Inter-Federation Competition:

Sales Tax Externalities with Multiple Federations

David R. Agrawal*

University of Georgia

May 2012; Preliminary: Not for Quotation

Abstract

Existing models of tax competition focus on intra-federation competition. This paper analyzes how introducing inter-federation competition affects the strategic nature of tax competition. In the context of a Nash equilibrium, the paper shows that lower levels of government will respond differently to higher levels of government depending on the local government's position within the federation – specifically whether the locality is internal or peripheral relative to the federation's borders. Inter-federation competition will also introduce "diagonal externalities," which are fiscal externalities between neighboring jurisdictions, but that occur between different levels of government. A diagonal externality will have similar effects as a horizontal externality. The paper uses two unique data sets, a cross-section of all local tax rates in the United States and spatial proximity data, to test how local governments react to horizontal, vertical, and diagonal competitors. The empirical specifications allow for vertical and horizontal externalities to have interaction effects and allow for strategic reactions that vary based on proximity to neighboring federations. If interaction effects and distance based effects are omitted from the estimating equation, the vertical strategic reaction will be approximately 30% too large in absolute value. After including these terms, a one percentage point increase in the county tax rate implies that municipal tax rates in the county will be approximately 0.60 percentage points lower; every one mile increase in distance from the county border mutes this effect by 0.11 percentage points. Taxes in neighboring jurisdictions have a large and positive effect.

JEL: H25, H71, H73, H77

Keywords: Commodity Taxation, Tax Competition, Fiscal Federalism

1 Introduction

The literature on fiscal competition defines two types of fiscal externalities. Horizontal externalities occur between neighboring jurisdictions with separate tax bases. Vertical externalities occur between different levels of government that share part of the same tax base. Horizontal externalities arise because jurisdictions do not account for the effect that a tax rate change will have on a neighboring jurisdiction's revenue. Starting from equal tax rates, suppose one jurisdiction raises its consumption tax rate. An increase in the tax rate of this jurisdiction will result in cross-border shopping in the direction of all its neighbors. As a result, tax revenues will rise in the neighboring jurisdictions – improving social welfare in these jurisdictions under some assumptions. However, when that government raised its tax rate, it did not account for this social benefit to its neighbors – and it overestimates the *social* (global) marginal cost of public funds from raising taxes. As a result, taxes are too low in equilibrium.

Vertical externalities arise between different levels of government that share the same tax base. The state of New York and New York City both realize tax revenue from consumers in New York City. When New York City raises its tax rate, a vertical externality arises because total consumption of goods in New York City declines. The state's tax revenue falls, even though the state tax rate remains unchanged. Although state revenues fall, city only internalizes the effect of the rate change on its own tax base. Vertical externalities result in governments underestimating the social marginal cost of public funds. Thus, taxes are too high in equilibrium.

Existing models of sales tax competition have studied multiple levels of government in the context of a unitary federation. Systems of government such as the United States, however, are often characterized by multiple federations (i.e., fifty states in the United States with many sub-federal governments likes towns and counties). Introducing multiple competing federations that are possibly asymmetric results in the United States sales tax system being characterized by exogenously given lines in the tax system – borders. These lines are not a result of intended tax policy, but differences in preferences for public goods across borders result in discontinuous changes in tax rates as an individual crosses state lines. Sub-state governments are, therefore, heterogeneous with respect to how far they are located from the discontinuous change in the tax rates. As a result of heterogeneity in distance to the border, the nature of the strategic response to the state government and with respect to neighboring states will inevitably not be uniform within a state.

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? In a world

where the federal government has no horizontal competitors of its own (other counties), federal taxes will (under most assumptions) be too high as the federal government only faces a vertical externality. Inter-federation competition inevitably constrains the federal government by inducing horizontal externalities at the federal level, while also triggering additional competition at the sub-federal level across federation borders. As a result of the discontinuity induced by the federation's border, this cross-federation competition is inevitably different than horizontal competition within the federation. These differences have strong implications for estimating the strategic reaction functions in a federation.

Inter-federation competition is traditionally used to mean competition among nations. However, I use the term inter-federation competition to highlight competition among the higher levels of government. As such, counties are federal to towns just as the national government is federal to the state governments. Throughout this chapter, 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 reltionship as well.

I argue that inter-federation competition is essential to fully understand horizontal and vertical externalities. Modeling and estimating vertical externalities with only one "federal" government ignores horizontal competition among neighboring federal governments – which will put downward pressure on rates. An assumption of a unitary federation may be valid when the federal government is the national government – as it realizes 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 and this will constrain its taxes from being too high. 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.

Introducing multiple federations will also allow me to develop a more realistic model where sub-federal 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 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 affect members of the federation that are located at its periphery in a manner differently

than they affect members of the federation that are located internal to the federation. I will also show that "diagonal externalities" will arise. A diagonal externality is the effect of a county's tax rate on a municipality in the neighboring county. For a peripheral jurisdiction, it is identical in the nature of the strategic response to a horizontal externality. The results are derived in a model where consumers have a downward sloping demand function.

In addition to 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. 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 externalities. Additionally, when the federal government has horizontal competitors, externalities induced by the federal government on its competitor must also be considered. The existing literature estimates a strategic reaction function for the average jurisdiction assuming that all externalities are of the same magnitude – no matter the spatial composition of the federation. The empirical results suggest that, with multiple federations, in order to obtain consistent 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.

This paper will focus on two dimensions of "spatial economics" – "spatial interdependence or contagion" and "spatial location." One, 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. Two, spatial location is the process by which distance from or proximity to a particular geographic feature influences tax setting behavior. Example of spatial location are that tax rates may be a function of proximity to a border or to an amenity. The evidence will show that the strength of spatial interdependence can be heterogeneous with respect to a jurisdiction's spatial location. Adding elements of spatial location will complicate the analysis, but will provide me with a unique and convincing identification strategy to identify tax competition where I rely on demonstrating that strategic interaction is heterogeneous with respect to spatial location in a manner predicted by theory.

The baseline empirical results show that a one percentage point increase in county sales tax rates reduces municipal sales tax rates by .640 percentage points. Alternatively, a one

4

percentage point increase in the average neighboring tax rate will increase a municipality's local sales tax by .399 percentage points. The results are similar in spirit but smaller in magnitude when the definition of neighborliness is made more restrictive. Relative to a specification accounts for horizontal and vertical interaction effects, diagonal externalities, and distance based effects, the estimates from the specification currently estimated in the literature are approximately 30% too large in absolute value. The implications of the estimating strategy are threefold. (1) Omitting interaction effects of horizontal and vertical externalities will induce a substantial bias (between 30% and 60% depending on the definition of neighborliness) that over-estimates the magnitude of the true strategic vertical and horizontal interactions. This bias will arise even if the federal government has no horizontal competitors. (2) If a "federal" government has horizontal competitors as is the case of states and counties, then the neighboring federation's tax rate and the proximity to the neighboring federation must be accounted for in order to obtain an unbiased estimate of the true strategic reactions. (3) The strategic nature of local governments is vastly different than that the strategic reaction of states. The existing state-based empirical literature indicates that most U.S. states consider federal tax rates as strategic complements (although the regression estimates are often not significantly different from zero). The results in this paper suggest that municipalities consider county sales tax rates as strong strategic substitutes. Because the theoretical model in this paper suggests that the direction of the vertical externality depends on the relative magnitudes of the elasticity of demand and the elasticity of cross-border shopping, the empirical results also shed light on the relative magnitudes of these elasticities. The empirical finding that county and town tax rates are strategic substitutes is consistent with municipalities having a much larger elasticity of cross-border shopping than states. Such a conclusion is not unreasonable given that the elasticity of cross-border shopping is often decreasing as state size increases.

I proceed by outlining the the existing fiscal federalism literature and then section three will develop a model that includes inter-federation competition where jurisdictions are not symmetric. Section four discusses empirical concerns and the data used and section five outlines the empirical methodology. Section six details the results.

2   Sales Tax Competition in Federations

2.1   Theoretical Models of Fiscal Federalism

In the United States, Canada, and India, multiple levels of government share the same sales tax base. The same is true of income and capital taxes in many other countries as well.

Despite federalism being a part of many world-wide tax systems, the theoretical literature on sales tax competition has often ignored its role. When, and if, a federal government does appear in models of tax competition, its sole purpose is usually to deal with inefficiencies arising from horizontal competition – without any purpose in its own right to raise tax revenue or maximize welfare.

Models of sales tax competition giving authority in its own right to the federal government are Keen (1998) and Devereux, Lockwood and Redoano (2007).[1] 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.

Keen (1998) develops a model in which states and the federal government compete over sales tax rates to maximize revenue. Under simplifying assumptions, the equilibrium tax rate follows and inverse elasticity rule where jurisdictions set tax rates inversely proportional to the elasticity of demand. If the elasticity of demand is constant, an increase in the federal tax rate will result in states raising tax rates. On the other hand, if the demand curve is linear, increases in the federal tax rate result in states lowering the state tax rates. The necessary and sufficient condition for state and federal tax rates to be strategic complements is for the demand function to be log-convex in prices. Devereux, Lockwood and Redoano (2007) generalizes some of the assumptions from Keen (1998). Because the nature of the strategic relationship between state and federal tax rates is ambiguous in the simple model of Keen (1998), it is not unexpected that it remains ambiguous in the more complicated model of Devereux, Lockwood and Redoano (2007).

2.2   Empirically Estimating Reaction Functions

In the first serious attempt to estimate vertical externalities, Besley and Rosen (1998) uses panel data and regresses the state cigarette tax rate on the federal cigarette tax rate. Such a regression fails to account for horizontal externalities and the simultaneity in the tax rates. Esteller-Moré and Solé-Ollé (2001) corrected for these problems by instrumenting for federal tax rates and including (instrumented) average neighboring tax rates on the right-hand side of the equation. However, because the federal tax rate does not vary in any cross-section, such a procedure does not allow for the inclusion of time effects in any panel. Other authors have followed approaches different to the instrumental variable approach of Esteller-

---

[1]In addition to these articles, other papers – for example, Keen and Kotsogiannis (2002) and Keen and Kotsogiannis (2003) – have focused on federalism in the context of capital tax competition. The literature on sales tax competition in federations is much more sparse.

Moré and Solé-Ollé (2001). Fore example, Hayashi and Boadway (2001) claim to avoid the endogeneity problem altogether by assuming that the interaction between different levels of government occurs with a time lag so that the values of the federal tax rates are no longer simultaneous. Revelli (2001) argues that estimating the reaction function in first differences and instrumenting for the differenced county tax rates with the lagged level of the tax rates eliminates the endogeneity problem.[2] In addition to these issues, Fredriksson and Mamun (2008) point out that the assumption of putting the sub-national tax rate on the left-side of the equation is completely arbitrary. It is conceivable that the vertical externality works in the opposite direction – that state tax rates cause the federal government to respond.

Brülhart and Jametti (2006) tries to estimate the vertical and horizontal externalities without identifying the effects off of the slopes of reaction functions. Brülhart and Jametti (2006) theoretically shows that when horizontal externalities dominate, an increase in the number of sub-national governments will lower equilibrium tax rates. If vertical externalities dominate, an increase in the number of sub-national governments will raise equilibrium tax rates. Using a cross-section of Swiss cantons, the paper regresses local tax rates on canton tax rates and the number of localities in a canton. By doing so, they implicitly assume that there is no horizontal competition between localities of different cantons – even if these localities are contiguous. Under this assumption, the coefficient on the number of localities is positive, which suggests that vertical externalities dominate horizontal externalities.

To summarize, the existing literature faces two main threats to identification. One, because tax rates are set simultaneously, reaction functions cannot be estimated by OLS because of the endogeneity problem. Second, inclusion of horizontal externalities in the regression is essential to account for the spatial relationship of nearby jurisdictions, but it is important to identify strategic reaction rather than spatial shocks.

2.3   Relationship to Industrial Organization

The literature on tax competition has several interesting parallels with the industrial organization literature. In industrial organization, the upstream firm is the firm that supplies the inputs in production process and the firms that produce the good are downstream firms. The theoretical analysis of upstream and downstream firms of 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. This literature explores externalities between the retailers and highlights the double marginalization effect. Double marginalization is the result

---

[2]This is the only paper measuring vertical externalities with data at the sub-national level of government.

of upstream and downstream firms independently setting prices without accounting for the vertical externality between firms; double marginalization can be eliminated by vertical integration. Of course, integration in the tax competition literature would require lower levels of government surrendering political authority.

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. Two differences are that the number of downstream firms in Saggi and Vettas (2002) is endogenous and upstream firms have access to a two-part tariff. In the tax competition model to follow, counties cannot charge municipalities a fixed fee and the number of jurisdictions is exogenous. Another way to eliminate double marginalization that might be more relevant to a tax competition setting is resale price maintenance, under which an upstream firm regulates the price setting behavior by its own downstream firm.

Recently, the industrial organization literature has focused on empirically estimating vertically integrated and separated markets. In many cases, the approach taken is vastly different to the literature on tax competition for several reasons: data availability at the firm level is quite different and in a single market, firms may be either vertically integrated or independent. In one example, Hastings (2004) demonstrates that price competition is weakened if independent gas stations are replaced by integrated gas stations. For a second example, Villas-Boas (2007) has ruled out double marginalization in a particular industry and instead finds that "manufacturers are pricing at marginal cost and that retail prices are the unconstrained profit maximizing price." Although the methods are different, ruling out double marginalization has consequences as to whether prices are too high or too low, which is analogous to the debate over taxes in the presence of both vertical and horizontal competition. Although the industrial organization literature has emphasized endogenous entry of firms and various pricing schemes not available to the tax authority, the literature above highlights that some of the added complexities discussed in this paper have also recently arisen in the industrial organization literature. Both the industrial organization and tax competition literature would benefit from more familiarity with each other.

3 Model

The following 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 $j = 1, 2, ...M$ and are 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 $i = 1, 2, ...M(m + \frac{1+M}{2})$. The M federations in the model compete with each other and with the towns in the model. All of the federations are within a common union (state). Towns may differ in size (both population and length). 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 ϕ_i denotes the population density. Consumers and producers are located at every point along the continuous line segment.

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 and constant at p_i across all towns and normalize it to one. This assumption follows from having a continuum of producers in the model. Producers cannot manipulate cross-border shopping by changing the pre-tax price. Rather, they are perfectly competitive and set prices equal to marginal cost no matter the location of the firm along the line segment. Firms enter or exit to meet additional or reduced demand.

Demand for the taxed good is denoted x. The untaxed good is assumed to be the numéraire and is used as payment for purchase of the consumption good. 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 good.

Every level of government can set a specific commodity tax on the consumption good. Taxes are levied under the origin principle, which implies that the location of the transaction – not the consumer's residence – determines the tax rate.[3] Towns (sub-federal governments) set a local tax rate t_i, counties (federal governments) set a county tax rate τ_j that applies to all towns within the county ($i \epsilon j$), and the state government (the union government) sets a state tax rate T.[4] The state tax rate applies to all locations along the infinite line segment. In any specific town, the after tax price q_i is equal to $1 + t_i + \tau_j + T$, which can be interpreted as an ad valorem tax because the pre-tax price is normalized to one.

Individuals face a choice regarding how much of the consumption good to purchase, as

---

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

[4]To answer the question posed in this paper concerning inter-federation competition, a state government is not strictly necessary. However, when I take this question to the data, county governments will inevitably fall under the jurisdiction of a third level of government – the state – and will compete with localities and the unifying state government.

well as whether to purchase the good at home or abroad. When purchasing the good in the home town of residence, the consumer goes to the store located at the same point of the line segment that she lives on. If this is the case, 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 $C_i(d)$ is assumed to be linear in distance to the border, d, such that $C_i(d) = c_i d$. Note c_i denotes a constant per unit of distance cost for traveling to the border and is independent of x, so that the amount of the good purchased does not change the transportation cost of the buyer. All cross-border shoppers will purchase the good from the first store in the neighboring jurisdiction and are constrained from shopping multiple towns over.[5]

Individuals will cross-border shop if the utility benefit from shopping abroad is larger than the utility received from purchasing the good at home. Denote $v(q) = \max [u(x) - qx]$ as the indirect utility from the taxed good. Denote $x(q) = \text{argmax}\,[u(x) - qx]$ 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 (the benefit) with the transportation cost function (cost), it is easy to see that a consumer living in i will only shop in $k \neq i$ if $q_i > q_k$ and if she lives at a distance of

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

from the border of town k.

Governments compete in a Nash game. Counties and towns are considered simultaneous movers in the game.[6] The state tax rate is exogenously fixed at T under the assumption that any individual county or town is small and cannot affect the state tax rate. Governments are assumed to be Leviathans and the objective function of governments is to maximize revenue

$$R_i = t_i B_i \tag{2}$$

where B_i is the tax base defined below.[7] The tax base for a town is influenced by tax rates in neighboring towns $i + 1$ and $i - 1$, where these neighboring towns may be in the same

---

[5]Relaxing the assumption of shopping in only one town over is very difficult. It would require a set of inequalities governing every possible cross-border shopping possibility. Doing so would effectively make characterizing the equilibrium excessively complex. The assumption of shopping one town over will create stark results for contiguous neighbors. Relaxing this assumption would likely imply that these stark results will hold to a lesser degree for any neighbor that is proximate enough to shop within a particular region.

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

[7]Maximizing revenue is a simplifying assumption that allows for explicit characterization of the Nash equilibrium in the model. Revenue maximization is equivalent to welfare maximization when individuals place a high marginal valuation on the public good in comparison to private consumption.

county or in a different county. Therefore, the tax base will be a function of a jurisdiction's tax rate as well as its neighbors' rates.

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. Because I am introducing asymmetric federations in the model, county tax rates will be different in equilibrium because counties differ in size. With an infinite line segment, there are an infinite number of possible cases to consider, so simplifying assumptions need to be placed on the problem. Recall that counties are indexed $j = 1, 2, ...M$ and contain $m + j$ towns. I assume that the length of a town is identical for all towns in the model. 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 can conclude that the Nash equilibrium of county tax rates will follow the following pattern: $\tau_1 < \tau_2 < ...\tau_M$. This assumption places no restriction on the pattern of local taxes within a county. 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.[8] 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.

Defining $\rho_i = \frac{\phi_i}{c_i}$, the tax base for towns can now be written as

$$B_i = x(q_i)[n_i + \rho_{i+1}(v(q_i) - v(q_{i+1})) - \rho_i(v(q_{i-1}) - v(q_i))] \qquad (3)$$

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

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. The key is to see how the tax rate of a competing jurisdiction affects marginal revenue in the region of marginal revenue equal

---

[8]Relaxing this assumption would not change the functional form of the elasticities, but it would distort the pattern of the subscripts on these equations. The presence of a tax gradient allows for a pattern in the perceived elasticities.

to zero. The sign of the slope of the reaction functions tells whether the vertical externality creates upward or downward pressure on tax rates through a competitive process. The steepness of this function informs the researcher as to how responsive towns are to the externality. 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.

3.1   Nash Equilibrium

Under certain conditions, a Nash equilibrium is guaranteed to exist.[9] The solution to this game can be solved locally by considering arbitrary counties and towns at the interior of the line. The solution characterizing this random county will hold for all other interior counties because the towns follow the same pattern within a county. For ease of notation, write $x(q_i)$ as x_i. Derivatives are denoted with a prime. The local tax rates are implicitly defined by the reaction function:

$$\frac{\partial R_i}{\partial t_i} = B_i + t_i \frac{\partial B_i}{\partial q_i} = x_i s_i + t_i s_i x_i' - t_i x_i^2 (\rho_i + \rho_{i+1}) = 0, \tag{4}$$

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

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

$$\frac{t_i}{q_i} = \frac{1}{-\frac{q_i}{B_i}\frac{\partial B_i}{\partial q_i}} = \frac{1}{\varepsilon_i + \theta_i}, \tag{5}$$

where $\varepsilon_i = -\frac{q_i x_i'}{x_i}$ is the elasticity of demand for the consumption good and $\theta_i = \frac{q_i x_i (\rho_i + \rho_{i+1})}{s_i}$ is the elasticity of the number of shoppers in town i accounting for both the in- and out-flows. Both elasticities are defined to be positive numbers under the assumption that demand curves slope downward.

---

[9]Some of the conditions for existence of a Nash equilibrium include that municipalities must be sufficiently large in size such that tax rates are positive, towns are not composed of swing-shoppers who purchase goods multiple towns away, and all towns are composed of at least some shoppers with non-zero demand. This paper is not focused on characterizing the existence and uniqueness of an equilibrium.

Similarly for counties, the reaction function can be similarly defined. However, as I will focus on municipalities in the empirical analysis to follow, I will not emphasize the slope of the reaction functions for county governments.

3.2 Strategic Interaction

The two equations above implicitly determine tax rates as a function of the county and state tax rates. In measuring the response to vertical externalities, it is traditionally assumed that the reaction is top-down. Using the implicit function theorem and Roy's identity, the slopes of the reaction functions can be calculated. Devereux, Lockwood and Redoano (2007) – in a two town, one federation model – prove that the slope of the local government's reaction function may be positive or negative depending on the relative sizes of ε_i, θ_i, and the curvature of the demand function. Devereux, Lockwood and Redoano (2007) prove this for a symmetric Nash equilibrium – where there is no cross-border shopping in equilibrium. I relax the symmetry assumption.

It is useful to state the conditions found in Devereux, Lockwood and Redoano (2007) for upward and downward sloping reaction functions. Let $\eta_i = \frac{q_i x_i''}{x_i'}$ denote the curvature of the demand function. If $\theta_i - \varepsilon_i - \eta_i > 0$, the reaction function 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. θ_i represents the strength of horizontal tax competition and can increase if transportation costs fall or the population density increases near the border.

In the subsequent sections, I consider how inter-federation competition affects the strategic response of local governments. For all the following proofs, I assume that demand is either iso-elastic or log-linear. This assumption simplifies the intuition by holding x_i'' and x_i' constant along the demand curve, which is likely a good assumption in a local region of a Nash equilibrium. Propositions one and two assume θ_i is constant within a federation, but proposition 3 seeks to relax this assumption.

3.2.1 Slopes With Respect to County Tax Rates

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.

Proposition 1. *In the neighborhood of a 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.

First, all of the slopes of the reaction functions are "partial" derivatives in the sense that they do not account for the fact that changes in the county rate induce a town's neighbors to change their tax rate as well. The derivatives derived below only account for the direct effect of a change in the county tax rate and, therefore, can be interpreted as a short-run response before other jurisdictions have the chance to respond.

Consider the case of a homogeneous ρ for all towns. Because the slope of the reaction function measures how responsive localities are to county tax rates, a larger slope implies the vertical externality is more likely to generate upward sloping reaction functions. 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 of this town via three channels. It will change the demand function for individuals because x is a function of q. Changes in the county tax rate will also distort the number of individuals living in i and purchasing goods in $i-1$.[10] It will also distort the inflow in cross border shopping from town $i+1$. However, because the price rises by the same amount in both locations, it mitigates the change in cross-border shopping relative to if prices rose in only one town. 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. However, the post-tax price remains unchanged on one side of the town's border. Therefore, for a town neighboring a high-tax county, an increase in the county rate will substantially reduce inflows. For a town neighboring a low-tax county, an increase in the county tax rate will substantially increase outflows because the rate changes on only one side of the border. Either way, the tax base becomes much smaller relative to an interior town and the elasticity becomes larger. Therefore, the peripheral town has more incentive to lower its rate in response to a higher county rate relative to the interior neighbor.

Another way to say this is that being located on the periphery makes it more likely that the reaction function is downward sloping. The intuition is that periphery towns want to

---

[10]Mathematically, this is a subtle point. Recall that the number of cross-border shoppers leaving a jurisdiction is given by $v(q_{i-1}) - v(q_i)$. In this case, both q_{i-1} and q_i increase by the same amount. The implication of this is that $v(q_{i-1}) - v(q_i)$ changes because v is not linear in q.

capture additional cross-border shoppers from the neighboring county if they are in a low-tax county or they want to discourage their residents from leaving the county if they are in a high-tax county. Unlike interior towns, this leakage or capture is exaggerated when the county tax rate increases. For interior towns, the leakage changes somewhat, but remains relatively similar.

Mathematically, define the following cross-price elasticity

$$\theta_{i,k} = \frac{q_k}{B_i} \frac{\partial B_i}{\partial q_k} \frac{q_i}{q_k} = \begin{cases} \frac{\rho_k x_k q_k}{s_i} \frac{q_i}{q_k} & \text{for } i < k = i+1 \\ \frac{\rho_i x_k q_k}{s_i} \frac{q_i}{q_k} & \text{for } i > k = i-1 \end{cases} \tag{6}$$

Recall $\theta_i = \frac{q_i x_i (\rho_i + \rho_{i+1})}{s_i}$ is the response of town i's number of shoppers with respect to changes in the town's own price, q_i. 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. For ease of interpretation, the elasticity is scaled by the price ratios to account for the fact that the equilibrium is asymmetric. Note that if in a symmetric equilibrium, θ_i is approximately two times $\theta_{i,k}$ because the town's own price influences two borders.

Define $D_i = -\frac{\partial^2 R_i}{dt_i^2} = 2\varepsilon_i^2 + 2\theta_i^2 + \varepsilon_i \eta_i + \theta_i \varepsilon_i$ as the negative of the second derivative of the town's revenue function. The concave shape of the Laffer curve guarantees that D_i is positive.

I can then characterize the slope of the reaction functions for periphery and internal towns as follows. The slope of the reaction function for towns at the periphery is given by

$$[\varepsilon_i(\theta_i - \varepsilon_i - \eta_i) + \theta_i(\theta_{i,k} - \theta_i)]/D_i \tag{7}$$

where $k = i+1$ if i is the left-most town in a county (i.e. a town with outward cross-border shopping across county lines) or where $k = i-1$ if i is the right-most town in a county (i.e. a town experiencing inward cross-border shopping across the county line). The slope of the reaction function for the internal towns are give by:

$$[\varepsilon_i(\theta_i - \varepsilon_i - \eta_i) + \theta_i(\theta_{i,i+1} + \theta_{i,i-1} - \theta_i)]/D_i \tag{8}$$

Note that for peripheral towns, $\theta_{i,k}$ will enter only once into the slope of the reaction function because a change in its county rate only changes the neighboring rate on one border. And $\theta_{i,k}$ is more inelastic than θ_i for small changes in price. Intuitively this is because θ_i accounts for changes in the number of shoppers resulting from changes across two borders. An increase in the price in your own town causes a distortion to the amount of cross-border

shopping to both the left and right.[11] However, the cross-elasticity only influences your tax base via one border. An increase in the neighbor's tax rate will cause less leakage out of your town or more entry in to your town, but not both. How large this distortion is will depend on ρ and increases as density increases or if costs of shopping decline. More generally, Proposition 1 can be presented as:

Corollary 1. *In the neighborhood of a Nash Equilibrium, the slope of a town's reaction function with respect to county tax rates is increasing in the number of within county neighbors if all towns have the same number of total neighbors. Keeping in mind that this slope may be positive or negative, the more within county neighbors a town has, the more likely the reaction function is upward sloping.*

3.2.2 The Diagonal Externality

Towns at the periphery of a county are also affected by the county tax rate in the neighboring county. 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 externality 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 this externality a diagonal externality.

This raises the question of whether the slope of the reaction function with respect to the neighboring county tax rate is any different than with respect to the neighboring town's tax rate. Or is this just another horizontal externality?

Proposition 2. *In the neighborhood of a Nash equilibrium, the slope of a peripheral town's reaction function with respect to the neighboring county tax rate is identical to the slope of the reaction function with respect to the neighboring town's tax rate.*

The implication of this proposition is that the diagonal externality is exactly equivalent to a horizontal externality.[12] Consider a town that has only one neighbor and each town is in a separate county. If the neighboring town raises its tax rate, this gives rise to a classical horizontal externality. The county governments competing with each other also generate a horizontal externality on each town's base as well. 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

---

[11]Mathematically, $v(q_i)$ enters twice into the tax base.

[12]If the assumption of shopping one town over were relaxed, diagonal and horizontal externalities would not be identically equivalent, but would still look similar to each other in sign.

16

the town raising its rate or a higher level-of government (which has no jurisdiction over her home rate) raising its rate.

This may seem intuitive and simple, but the theoretical and empirical literature have ignored this externality. Failing to account for this externality in any regression will yield to biased estimates of the horizontal externality if the researcher specifically desires to estimate the same level of government horizontal effect.

Mathematically, consider town i's reaction function. Town i is located at the periphery of its own county. The slope of the reaction function with respect to the neighboring county's tax rate is

$$[\theta_i \theta_{i,k}]/D_i \tag{9}$$

where $k = i+1$ if in a relatively low-tax county or $k = i-1$ if in a high-tax county relative to the neighbor. Of course, this slope is in a local region of the Nash equilibrium and, therefore, the low-tax county is not allowed to become the high-tax county. Differentiating the reaction function with respect to the neighboring town yields the same expression.

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. Also recall that the elasticity θ increase with the intensity of horizontal tax competition – it increase when density near the border increases or when transportation costs fall.

This result is important because it demonstrates that the diagonal externality behaves in a manner identical to the horizontal externality. Empirically, any regression specification attempting to estimate the true vertical externality must account for this by including the effective total neighboring tax rate on the right hand side. Surprisingly, the few empirical specifications involving multiple federations have ignored this diagonal externality and have only included the neighboring local tax rate on the right side of the equation.

3.2.3 How Do Tax Differentials Change the Nature of the Strategic Interaction?

In this section, I relax the assumption that θ is constant within a federation by considering how the size of the discontinuity in tax rates influences the strategic interaction.

Proposition 3. *The closer the tax rate of an interior town i is to the tax rate of its high-tax neighbor (relative to the low-tax neighbor), the more likely the reaction function of town i will be upward sloping with respect to the county tax rate.*

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.[13] The implication of this is that the elasticity of the tax base with respect to the town's own price θ_i is small relative to the cumulative cross 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 $(\theta_{i,i+1} + \theta_{i,i-1} - \theta_i) < 0$.

Consider Figure 3.2.3. 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.

The difference in the indirect utilities above are the level-changes – under the assumption of uniform ρ. The elasticities are the percentage changes and are proportional to this and can be generalized for heterogeneous costs.

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.

---

[13]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.

3.2.4 Interaction With State Tax Rates

Changes in the state tax rate will also prompt a response in the town jurisdictions. Here, the slope of the reaction function with respect to the state rate will have an identical form for all towns by assumption. In the empirics, diagonal externalities resulting from state tax rates are also likely to exist. As a result, I will control for tax differentials and a functional form of distance to the state borders. As a robustness, check I also drop border counties. However, I am unable to simply incorporate these state interactions into the simple model above. Intuitively, the state interactions with towns should be similar to the county interactions with town if people view changes in county rates similar to changes in state rates. Of course, state rates may be more salient and in the model above, the state acts exogenously.

4   Implications for Empirical Analysis

The model above highlights several implications for the empirical analysis. The traditional empirical analysis is to run a regression given by the following equation:

$$t_{ij} = \alpha_0 + \alpha_1 t_{-i} + \alpha_2 \tau_j + \delta X_{ij} + \gamma F_j + \varepsilon_{ij} \tag{10}$$

where X are local controls and F are federal controls or dummies. Define t_{-i} as the weighted average of other neighboring tax rates such that

$$t_{-i} = \sum_{k \neq i} w_{ik} t_k. \tag{11}$$

Denote w_{ik} as exogenous weights normalized such that $\Sigma_{k \neq i} w_{ik} = 1$.[14] Tax rates on the right-hand side are usually instrumented for with the weighted average of the neighbors' X. The specification above is the one implemented in Devereux, Lockwood and Redoano (2007). The coefficient on t_{-i} identifies the strategic reaction in response to a jurisdiction's neighboring tax rates. The coefficient on τ_j estimates the strategic reaction to the vertical externality or $\frac{\partial t}{\partial \tau}$.

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.[15] Two, the right side of the regression equation must include the neighboring county rate in addition to the neighboring town rate. This variable may

---

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

[15]This will help to separate the strategic reaction from the fact that similar jurisdictions are being hit by common unobserved shocks – as these shocks are unlikely to be correlated with distance to the border.

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 relevant variables on the right side. For example, I show that increases in horizontal externalities result in a steeper vertical reaction function if $\varepsilon > \theta$. This implies a systematic correlation between county rates and neighboring town rates. Therefore, the true measure of the vertical externality must be α_2 plus an interaction with horizontal externalities. In fact, Devereux, Lockwood and Redoano (2007) acknowledge the importance of this interaction theoretically, but omit it empirically. I explain this point in further detail below.

4.1 Why Interaction Effects Are Essential

The standard literature has estimated vertical externalities using only the level of the local tax rate and the federation tax rate. From the slopes of the reaction functions, it is evident that θ_i captures the strength of the horizontal externality. Recall that θ_i represents the response in the number of cross-border shoppers and that it is a function of how costly cross-border shopping is and the density at the border of the town. Because the slopes of the reaction functions depend on θ_i, the reaction to the vertical externality becomes more intense as the horizontal externality 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 of neighboring local rates with the federation rate. Any specification omitting this interaction will suffer an omitted variable bias.

To see this empirically, it is useful to consider a multi-level model[16] of tax competition.[17] 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 – which omits the additional controls of equation (10) – on neighboring tax rates

$$t_{ij} = \alpha_{0j} + \alpha_{1j}t_{-i} + \varepsilon_{ij}, \tag{12}$$

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 τ_j, which only vary across the j level and not within the j level of the model, affect t_{ij} with some error. The

---

[16]For a summary of multi-level modeling, see Franzese (2005).

[17]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, where all of the multiple levels are substituted into the estimating equation. The reason for this is that estimating the model using hierarchical linear modeling will place stricter assumptions on the nature of the error term. Therefore, it is preferable to estimate the nature of the strategic interaction as a single equation.

following equation adds this effect to (12):

$$\alpha_{0j} = \gamma_{00} + \gamma_{01}\tau_j + u_{0j}. \tag{13}$$

It is clear that substituting (13) into (12) will yield (10) sans controls. This is where the literature on tax competition within federations stops. However, the theoretical results in this paper and in Devereux, Lockwood and Redoano (2007) 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 τ_j and vice versa. Assuming that this interaction also occurs with error, this implies

$$\alpha_{1j} = \gamma_{10} + \gamma_{11}\tau_j + u_{1j}, \tag{14}$$

and substituting (13) and (14) into (12) yields

$$t_{ij} = \gamma_{00} + \gamma_{01}\tau_j + (\gamma_{10} + \gamma_{11}\tau_j)t_{-i} + u_{1j}t_{-i} + u_{0j} + \varepsilon_{ij}, \tag{15}$$

which is not the same as equation (10). The implication is that the existing literature has estimated the strategic reaction as $\frac{\partial t}{\partial \tau} = \gamma_{01}$ despite the true reaction being $\frac{\partial t}{\partial \tau} = \gamma_{01} + \gamma_{11}t_{-i}$ and where the estimates of γ_{01} are different across the two specifications because they are derived from different models. Failure to account for this interaction effect will yield biased estimates of γ_{01} where the bias is given by $\hat{\gamma_{11}}\frac{cov(t_{-i}, t_{-i}\tau_j)}{var(t_{-i})}$ with $\hat{\gamma_{11}}$ expected to be positive.

5 Empirical Methodology

I propose estimating the nature of the strategic interaction in a cross-sectional context. The following sections outline the data available to me along with the proposed methodology.

5.1 Data

I have a unique cross-sectional data set from April 2010 that includes local, county, and state tax rates for all jurisdictions in the United States.[18] In addition to the tax data, I have also generated several comprehensive data sets concerning the spatial proximity of jurisdictions. Using ArcGIS software, I have generated the following variables: the driving distance from

---

[18]The data is proprietary but was provided to me for free. For a complete description of the data see http://www.prosalestax.com/

the population weighted centroid of each Census Place[19] to the nearest intersection of a major road crossing at state and county borders – denoted D; measures of neighborliness of Census Places with respect to other places, including the distance to every other Census place within a fifty mile region of every town; and the length of each jurisdiction's perimeter and area.

Driving distances 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.[20] Population weighted centroids are calculated as the balance point at which an imaginary flat surface of the Census place would balance given the population distribution of Census blocks within the place. The driving distance calculated is the distance that minimizes the time to drive to the closest border. For a detailed description of this calculation, see Agrawal (2011).

Contiguity is often not a satisfactory measure of horizontal competition because many places have zero contiguous neighbors if completely surrounded by unincorporated places. To calculate broader measures of neighborliness, I define a jurisdiction as neighbors if they are within a twenty-five or fifty mile buffer of each other.[21] To measure the diagonal externality, I define the neighboring county as the county that is closest to the town based on the distance criteria above.

Control variables are from the 2000 United States Census plus geographic and political controls that I have created myself. The set of possible control variables include a municipality's area, perimeter, number of contiguous neighbors, population, the fraction of individuals who work in state or in county, average income, education, the fraction of seniors, whether the town is near the ocean or an international border,[22] and the Obama vote share from 2008. 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

---

[19]A Census Place is generally an incorporated place with an active government and definite geographic boundaries such as a city, town, or village. In many western states, a Census Place may be an unincorporated place that has no definite boundaries or government. 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.

[20]A major road is a Census classification including most non-residential roads. I omit residential roads for computational ease as well as a result of the fact that little cross-border shopping will occur on residential roads because they will not be necessarily proximate to shopping facilities.

[21]I define a buffer as a region that is twenty miles from the exterior perimeter of a town. Because this region is defined from the exterior rather than the centroid, it need not be a circle. Then, any place that intersects this buffer region (either completely or only partially) is defined as a neighbor. I also know the exact distance between each jurisdiction in case weighting by distance within this region is desirable.

[22]These dummy variables take on the value of one if the nearest state border would be an international border or a major body of water. Observations where the nearest county border is also an international border are dropped from the sample because the IV procedure outlined below would not be able to be used for these jurisdictions.

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 distance to the nearest state border and the complete set of interactions of these variables.[23] Including these terms will help to control for diagonal externalities across state lines, but which are not the focus of this paper.

5.2   Estimation in a Cross-Section

To make the empirical estimation feasible, I assume that tax rates have been in equilibrium for a sufficient amount of time or put differently, that the national data set of sales taxes in 2010 is out of equilibrium due to large shocks. I will focus on how municipalities react to the county level rates, neighboring county rates, and neighboring town rates. I propose estimating

$$
\begin{aligned}
t_{ij} = \quad & \alpha_0 + \underbrace{\alpha_1 \tau_{i,j}}_{Vertical} + \underbrace{\alpha_2 t_{-i}}_{Horizontal} \underbrace{\alpha_3 t_{-i}\tau_j}_{Interaction} + \underbrace{\alpha_4 \tau_{i,-j}}_{Diagonal} \\
& + \underbrace{\alpha_5 \tau_{i,j} d_i}_{Vertical-Distance} + \underbrace{\alpha_6 \tau_{i,-j} d_i}_{Diagonal-Distance} \\
& + X_{ij}\beta + S_s + \varepsilon_{ij}
\end{aligned} \tag{16}
$$

where t_{ij} is municipal plus sub-municipal local option taxes and t_{-i} is defined in equation 11. 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.[24] 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

$$
w_{ik} = \begin{cases} \frac{1}{n_i} & \text{if } k \in N_i \\ 0 & \text{if } k \notin N_i \end{cases} \tag{17}
$$

where n_i is the number of towns in N_i.

The matrix X_{ij} are the local controls listed above and S_s are state fixed effects – that control for the level of the state tax rate in a state along with other within state policies. Then, $\tau_{i,j}$ is defined as the county tax rate that town i is located in[25] and $\tau_{i,-j}$ is the nearest

---

[23]These are the same terms in Agrawal (2011) except that I impose $\log(d)$ as the distance function rather than the quintic function of distance in that paper. Imposing this log functional form is consistent with the results in Agrawal (2011) where the marginal effects of distance are steepest near the border and decreasing to zero in absolute value. Additionally, because this paper does not focus on diagonal externalities at state borders – and instead focuses on county borders – the precision of the quintic polynomial is not necessary.

[24]I will show the results are robust to various weighting schemes in future sections.

[25]Towns that cross multiple county borders are dropped from the sample.

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 the diagonal externality discussed above only manifests itself for the nearest county neighbor. One reason for this assumption is that I would like to test how the diagonal externality varies with distance, which would not be feasible if multiple counties are considered as neighbors. Making this assumption allows me to reduce what would me a multi-dimensional 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 either linear in distance to the county border, D_i, or the log of it, $\log(D_i)$. The inclusion of this distance function is driven by the theory and the use of the log specification implies that the effect of distance will be close to zero for very large distances. Proposition 1 states that the strategic reaction to one's own county tax rate is more likely to be positive for towns closer to the border.

Note that the strategic reactions are now given by the mean analytic derivatives of the estimating population:

$$
\begin{aligned}
E[\tfrac{\partial t}{\partial \tau_j}] &= \frac{1}{M} \sum_{i=1}^{M} (\alpha_1 + \alpha_3 t_{-i} + \alpha_5 d_i) \\
E[\tfrac{\partial t}{\partial t_{-i}}] &= \frac{1}{M} \sum_{i=1}^{M} (\alpha_2 + \alpha_3 \tau_{i,j}) \\
E[\tfrac{\partial t}{\partial \tau_{-j}}] &= \frac{1}{M} \sum_{i=1}^{M} \alpha_4 + \alpha_6 d_i,
\end{aligned}
\tag{18}
$$

where M is the total number of observations in the estimating sample. Standard errors for mean derivatives are calculated using the Delta Method. The marginal effects given by Equation (18) are a consistent estimate of the mean derivative in the conditional population.

In addition to the spatial lags of the municipal tax variables, the error term is allowed to follow a spatial process such that ε is allowed to be auto-correlated across towns. The error is assumed to have the following first-order spatial auto-regressive form:

$$
\varepsilon_{ij} = \rho w_{ik} \varepsilon + v_{ij} \tag{19}
$$

where ε is i.i.d over space and ρ is the spatial auto-regressive coefficient.

The presence of the spatial lags of tax rates on the right hand side of the equation implies that standard ordinary least squares results will be biased because neighboring town and county tax rates are endogenous; tax rates are determined simultaneously. Two possible solutions exist – maximum likelihood estimation (Case, Hines and Rosen 1993) and instrumental variables estimation (Figlio, Koplin and Reid 1999). The presence of both

horizontal, vertical, and diagonal externalities in the estimating equation along with spatial auto-correlation in the error term make use of the maximum likelihood method extremely difficult. 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 for t_{-i} are the weighted average of several control variables. Simply put, the standard instrument for neighboring tax rates is $\Sigma_{k \in N_i} w_{ik} x_{kj}$, where x_{kj} is a column 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. I, however, am less inclined to believe that the neighbors' X's (for example, population) have no direct effect on town i's tax rate. In fact, standard theoretical models imply that this is exactly the case. Although population directly influences town k's tax rate, which then indirectly changes town i's tax rate, town k's population – and density – directly determine the elasticity of demand that town i realizes.

Instead of using the entire subset of the X's as instruments, I will only use geographic variables (which are not often controls in previous studies) as instruments. Specifically, I will use area and perimeter 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 $t_{-i} = \Sigma_{k \in N_i} w_{ik} t_{kj}$, I use $\text{area}_{-i} = \Sigma_{k \in N_i} w_{ik} \text{area}_{kj}$ and $\text{perimeter}_{-i} = \Sigma_{k \in N_i} w_{ik} \text{perimeter}_{k,j}$ as instruments. Of course, the regression specifications above also include interaction terms, in which case they are instrumented for with the interactions of the instruments and the respective term.

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. 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. 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 helps to make them exogenous. As a contrary point, 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.[26]

One final point is that omitted controls may be correlated with the error term – in which case the estimates from the IV approach will be inconsistent. Such a correlation could arise if endogenous sorting based on tax rates occurs. This correlation could also arise from the fact that control variables in the cross-sectional data set are constrained to variables only from the long form of the United States Census. One solution is to exploit a panel data and first-difference the equation. First differencing the data would eliminate any omitted variables that are relatively independent of time. Unfortunately, no publicly available panels of local tax rates exist for the national sample. Panel first-differencing would also increase inefficiency and would weaken the instrumental variables, perhaps dramatically.

6 Empirical Results

6.1 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 16. The coefficient representing the horizontal interaction, t_{-i}, is expected to be positive because towns mimic their neighbors. The diagonal externality represented by $\tau_{i,-j}$ should be positive because the diagonal externality is identical to a horizontal externality in the local region of the border. The vertical externality, τ_j, is ambiguous as the relative magnitudes of θ and ε determine the sign of the vertical externality. The interaction effect $t_{-i}\tau_j$ may also be ambiguous. but if $\varepsilon > \theta$, then horizontal externalities will fuel vertical externalities and the coefficient should be positive. The vertical externality will be affected by distance $d_{ij}\tau_j$ in a positive manner as interior towns are more likely to mimic the federation's tax rate because the leakage from a county government increase is mitigated by being far from the border. Finally, the diagonal externality will be affected by distance through $d_{ij}\tau_{i,-j}$ and the effect is expected to be negative; interior towns are less likely to mimic the neighboring federation's tax rate because they are far from the border.

---

[26]I conduct a Hansen J test of over-identification in each specification. The results of this test – a failure to reject the null hypothesis that all the instruments are uncorrelated with the error – is suggestive that in the presence of one valid instrument, the other instrument is also valid.

6.2 Main Results

Table 2 presents the baseline results using GMM-IV estimation, where the measure of the horizontal interaction is the average tax rates within a fifty mile region (i.e., a buffer) of the town of interest. All of the regression equations in the table contain state fixed effects and the control variables discussed above. The coefficients on the endogenous regressors from the second stage are reported in addition to the mean derivatives, which are the slopes of the reaction functions after accounting for interaction effects and distance. 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 strategic interaction and, therefore, will be the focus of the following discussion. Each of the columns from 1 to 7 build sequentially on what the current literature has estimated, with column 7 being the specification that the theory suggests is accurately specified.

Column 1 presents the results where the vertical reaction 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 largely negative result that is consistent with the regression discontinuity results in Agrawal (2011). Several explanations exist for the opposite finding. Towns may react in 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, θ, is more likely to be larger for local governments than for state governments – and the theoretical model implies the larger θ is relative to ε, 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, perhaps indicative of yardstick competition.

Column 3 presents a similar specification to Devereux, Lockwood and Redoano (2007), where the vertical and horizontal reactions are estimated jointly, but without any interaction effects. In general, Devereux, Lockwood and Redoano (2007) find positive coefficients on the sign of the horizontal interaction and positive but insignificant results for the vertical interaction. 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 .836 percentage points. Contrarily, a 1 percentage point increase in the average of the neighbors tax rate, increases a municipality's local tax rate by .456 percentage points. Keeping in mind that this is the most robust specification currently estimated in the

literature, it is useful to compare the future results to these benchmark numbers.

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. In the case of two 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%.[27] In most every specification where critical values are tabulated, the bias is less than 5%. 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 externalities and specification 5 accounts for the diagonal externality. Looking at the mean derivatives, the slopes of the reaction function fall by almost .15 percentage points. Although the diagonal externality is of an unexpected sign, the result is not significant at traditional levels in most specifications. Specification 6 and 7 (the preferred specification) add in the interaction of the vertical and diagonal externalities with a distance function. Notice that the slope of the reaction function with respect to the county tax rate falls to -.640 and the slope of the horizontal reaction function falls to .399. The estimates in column 3 are 30% and 15% larger than the vertical and horizontal reactions in the complete specification of column 7. 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.

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 six endogenous regressors. With many variables in need of instrumenting, the concern of weak instruments become more worrisome. As an alternative, I estimate specification 7 by limited information maximum likelihood (LIML). If the estimating equation is correctly specified and the distributional assumptions hold, LIML will produce unbiased coefficient

---

[27]In some cases, Stock and Yogo (2005) did not calculate critical values for the number of endogenous regressors. For these specifications, I will discuss an alternative method to determine if the estimates suffer a weak instrument problem.

estimates even if weak instruments are present. Because the point estimates in column 10 are similar to column 7 and because the mean derivatives are still smaller than specification 3, it suggests that the absence of weak instruments in the GMM estimates.

Before proceeding to the robustness checks, it is worth summarizing the implications of the results above. Possible misspecifications of the identifying equation by failing to account for diagonal externalities, the interaction of horizontal and vertical externalities, and the distance to federation borders will yield the researcher to believe that the slopes of reaction functions are larger in absolute value than they actually are. Therefore, when estimating vertical and horizontal reaction functions it is essential to include these terms in the regression equation. Even if the federation has no horizontal competitors of its own, the the interaction effects of neighboring jurisdictions and higher levels of jurisdictions are essential variables. Lastly, if the federation has multiple competitors of its own, diagonal reactions and reactions contingent on distance to the nearest federation should be included in the estimating equation if possible.

Finally, the magnitudes of the strategic interactions in this paper are different and, for the vertical externalities, opposite in sign to the traditional literature. This suggests that using comprehensive local data will produce different strategic interactions than using state level data. Again, 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. Relatively few papers have analyzed local reaction functions and those papers (Revelli 2001; Büttner 2003) that have analyzed local data have often used data from only one state within a country. 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.

6.3 How Robust Are The Results?

In the following section, I will discuss three sets of robustness checks: redefining what constitutes a neighboring jurisdiction, specifying a linear distance function rather than a log distance function, and cutting the data set in a variety of ways.

6.3.1 Alternative Definitions of Neighborliness

In the results presented above, I assumed that jurisdictions competed with neighbors in a fifty mile region of the town borders. Such a measure of neighborliness is quite large. As an alternative, I use the average tax rate of jurisdictions within 25 miles as a measure of

neighborliness.[28] Table 3 presents the results. Columns 1 - 7 are identical in all respects to Table 2 except the measure of neighborliness is restricted to jurisdictions in a smaller vicinity of the town. Column 8 reproduces the previous column using LIML instead of GMM.

Again, column 3 is the specification currently estimated in the literature. The slope of the vertical reaction function implies a 1 percentage point increase in the county tax rate lowers municipal tax rates by .705 percentage points. A 1 percentage point increase in the average neighbor tax rate raises municipal tax rates by .387 percentage points. The sign and significance of these results is similar to Table 2, but the magnitude of the marginal derivatives is slightly smaller. In the preferred specification (including interaction effects, distance effects, and diagonal externalities), the slope of the vertical reaction function is -.439 and the slope of the horizontal reaction function is .374. With regard to the estimate of the reaction function with respect to the county rate, the slope is smaller than before. However, the results in column 3 are now 60% larger than the results from the complete specification in column 7, suggesting that the incomplete specification of column 3 potentially will induce a larger bias when the definition of one's neighbors is more restrictive.

6.3.2 Alternative Distance Functions

In all of the previous specifications, I have assumed that the distance function to county borders is $\log(D)$. Such a parametrization imposes that as distances become very large, the effect of distance converges to zero. Although some counties are very large in size, most counties are small in size. In fact, the average town is located approximately 8 miles from the nearest county border. Because distances to county borders are relatively small, the effect of distance on the strategic interactions may actually be linear in distance. In Table 4, I report the results when the distance function is linear. Columns 1, 2, and 3 correspond to the results in Columns 6, 7 and 10 from Table 2. Columns 4, 5, and 6 of Table 4 correspond to the results from columns 6-8 in Table 3.

When the distance function is linear, the point estimates change slightly but the interpretation of the results is similar in spirit. In some specifications, the interaction of the linear distance function and the county tax rate become significant at the 90% confidence level. For example, in Column 2, $\tau_{i,j}d_i$ has a positive and significant coefficient of .012. This implies that holding the county tax rate fixed, a one mile increase in distance from the nearest county border increases the slope of the reaction function by .012 percentage points. Such a result is consistent with the theory – interior jurisdictions are more likely to have upward sloping reaction function. However, the small magnitude of this coefficient suggests that

---

[28]I also try using jurisdictions within ten miles, but many jurisdiction then have no neighbors and the instrument validity comes into question.

the strategic reaction to the county rate will remain significantly negative for all reasonable distances.

6.3.3 Various Model Specifications and Sample Restrictions

Table 5 reports robustness checks for various sample restrictions and weighting schemes. Column 1 presents the baseline results. Column 2 drops jurisdictions for which the nearest county border is also a state border. This specification reduces the possibility that the neighboring state is also producing a diagonal externality on towns proximate to its borders and allows me to cleanly identify the diagonal externality resulting from neighboring counties. The measure of the diagonal interaction becomes larger in absolute value and the measures of the horizontal and vertical interactions become slightly smaller in absolute value. As a second robustness check, column 3 drops jurisdictions where the nearest state border would be an ocean or a major body of water. One reason for this is that theoretical models of tax competition using a Hotelling line segment where some jurisdictions have one neighbor and other jurisdictions have two neighbors produce very different results compared to if all jurisdictions have the same number of neighbors. Specifically, jurisdictions at the end of the line segment – i.e., proximate to the ocean – have incentives to raise tax rates because the elasticity of cross-border shopping is perceived as smaller. Dropping jurisdictions near water amplifies the strategic reaction of both the vertical and horizontal interactions. Such a result is suggestive that the elasticity of cross-border shopping is, in fact, larger for jurisdictions not near the ocean, which is consistent with the theoretical model.[29]

In columns 4 and 5, I restrict the sample to towns within ten or five miles of a county border, respectively. This specification would be most applicable for identifying the diagonal externality. The theory predicts that the diagonal externality is most salient and positive for a local region of county borders. Even in these specifications, where the effect is expected to be salient, it is still negative. However, in the more localized region of the border the diagonal strategic reaction shrinks in absolute value. This suggests that for towns relatively proximate to the county border, the diagonal externality is more likely, but not consistently, positive. Additionally, for towns in a local region of the border, the interaction of horizontal and vertical externalities plays a more important role.

Column 5 weights each observation in the sample by population such that cities are given more weight in the data set. For the weighted sample, the interaction effects of the vertical and horizontal interaction are strongest. However, the mean derivative with respect

---

[29]Such a statement requires that the demand function is identical for interior and ocean jurisdictions as differences in the magnitude of the vertical reaction could also be driven by various demand functions and the parameter η.

to the county tax remains negative and significant. Although the theory suggests that horizontal interactions are unambiguously positive, the weighted results suggest that the mean derivative with respect to neighboring tax rates is negative. This results fits with the existing empirical evidence that cities set much higher sales tax rates than their local neighbors, perhaps because the tax base is so large in cities that the benefit of raising city sales taxes outweighs the loss of cross-border shopping to the suburbs.

Finally, columns 7-9 alter the exogenous weights that determine t_{-i}. Define N_i as the set of towns within a fifty mile radius of town i, s_{ik} as the distance between town i and town k, and ϕ_k as the population of town k. Recall that $t_{-i} = \Sigma_{k \neq i} w_{ik} t_k$. Then column 7 specifies exogenous weights that are normalized to sum to 1 given by Equation 20:

$$
w_{ik} = \begin{cases} \dfrac{1/s_{ik}}{\sum\limits_{k \in N_i} \left(1/s_{ik}\right)} & \text{if } k \in N_i \\ 0 & \text{if } k \notin N_i \end{cases} ,
\tag{20}
$$

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 8 estimates the equation using the weights from Equation 21:

$$
w_{ik} = \begin{cases} \dfrac{\phi_k}{\sum\limits_{k \in N_i} \phi_k} & \text{if } k \in N_i \\ 0 & \text{if } k \notin N_i \end{cases} ,
\tag{21}
$$

such that neighboring towns within the fifty mile radius of town i are given more weight if they have a larger population. Lastly, Equation 22 specifies the weights used in Column 9:

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

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. The weights above and the weights given by equation 17 are the most common weights in the literature. The weights used in equation 17 are most likely to be interpreted as spatial weights, while the weights given by the three equations above are more indicative of economic flows.

When using distance based weights, the strategic interaction with the county tax rate remains approximately the same magnitude as the baseline specification, but the horizon-

tal interaction shrinks in absolute value. Population based weighting schemes shrinks both the slopes of the vertical and horizontal reaction functions in absolute value. Finally, the population-distance weights shrink the vertical interaction but increase the horizontal interaction relative to the baseline specification. Although I present these results as standard weighting schemes in the literature on tax competition, I believe they are less accurate for local competition than a simple average of neighboring tax rates. A recent theoretical model and survey results of local governments (Janeba and Osterloh 2011) 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. For this reason and because of its ease in interpretation, I prefer using the unweighted average of tax rates.

7 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 a diagonal externality – an externality induced by a different level of government that does not share the same tax base – that has similar consequences as a horizontal externality.

The theoretical predictions of the model shed light on the appropriate estimation strategy. Empirically estimating the nature of the strategic competition in a federation requires neighboring tax rates to be interacted with higher level of government's tax rate. Intuitively, this is driven by the fact that increases in horizontal competition have the potential to trigger additional vertical competition. In the presence of multiple federations, the neighboring federation's tax rate must also be included. Furthermore, because vertical externalities vary depending on whether a town is interior or peripheral to the federation, it is appropriate to look for the presence of a vertical externality that is heterogeneous with respect to distance to the border.

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 equilibrium tax rate. In specifications omitting these terms for local sales tax data, the bias in the estimate of the vertical externality is approximately 30% of the correctly specified reaction function.

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 and unambiguous in theory. 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 .60 percentage points; this suggests that county and municipal taxes are strategic substitutes. The result is inconsistent with 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, University of Michigan.

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.

Büttner, Theiss. 2003. "Tax Base Effects and Fiscal Externalities of Local Tax Competition: Evidence from a Panel of German Jurisdictions." *Journal of Urban Economics*, 54: 110–128.

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.

Hastings, Justine S. 2004. "Vertical Relationships and Competition in Retail Gasoline Markets: Empirical Evidence from Contract Changes in Southern California." *American Economic Review*, 94(1): 317–328.

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. 2011. "Tax and the City - A Theory of Local Tax Competition and Evidence for Germany." Working Paper, University of Mannheim.

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.

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.

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

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

Revelli, Federico. 2001. "Spatial Patterns in Local Taxation: Tax Mimicking or Error Mimicking?" *Applied Econometrics*, 33(9): 1101–1107.

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

Saggi, Kama, 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.

Villas-Boas, Sofia Berto. 2007. "Vertical Relationships between Manufacturers and Retailers: Inference with Limited Data." *The Review of Economic Studies*, 74(2): 625–652.

Figure 1: Reaction Functions
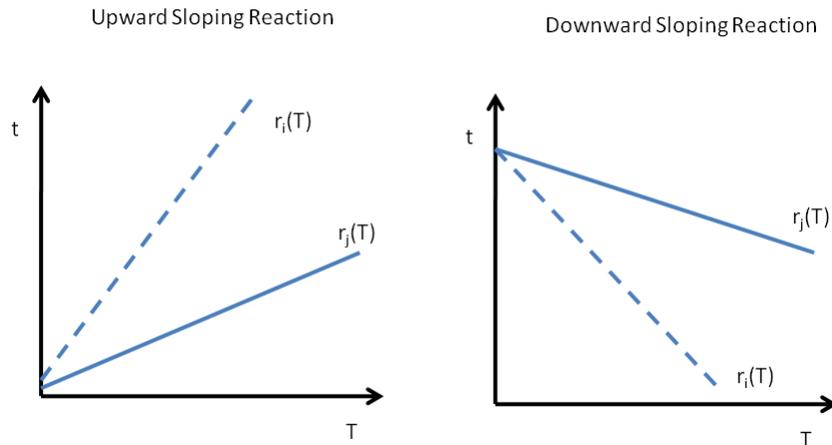
Figure 2: Change in Cross-Border Shopping due to a Change in the County Rate
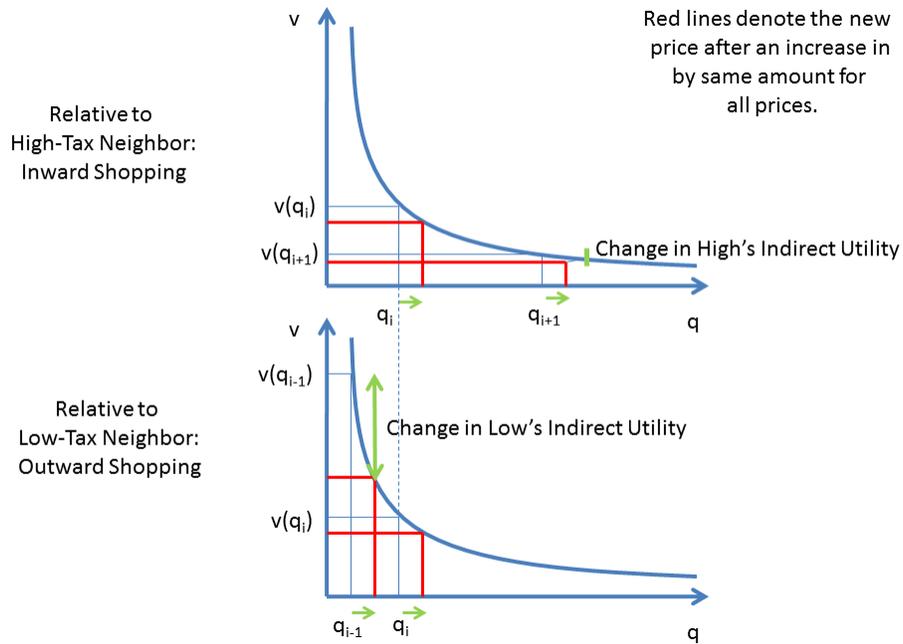
Table 1: Summary Statistics
Averages of Variables by Type
Standard Deviations in ( )

Type	Variable	Mean
Dependent	Local Tax Rate	.805
		(1.172)
Other	Distance to County Border	7.810
		(7.028)
	County Tax Rate	1.035
		(1.158)
Endogenous	Average Neighbors' Rate	.696
Regressors	(50 Mile Radius)	(.834)
	Average Neighbors' Rate	.698
	(25 Mile Radius)	(.906)
	Neighboring County Tax Rate	.980
		(1.123)
	Number of Neighbors	1.761
		(.678)
	Town Area	5.105
		(15.181)
	Town Perimeter	13.166
		(19.456)
	Population	7923
		(83,329)
	Senior (%)	16.006
		(7.872)
Control	College (%)	21.702
Variables		(14.078)
	Income	36,877
		(18,558)
	Near International Border	.082
		(.276)
	Near Ocean	.147
		(.354)
	Work in County (%)	68.980
		(20.462)
	Work in State (%)	96.109
		(43.325)
	Obama Vote Share	43.325
		(13.842)
	Sample Size	12,994

The log of distance 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 – Neighbors Within 50 Miles

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\tau_{i,j}$	-.912***		-.836***	-.655***	-.649***	-.571***	-.743**	-.803***	-.895***	-.860**
	(.114)		(.106)	(.159)	(.209)	(.177)	(.301)	(.112)	(.216)	(.355)
t_{-i}		.561***	.456***	.545***	.424***	.474***	.464***			.406**
		(.069)	(.070)	(.097)	(.161)	(.162)	(.167)			(.197)
$t_{-i}\tau_j$				-.112	-.051	-.082	-.066			-.023
				(.085)	(.118)	(.133)	(.141)			(.169)
$\tau_{i,-j}$					-.379	-.309	-.124	-.472***	-.198	-.148
					(.243)	(.239)	(.197)	(.164)	(.165)	(.246)
$\tau_{i,j}d_i$						.001	.090		.114	.119
						(.009)	(.079)		(.079)	(.092)
$\tau_{i,-j}d_i$							-.094		-.131	-.123
							(.085)		(.081)	(.098)
$E[\frac{\partial t}{\partial \tau_j}]$	-.912***		-.836***	-.745***	-.691***	-.638***	-.640***	-.803***	-.695***	-.672***
	(.114)		(.106)	(.116)	(.136)	(.114)	(.120)	(.112)	(.117)	(.138)
$E[\frac{\partial t}{\partial t_{-i}}]$		.561***	.456***	.431***	.374***	.393***	.399***			.383***
		(.069)	(.070)	(.070)	(.079)	(.075)	(.074)			(.078)
$E[\frac{\partial t}{\partial \tau_{-j}}]$					-.379	-.309	-.289	-.472***	-.428***	-.363
					(.243)	(.239)	(.227)	(.164)	(.143)	(.287)
Buffer	-	50	50	50	50	50	50	-	-	50
Distance	-	-	-	-	-	$\log(D)$	$\log(D)$	-	$\log(D)$	$\log(D)$
Method	GMM	GMM	GMM	GMM	GMM	GMM	GMM	GMM	GMM	LIML
Over-id†	.983	.748	.406	.605	.764	.691	.560	.511	.436	.5817
Weak Inst.‡	33.074	611.2	18.294	23.631	4.848	4.408	3.985	4.852	2.050	-
Critical‡	19.93	19.93	7.56	7.77	-	-	-	7.56	-	-
R^2	.598	.619	.626	.632	.601	.619	.621	.558	.579	.604
N	12,977	12,977	12,977	12,977	12,498	12,498	12,498	12,498	12,498	12,498

***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 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 3: Slopes of Reaction Functions – Neighbors Within 25 Miles

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\tau_{i,j}$	-.912***		-.705***	-.781***	-.366	-.362**	-.464*	-.422
	(.114)		(.124)	.235	(.247)	(.181)	(.269)	(.346)
$t_{-i,j}$		.602***	.387***	.312**	.487**	.462**	.471**	.419
		(.081)	(.101)	(.154)	(.233)	(.207)	(.192)	(.250)
$t_{-ij}\tau_j$				.065	-.119	-.089	-.090	-.030
				(.142)	(.138)	(.134)	(.130)	(.070)
$\tau_{i,-j}$					-.158	-.183	-.071	-.156
					(.253)	(.243)	(.187)	(.255)
$\tau_{i,j}d_i$						-.002	.057	.044
						(.009)	(.074)	(.089)
$\tau_{i,-j}d_i$							-.062	-.050
							(.078)	(.093)
$E[\frac{\partial t}{\partial \tau_j}]$	-.912***		-.705***	-.728***	-.463***	-.439***	-.438***	-.370**
	(.114)		(.124)	(.149)	(.157)	(.119)	(.118)	(.146)
$E[\frac{\partial t}{\partial t_{-i}}]$		.602***	.387***	.379***	.370***	.374***	.382***	.389***
		(.081)	(.101)	(.010)	(.132)	(.117)	(.111)	(.128)
$E[\frac{\partial t}{\partial \tau_{-j}}]$					-.158	-.183	-.179	-.244
					(.253)	(.243)	(.219)	(.311)
Buffer	-	25	25	25	25	25	25	25
Distance	-	-	-	-	-	$\log(D)$	$\log(D)$	$\log(D)$
Method	GMM	GMM	GMM	GMM	GMM	GMM	GMM	LIML
Over-id†	.983	.499	.463	.518	.225	.269	.315	.323
Weak Inst.‡	33.074	103.8	19.212	14.934	3.628	2.989	2.572	-
Critical‡	19.93	19.93	7.56	7.77	-	-	-	-
R^2	.598	.645	.650	.647	.659	.659	.659	.652
N	12,977	12,933	12,933	12,933	12,459	12,459	12,459	12,459

***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 Table 2 except that horizontal neighborliness is defined within a 25 mile buffer region of the town rather than a 50 mile region.

†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 Linear

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\tau_{i,j}$	-.599***	-.737***	-.781***	-.308	-.344	-.337	-.915**
	(.196)	(.247)	(.273)	(.211)	(.242)	(.292)	(.173)
$t_{-i,j}$	.458***	.458***	.411***	.516**	.547***	.456*	
	(.153)	(.155)	(.177)	(.206)	(.196)	(.248)	
$t_{-ij}\tau_j$	-.063	-.067	-.032	-.110	-.131	-.047	
	(.119)	(.119)	(.139)	(.127)	(.122)	(.156)	
$\tau_{i,-j}$	-.331	-.163	-.218	-.130	-.032	-.160	-.271
	(.224)	(.208)	(.258)	(.226)	(.203)	(.284)	(.171)
$\tau_{i,j}d_i$	.001	.012*	.012	.000	.006	.003	.014**
	(.001)	(.007)	(.008)	(.001)	(.006)	(.007)	(.007)
$\tau_{i,-j}d_i$		-.012*	-.012		-.006	-.004	-.015**
		(.007)	(.008)		(.006)	(.007)	(.008)
$E[\frac{\partial t}{\partial \tau_j}]$	-.648***	-.695***	-.711***	-.398***	-.406***	-.348**	-.796***
	(.128)	(.143)	(.158)	(.136)	(.143)	(.172)	(.136)
$E[\frac{\partial t}{\partial t_{-i}}]$	.395***	.392***	.380***	.407***	.417***	.410***	
	(.076)	(.076)	(.081)	(.121)	(.118)	(.135)	
$E[\frac{\partial t}{\partial \tau_{-j}}]$	-.331	-.259	-.314	-.130	-.079	-.190	-.298*
	(.224)	(.212)	(.26)	(.226)	(.210)	(.288)	(.166)
Buffer	50	50	50	25	25	25	-
Distance	linear	linear	linear	linear	linear	linear	linear
Method	GMM	GMM	LIML	GMM	GMM	LIML	GMM
Over-id†	.729	.664	.680	.264	.329	.333	.578
Weak Inst.‡	4.644	3.946	-	2.996	2.324	-	1.959
Critical‡	-	-	-	-	-	-	-
R^2	.615	.618	.609	.666	.667	.661	.573
N	12,498	12,498	12,498	12,459	12,459	12,459	12,498

***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 2 and 3 except the interaction with the distance function uses a linear distance term, rather than the log of distance.

†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

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\tau_{i,j}$	-.743**	-.727**	-1.202***	-.491**	-.388	-.045	-.802***	-.390	-.152
	(.301)	(.286)	(.433)	(.250)	(.279)	(.139)	(.306)	(.254)	(.298)
t_{-i}	.464***	.446***	.399***	.676***	.972***	.065	.105	.214	.737
	(.167)	(.153)	(.129)	(.187)	(.244)	(.136)	(.202)	(.237)	(.263)
$t_{-i}\tau_j$	-.066	-.058	.037	-.211	-.411**	-.500***	.106	-.063	-.200***
	(.141)	(.123)	(.156)	(.151)	(.211)	(.106)	(.142)	(.054)	(.063)
$\tau_{i,-j}$	-.124	-.181	.338	-.193	-.063	.109	-.183	.032	.388
	(.197)	(.183)	(.442)	(.264)	(.282)	(.139)	(.195)	(.194)	(.260)
$\tau_{i,j}d_i$	.090	.112	.139	-.080	-.063	.030	.088	.032	.080
	(.079)	(.087)	(.174)	(.121)	(.103)	(.036)	(.087)	(.076)	(.086)
$\tau_{i,-j}d_i$	-.094	-.122	-.136	.075	.058	-.047	-.108	-.049	-.089
	(.085)	(.091)	(.171)	(.119)	(.281)	(.040)	(.083)	(.078)	(.088)
$E[\frac{\partial t}{\partial \tau_j}]$	-.640***	-.574***	-.937***	-.766***	-.635***	-.400***	-.560***	-.412***	-.263*
	(.120)	(.107)	(.275)	(.181)	(.159)	(.076)	(.124)	(.123)	(.149)
$E[\frac{\partial t}{\partial t_{-i}}]$	.399***	.389***	.435***	.461***	.543***	-.428***	.209*	.151	.540***
	(.074)	(.085)	(.093)	(.094)	(.128)	(.152)	(.126)	(.184)	(.203)
$E[\frac{\partial t}{\partial \tau_{-j}}]$	-.289	-.396*	.110	-.089	-.114	.027	-.371	-.054	.232
	(.227)	(.225)	(.264)	(.234)	(.269)	(.103)	(.234)	(.151)	(.201)
Buffer	50	50	50	50	50	50	50	50	50
Distance	$\log(D)$	$\log(D)$	$\log(D)$	$\log(D)$	$\log(D)$	$\log(D)$	$\log(D)$	$\log(D)$	$\log(D)$
Restriction	No	Border	Ocean	$D < 10$	$D < 5$	Weights	$1/D$	Population	Pop/D
Method	GMM	GMM	GMM	GMM	GMM	GMM	GMM	GMM	GMM
Over-id[†]	.560	.418	.313	.575	.646	.714	.0941	.059	.144
Weak Inst.[‡]	3.985	4.219	1.018	2.554	1.600	2.974	2.145	4.101	3.539
Critical[‡]	-	-	-	-	-	-	-	-	-
R^2	.621	.744	.615	.606	.616	.779	.609	.643	.610
N	12,498	11,552	10,725	9050	4811	12,498	12,498	12,498	12,498

***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. (1) presents the baseline results from Table 2. (2) drops towns where the nearest county border is also a state border. (3) drops towns where the nearest state border would be an ocean or a major lake. (4) and (5) restrict the estimating sample to towns within ten and five miles of a county border, respectively. (6) weights each observation in the regression equation by the population of the jurisdiction. (7) weights each town in t_{-i} by the inverse distance to the neighbor. (8) weights each town in t_{-i} by the population of the neighbor. (9) weights each town in t_{-i} by the population of the neighbor neighbor and then by the inverse of distance to the neighbor.

[†]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.