.589** | -.253 | -.327 | -.552** |
| | (.226) | (.187) | (.232) | (.630) | (.215) | (.248) |
| $t_{-i}$ | .381*** | .510*** | .439*** | .700*** | .547*** | .505*** |
| | (.119) | (.132) | (.123) | (.237) | (.090) | (.100) |
| $t_{-i}\tau_j$ | | .006 | | -.298 | | .017 |
| | | (.077) | | (.336) | | (.086) |
| $\tau_{i,-j}$ | | -.092 | | .750*** | | |
| | | (.210) | | (.248) | | |
| $\tau_{i,j}d_i$ | | .0008 | | -.002 | | .001*** |
| | | (.001) | | (.004) | | (.0003) |
| $\tau_{i,-j}d_i$ | | -.0004 | | -.002 | | |
| | | (.005) | | (.003) | | |
| $E[\frac{\partial t}{\partial \tau_j}]$ | -.152 | -.331* | -.589** | -.594** | -.327 | -.525** |
| | (.226) | (.171) | (.232) | (.274) | (.215) | (.215) |
| $E[\frac{\partial t}{\partial t_{-i}}]$ | .381*** | .515*** | .439*** | .461*** | .547*** | .520*** |
| | (.119) | (.101) | (.123) | (.133) | (.090) | (.086) |
| $E[\frac{\partial t}{\partial \tau_{-j}}]$ | | -.083 | | .716*** | | |
| | | (.187) | | (.221) | | |
| Buffer | 50 | 50 | 50 | 50 | 50 | 50 |
| Distance | | linear | | linear | | linear |
| Restriction | High-Tax County | High-Tax County | Low-Tax County | Low-Tax County | Same-Tax County | Same-Tax County |
| Method | GMM | GMM | GMM | GMM | | GMM |
| Over-id[†] | .582 | .382 | .111 | .363 | .127 | .113 |
| Weak Inst.[‡] | 8.896 | 1.884 | 8.231 | 1.787 | 17.144 | 12.282 |
| Critical[‡] | 7.56 | - | 7.56 | - | 7.56 | - |
| $R^2$ | .634 | .634 | .689 | .637 | .723 | .724 |
| N | 3155 | 3155 | 2850 | 2850 | 6990 | 6990 |

***99%, **95%, *90%, standard errors are robust, standard errors for mean derivatives are calculated using the Delta Method All the specifications include the same set of controls and fixed effects as the previous tables.The instruments remain the same. This table tests whether the effects depend on whether a town is located in a high-tax, low-tax or same-tax county relative to the nearest neighboring county. (1)-(2) restrict the sample to towns where the nearest neighboring county sets a lower tax rate than its own county. (3)-(4) restrict the sample to towns where the nearest neighboring county sets a higher tax rate than its own county. (5)-(6) restrict the sample to towns where the nearest neighboring county sets the same tax rate as its own county. $E[\frac{\partial t}{\partial \tau_j}]$ represents the effect of the federation's tax rate. $E[\frac{\partial t}{\partial t_{-i}}]$ represents the effect of the average neighboring tax rate. $E[\frac{\partial t}{\partial \tau_{-j}}]$ represents the effect of the neighboring town tax rate.

[†]The test of over-identification reports the p-value of the Hansen J test.

[‡]The test for weak instruments is the the robust Kleibergen-Paap Wald rk F statistic. I report the critical values (when available) for which the bias induced by weak instruments is less than 10% if the test statistic is greater.

# 7 Appendix

## 7.1 Derivation of Equation 8

Define the expression in equation 3 as $F \equiv \frac{\partial R_i}{\partial t_i} = B_i + t_i \frac{\partial B_i}{\partial q_i} = x_i s_i + t_i s_i x_i' - t_i x_i^2(\rho_i + \rho_{i+1}) = 0$. Note that $\frac{\partial s_i}{\partial q_i} = (\rho_i + \rho_{i+1})v'(q_i) = -x_i(\rho_i + \rho_{i+1})$. Then I can calculate:

$$\frac{\partial F}{\partial t_i} = 2\frac{\partial B_i}{\partial q_i} + t_i \frac{\partial^2 B_i}{\partial q_i^2} = 2(x_i' s_i - x_i^2(\rho_i + \rho_{i+1})) + t_i[x_i'' s_i - 3x_i x_i'(\rho_i + \rho_{i+1})]$$

For a interior town the derivative of $F$ with respect to its own county's tax rate is

$$\frac{\partial F}{\partial \tau_j} = \frac{\partial B_i}{\partial q_{i-1}} + \frac{\partial B_i}{\partial q_i} + \frac{\partial B_i}{\partial q_{i+1}} + t_i\left(\frac{\partial^2 B_i}{\partial q_{i-1}^2} + \frac{\partial^2 B_i}{\partial q_i^2} + \frac{\partial^2 B_i}{\partial q_{i+1}^2}\right)$$
$$= x_i' s_i - x_i^2(\rho_i + \rho_{i+1}) + x_i x_{i-1}\rho_i + x_i x_{i+1}\rho_{i+1} + t_i[x_i'' s_i - 3x_i x_i'(\rho_i + \rho_{i+1}) + x_i' x_{i-1}\rho_i + x_i' x_{i+1}\rho_{i+1}]$$

Then the vertical reaction function of town $i$ with respect to county $j$'s tax rate is given by $-\frac{\partial F}{\partial \tau_j}/\frac{\partial F}{\partial t_i}$. The above two equations can be converted to elasticities by multiplying both the numerator and denominator of this expression by $\frac{q_i}{x_i s_i}$. They can further be simplified using equation 4. Doing so yields:

$$\frac{\partial t_i}{\partial \tau_j} = -\frac{-\varepsilon - \theta_i + \theta_{i,k} - \frac{1}{\varepsilon+\theta_i}\varepsilon(\eta - 3\theta_i + \theta_{i,k})}{-2\varepsilon - 2\theta_i - \frac{1}{\varepsilon+\theta_i}\varepsilon(\eta - 3\theta_i)} = \frac{-\varepsilon^2 - \theta_i^2 + \theta_i\varepsilon + \theta_i\theta_{i,k} - \varepsilon\eta}{2(\varepsilon + \theta_i)^2 + \varepsilon\eta - 3\varepsilon\theta_i} \tag{24}$$

and simplifying a bit further (and evaluating the expression at $q_i = q_{i+1} = q_{i-1}$) yields the upper branch of the equation in the text. A similar derivative can be calculated for the peripheral town by omitting the appropriate $\frac{\partial B_i}{\partial q_k}$ and $\frac{\partial^2 B_i}{\partial q_k^2}$ from the expression to $\frac{\partial F}{\partial \tau_j}$.

Diagonal and horizontal reaction functions can be derived using the same procedure. Notice that the denominator will remain the same. Only the numerator will change. Because of the similarity in the procedure, the derivations are omitted for simplicity.

## 7.2 Asymmetric Model

In the text, each federation was composed of the same number of sub-federal governments. I now relax this assumption. Federations (counties) are indexed $j = 1, 2, ....M$ and each county is composed of $m + j$ towns each such that each sequential county has one more town than the previous; county 1 will have $m+1$ towns, but county $M$ will have $m+M$ towns. The ordering of counties in this manner is not important, but it will allow me to characterize the model's solution using a simpler notation than if counties were organized arbitrarily along a line segment. Towns are indexed $i = 1, 2, ....M(m + \frac{1+M}{2})$ and the index does not reset across counties. The $M$ federations in the model compete with each other and with the towns in the model. The towns compete with other towns and with the federations. The rest of the set up and assumptions of the model remains the same. I maintain assumption 1 and 2, but because counties now differ in the their size, the equilibrium will no longer be symmetric.

Because each county is ordered such that it has one more identical town than the previous county along the line, the length of each county increases as $j$ increases, which implies that $n$ is also increasing in $j$. Kanbur and Keen (1993) and Nielsen (2001) show that tax rates are increasing as the size (population or geographic size) increases. Using the intuition from Kanbur and Keen (1993) and Nielsen (2001) that the perceived elasticity of cross-border shopping in a big county (one with more identical towns) is inelastic relative to a small county, I assume conclude that the Nash equilibrium of county tax rates will follow the following pattern: $\tau_1 \leq \tau_2 \leq ...\tau_M$. This assumption places no restriction on the pattern of local taxes within a county. However, Agrawal (2011) shows that revenue maximizing towns of identical size will set higher rates the closer to a high-tax county neighbor and lower rates closer to a low-tax neighbor. I assume this pattern holds when evaluating the reaction functions below because the models differ only in whether demand is perfectly inelastic or not.[22] Under this assumption, the tax rate in town $i-1$ will always be lower than the tax rate in town $i$. Therefore, residents on the west portion of town shop abroad, while additional entry occurs on the eastern side of the town. I assume the asymmetric Nash equilibrium exists and is characterized by

$$\begin{array}{cc} \tau_1 \leq \tau_2 \leq ...\tau_M & \& \\ t_i \leq t_{i+1} \leq ...t_{m+j} & \text{for town}\, i \text{ in county}\, j \\ t_i + \tau_j \leq t_{i+1} + \tau_{j+1} & \text{for all county borders.} \end{array} \tag{25}$$

and 4 still characterizes the Nash equilibrium for counties and towns. Notice that 25 is the analog to assumption 3 in the text.

In a symmetric Nash equilibrium, the elasticity of cross-border shopping is constant within a federation. This is no longer true. I consider how differences in county tax rates resulting from the asymmetric equilibrium defined by equation 25 influences the strategic interaction of the local governments. Recall that if the elasticity of demand is constant and local tax rates are increasing from low-tax counties toward high-tax counties, equation 4 implies that all of the local tax differences are due to differences in the elasticity of cross-border shopping. Specifically, for jurisdictions closest to a low tax county, $\theta_i$ is relatively high. For jurisdictions closest to the high-tax county, $\theta_i$ is relatively low. Thus, following Agrawal

[22] Relaxing this specific characterization of the equilibrium would not change the functional form of the reaction functions or the revenue function, but it would distort the pattern of the subscripts on the $\rho$'s in these equations. The presence of a tax gradient allows for a pattern in the perceived elasticities.

(2011), the elasticity $\theta_i$ should be monotonically decreasing from the left-most town to the right-most town along the line segment. The natural question is how the strategic reaction changes when the elasticity of cross-border shopping increases. Define $\theta_{i,i+1} + \theta_{i,i-1} \equiv \gamma_k$ for an internal town and $\theta_{i,k} \equiv \gamma_k$ for a peripheral town. Differentiating the part of equation 8 for the interior town that was not evaluated in a symmetric equilibrium and assuming that the cross-price elasticities $\theta_{i,k}$ are not a function of the own-price elasticity $\theta_i$ or if they are that these effects are sufficiently small to be ignored,[23] provides some important evidence. For an increase in $\theta_i$, the change in a town's strategic reaction to its own county tax rate will be proportional to:

$$\frac{\varepsilon^3 + \varepsilon^2(\gamma_k - 2\theta_i) - (3\varepsilon + 2\gamma_k)\theta_i^2}{D_i^2} \tag{26}$$

Equation 26 is derived in the proof below and the necessary conditions for this equation to be positive are also derived.

**Proposition 3.** *In the neighborhood of an asymmetric Nash equilibrium, the heterogeneity in the strategic reaction function between internal and peripheral towns critically depends on whether the reaction function is increasing or decreasing in the elasticity of cross border shopping in addition to whether the jurisdiction is internal or peripheral to a low-tax or high-tax county border.*

*Proof.* The question is how does $\varepsilon_i(\theta_i - \varepsilon_i - \eta_i) + \theta_i(\theta_{i,i+1} + \theta_{i,i-1} - \theta_i)]/D_i$ change with $\theta_i$. Recall $\varepsilon$ and $\eta$ are constant along a demand curve. I also assume that $\theta_{i,k}$ is not a function of $\theta_i$ because I want to isolate the change in the reaction for a change in $\theta_i$, all else equal. Define $\theta_{i,i+1} + \theta_{i,i-1} \equiv \gamma_k$ where $\theta_{i,i-1}$ is equal to zero for the left-most town in the county and $\theta_{i,i+1}$ is equal to zero for the right-most town in the county. Differentiating with respect to $\theta_i$ using the product rule yields:

$$\frac{(\varepsilon + \gamma_k - 2\theta_i)D_i - (4\theta_i + \varepsilon)(\varepsilon\theta_i - \varepsilon^2 - \varepsilon\eta - \theta_i^2 + \theta_i\gamma_k)}{D_i^2}.$$

Expanding terms and grouping liker terms yields:

$$\frac{3\varepsilon^3 - 3\varepsilon\theta^2 + 2\varepsilon^2\eta + 2\theta_i\varepsilon\eta - 2\gamma_k\theta_i^2 + \varepsilon\eta\gamma_k + 2\varepsilon^2\gamma_k}{D_i^2},$$

Using the fact that $\eta_i = -(1 + \varepsilon_i)$ for iso-elastic demand, yields the equation in the text.

[23] Relaxing this assumption can easily be incorporated in to the expression below by differentiating $\gamma_k$ with respect to $\theta_i$.

This expression will definitely be positive if:

$$\varepsilon^3 + \varepsilon^2\gamma - \varepsilon^2 2\theta_i - (3\varepsilon + 2\gamma_k)\theta_i^2 > 0,$$

Recalling that I have assumed that $\gamma_k$ is not a function of $\theta_i$, this expression is a quadratic with in the elasticity of cross-border shopping that has roots:

$$\theta_i = \frac{-\varepsilon^2 \pm \chi^{1/2}}{2\gamma_k + 3\varepsilon}$$

where $\chi = 4\varepsilon^4 + 5\varepsilon^3\gamma_k + 2\varepsilon^2\gamma_k$. Noting that all variables are positive, it is clear that for reasonable values of the elasticity of demand, $\varepsilon^2 < \chi^{1/2}$. For this reason, I can eliminate the negative root from the quadratic formula. The derivative of the $\varepsilon^3+\varepsilon^2\gamma-\varepsilon^2 2\theta_i-(3\varepsilon+2\gamma_k)\theta_i^2$ with respect to $\theta_i$ is

$$-2\varepsilon^2 - 2\theta_i(2\gamma_k + 3\varepsilon),$$

which evaluated at the only positive root simplifies nicely to $\chi^{1/2} > 0$. Therefore, I know that Equation 26 is positive if

$$\theta_i > \frac{-\varepsilon^2 + \chi^{1/2}}{2\gamma_k + 3\varepsilon}.$$

□

Notice that if expression 26 is positive, movement from the peripheral town at the high-tax (eastern most) county border to an interior town unambiguously makes the strategic reaction more likely to be positive. Moving from the peripheral town to the interior town implies that $\theta_i$ rises. An increase in $\theta_i$ then has a positive effect via equation 26. Furthermore, the strategic reaction given by equation 8 now contains an additional cross price elasticity $\theta_{i,k} > 0$ relative to equation 8, which reinforces that positive effect. If expression 26 is negative, then the negative influence of the higher $\theta_i$ must be counter-balanced with the additional positive pressures of the cross-price elasticity. On the other hand, for a peripheral town at the low-tax county border (the western-most town in the county), movement to the interior of the state implies that $\theta_i$ is falling. If expression 26 is positive, the decline in this elasticity has a negative effect on the strategic interaction. But at the same time, the interior town worries about two borders and realizes and additional $\theta_{i,k} > 0$. Thus, the effect is possibly ambiguous for towns near low-tax borders. If expression 26 is negative, then movement to the interior of the state from this direction unambiguously increases the strategic reaction. However, if changes in $\theta_i$ across neighboring jurisdictions are small, then the effect noted in proposition 1 is likely to dominate the fact that $\theta_i$ decreases from low-tax counties toward high-tax counties.

Of course, one may also wonder how the strategic reaction changes for two internal towns in an asymmetric equilibrium. In this case, the researcher needs to determine the sign of equation 26. However, tax differentials at town borders are also informative of the slope if elasticities cannot be observed directly or are of unknown magnitude. Signing $\theta_i(\theta_{i,i+1} + \theta_{i,i-1} - \theta_i)$ is especially important for the case of iso-elastic demand as this term is the only term that puts downward pressure on the reaction function.

**Proposition 4.** *In the neighborhood of an asymmetric Nash equilibrium, the closer the tax rate of an interior town $i$ is to the tax rate of its high-tax neighboring town (relative to the low-tax neighboring town), the more likely $\theta_i(\theta_{i,i+1} + \theta_{i,i-1} - \theta_i)$ is positive.*

*Proof.* Recalling that I assume that $\rho_i$ does not differ across towns and rewrite the expression $(\theta_{i,i+1} + \theta_{i,i-1} - \theta_i)$ as

$$\frac{\rho x_{i+1} q_i}{s_i} + \frac{\rho x_{i-1} q_i}{s_i} - \frac{q_i x_i (2\rho)}{s_i} = \frac{\rho q_i}{s_i}(x_{i+1} + x_{i-1} - 2x_i).$$

Then use Roy's identity to yield:

$$\frac{\rho q_i}{s_i}(-\frac{\partial v_{i+1}}{\partial \tau} - \frac{\partial v_{i-1}}{\partial \tau} + 2\frac{\partial v_i}{\partial \tau}),$$

which is likely to be greater than zero when:

$$2\frac{\partial v_i}{\partial \tau} > \frac{\partial v_{i+1}}{\partial \tau} + \frac{\partial v_{i-1}}{\partial \tau}$$

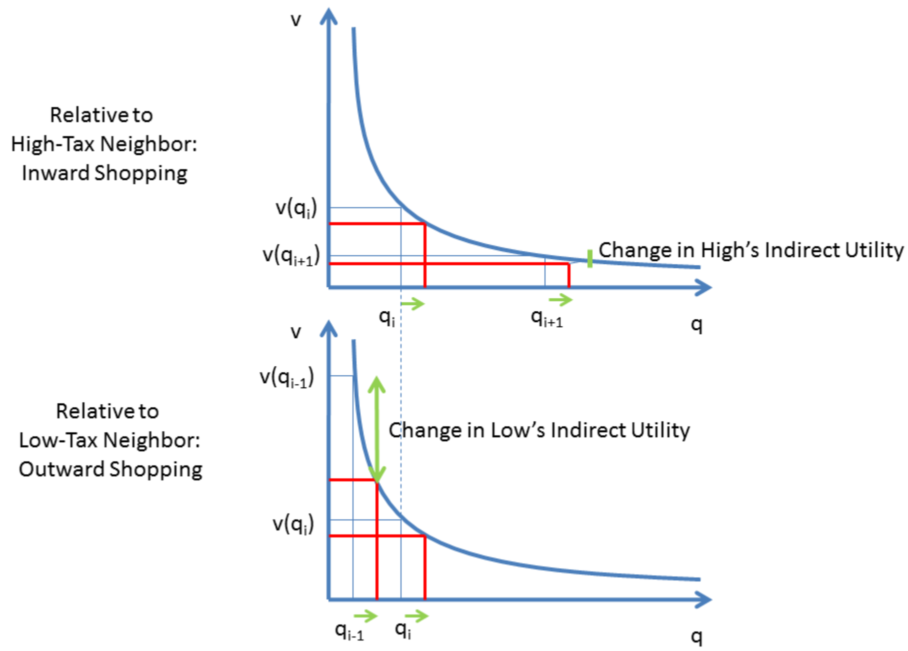
from Figure 4 it is evident that $\frac{\partial v_{i+1}}{\partial \tau} > \frac{\partial v_i}{\partial \tau} > \frac{\partial v_{i-1}}{\partial \tau}$. Thus, the closer $\frac{\partial v_i}{\partial \tau}$ is to $\frac{\partial v_{i+1}}{\partial \tau}$, the more likely the above expression will be positive. □

Recall $i-1$ is the low-tax neighbor and $i+1$ is the high-tax neighbor. When a town's rate is close to the high-tax neighboring town and the county rate increases in both jurisdictions, the change in $v(q_i)$ and $v(q_{i+1})$ are relatively similar – because the prices are similar in both jurisdictions. However, when the county tax rate changes, the change in $v(q_{i-1})$ is much larger than the change of $v(q_i)$ – because the indirect utility function is downward sloping and convex with respect to prices.[24] The implication of this is that the elasticity of the tax base with respect to the town's own price $\theta_i$ is small relative to the cumulative cross

[24] The FOC of the consumer's maximization problem implies that $u'(x) = q$. Totally differentiating this with respect to prices implies that $x'(q) = \frac{1}{u''(x)} > 0$ by concavity of the utility function. Totally differentiating the indirect utility function with respect to prices implies $v' = -x(p) < 0$. Totally differentiating again yields $v'' = x'(p) > 0$. Taken together, these imply that the indirect utility function is convex with respect to prices.

price elasticities. As a result, $(\theta_{i,i+1} + \theta_{i,i-1} - \theta_i) > 0$. Conversely, if relatively close to the neighboring low-tax rate, then the change in $v(q_i)$ and $v(q_{i-1})$ are relatively similar, but large. Also recall the change in $v(q_i)$ needs to be accounted for twice. However, the change in $v(q_{i+1})$ is small relative to these other two changes. The implication is that $(\theta_{i,i+1} + \theta_{i,i-1} - \theta_i) < 0$.

Figure 4: Change in Cross-Border Shopping due to a Change in the County Rate




The curve above represents the indirect utility function of a consumer. The tax rates increase from town $i-1$ to town $i+1$ such that town $i$ has an after tax price in the middle. All towns are in the same county and the initial light straight lines represent the starting positions. Then consider what happens when the county raises its tax rate. Because each town is in the same county, this raises the after tax price $q$ by the same amount in each county. The new positions are represented by the bold lines. However, because of the shape of the indirect utility function notice that the change in indirect utility is largest in the low-tax town and smallest in the high-tax town.

Consider Figure 4. In the figure, the prices start at $q_i$, $q_{i-1}$ and $q_{i+1}$. Assume that town $i$ is internal to the county. Then, if the county rate increases, all prices rise by a constant amount to the bolder lines on the graph. The amount of cross-border shopping is proportional to the difference between the indirect utilities. In the top graph, the neighbor has a higher price so cross-border shopping is inward to $i$. In the second graph, the neighbor has a lower price, so cross-border shopping is outward from $i$. Note that $v(q_i) - v(q_{i+1})$ becomes smaller after the tax increase and $v(q_{i-1}) - v(q_i)$ also becomes smaller – but it falls by a much larger amount because of the convexity of the indirect utility function. The closer $q_i$ is to its high-tax neighbor, the smaller the change in $v(q_{i-1}) - v(q_i)$ and vice-versa. This mitigates the change in outward cross-border shopping and amplifies the change in inward cross-border shopping.

Intuitively, the transportation cost function does not depend on the tax rate. Therefore, no matter the tax rate, an individual must pay $cd$ to cross-border shop. When the price increases because the county tax rate increases, demand will decrease. As a result, lower demand at a higher price implies the total benefit of cross-border shopping will fall, but the total cost remains the same. The amount of cross-border shopping will change as a result – but how much it changes by will depend on the relative local tax rates in both jurisdictions.