Inter-Federation Competition:
Sales Tax Externalities with Multiple Federations

David R. Agrawal*
University of Georgia

This Version: January 2013

Abstract
Existing models of tax competition focus on intra-federation competition. This paper analyzes how introducing inter-federation competition affects the strategic nature of sub-federal tax competition. In the context of a Nash equilibrium, the paper shows that lower levels of government will react heterogeneously to the tax rate of the higher level of government depending on the local government’s proximity to the federation borders. Inter-federation competition will also introduce “diagonal externalities,” which are fiscal externalities between neighboring jurisdictions that are of a different level of government. A diagonal externality will have a similar effect on local tax rates as competition between neighboring jurisdictions of the same level. The paper uses two unique data sets to test for local fiscal competition in the presence of inter-federation competition: a cross-section of all local tax rates in the United States and spatial proximity data. The empirical specifications allow for vertical and horizontal externalities to have interaction effects and allow for strategic reactions that vary based on proximity to the nearest neighboring federation. Accounting for federal competition increases in absolute value the vertical strategic reaction by approximately 50%. A one percentage point increase in the county tax rate implies that municipal tax rates in that county will be approximately 0.40 percentage points lower. The results also indicate the sign and magnitude of the horizontal and vertical interactions are heterogeneous across localities. Diagonal externalities are found to have the same sign as horizontal externalities.

JEL: H25, H71, H73, H77, R50

Keywords: Commodity Taxation, Tax Competition, Fiscal Federalism

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

As a result of fiscal federalism, two types of fiscal externalities exist. 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 Connecticut raises its commodity tax rate. An increase in the tax rate of this jurisdiction will result in cross-border shopping toward the neighboring state of Rhode Island. As a result, tax revenues will rise in the neighboring states. However, when Connecticut raised its tax rate, it did not account for this social benefit to its neighbors – overestimating 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 obtain 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, the 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). In such a context, how does introducing inter-federation competition into a model of sales taxation change the strategic interaction between governments of different levels? Inter-federation competition is traditionally used to mean competition among nations. However, I use the term inter-federation competition to highlight competition among the higher levels of government within the federalist structure. As such, counties are federal to towns just as the national government is federal to the state governments. Throughout this paper, inter-federation competition will refer to competition among multiple counties that are composed of multiple towns, but the theoretical models in the paper could similarly apply to the national-state relationship as well. The model presented in this paper is the first analysis of sales tax competition allowing for both inter-federation competition (counties competing with other counties), intra-federation competition (competition of towns within the federation), along with com-
petition across federal borders (competition with towns in neighboring federations and with neighboring counties). The inclusion of inter-federation competition results in a model of spatial tax competition that is highly applicable to the local context – where tax competition over sales taxes is arguably most intense.

Inter-federation competition is essential to fully understand horizontal and vertical externalities, especially at the local level. Modeling and estimating vertical externalities with only one “federal” government ignores horizontal competition among neighboring federal governments – which will put downward pressure on tax rates. Inter-federation competition inevitably constrains the federal governments by inducing horizontal externalities at the federal level, while also triggering additional competition at the sub-federal level across federation borders. These differences have strong implications for estimating the strategic reaction functions in a federation. An assumption of a unitary federation may be valid when the federal government is the national government – as any leakage out of the United States boundaries is relatively small. However, when the “federal” governments are state or county governments, such an assumption is no longer valid. The more decentralized the “higher” level of government, the more likely it will have significant horizontal competitors of its own 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.

Allowing for multiple competing federations will result in “spatial economics” taking on two unique dimensions – “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. For example, tax rates may be a function of proximity to a border or to an amenity. In the presence of multiple federations, the strength of spatial interdependence can be heterogeneous with respect to a jurisdiction’s spatial location. 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.
Spatial location effects arise in a model with multiple federations because governments have multiple borders. In traditional single federation models, there are usually only two-sub federal governments. With only two sub-federal governments, both governments have only one border (one competitor). In such a model, all members of the federation are peripheral (they are located at the federation’s borders). In reality, not all members of a federation are located at the border. This paper shows that vertical externalities make it so that the strategic reaction of a locality to its county tax rate is more likely to be negative for jurisdictions located at the county’s periphery. In addition to this heterogeneity, the presence of fluid federal borders will give rise to a new type of externality not yet discovered in the literature: “diagonal externalities.” A diagonal externality is the effect of a county’s tax rate on a municipality in the neighboring county. For a peripheral town, it is identical in nature to the strategic response of horizontal competition. One example of a diagonal externality is that if the state of California increases its tax rate, it increases the tax base in Clark County in Nevada. However, the channel by which this externality occurs is very different than the horizontal externality imposed on Clark County in Nevada if San Bernadino County in California had increased its tax rate.

Because of these theoretical considerations, the estimation strategy for determining the strategic response to vertical and horizontal externalities becomes more complex in the presence of inter-federation competition. 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. These modifications to the methods complicate the analysis, but they also provide the researcher with the ability to identify heterogeneous strategic interactions. If these heterogeneities match the theoretical predictions, the researcher can more convincingly argue that the identified effects are indeed the result of tax competition rather than the result of common shocks.

The baseline empirical results show that a one percentage point increase in county sales tax rates reduces municipal sales tax rates by .40 percentage points. For the horizontal
externality, a one percentage point increase in the average of a town’s neighbors’ tax rates will increase a municipality’s local sales tax by .46 percentage points. Accounting for horizontal and vertical interaction effects, diagonal externalities, and distance based effects resulting from inter-federation competition increases (in absolute value) the vertical strategic reaction by approximately 50% relative to the baseline specification estimated in the literature. To summarize (1) Omitting interaction effects of horizontal and vertical externalities will induce a bias into the strategic reactions and 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 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 often decreases as jurisdiction size increases.

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 purpose is usually to deal with inefficiencies arising from horizontal competition – without any purpose of raising 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). 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 an 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) generalize 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-More 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-More and Solé-Ollé (2001). For example, Hayashi and Boadway (2001) claim to avoid the endogeneity problem 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

---

1In 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. A working paper by Becker and Büttner (2012) study capital tax competition in the presence of competing federations. The literature on sales tax competition in federations is much more sparse. However, there is an extensive literature on sales tax competition with one level of government: Mintz and Tulkens (1986), Kanbur and Keen (1993), Trandel (1994), Nielsen (2001). For a summary of the literature see Keen and Konrad (2013).
for the differenced county tax rates with the lagged level of the tax rates eliminates the endogeneity problem. 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. Many of these studies focus on competition within a federation. However, Geys and Osterloh (2012) show that government officials near borders do not view tax competition as being constrained to neighbors within a federation.

Brühlhart and Jametti (2006) tries to estimate the vertical and horizontal externalities without identifying the effects off of the slopes of reaction functions. Brühlhart 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. 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 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 literatures would benefit from more familiarity with each other.

3 Model

The model expands Devereux, Lockwood and Redoano (2007). The geographic setup of the model features sub-federal jurisdictions (towns) located on a (possibly) infinite length horizontal line segment. Federations (counties) are indexed $j = 1, 2, \ldots, M$ and each county is composed of $m + j$ towns each such that each sequential county has one more town than the previous; county 1 will have $m + 1$ towns, but county $M$ will have $m + M$ towns. The
ordering of counties in this manner is not important, but it will allow me to characterize the
model’s solution using a simpler notation than if counties were organized arbitrarily along
a line segment. Towns are indexed \( i = 1, 2, \ldots, M(M + 1 + M) \) and the index does not reset
across counties. The \( M \) federations in the model compete with each other and with the
towns in the model. The towns compete with other towns and with the federations. All
of the federations are within a common union (state). Each town has \( n_i \) residents and is
\( l_i \) units long on the line segment. The relationship between length and population is that
\( n_i = \phi_i l_i \) so that \( \phi_i \) denotes the population density. Consumers and producers are located
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.\(^2\) Towns (sub-federal governments)
set a local tax rate \( t_i \), counties (federal governments) set a county tax rate \( \tau_j \) that applies to
all towns within the county, and the state government (the union government) sets a state
tax rate \( T \).\(^3\) The state tax rate applies to all locations along the infinite line segment. In
any specific town, the after tax price \( q_{ij} \equiv q_i \) is equal to \( 1 + t_i + \tau_j + T \).

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

\(^2\)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.

\(^3\)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.
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.\footnote{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.}

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 $x(q) = \arg \max [u(x) - qx]$ 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 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

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

from the border of town $k$.

Governments compete in a Nash game. Counties and towns are considered simultaneous movers in the game.\footnote{While a leader-follower assumption is realistic for higher levels of government (national), it seems plausible that towns and counties are simultaneous movers.} 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

$$R_i = t_i B_i$$

where $B_i$ is the tax base defined below.\footnote{Maximizing revenue is a simplifying assumption that allows for explicit analysis 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.} 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 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 neighboring jurisdiction’s tax rate includes the local rate and the county tax rate, which may or may not be the same as jurisdiction $i$.

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 \( j = 1, 2, \ldots, M \) 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: \( \tau_1 \leq \tau_2 \leq \ldots \tau_M \). This assumption places no restriction on the pattern of local taxes
within a county. However, Agrawal (2011) shows that revenue maximizing towns of identical
size will set higher rates the closer to a high-tax county neighbor and lower rates closer to a
low-tax neighbor. I assume this pattern holds when evaluating the reaction functions below
because the models differ only in whether demand is perfectly inelastic or not.\(^7\) 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. The asymmetric Nash equilibrium is characterized by

\[
\begin{align*}
\tau_1 & \leq \tau_2 \leq \ldots \tau_M \\
t_i & \leq t_{i+1} \leq \ldots t_{m+j} \quad \text{for town } i \text{ in county } j.
\end{align*}
\]

(3)

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

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

(4)

where the term in [ ] of the town tax base is defined as \( s_i \) and \( \rho_i \) is interpreted as the
intensity of the horizontal competition. It is clear that the only consequence of conjecturing
that the equilibrium is given by equation 3 is that the direction of cross-border shopping tells
which \( \rho_i \) enters on each indirect utility differential – specifically that the differential is always

\(^7\)Relaxing this specific characterization of the equilibrium would not change the functional form of the
reaction functions or the revenue function, but it would distort the pattern of the subscripts on these
equations. The presence of a tax gradient allows for a pattern in the perceived elasticities.
multiplied by the $\rho_i$ corresponding to the relatively high-tax jurisdiction. I reiterate that guessing the equilibrium is characterized by equation 3 acts only to simplify the subscript notation and reduces the problem from several cases to a single case.

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 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, the Nash equilibrium above is guaranteed to exist.\textsuperscript{8} 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 $x_i$. The local tax rates are implicitly defined by the reaction function:

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

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

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

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

\textsuperscript{8}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.
where
\[ \varepsilon_i = -\frac{q_i x'_i}{x_i} \]  \hspace{1cm} (7)
is the elasticity of demand for the consumption good and
\[ \theta_i = \frac{q_i x_i (\rho_i + \rho_{i+1})}{s_i} \]  \hspace{1cm} (8)
is the elasticity of the number of shoppers in town \( i \) and it differs from Devereux, Lockwood and Redoano (2007) by accounting for both the in- and out-flows. Both \( \varepsilon_i \) and \( \theta_i \) are defined to be positive numbers under the assumption that demand curves slope downward. Similarly for counties, the reaction function can also be defined.

### 3.2 Strategic Interaction

The equations above implicitly determines 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 \( \varepsilon_i, \theta_i, \) 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 work to 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. In a unitary federation with a symmetric equilibrium, 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. \( \theta_i \) represents the strength of horizontal tax competition and can increase if transportation costs fall or the population density increases near the border.

In the subsequent sections, I consider how inter-federation competition affects the strategic response of local governments. In order to make some progress on the question, I need to make some simplifying assumptions.

**Assumption 1:** Demand is iso-elastic. This assumption simplifies the intuition by holding \( x''_i \) and \( x'_i \) constant along the demand curve such that the demand elasticity \( \varepsilon_i = \varepsilon \) and curvature \( \eta_i = \eta \) are constant, which is likely a good assumption in a local region of a Nash
equilibrium. I keep the subscripts on $\varepsilon_i$ and $\eta_i$ – even though the assumption of iso-elastic demand means they are constant even in an asymmetric equilibrium – so that the reader can evaluate how this assumption influences the results.

**Assumption 2:** The ratio $\rho_i$ is homogenous across all towns. This assumption reduces the model so that the only forms of heterogeneity are that some counties are bigger than other counties and towns are heterogeneous in their spatial proximity to a county border. The assumption will make the intuition of the propositions clearer. Again, I maintain the subscripts in the interest of transparency.

All of the equations to follow characterize the asymmetric equilibrium where $\theta_i$ is allowed to vary across towns by virtue of differences in county tax rates. Notice that all of the differences in the pattern of the equilibrium from equation 3 are driven by differences in $\theta_i$ in equation 6.

### 3.2.1 Reaction 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. I am interested in showing whether the strategic reaction by a town to its own county’s tax rate ($\frac{\partial t_i}{\partial \tau_j}$ for town $i$ in county $j$) varies by whether a town is internal or peripheral to the county borders.

In order to answer this question, define the following cross-price elasticity, which is scaled by the price ratios:

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

(9)

Recall $\theta_i = \frac{q_i x_i (\rho_i + \rho_{i+1})}{s_i}$ is the response of the number of shoppers of town $i$ 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. Note that if in a symmetric equilibrium, $\theta_i$ is approximately two times $\theta_{i,k}$ because the town’s own price influences two borders.

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

I can then characterize the slope of the reaction functions for peripheral and internal towns as follows. The slope of the reaction function for towns at the periphery is calculated by differentiating equation 6 after appropriately accounting for how the tax base changes as a result of the town being at the periphery.\(^9\) The slope of the reaction function in the

\(^9\)Once the researcher accounts for the number of borders, the derivation of this equation follows similar
asymmetric equilibrium is given by:

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

(10)

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 given by:

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

(11)

To build intuition, assume that that $\theta_i$ is constant within a federation, such that the equilibrium is symmetric. I make this assumption only to derive Proposition 1. This assumption could arise if larger counties have lower preferences for public goods (Haufler 1996) that exactly offset the larger size effect. In such a world where counties set identical tax rates, no differences in town tax rates will emerge. The reason for this simplifying assumption initially is to benchmark the results in this paper (with multiple federations) relative to otherwise identical models within a unitary federation. In the subsequent sections, I relax the symmetric assumption and demonstrate how the asymmetric equilibrium defined above changes the underlying results.

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

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

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

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 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. 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 steps as in the appendix of Devereux, Lockwood and Redoano (2007), so the proof is omitted.
rate only changes the neighboring rate on one border. And \( \theta_{i,k} \) is more inelastic than \( \theta_i \) for small changes in price. Intuitively this is because \( \theta_i \) 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. 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 \( \rho \) and increases as density increases or if costs of shopping decline. This can be seen very easily by evaluating Equation 10 at a symmetric equilibrium, which yields \( [\varepsilon_i(\theta_i - \varepsilon_i - \eta_i) - \theta_i^2/2]/D_i \). Equation 11 reduces to \( [\varepsilon_i(\theta_i - \varepsilon_i - \eta_i)]/D_i \).

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. (1) It will change the demand function for individuals because \( x \) is a function of \( q \). (2) Changes in the county tax rate will also distort the number of individuals living in \( i \) and purchasing goods in \( i - 1 \). (3) 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, the change in the base is mitigated 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 at one of the town’s borders. 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 for a periphery town 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 capture additional cross-border shoppers from the neighboring county if they are in a low-

---

10Mathematically, 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 \( q \).
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. Now I proceed by returning to the asymmetric equilibrium.

3.2.2 The Diagonal Externality

Because there are multiple federations, towns at the eastern border of county \( j \) directly compete with the border town in county \( j + 1 \) as well as county \( j + 1 \). Thus, 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. Thus, a diagonal externality on jurisdiction \( i \) is the externality imposed by a neighboring jurisdiction that does not cohabit the same tax base as jurisdiction \( i \), but that is not at the same level of governance. One example of a diagonal externality is the relationship between Clark County, Nevada and the state of California. Clark County borders California but is not in California. However, if the state of California raises its tax rate, Clark County will see an increase its tax base. This effect would be the same if San Bernadino County, California (which borders Clark County, Nevada) had raised its tax rate except for the fact that the channel through which the externality occurs is at a different level of government and California’s externality on Clark County should not be mistaken for a horizontal relationship. Formally:

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

Notice how the diagonal externality is unlike a vertical externality. In contrast, a vertical externality arises when different levels of government share the same tax base and an increase in taxes levied by one level reduces the tax base of the other level. It is also different from a horizontal fiscal externality. A horizontal externality arises when the governments are of the same level and therefore do not share the same tax base (rather they compete for it), which implies that an increase in taxes levied by one government increases the the tax base of the other government. Thus, the consequences of the diagonal externality is identical to a horizontal externality, but the effect results from different levels of government in the fiscal hierarchy. The researcher wants to be sure to identify the effect of competition with
neighboring towns separately from the effect of competition with neighboring counties.

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.

**Proposition 2.** In the neighborhood of an asymmetric 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 observationally equivalent to a horizontal externality in the data. 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 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 literatures have ignored this externality. Failing to account for this externality in any regression will yield biased estimates of the horizontal externality if the researcher specifically desires to estimate the same-level-of-government horizontal effect.

To prove proposition 2, derive the reaction function for town $i$ in county $j$ with respect to county $j + 1$’s tax rate; the exercise can be repeated for a town at the western edge of the county with respect to county $j + 1$’s tax rate. When town $i$ is located at the eastern periphery of its own county, this strategic reaction can be found by differentiating equation 6 with respect to $\tau_{j+1}$. The slope of the reaction function with respect to the neighboring county’s tax rate is

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

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. Note in a symmetric equilibrium this becomes $\frac{\theta_i^2}{2} / D_i$.

---

11If 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.
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 $\theta$ 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 $\theta$ is constant within a federation used to derive Proposition 1. I consider how differences in county tax rates resulting from the asymmetric equilibrium defined by equation 3 influences the strategic interaction of the local governments. Recall that if the elasticity of demand is constant and local tax rates are increasing from low-tax counties toward high-tax counties, equation 6 implies that all of the local tax differences are due to differences in the elasticity of cross-border shopping. Specifically, for jurisdictions closest to a low tax county, $\theta_i$ is relatively high. For jurisdictions closest to the high-tax county, $\theta_i$ is relatively low. Thus, following Agrawal (2011), the elasticity $\theta_i$ should be monotonically decreasing from the left-most town to the right-most town along the line segment. The natural question is how the strategic reaction changes when the elasticity of cross-border shopping increases. Define $\theta_{i,i+1} + \theta_{i,i-1} \equiv \gamma_k$ for an internal town and $\theta_{i,k} \equiv \gamma_k$ for a peripheral town. Differentiating equation 11, assuming that the cross-price elasticities $\theta_{i,k}$ are not a function of the own-price elasticity $\theta_i$ or if they are that these effects are sufficiently small to be ignored,\textsuperscript{12} provides some important evidence. For an increase in $\theta_i$, the change in a town’s strategic reaction to its own county tax rate will be proportional to:

$$\frac{\varepsilon_i^2 + \varepsilon_i^2 (\gamma_k - 2\theta_i) - (3\varepsilon_i + 2\gamma_k)\theta_i^2}{D_i^2}$$  \hspace{1cm} (13)

Equation 13 is derived in the Appendix and the necessary conditions for this equation to be positive are also derived.

\textsuperscript{12}Relaxing this assumption can easily be incorporated in to the expression below by differentiating $\gamma_k$ with respect to $\theta_i$.\hspace{1cm}19
Proposition 3. In the neighborhood of an asymmetric Nash equilibrium, heterogeneity in the strategic reaction function between internal and peripheral towns critically depends on whether the reaction function is increasing or decreasing in the elasticity of cross border shopping in addition to whether the jurisdiction is internal or peripheral to a low-tax or high-tax county border.

Notice that if expression 13 is positive, movement from the peripheral town at the high-tax (eastern most) county border to an interior town unambiguously makes the strategic reaction more likely to be positive. Moving from the peripheral town to the interior town implies that \( \theta_i \) rises. An increase in \( \theta_i \) then has a positive effect via equation 13. Furthermore, the strategic reaction given by equation 11 now contains an additional cross price elasticity \( \theta_{i,k} > 0 \) relative to equation 10, which reinforces that positive effect. If expression 13 is negative, then the negative influence of the higher \( \theta_i \) must be counter-balanced with the additional positive pressures of the cross-price elasticity. On the other hand, for a peripheral town at the low-tax county border (the western-most town in the county), movement to the interior of the state implies that \( \theta_i \) is falling. If expression 13 is positive, the decline in this elasticity has a negative effect on the strategic interaction. But at the same time, the interior town worries about two borders and realizes and additional \( \theta_{i,k} > 0 \). Thus, the effect is possibly ambiguous for towns near low-tax borders. If expression 13 is negative, then movement to the interior of the state from this direction unambiguously increases the strategic reaction. However, if changes in \( \theta_i \) across neighboring jurisdictions are small, then the effect noted in proposition 1 is likely to dominate the fact that \( \theta_i \) decreases from low-tax counties toward high-tax counties.

Of course, one may also wonder how the strategic reaction changes for two internal towns in an asymmetric equilibrium. In this case, the researcher needs to determine the sign of equation 13. However, tax differentials at town borders are also informative of the slope if elasticities cannot be observed directly.

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

I give the intuition graphically and show the math demonstrating this in the appendix. 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. The implication of this is that the elasticity of the tax base with respect to the town’s own price $\theta_i$ is small relative to the cumulative cross-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 2. In the figure, the prices start at $q_i$, $q_{i-1}$ and $q_{i+1}$. Assume that town $i$ is internal to the county. Then, if the county rate increases, all prices rise by a constant amount to the bolder lines on the graph. The amount of cross-border shopping is proportional to the difference between the indirect utilities. In the top graph, the neighbor has a higher price so cross-border shopping is inward to $i$. In the second graph, the neighbor has a lower price, so cross-border shopping is outward from $i$. Note that $v(q_i) - v(q_{i+1})$ becomes smaller after the tax increase and $v(q_{i-1}) - v(q_i)$ also becomes smaller – but it falls by a much larger amount because of the convexity of the indirect utility function. The closer $q_i$ is to its high-tax neighbor, the smaller the change in $v(q_{i-1}) - v(q_i)$ and vice-versa. This mitigates the change in outward cross-border shopping and amplifies the change in inward cross-border shopping.

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

### 3.2.4 Interaction With State Tax Rates

Changes in the state tax rate will also prompt a response in the town jurisdictions – especially if the town is peripheral to the state. 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. Intuitively, the state interactions with towns should be similar to the county interactions with town if people view changes in county rates

---

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.
similar to changes in state rates. Of course, state rates may be more salient but 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:

\[ t_{ij} = \alpha_0 + \alpha_1 t_{-i} + \alpha_2 \tau_j + \delta X_{ij} + \varepsilon_{ij} \] (14)

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

\[ t_{-i} = \sum_{k \neq i} w_{ik} t_k. \] (15)

Denote \( w_{ik} \) as exogenous weights normalized such that \( \Sigma_{k \neq i} w_{ik} = 1 \). Tax rates on the right-hand side are usually instrumented for with the weighted average of the neighbors’ \( X \). The coefficient on \( t_{-i} \) identifies the strategic reaction in response to a jurisdiction’s neighboring tax rates. The coefficient on \( \tau_j \) 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 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, 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 \( \alpha_2 \) plus an interaction with horizontal externalities.

\(^{14}\)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.
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 \( \theta_i \) captures (in part) the strength of the horizontal externality. Recall that \( \theta_i \) 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 \( \theta_i \), 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 of tax competition. 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 (14) – on neighboring tax rates

\[
t_{ij} = \alpha_{0j} + \alpha_{1j}t_{-i} + \varepsilon_{ij},
\]

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

\[
\alpha_{0j} = \gamma_{00} + \gamma_{01}\tau_j + u_{0j}.
\]

It is clear that substituting (17) into (16) will yield (14) without 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 \( \tau_j \) and vice versa. Assuming that this interaction

\footnote{For a summary of multi-level modeling, see Franzese (2005). It is useful to think of this empirical problem in the context of a multi-level model. However, the empirical strategy that will follow this section will use a reduced form setup to the problem, 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.}
also occurs with error, this implies

$$\alpha_{1j} = \gamma_{10} + \gamma_{11} \tau_j + u_{1j},$$  \hfill (18)

and substituting (17) and (18) into (16) yields

$$t_{ij} = \gamma_{00} + \gamma_{01} \tau_j + (\gamma_{10} + \gamma_{11} \tau_j) t_{-i} + u_{1j} t_{-i} + u_{0j} + \varepsilon_{ij},$$  \hfill (19)

which is not the same as equation (14). 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 \( \gamma_{01} \) 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 \( \gamma_{01} \) where the bias is given by \( \hat{\gamma}_{11} \frac{\text{cov}(t_{-i}, \tau_j)}{\text{var}(t_{-i})} \) 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.\(^{17}\) 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 time from the population weighted centroid of each Census Place\(^{18}\) 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 each jurisdiction’s perimeter and area.

Driving times are measured from the population weighted centroid of each Census Place to the nearest intersection of a major road and a state or county border crossing.\(^{19}\) Population

---

17 The data is proprietary but was provided to me for free. For a complete description of the data see http://www.prosalestax.com/
18 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.
19 A major road is a Census classification including most non-residential roads. I omit residential roads for
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 time calculated is the time 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 are small and may have zero contiguous neighbors. 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. To measure the diagonal externality, I define the neighboring county as the county that is closest to the town based on the driving time criteria.

Control variables are from the 2010 United States Census plus geographic and political controls that I have generated myself. The set of possible control variables include a municipality’s area, perimeter, number of contiguous neighbors, population, the fraction of seniors, individuals with less than college education, income, the fraction male, the percent on public assistance, age, the ratio of private to public school students, the number of rooms in the house, the age of houses in the town, the fraction of non-citizens, 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 border is an essential determinant of local tax rates. Therefore, as controls, I include the tax differential at the nearest state border, a dummy for whether the town is in a high-or same-tax state relative to the nearest neighbor, the log of driving time to the nearest state border and the complete set of interactions of these variables. 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

For empirical estimation in the cross-section to be feasible, I assume that tax rates are in equilibrium, or put differently, that the national data set of sales taxes in 2010 did not experienced large shocks. Given the large number of observations, such an assumption is 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.

I define a buffer as a region that is twenty miles from the center of a town. Then, any place that intersects this buffer region is defined as a neighbor. I also know the exact distance between each jurisdiction in case weighting by distance within this region is desirable.

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.
realistic and implies that the strategic reactions represent long-run reactions rather than a short-run response. I will focus on how municipalities react to the county level rates, neighboring county rates, and neighboring town rates. I estimate

$$t_{ij} = \alpha_0 + \alpha_1 \tau_{i,j} + \alpha_2 t_{-i} + \alpha_3 t_{-i} \tau_{i,j} + \alpha_4 \tau_{i,-j} + \alpha_5 \tau_{i,j} d_i + \alpha_6 \tau_{i,-j} d_i + X_{ij} \beta + S + \varepsilon_{ij}$$

(20)

where $X_{ij}$ are the local controls listed above and $S$ are state fixed effects – that control for the level of the state tax rate in a state along with other within state policies. The tax rate $t_{ij}$ is municipal plus sub-municipal local option taxes and $t_{-i}$ is defined in equation 15. Specifically, if town $k$ is within fifty miles of town $i$, the weights in the main specification are equal to one divided by the number of jurisdictions within fifty miles of town $i$ and zero otherwise. Thus, the interpretation of $t_{-i}$ is the average tax rate of town $i$’s neighbors. Defining $N_i$ as the set of towns within a fifty mile region of town $i$, then

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

(21)

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

Define $\tau_{i,j}$ as the county tax rate that town $i$ is located in$^{22}$ and $\tau_{i,-j}$ is the nearest neighboring county’s tax rate to town $i$. Because $\tau_{i,-j}$ is not a weighted average of all the neighboring counties, I implicitly assume that 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 multidimensional problem into a single dimension. Finally, in a similar spirit, $d_i$ is a measure of distance from the town centroid to the nearest county border and is linear in the driving time to the county border, $d_i$. The assumption that the reactions decline in a linear manner from the border is realistic given that the average town only 12 minutes away from the closest border. I show the results are robust to more non-parametric assumptions. The inclusion of this distance function is driven by Proposition 1.

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

$^{22}$Towns that cross multiple county borders are defined as being in the county that has the majority of the town’s population.
estimating population:

\[
E[\frac{\partial t}{\partial \tau}] = \frac{1}{M} \sum_{i=1}^{M} (\alpha_1 + \alpha_3 t_{-i} + \alpha_5 d_i) \\
E[\frac{\partial t}{\partial \tau_j}] = \frac{1}{M} \sum_{i=1}^{M} (\alpha_2 + \alpha_3 \tau_{i,j}) \\
E[\frac{\partial t}{\partial \tau_{-j}}] = \frac{1}{M} \sum_{i=1}^{M} \alpha_4 + \alpha_6 d_i,
\]

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 (22) are a consistent estimate of the mean derivative in the population.

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 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). Thus, the results will be consistent even if the error the following first-order spatial auto-regressive form:

\[
\varepsilon_{ij} = \tau \sum_{k \neq i} w_{ik} \varepsilon_k + \nu_{ij}
\]

where \(\nu_{ij}\) is i.i.d. over space and \(\tau\) is the spatial auto-regressive coefficient.\(^{23}\)

For neighboring tax rates, the standard instruments in the literature for \(t_{-i}\) are the weighted average of several control variables; the method of Kelejian and Prucha (1998) requires using a subset of the neighbors’ control variables. Simply put, the standard instrument for neighboring tax rates is \(\sum_{k \in N_i} w_{ik} x_k\), where \(x_k\) is a variable in \(X\). Although this instrument – using the weighted average of several \(x\)’s – is used mostly for state level tax

\(^{23}\)The standard errors reported below are robust to heteroskedasticity but they are not clustered at the county level because in the spatial context, the i.i.d. assumption can be violated because towns are in the same county and because they are near each other but in different counties. For this reason, clustering places strong assumptions on how the i.i.d. assumption is violated. Because I will use a two-stage GMM estimator, making a strong assumption regarding clustering could influence the coefficient estimates in a manner not consistent with the true spatial process on the error terms. Notice in the equation above, the i.i.d. assumption is violated because the weighted averages of the neighboring jurisdictions enter into the error term – but that these neighboring jurisdictions need not be in the same county (they can be across federations). Because the IV estimates are consistent using the spatial IV approach discussed above, I elect not to cluster the standard errors at the county level.
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 when towns are small and populations may cluster near town borders. Thus, the population of small jurisdictions may directly influence the neighboring town.

Instead of using the entire subset of the $X$’s as instruments, I will only use geographic variables as instruments. Specifically, I will use area and perimeter of the town as instruments. For the county tax rate, I will use the county area and perimeter as instruments. For the neighboring county, I will use its county perimeter and area as instruments. To instrument for $t_{-i} = \sum_{k \in N_i} w_{ik} t_k$, I use area$_{-i} = \sum_{k \in N_i} w_{ik} \text{area}_k$ and perimeter$_{-i} = \sum_{k \in N_i} w_{ik} \text{perimeter}_k$ as instruments. Of course, the regression specifications above also include interaction terms, in which case they are instrumented for with the interactions of the 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. The theoretical tax competition literature implies that area and perimeter are important determinants of a jurisdiction’s own tax rates. Further, county borders were likely to be historically drawn on latitudes and longitudes or broader geographic features. The area and perimeter of a county depend on a county’s characteristics such as whether along a body of water, broader geographic features, and how counties were divided historically. Because area and perimeter are historically drawn, the evolution of time with these variables strengthens the case for their exogeneity. Similarly, the town’s area and perimeter often depend on how municipalities were historically formed within the county and the characteristics within the county when the town borders were historically drawn – which in most cases were not at the same time county lines were delineated. Because jurisdiction boundaries were often randomly drawn, I argue that they are exogenous instruments. More simply put, the selection of these instruments can be viewed as an application of Kelejian and Prucha (1998) where area and perimeter are the subset of the $X$’s that I use as instruments.
5.3 Hypotheses to Test

Before presenting the results, recall that the theoretical model provides the following testable hypotheses regarding the sign on the coefficients from equation 20. 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, \( \tau_j \), is ambiguous as the relative magnitudes of \( \theta \) and \( \varepsilon \) 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 \( d_{i} \tau_j \) 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 \( d_{i} \tau_{i,-j} \) 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.

6 Empirical Results

6.1 Main Results

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

Table 3 presents the baseline results using spatial 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 large negative result that is consistent with the regression discontinuity results in Agrawal (2011). Several explanations exist for the opposite finding. Towns may react in a different manner to county rates than states will react to the federal government as the institutional structure of lower level governments is different. Alternatively, municipalities face a much more mobile tax base relative to states. The increased mobility of the tax base implies that the elasticity of cross-border shopping, $\theta$, is more likely to be larger for local governments than for state governments – and the theoretical model implies the larger $\theta$ is relative to $\varepsilon$, the more likely that county and local rates will be strategic substitutes. Column 2 estimates the horizontal reaction function and finds a significant positive relationship with neighboring jurisdictions’ tax rates; this result is standard in the literature – but as the results below will indicate, need not be true for all towns.

Column 3 presents a similar specification to 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 .251 percentage points. Contrarily, a 1 percentage point increase in the average of the neighbors tax rate, increases a municipality’s local tax rate by .536 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. Comparing this to the OLS coefficients shows that the bias from the OLS estimations exists but is not large and does not change the sign of the estimates.

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

Specification 4 adds the interaction of the vertical and horizontal externalities and specification 5 accounts for the diagonal externality. Looking at the mean derivatives, the slopes of the horizontal reaction function fall by almost .05 percentage points. Although the diagonal externality is of the expected sign, the result is not the same magnitude as the horizontal reaction as suggested by the theory. Specification 6 and 7 (the preferred specification) add in the interaction of the vertical and diagonal externalities with a linear distance function. Notice that the slope of the reaction function with respect to the county tax rate falls to -.394 and the slope of the horizontal reaction function falls to .458. The estimate in column 3 is 15% smaller for the horizontal reaction function in the complete specification of column 7. The vertical reaction function is 55% larger in absolute value than the result estimated using the baseline specification in the literature. This suggests that failing to account for inter-federation competition attenuates the vertical reaction closer to zero. The results from columns 6 and 7 suggest that the spatial location effects are extremely important to calculating the mean derivatives. Columns 8 and 9 include only a sub-set of the endogenous regressors in column 7 in order to demonstrate what the reaction functions would look like if only the vertical and diagonal elements were included. Notice that in column 9, which adds only spatial location effects on the vertical reaction function, increases the intensity of the strategic reaction in absolute value relative to column 3.

Finally, column 10 provides an important verification that the instruments in column 7 are not weak because the Stock and Yogo (2005) critical values are not tabulated for the case of that many endogenous regressors. With many variables in need of instrumenting, the concern of weak instruments become more worrisome. As an alternative, I estimate the equation by limited information maximum likelihood (LIML). If the estimating equation is correctly specified and the distributional assumptions hold, LIML will produce unbiased

---

24In 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.
coefficient estimates even if weak instruments are present. Because the point estimates in column 10 are similar to column 7 and because the mean derivatives change in the same manner, it suggests that weak instruments in the GMM estimates should not be a concern. Of course, I acknowledge that instrumenting for many variables is not ideal, but I need to weigh this with estimating an incorrectly specified estimating equation. The results comparing the full specification with the baseline specification in the literature suggest not allowing for inter-federation competition can significantly bias strategic reaction estimates at the local level as a result of a misspecified estimating equation.

6.1.1 Heterogeneity in Reaction Functions

Up until now, I have focused on how correctly specifying the reaction function impacts the mean derivatives without discussing the heterogeneous responses that arise. The horizontal reaction function of town $i$ with respect to the average tax rates of its neighbors will depend on the tax rate of the county it is in. Figure 3 depicts this heterogeneity – keeping in mind that the instrumental variable strategy limits me to imposing that the reaction function interacts linearly with the county tax rate. Notice that neighboring tax rates are strategic complements for jurisdictions in counties with less than a 3.75% tax rate. However, neighboring tax rates become strategic substitutes after the county tax rate become sufficiently high. Approximately 7% of towns are in counties with a sufficiently high county taxes to imply that town $i$’s tax rate is a strategic substitute with its neighbor’s tax rates. Such a result of strategic substitutes is not traditionally found in the literature, but no studies allow for the interaction effect that drives this heterogeneity.\footnote{One study that empirically finds downward sloping reaction functions is Parchet (2012). However, Wildasin (1988) and Vrijburg and de Mooij (2012) provide theoretical justification for taxes as strategic substitutes.}

25 I find that for a subset of towns in the sample, neighboring jurisdictions’ rates are likely to be strategic substitutes.

The empirical specification allows the vertical reaction function of town $i$ with respect to county $j$ to be heterogeneous along two dimensions. It will depend on distance to the county border and on the average of the neighboring town tax rates. Figure 4 shows these effects under the limitation that the IV approach requires I impose that heterogeneous responses are linear in each of these effects. The left panel indicates that the strategic reaction of a town to its own county’s tax rate is most negative when the town has high-tax neighbors and when that town is near the county border, which is consistent with the theory. The contour map illustrates that holding fixed your neighbors’ tax rate, towns that are internal to the county have reaction functions that are mildly closer to zero. Furthermore, holding distance to the county border constant, each increase in your average neighbors’ tax rate
dramatically intensifies the negative strategic reaction to the county rate.

6.2 How Robust Are The Results?

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

6.2.1 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 may seem large, but corresponds to the spatial reach of tax competition that Agrawal (2011) found in response to state borders. My preference is also to use a broader measure of neighborliness because it allows for towns to compete with more heterogeneous types of jurisdictions that may have stores not available in the local area. As an alternative, I use the average tax rate of jurisdictions within 25 miles as a measure of neighborliness. Table 4 presents the results. Columns 1 - 7 are identical in all respects to Columns 1-7 of Table 3 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 .175 percentage points. A 1 percentage point increase in the average neighbor tax rate raises municipal tax rates by .401 percentage points. The sign and significance of these results is similar to Table 3, but the magnitude of the derivatives are slightly smaller in absolute value. In the preferred specification (including interaction effects, distance effects, and diagonal externalities), the slope of the vertical reaction function is -.222 and the slope of the horizontal reaction function is .559. With regard to the estimate of the reaction function with respect to the neighboring county tax rate, the slope is insignificant but of the opposite sign compared to the previous specification.

6.2.2 Alternative Distance Functions

In all of the previous specifications, I have assumed that the distance function to county borders is linear. Such a parametrization imposes that as driving times become very large, the effect of distance on the strategic reaction remains constant. Although some counties are very large in size, most counties are small in size. In fact, the average town is located
approximately 12 minutes 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 for most observations in the sample. However, I also want to treat
distance in a non-parametric manner. I do this by defining a dummy variable that is equal
to one for towns that are in the 90th percentile of driving time to the county border. Thus,
the dummy variable equals one for towns that are more than 22 minutes from the nearest
county border (only twice the distance of the average town to a border) and zero otherwise.
In Table 5, I report the results when the distance function is a dummy variable. Columns
1, 2, and 3 correspond to the results in Columns 6, 7 and 10 from Table 3. Columns 4, 5,
and 6 of Table 5 correspond to the results from columns 6-8 in Table 4.

When the distance function is non-parametric, the point estimates of the mean derivatives
change slightly but the interpretation of the results is similar in spirit. In Column 2, \( \tau_{i,j}d_i \)
has a positive coefficient of .126. Notice this estimate is much larger than the coefficient on
the linear distance interaction, which implied that a one minute increase in distance from
the border increased the strategic reaction by at a very small gradient. The result of the
dummy variable approach suggests that for towns that are extremely far from borders, the
town government cares much less about leakages across the county border, which creates
upward pressure on the strategic reaction. Such a result is consistent with the theory –
interior jurisdictions are more likely to have upward sloping reaction functions with respect
to the county rate.

6.2.3 Various Model Specifications and Sample Restrictions

Table 6 reports robustness checks for various sample restrictions. Column 1 and 2 drop
jurisdictions for which the nearest county border is also a state border. This specification
reduces the possibility that the neighboring state is also producing a diagonal externality on
towns proximate to its borders. The measure of the diagonal interaction remains insignificant
but of the opposite sign. Moving from the specification currently used in the literature to the
complete specification makes the vertical reaction more negative and the horizontal reaction
less positive. This is consistent with the first set of results, but the change in the estimates
is much smaller when using the restricted sample, which suggests that some of the initial
results may arise from state borders.

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

Neighbor weights may also be varied. Table 7 specifies different weighting scheme to cal-
culate the weighted average of the neighboring jurisdictions’ rates. Thus the only difference
in this table is that each set of columns alters the exogenous weights that determine $t_{-i}$.
Define $N_i$ as the set of towns within a fifty mile radius of town $i$, $s_{ik}$ as the distance between
town $i$ and town $k$, and $\phi_k$ as the population of town $k$. Recall that $t_{-i} = \sum_{k \neq i} w_{ik} t_k$. Then
columns 1-2 use exogenous weights that are normalized to sum to 1 given by Equation 24:

$$w_{ik} = \begin{cases} 
\frac{1}{s_{ik}} & \text{if } k \in N_i \\
\sum_{k \in N_i} \frac{1}{s_{ik}} & \text{if } k \notin N_i \\
0 & \text{if } k \notin N_i 
\end{cases},$$

which can be interpreted as inverse distance weights. In this specification towns closer to
town $i$ are given more weight than towns far away. Column 3-4 estimate the equation using
the weights from Equation 25:

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

such that neighboring towns within the fifty mile radius of town $i$ are given more weight if
they have a larger population. Lastly, Equation 26 specifies the weights used in Columns
5-6:

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

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 21 are the most common weights in the literature and are the
simplest to interpret (as a simple average). The weights used in equation 21 are most likely
to be interpreted as purely spatial, while the weights given by the three equations above are more indicative of economic flows or some other factor influencing the weights.

When using inverse distance based weights, the strategic interaction with the county tax rate remains approximately the same magnitude as the baseline specification, although both interactions shrink in absolute value. This is consistent with the results where the size of the buffer region was shrunk. The diagonal interaction also shrinks but remains positive. The population-distance weights shrink the vertical interaction but increase the horizontal interaction relative to the baseline specification. The population weights and population inverse distance weights result in more intense vertical strategic interactions relative to the binary weights. The sizes of the horizontal interactions shrink slightly. When using these weighting schemes, the diagonal interaction intensifies in magnitude and becomes highly significant. 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 who may compete with other cities. However, small cities may compete with larger city centers if they view their consumers as deciding between purchasing a good in a small hometown store or in a big city where many more prominent and perhaps luxurious shopping opportunities may exist. For these reasons and because of its ease in interpretation, I prefer using the unweighted average of tax rates.

6.2.4 By Type of Border

As a final robustness check, Table 8 breaks down the results by whether the town is in a county that is a high-tax county, low-tax county, or a county that sets the same tax rate as the neighboring county. When doing this exercise, the strategic reaction to the town’s own county tax rate becomes closer to zero to zero for towns that are in high-tax counties. This suggests that for a town on the high-tax side of the border, if its county raises its tax rate, the town only partially offsets the county’s tax increase. On the other hand, for a town on the low-tax side of the border, when its county raises its tax rate, the town more than offsets the county’s tax increase – although the result is not significantly different from zero perhaps as a result of the small sample size.
6.3 Discussion

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

While spatial location effects explain much of the bias, allowing vertical and horizontal reactions to interact results in large heterogeneity in the strategic reaction functions. This interaction effect is important for jurisdictions in unitary federations as well. Towns that are in high-tax counties are much more likely to perceive their neighboring tax rates as strategic substitutes than a town in a low-tax county. In fact, for towns in counties with a tax rate above 3.75%, the regressions predict that the horizontal interactions are strategic substitutes rather than strategic complements. These interaction effects also make the vertical interactions extremely heterogeneous across the population.

Given that the tax competition literature often has difficulty distinguishing between strategic interaction and common spatial shocks, exploiting spatial location attributes is important to helping the researcher identify tax competition. In addition to this, the researcher must account appropriately for inter-federation competition when claiming to identify unbiased estimates of strategic interaction. Accounting for inter-federation competition can increase vertical strategic interactions by approximately 50% of the baseline estimated in the literature without the effects of inter-federation competition. Although the methods in this paper have their limitations, what is clear is that failing to account for competition across federation boundaries and for heterogeneity in strategic reaction functions can result in substantial omitted variable bias. Inter-federation competition is most transparent at the local level where county borders are relatively close to every town. However, inter-federation
competition may also arise at state and national borders if these borders are open.

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. 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.

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. 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.

With respect the nature of the strategic interaction, local sales taxes are strategic complements with neighboring sales taxes – a finding consistently found in the literature. However,
for a subset of towns in high-tax counties, local tax rates may be strategic substitutes with their neighbors. With respect to vertical interactions, the theoretical literature suggests the interaction may be positive or negative. I find that a one percentage point increase in county tax rates lowers municipal tax rates by about .40 percentage points; this suggests that county and municipal taxes are strategic substitutes. The result is 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.

8 Appendix

8.1 Derivation of Equation 13

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

\[
\frac{(\varepsilon_i + \gamma_k - 2\theta_i)D_i - (4\theta_i + \varepsilon_i)(\varepsilon_i\theta_i - \varepsilon_i^2 - \varepsilon_i\eta_i - \theta_i^2 + \theta_i\gamma_k)}{D_i^2}.
\]

Expanding terms and grouping liker terms yields:

\[
\frac{3\varepsilon_i^3 - 3\varepsilon_i\theta_i^2 + 2\varepsilon_i^2\eta_i + 2\theta_i\varepsilon_i\eta_i - 2\gamma_k\theta_i^2 + \varepsilon_i\eta_i\gamma_k + 2\varepsilon_i^2\gamma_k}{D_i^2},
\]

Using the fact that \( \eta_i = -(1 + \varepsilon_i) \) for iso-elastic demand, yields the equation in the text. This expression will definitely be positive if:

\[
\varepsilon_i^3 + \varepsilon_i^2\gamma - \varepsilon_i^22\theta_i - (3\varepsilon_i + 2\gamma_k)\theta_i^2 > 0,
\]
Recalling that I have assumed that $\gamma_k$ is not a function of $\theta_i$, this expression is a quadratic with in the elasticity of cross-border shopping that has roots:

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

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

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

which evaluated at the only positive root simplifies nicely to $\chi^{1/2} > 0$. Therefore, I know that Equation 13 is positive if

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

### 8.2 Mathematics for Proposition 4

Assume that $\rho_i$ does not differ across towns and rewrite the expression $(\theta_{i,i+1} + \theta_{i,i-1} - \theta_i)$ as

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

Then use Roy’s identity to yield:

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

which is likely to be greater than zero when:

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

from Figure 2 is is evident that $\frac{\partial v_{i+1}}{\partial \tau} > \frac{\partial v_i}{\partial \tau} > \frac{\partial v_{i-1}}{\partial \tau}$. Thus, the closer $\frac{\partial v_i}{\partial \tau}$ is to $\frac{\partial v_{i+1}}{\partial \tau}$, the more likely the above expression will be positive.

### References


Figure 1: Reaction Functions

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

Figure 2: Change in Cross-Border Shopping due to a Change in the County Rate

The curve above represents the indirect utility function of a consumer. The tax rates increase from town \(i - 1\) to town \(i + 1\) such that town \(i\) has an after tax price in the middle. All towns are in the same county and the initial light straight lines represent the starting positions. Then consider what happens when the county raises its tax rate. Because each town is in the same county, this raises the after tax price \(q\) by the same amount in each county. The new positions are represented by the bold lines. However, because of the shape of the indirect utility function notice that the change in indirect utility is largest in the low-tax town and smallest in the high-tax town.
The vertical axis represents the strategic reaction of a town with respect to its neighboring towns from column (7) in table 3. The strategic reaction of the horizontal reaction function is a function of the county tax rate because the regression specification allows for horizontal and vertical reactions to interact. Notice that towns in relatively low tax counties perceive their neighbors' tax rates as strategic complements. However, towns in very high tax counties view their neighbors' tax rates as strategic substitutes.

In the left panel, the vertical axis represents the strategic reaction of a town with respect to its county's tax rate from column (7) in table 3. The strategic reaction of the vertical reaction function is a function of the neighboring town tax rates because the regression specification allows for horizontal and vertical reactions to interact. The regression specification also allows the strategic reaction to differ on the basis of proximity to the nearest county border. For simplicity I only show the reaction functions for towns within an hour of the nearest county border and for neighboring tax rates that average between zero and five percent. The left panel illustrates the heterogeneity in the reaction functions using a three dimensional figure. The right panel condenses the left figure into a contour plot where the darker colors correspond to strategic reactions that are more intense – which for downward sloped reaction functions evidenced above implies that darker colors imply the reaction function slopes are further from zero.
Table 1: Summary Statistics  
Averages of Variables by Type  
Standard Deviations in ()

<table>
<thead>
<tr>
<th>Type</th>
<th>Variable</th>
<th>Mean</th>
<th>( )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent</td>
<td>Local Tax Rate</td>
<td>.768</td>
<td>(1.164)</td>
</tr>
<tr>
<td>Other</td>
<td>Distance to County Border (minutes)</td>
<td>12.099</td>
<td>(19.859)</td>
</tr>
<tr>
<td>County Tax Rate</td>
<td></td>
<td>1.037</td>
<td>(1.188)</td>
</tr>
<tr>
<td>Endogenous Regressors</td>
<td>Unweighted Average Neighbors' Rate (50 Mile Radius)</td>
<td>.767</td>
<td>(.876)</td>
</tr>
<tr>
<td></td>
<td>Unweighted Average Neighbors' Rate (25 Mile Radius)</td>
<td>.764</td>
<td>(.952)</td>
</tr>
<tr>
<td></td>
<td>Neighboring County Tax Rate</td>
<td>.991</td>
<td>(1.173)</td>
</tr>
<tr>
<td>Town Perimeter</td>
<td></td>
<td>14.311</td>
<td>(23.799)</td>
</tr>
<tr>
<td>Town Area</td>
<td></td>
<td>5.550</td>
<td>(17.751)</td>
</tr>
<tr>
<td>Number of Neighbors</td>
<td></td>
<td>1.761</td>
<td>(1.768)</td>
</tr>
<tr>
<td>Population</td>
<td></td>
<td>9032</td>
<td>(58,169)</td>
</tr>
<tr>
<td>Senior (%)</td>
<td></td>
<td>16.035</td>
<td>(8.160)</td>
</tr>
<tr>
<td>Control Variables</td>
<td>Less than College (%)</td>
<td>81.802</td>
<td>(14.113)</td>
</tr>
<tr>
<td></td>
<td>Income</td>
<td>56.665</td>
<td>(30.817)</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>49.129</td>
<td>(5.447)</td>
</tr>
<tr>
<td>Percent on Public Assistance</td>
<td></td>
<td>2.434</td>
<td>(3.520)</td>
</tr>
<tr>
<td>White (%)</td>
<td></td>
<td>84.442</td>
<td>(19.986)</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td>39.352</td>
<td>(8.023)</td>
</tr>
<tr>
<td>Ratio of Private School : Public School Students</td>
<td></td>
<td>.131</td>
<td>(.487)</td>
</tr>
<tr>
<td></td>
<td>Number of Rooms</td>
<td>5.564</td>
<td>(.733)</td>
</tr>
<tr>
<td></td>
<td>Age of Structure</td>
<td>44.245</td>
<td>(15.712)</td>
</tr>
<tr>
<td></td>
<td>Non-citizen (%)</td>
<td>3.149</td>
<td>(5.817)</td>
</tr>
<tr>
<td></td>
<td>Obama Vote Share</td>
<td>43.747</td>
<td>(13.773)</td>
</tr>
<tr>
<td>Sample Size</td>
<td></td>
<td>12,996</td>
<td></td>
</tr>
</tbody>
</table>

The log of driving time to the state border, a dummy for the relatively high-tax side of the border and same-tax side of a border, the tax differential at the border and a complete set of interactions is also included in every regression specification.
Table 2: Slopes of Reaction Functions – OLS

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau_{i,j}$</td>
<td>-0.440***</td>
<td>-0.177***</td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td>(0.024)</td>
</tr>
<tr>
<td>$t_{-i}$</td>
<td>0.499***</td>
<td>0.638***</td>
</tr>
<tr>
<td></td>
<td>(0.028)</td>
<td>(0.031)</td>
</tr>
<tr>
<td>$t_{-i}\tau_j$</td>
<td>-0.167***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.015)</td>
<td></td>
</tr>
<tr>
<td>$\tau_{i,-j}$</td>
<td>0.004</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td></td>
</tr>
<tr>
<td>$\tau_{i,j}d_i$</td>
<td>0.0009**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0003)</td>
<td></td>
</tr>
<tr>
<td>$\tau_{i,-j}d_i$</td>
<td>0.0003</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0004)</td>
<td></td>
</tr>
<tr>
<td>$E\left[\frac{\partial t}{\partial \tau_j}\right]$</td>
<td>-0.440***</td>
<td>-0.295***</td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td>(0.018)</td>
</tr>
<tr>
<td>$E\left[\frac{\partial t}{\partial t_{-i}}\right]$</td>
<td>0.499***</td>
<td>0.464***</td>
</tr>
<tr>
<td></td>
<td>(0.028)</td>
<td>(0.028)</td>
</tr>
<tr>
<td>$E\left[\frac{\partial t}{\partial \tau_{-j}}\right]$</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.014)</td>
</tr>
</tbody>
</table>

Buffer - 50
Distance - linear
Method OLS OLS
$R^2$ .689 .697
N 12,996 12,996

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

All of the specifications above include state fixed effects and the local control variables outlined in the text. The above results are estimated using OLS – thus ignoring endogeneity issues. Column (1) estimates the standard equation used in the literature. Column (2) accounts for the effects of inter-federation competition as suggested by the theoretical model in the paper. The mean derivatives represent the slope of the reaction functions.
Table 3: Slopes of Reaction Functions – Baseline Specification

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
<th>(10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau_{i,j}$</td>
<td>-0.280**</td>
<td>-0.251***</td>
<td>-0.282*</td>
<td>-0.192</td>
<td>-0.317**</td>
<td>-0.286*</td>
<td>-0.255***</td>
<td>-0.417***</td>
<td>-0.367*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.118)</td>
<td>(0.092)</td>
<td>(0.157)</td>
<td>(0.149)</td>
<td>(0.156)</td>
<td>(0.155)</td>
<td>(0.096)</td>
<td>(0.113)</td>
<td>(0.202)</td>
<td></td>
</tr>
<tr>
<td>$t_{i}$</td>
<td>0.550***</td>
<td>0.536***</td>
<td>0.506***</td>
<td>0.587***</td>
<td>0.582***</td>
<td>0.627***</td>
<td>0.527***</td>
<td>0.513***</td>
<td>0.558**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.069)</td>
<td>(0.066)</td>
<td>(0.092)</td>
<td>(0.154)</td>
<td>(0.147)</td>
<td>(0.144)</td>
<td>(0.066)</td>
<td>(0.065)</td>
<td>(0.251)</td>
<td></td>
</tr>
<tr>
<td>$t_{i} \tau_{j}$</td>
<td>0.026</td>
<td>-0.094</td>
<td>-0.116</td>
<td>-0.163</td>
<td>-0.075</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.083)</td>
<td>(0.194)</td>
<td>(0.186)</td>
<td>(0.182)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\tau_{i,j}d_{i}$</td>
<td>0.001**</td>
<td>0.001**</td>
<td>0.001**</td>
<td>0.002**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.0006)</td>
<td>(0.0006)</td>
<td>(0.0007)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\tau_{i}d_{i}$</td>
<td>-0.0002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0005)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$E\left[\frac{\partial t}{\partial \tau_{j}}\right]$ represents the slope of the vertical reaction function. $E\left[\frac{\partial t}{\partial t_{i}}\right]$ represents the slope of the horizontal reaction function. $E\left[\frac{\partial t}{\partial \tau_{i}}\right]$ represents the slope of the diagonal reaction function.

Buffer - 50 50 50 50 50 50 50 50 50 50
Distance - - - - - linear linear - linear linear
Method GMM GMM GMM GMM GMM GMM GMM GMM GMM LIML
Over-id† 0.193 0.077 0.203 0.356 0.291 0.564 0.517 0.154 0.585 0.517
Inst.‡ 19.93 19.93 7.56 7.77 - - 7.56 9.01 -
$R^2$ 0.659 0.651 0.682 0.679 0.690 0.694 0.690 0.682 0.579 0.694
N 12,996 12,996 12,996 12,996 12,996 12,996 12,996 12,996 12,996 12,996

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

All of the specifications above include state fixed effects and the local control variables outlined in the text. The instruments are always
the area and perimeter of the respective jurisdictions. In the case of horizontal externalities, the instruments are the average area and
perimeter in the buffer zone of fifty miles around each town. The instrument for interactions are the interaction of the instruments.

When the regression includes a tax term interacted with distance, the distance variable also enters the regression equation as a
stand-alone variable. The measure of distance is the time to the nearest county border and it enters all specifications in a linear
manner. Column (1) allows for vertical competition only. Column (2) allows for only horizontal competition. Column (3) is the
benchmark specification estimated in the literature. Column (4) adds interaction effects between the horizontal and vertical
externalities. Column (5) allows for interaction effects and diagonal interactions. Column (6) allows the vertical externality to vary
based on proximity to the county border. Column (7) allows the diagonal externality to also vary with respect to proximity to the
border. Columns (8) and (9) demonstrate the effect of only adding the diagonal externality or proximity effects on the vertical
externality to the baseline specification. Column (10) is the same as column (7) except that it is estimated using limited information
maximum likelihood. $E\left[\frac{\partial t}{\partial \tau_{j}}\right]$ represents the slope of the vertical reaction function. $E\left[\frac{\partial t}{\partial t_{i}}\right]$ represents the slope of the horizontal reaction function. $E\left[\frac{\partial t}{\partial \tau_{i}}\right]$ represents the slope of the diagonal reaction function.

†The test of over-identification reports the p-value of the Hansen J test.
‡The test for weak instruments is 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 – Neighbors Within 25 Miles

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau_{i,j}$</td>
<td>-0.280**</td>
<td>-0.175*</td>
<td>-0.414***</td>
<td>-0.246**</td>
<td>-0.374**</td>
<td>-0.341**</td>
<td>-0.401**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.118)</td>
<td>(.101)</td>
<td>(.167)</td>
<td>(.109)</td>
<td>(.144)</td>
<td>(.141)</td>
<td>(.171)</td>
<td></td>
</tr>
<tr>
<td>$t_{-i,j}$</td>
<td>0.436***</td>
<td>0.401***</td>
<td>0.276***</td>
<td>0.315***</td>
<td>0.277*</td>
<td>0.314***</td>
<td>0.244**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.069)</td>
<td>(.075)</td>
<td>(.105)</td>
<td>(.107)</td>
<td>(.112)</td>
<td>(.110)</td>
<td>(.124)</td>
<td></td>
</tr>
<tr>
<td>$t_{-ij} \tau_{j}$</td>
<td>-0.200**</td>
<td>-0.134</td>
<td>-0.158</td>
<td>-0.140</td>
<td>-0.224*</td>
<td>-0.135</td>
<td>-0.159</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.101)</td>
<td>(.103)</td>
<td>(.107)</td>
<td>(.105)</td>
<td>(.125)</td>
<td>(.135)</td>
<td>(.128)</td>
<td></td>
</tr>
<tr>
<td>$\tau_{i,j} d_i$</td>
<td>0.001*</td>
<td>0.001</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.0066</td>
<td>0.0006</td>
<td>0.0007</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.0006)</td>
<td>(.0006)</td>
<td>(.0006)</td>
<td>(.0006)</td>
<td>(.0006)</td>
<td>(.0006)</td>
<td>(.0007)</td>
<td></td>
</tr>
<tr>
<td>$\tau_{i,j} d_i$</td>
<td>-0.0002</td>
<td>-0.0003</td>
<td>-0.0003</td>
<td>-0.0003</td>
<td>0.0006</td>
<td>0.0006</td>
<td>0.0007</td>
<td></td>
</tr>
</tbody>
</table>

$E\left[\frac{\partial t}{\partial \tau_j}\right]$ represents the slope of the vertical reaction function. $E\left[\frac{\partial t}{\partial t_{-i,j}}\right]$ represents the slope of the horizontal reaction function. $E\left[\frac{\partial t}{\partial \tau_{i,j}}\right]$ represents the slope of the diagonal reaction function.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>linear</td>
<td>linear</td>
</tr>
<tr>
<td>Method</td>
<td>GMM</td>
<td>GMM</td>
<td>GMM</td>
<td>GMM</td>
<td>GMM</td>
<td>GMM</td>
<td>LIML</td>
</tr>
<tr>
<td>Over-id†</td>
<td>.193</td>
<td>.244</td>
<td>.173</td>
<td>.595</td>
<td>.269</td>
<td>.438</td>
<td>.305</td>
</tr>
<tr>
<td>Weak Inst.‡</td>
<td>48.661</td>
<td>208.7</td>
<td>29.752</td>
<td>28.534</td>
<td>4.608</td>
<td>3.192</td>
<td>2.332</td>
</tr>
<tr>
<td>Critical‡</td>
<td>19.93</td>
<td>19.93</td>
<td>7.56</td>
<td>7.77</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$R^2$</td>
<td>.659</td>
<td>.670</td>
<td>.690</td>
<td>.658</td>
<td>.659</td>
<td>.661</td>
<td>.666</td>
</tr>
<tr>
<td>N</td>
<td>12,996</td>
<td>12,996</td>
<td>12,996</td>
<td>12,996</td>
<td>12,996</td>
<td>12,996</td>
<td>12,996</td>
</tr>
</tbody>
</table>

***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 3 except that horizontal neighborliness is defined within a 25 mile buffer region of the town rather than a 50 mile region. All other controls and instruments remain the same. Columns (1)-(7) are estimated using two stage GMM. Column (8) is estimated using limited information maximum likelihood.

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

‡The test for weak instruments is 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 – When Distance Is a Dummy Variable

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau_{i,j}$</td>
<td>-.176</td>
<td>-.198</td>
<td>-.254</td>
<td>-.199</td>
<td>-.216</td>
<td>-.245</td>
<td>-.247**</td>
</tr>
<tr>
<td></td>
<td>(.134)</td>
<td>(.142)</td>
<td>(.164)</td>
<td>(.125)</td>
<td>(.134)</td>
<td>(.168)</td>
<td>(.121)</td>
</tr>
<tr>
<td>$t_{-i,j}$</td>
<td>.647***</td>
<td>.654***</td>
<td>.652**</td>
<td>.355***</td>
<td>.352***</td>
<td>.300***</td>
<td>.541***</td>
</tr>
<tr>
<td></td>
<td>(.137)</td>
<td>(.135)</td>
<td>(.321)</td>
<td>(.103)</td>
<td>(.102)</td>
<td>(.114)</td>
<td>(.065)</td>
</tr>
<tr>
<td>$t_{-ij}\tau_j$</td>
<td>-.178</td>
<td>-.195</td>
<td>-.197</td>
<td>.100</td>
<td>.100</td>
<td>.179</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.176)</td>
<td>(.174)</td>
<td>(.467)</td>
<td>(.098)</td>
<td>(.097)</td>
<td>(.115)</td>
<td></td>
</tr>
<tr>
<td>$\tau_{i,-j}$</td>
<td>.169</td>
<td>.211</td>
<td>.231</td>
<td>-.037</td>
<td>-.018</td>
<td>-.087</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.175)</td>
<td>(.182)</td>
<td>(.524)</td>
<td>(.125)</td>
<td>(.131)</td>
<td>(.165)</td>
<td></td>
</tr>
<tr>
<td>$\tau_{i,j}d_i$</td>
<td>.026</td>
<td>.126</td>
<td>.167</td>
<td>.004</td>
<td>.064</td>
<td>.115</td>
<td>.016</td>
</tr>
<tr>
<td></td>
<td>(.025)</td>
<td>(.143)</td>
<td>(.227)</td>
<td>(.024)</td>
<td>(.147)</td>
<td>(.180)</td>
<td>(.022)</td>
</tr>
<tr>
<td>$\tau_{i,-j}d_i$</td>
<td>-.103*</td>
<td>-.143</td>
<td>-.061</td>
<td>-.110</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.145)</td>
<td>(.217)</td>
<td>(.153)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$E[\frac{\partial t}{\partial \tau_j}]$ = -.310***, -.334**, -.389, -.123, -.133, -.097, -.245***

$E[\frac{\partial t}{\partial t_{-i}}]$ = .463***, .453***, .448**, .458***, .457***, .485***, .541***

$E[\frac{\partial t}{\partial \tau_{-j}}]$ = .169, .201, .216, -.037, -.024, -.098, .098

<table>
<thead>
<tr>
<th>Buffer</th>
<th>50</th>
<th>50</th>
<th>50</th>
<th>25</th>
<th>25</th>
<th>25</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method</td>
<td>GMM</td>
<td>GMM</td>
<td>LIML</td>
<td>GMM</td>
<td>GMM</td>
<td>LIML</td>
<td>GMM</td>
</tr>
<tr>
<td>Over-id†</td>
<td>.414</td>
<td>.435</td>
<td>.446</td>
<td>.220</td>
<td>.245</td>
<td>.261</td>
<td>.232</td>
</tr>
<tr>
<td>Weak Inst.‡</td>
<td>2.326</td>
<td>2.006</td>
<td>2.006</td>
<td>3.685</td>
<td>3.915</td>
<td>3.915</td>
<td>10.879</td>
</tr>
<tr>
<td>Critical‡</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7.77</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>.689</td>
<td>.684</td>
<td>.681</td>
<td>.667</td>
<td>.668</td>
<td>.639</td>
<td>.681</td>
</tr>
<tr>
<td>N</td>
<td>12,996</td>
<td>12,996</td>
<td>12,996</td>
<td>12,996</td>
<td>12,996</td>
<td>12,996</td>
<td>12,996</td>
</tr>
</tbody>
</table>

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

The estimating equations in this table are identical to Tables 3 and 4 except the interaction with the distance function uses a dummy variable interaction, rather than a linear distance term. All instruments and control variables remain the same. The dummy variable takes on the value of one if the driving time to the county border is more than 22 minutes and a value of zero otherwise. The cutoff of 22 minutes corresponds 90th percentile of towns with respect to time from the border. The expected sign on $\tau_{i,j}d_i$ remains positive while the expected sign on $\tau_{i,-j}d_i$ remains negative. Columns (1)-(3) corresponds to columns (6), (7) and (10) in table 3. Columns (4)-(6) correspond to columns (6)-(8) in table 4. Column (7) corresponds to column (9) in Table 3. $E[\frac{\partial t}{\partial \tau_j}]$ represents the slope of the vertical reaction function. $E[\frac{\partial t}{\partial t_{-i}}]$ represents the slope of the horizontal reaction function. $E[\frac{\partial t}{\partial \tau_{-j}}]$ represents the slope of the diagonal reaction function.

†The test of over-identification reports the p-value of the Hansen J test.
‡The test for weak instruments is the the robust Kleibergen-Paap Wald rk F statistic. I report the critical values (when available) for which the bias induced by weak instruments is less than 10% if the test statistic is greater.
Table 6: Slopes of Reaction Functions – Robustness Checks: Restrictions and Weighting

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \tau_{i,j} )</td>
<td>-.271***</td>
<td>-.280*</td>
<td>-.401</td>
<td>-1.15</td>
<td>-.317***</td>
<td>-.383***</td>
</tr>
<tr>
<td></td>
<td>(.097)</td>
<td>(.146)</td>
<td>(.259)</td>
<td>(1.24)</td>
<td>(.114)</td>
<td>(.223)</td>
</tr>
<tr>
<td>( t_{-i} )</td>
<td>.546***</td>
<td>.555***</td>
<td>.686***</td>
<td>1.15**</td>
<td>.482***</td>
<td>.498***</td>
</tr>
<tr>
<td></td>
<td>(.076)</td>
<td>(.105)</td>
<td>(.142)</td>
<td>(.503)</td>
<td>(.099)</td>
<td>(.223)</td>
</tr>
<tr>
<td>( t_{-i} \tau_j )</td>
<td>-.034</td>
<td>-.946</td>
<td>1.03</td>
<td>.295</td>
<td>.0004</td>
<td>.0001</td>
</tr>
<tr>
<td></td>
<td>(.085)</td>
<td>(.768)</td>
<td>(.117)</td>
<td>(.255)</td>
<td>(.004)</td>
<td>(.001)</td>
</tr>
<tr>
<td>( \tau_{i,j}d_i )</td>
<td>.0009</td>
<td>.205</td>
<td>.205</td>
<td>.292</td>
<td>.251</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.0005)</td>
<td>(.283)</td>
<td>(.283)</td>
<td>(.292)</td>
<td>(.251)</td>
<td></td>
</tr>
<tr>
<td>( \tau_{i,j}d_i )</td>
<td>-.0003</td>
<td>-.221</td>
<td>-.221</td>
<td>-.003</td>
<td>(.0003)</td>
<td>(.308)</td>
</tr>
<tr>
<td></td>
<td>(.0003)</td>
<td>(.308)</td>
<td>(.308)</td>
<td>(.308)</td>
<td>(.0007)</td>
<td></td>
</tr>
</tbody>
</table>

\[ E[\frac{\partial t}{\partial \tau_j}] \] represents the slope of the vertical reaction function.  
\[ E[\frac{\partial t}{\partial t_{-i}}] \] represents the slope of the horizontal reaction function.  
\[ E[\frac{\partial t}{\partial \tau_{i,j}}] \] represents the slope of the diagonal reaction function.

**All the specifications include the same set of controls and fixed effects as the previous tables. The instruments also remain the same.**

(1)-(2) drops towns where the nearest county border is also a state border in order to eliminate diagonal externalities between the towns and state governments. (3)-(4) restrict the estimating sample to towns within five miles of a county border. Columns (5)-(6) weight the observations in the sample so that each state receives equal weight and such that states with more towns do not give undue weight to the results. \( E[\frac{\partial t}{\partial \tau_j}] \) represents the slope of the vertical reaction function.  
\( E[\frac{\partial t}{\partial t_{-i}}] \) represents the slope of the horizontal reaction function.  
\( E[\frac{\partial t}{\partial \tau_{i,j}}] \) represents the slope of the diagonal reaction function.

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

**‡**The test for weak instruments is the robust Kleibergen-Paap Wald rk F statistic. I report the critical values (when available) for which the bias induced by weak instruments is less than 10% if the test statistic is greater.
Table 7: Slopes of Reaction Functions – Robustness Checks: Various Neighbor Weights

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \tau_{i,j} )</td>
<td>-0.200**</td>
<td>-0.368***</td>
<td>-0.623***</td>
<td>-0.475***</td>
<td>-0.645***</td>
<td>-0.432***</td>
</tr>
<tr>
<td></td>
<td>(0.099)</td>
<td>(0.133)</td>
<td>(0.068)</td>
<td>(0.125)</td>
<td>(0.067)</td>
<td>(0.121)</td>
</tr>
<tr>
<td>( t_{-i} )</td>
<td>0.451***</td>
<td>0.439***</td>
<td>0.032**</td>
<td>0.495***</td>
<td>0.044**</td>
<td>0.600***</td>
</tr>
<tr>
<td></td>
<td>(0.074)</td>
<td>(0.106)</td>
<td>(0.015)</td>
<td>(0.125)</td>
<td>(0.020)</td>
<td>(0.132)</td>
</tr>
<tr>
<td>( t_{-i} \tau_{j} )</td>
<td>-0.001</td>
<td>-1.263***</td>
<td>-1.623***</td>
<td>-1.263***</td>
<td>-1.623***</td>
<td>-1.623***</td>
</tr>
<tr>
<td></td>
<td>(0.102)</td>
<td>(0.031)</td>
<td>(0.035)</td>
<td>(0.031)</td>
<td>(0.035)</td>
<td>(0.035)</td>
</tr>
<tr>
<td>( \tau_{i,-j} )</td>
<td>0.052</td>
<td>0.195**</td>
<td>0.228***</td>
<td>0.228***</td>
<td>0.228***</td>
<td>0.228***</td>
</tr>
<tr>
<td></td>
<td>(0.114)</td>
<td>(0.083)</td>
<td>(0.084)</td>
<td>(0.084)</td>
<td>(0.084)</td>
<td>(0.084)</td>
</tr>
<tr>
<td>( \tau_{i,j}d_{i} )</td>
<td>0.001**</td>
<td>0.002***</td>
<td>0.002***</td>
<td>0.002***</td>
<td>0.002***</td>
<td>0.002***</td>
</tr>
<tr>
<td></td>
<td>(0.0006)</td>
<td>(0.0007)</td>
<td>(0.0007)</td>
<td>(0.0007)</td>
<td>(0.0007)</td>
<td>(0.0007)</td>
</tr>
<tr>
<td>( \tau_{i,-j}d_{i} )</td>
<td>-0.0003</td>
<td>0.0004</td>
<td>0.0003</td>
<td>0.0003</td>
<td>0.0003</td>
<td>0.0003</td>
</tr>
<tr>
<td></td>
<td>(0.0006)</td>
<td>(0.0006)</td>
<td>(0.0006)</td>
<td>(0.0006)</td>
<td>(0.0006)</td>
<td>(0.0006)</td>
</tr>
</tbody>
</table>

\( E[\frac{\partial t}{\partial \tau_{j}}] \) is the slope of the vertical reaction function. 
\( E[\frac{\partial t}{\partial t_{-i}}] \) is the slope of the horizontal reaction function. 
\( E[\frac{\partial t}{\partial \tau_{-j}}] \) is the slope of the diagonal reaction function.

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

All the specifications include the same set of controls and fixed effects as the previous tables. The instruments also remain the same.

Towns receive non-zero weight in the neighboring weight matrix if they are within the fifty mile buffer zone. (1)-(2) uses inverse distance weights to calculate the weighted average of neighboring tax rates such that it weights each town in \( t_{-i} \) by the inverse distance to the neighbor. (3)-(4) weights each town in \( t_{-i} \) by the population of the neighbor such that the most weight is given to the largest jurisdictions in the 50 mile radius. (5)-(6) weights each town in \( t_{-i} \) by the population of the neighbor and then by the inverse of distance to the neighbor such that the most weight is given to large towns that are also close. \( E[\frac{\partial t}{\partial \tau_{j}}] \) represents the slope of the vertical reaction function. \( E[\frac{\partial t}{\partial t_{-i}}] \) represents the slope of the horizontal reaction function. \( E[\frac{\partial t}{\partial \tau_{-j}}] \) represents the slope of the diagonal reaction function.

† The test of over-identification reports the p-value of the Hansen J test.
‡ The test for weak instruments is 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 8: Slopes of Reaction Functions – Robustness Checks: Type of Border

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \tau_{i,j} )</td>
<td>-.206</td>
<td>-.346*</td>
<td>-.589**</td>
<td>-.236</td>
<td>-.445**</td>
<td>-.708***</td>
</tr>
<tr>
<td></td>
<td>(.227)</td>
<td>(.183)</td>
<td>(.232)</td>
<td>(.628)</td>
<td>(.221)</td>
<td>(.294)</td>
</tr>
<tr>
<td>( t_{-i} )</td>
<td>.419***</td>
<td>.523***</td>
<td>.467***</td>
<td>.707***</td>
<td>.576***</td>
<td>.538***</td>
</tr>
<tr>
<td></td>
<td>(.123)</td>
<td>(.130)</td>
<td>(.133)</td>
<td>(.228)</td>
<td>(.097)</td>
<td>(.111)</td>
</tr>
<tr>
<td>( t_{-i} \tau_j )</td>
<td>.021</td>
<td>-.294</td>
<td>.012</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.079)</td>
<td>(.327)</td>
<td>(.081)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \tau_{i,-j} )</td>
<td>-.092</td>
<td>.723***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.210)</td>
<td>(.240)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \tau_{i,j}d_i )</td>
<td>.003</td>
<td>-.003</td>
<td>.002***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.002)</td>
<td>(.004)</td>
<td>(.003)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \tau_{i,-j}d_i )</td>
<td>-.00039</td>
<td>-.003</td>
<td>.0003</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.004)</td>
<td>(.004)</td>
<td>(.003)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[ E\left[ \frac{\partial t}{\partial \tau_j} \right] \]

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \tau_{i,j} )</td>
<td>-.206</td>
<td>-.321*</td>
<td>-.589**</td>
<td>-1.228</td>
<td>-.445**</td>
<td>-1.682***</td>
</tr>
<tr>
<td></td>
<td>(.227)</td>
<td>(.168)</td>
<td>(.232)</td>
<td>(.752)</td>
<td>(.221)</td>
<td>(.258)</td>
</tr>
<tr>
<td>( E\left[ \frac{\partial t}{\partial \tau_{-i}} \right] )</td>
<td>.419***</td>
<td>.556***</td>
<td>.467***</td>
<td>.128</td>
<td>.576***</td>
<td>.547***</td>
</tr>
<tr>
<td></td>
<td>(.123)</td>
<td>(.108)</td>
<td>(.133)</td>
<td>(.453)</td>
<td>(.097)</td>
<td>(.091)</td>
</tr>
<tr>
<td>( E\left[ \frac{\partial t}{\partial \tau_{i,j}} \right] )</td>
<td>-.082</td>
<td>.395</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.189)</td>
<td>(.279)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Buffer  | 50    | 50    | 50    | 50    | 50    |
| Distance Restriction | High-Tax County | High-Tax County | Low-Tax County | Low-Tax County | Same-Tax County |
| Method       | GMM   | GMM   | GMM   | GMM   | GMM   |
| Over-id†    | .557 | .387 | .091 | .330 | .1667 | .063 |
| Weak Inst.† | 8.187 | 2.097 | 7.938 | 1.718 | 16.162 | 14.574 |
| Critical‡  | 7.56 | - | 7.56 | - | 7.56 | - |
| \( R^2 \) | .639 | .634 | .689 | .641 | .724 | .723 |
| \( N \)  | 3155 | 3155 | 2850 | 2850 | 6990 | 6990 |

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

All the specifications include the same set of controls and fixed effects as the previous tables. The instruments remain the same. This table tests whether the strategic reactions depend on whether a town is located in a high-tax, low-tax or same-tax county relative to the nearest neighboring county. (1)-(2) restrict the sample to towns where the nearest neighboring county sets a lower tax rate than its own county. (3)-(4) restrict the sample to towns where the nearest neighboring county sets a higher tax rate than its own county. (5)-(6) restrict the sample to towns where the nearest neighboring county sets the same tax rate as its own county. \( E\left[ \frac{\partial t}{\partial \tau_j} \right] \) represents the slope of the vertical reaction function. \( E\left[ \frac{\partial t}{\partial \tau_{-i}} \right] \) represents the slope of the horizontal reaction function. \( E\left[ \frac{\partial t}{\partial \tau_{i,j}} \right] \) represents the slope of the diagonal reaction function.

†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.