Essays on the Theory and Empirical Evidence Regarding Spatial Tax Competition

by

David R. Agrawal

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in The University of Michigan
2012

Doctoral Committee:

Professor Joel B. Slemrod, Chair
Professor Robert J. Franzese Jr.
Professor James R. Hines Jr.
Assistant Professor David Y. Albouy

© David R. Agrawal 2012

All Rights Reserved

To Professor Dhammika Dharmapala for inspiring me to study public finance.

ACKNOWLEDGEMENTS

I am especially grateful to my dissertation committee chair, Joel Slemrod, for support of my research agenda, many hours of discussions, and an endless set of comments on my work whenever I needed it. The members of my committee – David Albouy, Robert Franzese, Jim Hines – provided invaluable office visits, advice, encouragement, and mentoring. I am especially thankful for the many difficult challenges that were asked of me and for the opportunity to learn from each of my committee members. I am happy to say that I strive to have the succes, intuition, and the technical grasp of economics that Joel, Jim, Rob and David have.

The first chapter of my dissertation – "The Tax Gradient" – benefited from comments and discussions with Claudio Agostini, Leah Brooks (discussant), Sebastien Bradley, Charles Brown, Paul Courant, Lucas Davis, Michael Devereux (discussant), Dhammika Dharmapala, Reid Dorsey-Palmateer, Michael Gideon, Makoto Hasegawa, William Hoyt (discussant), Ravi Kanbur, Sebastian Kessing, Miles Kimball, David Knapp, Michael Lovenheim (discussant), Yulia Paramonova, Emmanuel Saez, Stephen Salant, Daniel Silverman, Jeffrey Smith, Caroline Weber, and David Wildasin. I also wish to thank conference and seminar participants at the 2010 National Tax Association Conference, the 2010 Michigan Tax Invitational, the Michigan Public Finance Seminar, Oxford University's Centre for Business Taxation Doctoral Conference, the 2011 International Institute of Public Finance Congress, the 2011 National Tax Association Conference, the University of Alberta, Florida State University, the University of South Carolina, Georgia State University, the University of

Georgia, Cornell University, The Ohio State University John Glenn School of Public Affairs, the University of Toronto, Copenhagen Business School, Uppsala University, the University of Mannheim, and the Wharton School at the University of Pennsylvania. Special thanks to the International Institute of Public Finance for honoring this chapter with the Peggy and Richard Musgrave Prize.

The second chapter of my dissertation – "Games Within Borders" – benefited from comments and discussions with Paul Courant, Lucas Davis, Charles de Bartolomé, Dhammika Dharmapala, Reid Dorsey-Palmateer, Marcel Gérard (discussant), Michael Gideon, Makoto Hasegawa, Ravi Kanbur, Sebastian Kessing, Michael Lovenheim (discussant), Byron Lutz (discussant), Ben Niu, Stephen Salant, Nathan Seegert, Jeffrey Smith, Caroline Weber, and David Wildasin. Suggestions from conference and seminar participants at the 2011 International Institute of Public Finance Annual Congress, the 2011 Association of Public Economic Theory Conference, the Michigan Tax Invitational, and the Michigan Summer Seminar greatly improved the paper.

The final chapter of my dissertation – "Inter-Federation Competition" – benefited from comments and discussions with Charles Brown, Paul Courant, Michael Devereux, Michael Gideon, Makoto Hasegawa, Ben Lockwood, Jeffrey Smith and Caroline Weber. Suggestions from conference and seminar participants in the Michigan Public Finance Lunch, the 2011 National Tax Association Conference, and Oxford University's Centre for Business Taxation greatly improved the paper.

I owe special thanks to Caroline Weber, with whom I have shared an office in the Office of Tax Policy Research for four years. Over the course of those years, I have learned much from her and I am deeply appreciative of her mentoring as my senior classmate.

This research was made possible by the generous provision of data by ProSales Tax. Nicole Scholtz of the Spatial and Numeric Data Services provided excellent assistance with ArcGIS. Support from the Rackham Graduate School and the Office

of Tax Policy Research at the University of Michigan was greatly appreciated. Mary Ceccanese helped support the educational environment in the Office of Tax Policy Research.

Finally, I wish to thank my parents – Frances and Krishan Agrawal – and my grandparents – Robert and Irene Auliso – for their support of my education and career. I learned many things from them.

TABLE OF CONTENTS

DEDICATION ii

ACKNOWLEDGEMENTS iii

LIST OF FIGURES ix

LIST OF TABLES x

ABSTRACT xi

CHAPTER

I. Introduction 1

II. The Tax Gradient: Do Local Sales Taxes Reduce Tax Differentials at State Borders? 8

2.1 Introduction 8
2.2 Background 11
2.2.1 Comparison to the Existing Literature 12
2.2.2 Background on the Local Option Sales Tax 14
2.3 Theory of Local Competition with Heterogeneous States . . 16
2.3.1 Setup of the Model 16
2.3.2 Benchmark Solutions 20
2.3.3 Equilibrium with Three Heterogeneous States 21
2.4 Data and Evidence on the Levels of Local Tax Rates 28
2.4.1 Data 29
2.4.2 Graphical Analysis and Summary Statistics 31
2.4.3 Methodology: The Tax Level Effect 32
2.4.4 Result: Are Local Taxes Higher on One Side of the Border? 32
2.5 Estimating the Tax Gradient 33
2.5.1 Methodology: The Tax Gradient Effect 34

2.5.2 Results: How Steep Is the Tax Gradient? 40
2.5.3 Is the Tax Gradient Steepest Near the Border? . . . 44
2.5.4 Is the Tax Gradient Steepest for Large Tax Differentials? 46
2.5.5 Extension: Second Closest State Borders and County Borders 47
2.5.6 Extension: Who Adopts Local Taxes Within a State? 48
2.5.7 Discussion 50
2.6 Conclusion 52
2.7 Appendix 55
2.7.1 Proof of Uniqueness of the Equilibrium 55
2.7.2 Characterizing the Solution to an n Town Model . . 56
2.7.3 Methodology for Calculating Distance from the Border 59
2.7.4 An Alternative Empirical Specification 62

III. Games Within Borders: Are Geographically Differentiated Taxes Optimal? 85

3.1 Introduction 85
3.2 Tax Competition and Cross-Border Shopping 88
3.3 Model 89
3.3.1 Setup of the Model 89
3.3.2 Consumers and Firms 89
3.3.3 The State Planner's Problem 91
3.4 Results 93
3.4.1 Solving the Model 93
3.4.2 The Marginal Cost of Funds in a Federation 97
3.4.3 Characterizing a Pattern to the Optimal Tax Rates 99
3.5 What Determines the Outcome? 103
3.5.1 A State with Low Public Good Preferences 103
3.5.2 A State with High Public Good Preferences 106
3.6 Extensions and Discussion 108
3.6.1 Extensions 108
3.6.2 Comparison to the Existing Literature 109
3.7 Conclusion 111
3.8 Appendix 112
3.8.1 Derivation of MCF Terms 112
3.8.2 Derivation of Equation 3.19 113
3.8.3 Proof of Proposition 1 114
3.8.4 Proof of Corollary 1 114
3.8.5 Proof of Proposition 2 114
3.8.6 Proof of Corollary 2 115
3.8.7 The Characteristics of Lines and Optimal Tax Policy 115
3.8.8 Geographic Differentiation with Local Public Goods 120

IV. Inter-Federation Competition: Sales Tax Externalities with Multiple Federations 126

4.1 Introduction 126
4.2 Sales Tax Competition in Federations 131
4.2.1 Theoretical Models of Fiscal Federalism 131
4.2.2 Empirically Estimating Reaction Functions 133
4.2.3 Relationship to Industrial Organization 134
4.3 Model 136
4.3.1 Nash Equilibrium 141
4.3.2 Strategic Interaction 142
4.4 Implications of the Theory for the Empirical Analysis 150
4.4.1 Why Interaction Effects Are Essential for Estimation 152
4.5 Empirical Methodology 154
4.5.1 Data 154
4.5.2 Estimation in a Cross-Section 156
4.6 Empirical Results 161
4.6.1 Hypotheses to Test 161
4.6.2 Main Results 162
4.6.3 How Robust Are The Results? 166
4.7 Conclusion 171

V. Conclusion 180

BIBLIOGRAPHY 186

LIST OF FIGURES

Figure

2.1 Circular Model: Multiple States 66
2.2 Example of Equilibrium Tax Rates 66
2.3 Driving Distance from the Nearest Border 67
2.4 State Sales Tax Rates 67
2.5 County Tax Rates by Nearest Border Type 68
2.6 Missouri Example of Local Option Taxes 68
2.7 Regression Discontinuity Results 69
2.8 Local Linear Regression Versus Global Polynomial Regression . . . 70
2.9 Average Marginal Effects by Distance, Full Sample 71
2.10 Average Marginal Effects by Distance, States with Same Tax Rate . 71
2.11 Average Marginal Effects by Distance with 98 Mile Sample Restriction 72
2.12 Average Marginal Effects by Distance and Size of the Discontinuity 72
2.13 Predicted Values of Local Rates from a Global Polynomial Regression 73
2.14 Average Marginal Effects by Distance of the Probit Regression . . . 74
2.15 Simulation for a Multiple Town Model 74
2.16 Methodology for Calculating Distance 75
3.1 Geographic Layout of the Model 125
3.2 Summary of Possible Cases 125
4.1 Reaction Functions 174
4.2 Change in Cross-Border Shopping due to a Change in the County Rate 174

LIST OF TABLES

Table

2.1 Sales Tax Summary Statistics by State (April 2010) 76
2.2 Summary Statistics 77
2.3 Average Tax Rates by Driving Distance 78
2.4 RD Estimates 78
2.5 Regression Coefficients for the Estimating Equation 79
2.6 Mean Derivatives for Several Specifications 81
2.7 Mean Derivatives for Several Specifications 81
2.8 State by State Marginal Effects 82
2.9 Average Marginal Effects by Size of the Treatment/Discontinuity . 83
2.10 Mean Derivatives for Multiple Borders 84
4.1 Summary Statistics 175
4.2 Slopes of Reaction Functions – Neighbors Within 50 Miles 176
4.3 Slopes of Reaction Functions – Neighbors Within 25 Miles 177
4.4 Slopes of Reaction Functions – When Distance Is Linear 178
4.5 Slopes of Reaction Functions – Robustness Checks 179

ABSTRACT

The United States commodity tax system is fiscally decentralized: state, county, and municipal governments have taxing authority. Geographic borders create "lines," which are demarcations of goods into different tax rate categories; tax rates change discontinuously at lines. A discontinuous change in tax liabilities is a "notch" and results in deadweight loss from cross-border shopping and tax driven product innovation. This dissertation studies the consequences of lines and notches on tax competition and utilizes variation in local taxes to understand tax policy. The dissertation focuses on sales taxation, but the theories and empirical methodologies are informative for tax policy more generally and are applicable to issues of preferential taxation, optimal line drawing, and time and characteristic notches. Chapter II extends the standard model of sales tax competition to allow for tax rates to vary not only across states, but within states. I demonstrate that municipal sales taxes are lower in level terms on the high-tax side of a state border relative to municipal sales taxes on the low-tax side of a border. Using a new methodology for studying municipal responses to discontinuities, I show that municipal taxes are a function of distance to the nearest state border. After accounting for differences in tax rates within a state, discontinuous tax changes at borders become smaller. Chapter III answers whether geographic differentiation of tax rates is optimal from a state government's perspective. I demonstrate that under relatively weak assumptions, differentiation of tax rates is almost always optimal. Replacing a large distortion with many smaller distortions everywhere is optimal from a social welfare perspective. Chapter IV studies how municipal

governments react to county tax rates. Using an identification strategy grounded in a theoretical model, I identify strategic reaction functions. Jurisdictions mimic the tax rates of their neighbors. I also demonstrate – unlike state-federal reaction functions estimated in the literature – that a higher county tax rate implies strategically lower municipal tax rates. The results in this dissertation imply that fiscal federalism and both spatial location and spatial interaction must be considered in order to understand tax competition and optimal commodity taxation.

CHAPTER I

Introduction

Tax competition is the process of strategic interaction with other jurisdictions due to inter-jurisdictional mobility of the tax base. Tax competition can take two forms: interactions between governments of the same level and interactions between governments of different levels within a federation. Competition between different levels of government is called vertical competition and is present in fiscally decentralized countries. Raising taxes for one level of government implies a loss in the tax base to other levels of government; this is a negative fiscal externality. Fiscal federalism is the study of these inter-governmental relationships between the federal, state, and sub-state governments. Competition with neighboring jurisdictions is horizontal competition. Raising taxes in one jurisdiction implies a gain to the tax base of its neighbors; this creates a positive fiscal externality. A key element of horizontal competition is determining who the "neighbors" of a jurisdiction are. Geo-spatial economics is the study of geographic relationships between observations.

The subsequent chapters will focus on the nature and consequences of tax competition in public finance and will focus on unifying three main themes – tax competition, fiscal federalism, and geo-spatial economics – into a single and comprehensive framework.

The literature on tax competition focuses on how jurisdictions set tax rates when

they compete over a mobile tax base. In the case of commodity taxes, the mobile base is sales, which are subject to a retail sales tax or value added tax. The workhorse models of sales tax competition analyze how two asymmetric governments compete over tax rates. These models often do not have a federal government or if they do, the higher-level of government will enter passively. The traditional workhorse models are inherently "spatial" models because the consumers and firms are arranged at various distances from a jurisdiction's borders. Although the agents are arranged spatially, the models do not allow for spatial relationships among the jurisdictions. This dissertation will incorporate fiscal federalism and spatial relationships among jurisdictions into standard theoretical and empirical models of tax competition by allowing for tax rates to vary not only across states, but within states. I now discuss both of these innovations in turn.

First, fiscal federalism is an essential characteristic of many countries' tax systems. For example, retail sales taxation in the United States is decentralized to the states, which in turn have granted localities and counties taxation authority. In Canada, the provinces and territories have the authority to add additional taxes on top of the federal Goods and Services Tax. Federalism is not only a North American phenomenon. The Indian sales tax system grants states the authority to vary tax rates and tax bases. And even as the European Union has continued to discuss fiscal harmonization, national autonomy in setting the Value Added Tax rates remains an important issue. For the lowest levels of government within a federation (towns, counties, and districts within counties), the nature of the tax competition is inherently about spatial proximity – including its geographic relationship within the broader "federation's" borders. Is the town located in high- or a low-tax state or county? How many levels of government have taxing authority? Do states view a single municipality as a competitor or do they view that municipality as atomistic with respect to the state tax base? Is the jurisdiction internal or peripheral to the

federal government's borders?

Second, geographic and distance-based relationships between jurisdictions are essential to understanding the precise form of tax competition. If states have different tax rates within the state – and states will if the political system is characterized by federalism – it is inevitable that the geographic relationships among jurisdictions is essential in determining the tax equilibrium. Geographic borders create a discontinuous tax treatment of retail sales and encourage cross-border shopping by residents of high-tax states. These borders can result in otherwise homogenous jurisdictions setting non-uniform tax rates solely because the jurisdictions are heterogeneous in spatial proximity, which carries the implications that tax incidence and behavioral responses to sales taxes are heterogeneous within a state. Do jurisdictions compete with their neighbors or do they compete with other jurisdictions that have similar economic characteristics? How salient are borders between states and countries to lower levels of government? How far is someone willing to drive to save a few dollars in sales taxes? These are questions that – depending on the answer – will result in a municipality having a very different strategic reaction in setting tax rates.

This dissertation will focus on two dimensions of "spatial economics" – "spatial interdependence or contagion" and "spatial location." One, spatial interdependence is the process by which one jurisdiction has a contagion effect on another (perhaps neighboring) jurisdiction's tax rate. For example, when a jurisdiction sets a tax rate, it maximizes an objective function that aggregates the welfare of residents within the jurisdiction, but does so while competing with neighboring jurisdictions for a mobile tax base. This competitive process will influence the tax setting behavior of other geographically close and possibly overlapping jurisdictions. Two, spatial location is the process by which distance from or proximity to a particular geographic feature influences tax setting behavior. An example of spatial location is that tax rates may be a function of proximity to a border or to an amenity. As is highlighted by Chapter

IV, the strength of spatial interdependence can be heterogeneous with respect to a jurisdiction's spatial location. Previous studies of tax competition have focused on spatial interdependence, while ignoring spatial location. Adding elements of spatial location will complicate the analysis, but will provide me with a unique and convincing identification strategy to identify tax competition where I rely on demonstrating that strategic interaction is heterogeneous with respect to spatial location in a manner predicted by theory.

The study of tax competition in a federal and spatial context raises several fundamental questions that have been the subject of many public finance papers. Does strategic competition among localities and states result in a race to the bottom? Is coordination or harmonization of tax rates desirable? How decentralized should the tax system be and how many levels of government should have taxing authority? These are fundamental "big-picture" questions outlined in the tax competition literature. They are questions that are applicable to current policies and debates. For example, policymakers are currently discussing tax harmonization within the European Union, fundamental tax reform such as introducing a Value Added Tax in the United States, and the dire fiscal problems faced by many large urban cities inability to maintain a broad tax base. Yet, these are questions that the tax competition literature cannot completely answer without accounting for fiscal federalism and spatial relationships within states. I hope this dissertations is the start of such a discussion.

Chapter II will focus on the discontinuous tax treatment of retail sales resulting from state borders and will study how allowing for municipal/local option taxes has the potential to mute these discontinuities. In a model where towns within a federation maximize revenue and compete in a Nash game, equilibrium local tax rates decrease from the nearest high-tax border and increase from the nearest low-tax border. Using driving distance from the state border and data on all local sales tax rates in the United States for 2010, I empirically test whether tax rates follow the

pattern predicted by this theoretical model. Local tax rates on the low-tax side of the border are significantly higher than on the high-tax side of the border, reducing the differential in state tax rates at the border by more than half. Consistent with the model's prediction, a 100 mile increase in distance from the nearest high-tax border lowers local tax rates by 15 percent of the average local rate. Local taxes fall most rapidly closest to the border and when the differential in state tax rates is largest.

The results in Chapter II suggest that further fiscal decentralization can mute the distortions induced by state borders. However, allowing for municipal choice can potentially add additional distortions internal to the state. For this reason, it is not clear whether allowing for local option sales taxes are welfare enhancing even though the revenue consequences may appear clearer.

Chapter III studies whether it is optimal from a state perspective to levy geographically differentiated commodity taxes or preferential tax rates near borders. I show in a model where states within a federation maximize social welfare, a state's optimal commodity tax system is almost always geographically differentiated. The optimal pattern of geographic differentiation critically depends on fundamental parameters as well as whether a state has a preference for high or low tax rates. Under the assumption that utility is linear in consumption and that the elasticity of cross-border shopping is less than unity in absolute value, high-tax states should find it optimal to set a lower tax rate in the border region and low-tax states should set higher tax rates in the border region relative to the periphery region. High-tax states should set higher tax rates in the border region if the social welfare measure is sufficiently redistributive. Social welfare maximization makes it possible for taxes to be higher in the region near the state border compared to the region away from the border – an outcome that cannot arise when the government cares only about total tax revenue.

Although geographic differentiation is almost always optimal in Chapter III, fur-

ther fiscal decentralization can heighten horizontal tax competition between governments of the same level while also fostering additional vertical tax competition between governments of different levels within a federation. While Chapters II and III allow for horizontal externalities, the state governments view localities as atomistic or non-strategic. This is a realistic assumption for state governments, but it is unlikely that county governments view localities as atomistic. For county governments, potentially large vertical strategic reactions may exist.

Chapter IV considers the strategic interaction of towns and counties by analyzing how introducing inter-federation competition affects the strategic nature of tax competition. Inter-federation competition is traditionally used to mean competition among nations. However, I use the term inter-federation competition to highlight competition among the higher levels of government. As such, counties are federal to towns just as the national government is federal to the state governments. Throughout this chapter, inter-federation competition will refer to competition among multiple counties that are composed of multiple towns, but the theoretical models in the paper could similarly apply to the national-state reltionship as well. In the context of a Nash equilibrium, the paper shows that towns at the interior of a county will respond differently to the county tax rate in comparison to a town near the county's borders. Inter-federation competition will also introduce "diagonal externalities," which are fiscal externalities between neighboring jurisdictions, but that occur between different levels of government. A diagonal externality will have similar effects as a horizontal externality. The paper uses two unique data sets, a cross-section of all local tax rates in the United States and spatial proximity data, to test how local governments react to horizontal, vertical, and diagonal competitors. The empirical specifications allow for vertical and horizontal externalities to have interaction effects and allow for strategic reactions that vary based on proximity to neighboring federations. After including these terms, a one percentage point increase in the county tax rate implies

that municipal tax rates in the county will be approximately 0.60 percentage points lower for the average jurisdiction. Such a result contrasts with strategic reactions that are often insignificant or slightly positive for state-federal reaction functions.

The results in Chapter IV suggest that reaction functions for state-federal interactions are likely to be very different than town-county interactions. Geographic position within the federation is also essential to determining the magnitude of the strategic reaction. This highlights the main point of this dissertation: when studying tax competition, the federalist structure will imply that spatial interdependence, spatial location, and their interaction are all essential to analyzing tax setting behavior.

Chapter V will return to the fundamental "big-picture" questions in the literature and provide some answers. Does strategic competition among localities and states result in a race to the bottom? Is coordination or harmonization of tax rates desirable? How decentralized should the tax system be and how many levels of government should have taxing authority?

CHAPTER II

The Tax Gradient: Do Local Sales Taxes Reduce Tax Differentials at State Borders?

2.1 Introduction

The total sales tax rate in America ranges anywhere between 0% and 11%. Given these disparities in tax rates, consumers may have large incentives to engage in cross-border shopping, while firms may have large incentives to locate on the low-tax side of borders. Geographic borders create discontinuous changes in tax rates, which distort individual consumption and firm location decisions. Moreover, these discrete jumps in tax rates at borders may induce towns to levy their sales tax based on an approximately continuous function of distance from the border. This motivates my main question of interest: do localities assess local sales taxes as a function of distance from the nearest state border, where tax rates decrease for localities further from a higher-tax state, and increase for localities further from a lower-tax state?

Consider the example of high-tax California and low-tax Oregon. As a possible policy, California could change taxes continuously and set lower taxes closer to the Oregon border. This policy, in comparison to the uniform discontinuous case, would not change the effective price of the good for consumers (including taxes and transportation costs), but may keep extra revenue in California. However, a tax rate that is continuous in distance from the border may have high administrative costs. Use of the local option sales tax may be an administratively feasible way of obtaining a

geographically differentiated pattern of sales tax rates.

As a consequence of distortions from state tax competition, a municipality will set geographically differentiated tax rates depending on distance from the border. Therefore, the spatial arrangement of jurisdictions determines the strategic nature of local tax competition. Intuitively, the spatial arrangement matters because state borders influence the elasticity of demand with respect to prices. Local jurisdictions with a lower elasticity of demand will set higher tax rates. In the presence of discontinuous changes in state tax rates at borders, on the low-tax side, jurisdictions near the border realize a smaller elasticity of demand because the inflow of cross-state shoppers augments their tax base. On the high-tax side, the elasticity of demand is larger for jurisdictions near the border because the outflow of cross-state shoppers reduces the tax base.

In the context of a theoretical model that expands several elements from Kanbur and Keen (1993), I demonstrate that from the local government's perspective, the equilibrium tax rates in a decentralized system depend on the distance of each jurisdiction to a neighboring state. This contrasts with the existing tax competition literature, which has focused on the role of jurisdiction size as a determinant of tax rates and has generally concluded that if jurisdictions are the same in all characteristics, they will select the same tax rate. The model in this paper provides a novel insight: the spatial characteristics of towns – distance from the border and the size of the discontinuity in state tax rates at the border – within a federation determine tax rates. Given varying public good preferences across states, a broad uniform tax rate within the federation will not occur in equilibrium if sub-state governments exist – even if all of the sub-federal governments are identical in every respect except spatial arrangement.

The theoretical model yields four testable results. First, in a local region of the border, municipal sales taxes on the low-tax side of the border are higher than

municipal taxes on the high-tax side of the border. Second, on the relatively low-tax side of the border, municipal taxes decrease as towns are further from the border; on the relatively high-tax side of the border, local taxes increase as towns are further from the border. I refer to the pattern (slope) of local option taxes, moving away from the border, as the "tax gradient." Third, local taxes rise or fall most rapidly in the vicinity of the state border. This means that the tax gradient is steepest in the vicinity of the border. Fourth, the size of the discontinuity affects the slope of the tax gradient. The tax gradient is steeper when the differential in state tax rates is larger.

Empirically, the paper examines whether localities with the option to assess local sales taxes set their tax rate in a manner consistent with the four theoretical propositions above. The methodology outlined in the paper will be applicable to research on how jurisdictions respond to any policy (e.g., environmental, labor) that varies discontinuously at the state border. The paper uses a previously unused and comprehensive data set of all local sales tax rates in the United States – municipal, county, and district rates. I also have created an accurate and detailed data set of spatial proximity to borders. I find the shortest driving distance from the population centroid of each town to a state-border major road crossing, in order to accurately measure how towns set tax rates away from the border. This distance minimizes measurement error in the actual distance a consumer would travel to cross-border shop.

First, I find that local tax rates on the low-tax side of the border are significantly higher than on the high-tax side of the border, reducing the differential in state tax rates at the border by more than half. Ignoring local option taxes, the average differential in state tax rates is 1.9 percentage points. After accounting for all local option taxes, the tax differential at state borders decreases by 1.2 percentage points. Second, the average marginal effect of distance on the local (municipal plus sub-

district) tax rate in a jurisdiction is significantly negative and of the expected sign on the low-tax side. The marginal effect of distance is robust to controlling for distance from the second closest state border and the distance from county borders. A 100 mile increase in distance from the nearest high-tax border lowers local tax rates by 15% of the average local rate. Third, taxes fall most rapidly in a local region of the border. Fourth, taxes fall more rapidly if the tax differential in state tax rates is larger.

This paper will proceed as follows. After presenting background information on local option taxes, I develop a model where multiple town governments face pressures from tax differentials at multiple state borders and choose local tax rates in order to maximize revenue in the town. I show that in a Nash equilibrium, local tax rates depend on the distance from a state border. Given that the system of local option sales taxes in the United States approximates the theoretical model, I evaluate whether such a relationship between tax rates and distance from the state border exists in the data using a global polynomial regression design. Finally, I discuss the welfare implications of the policy and offer some concluding thoughts.

2.2 Background

In this section, I review the standard tax competition literature that my model will build upon. I also review current studies that have incorporated a role for distance from the border in their empirical or theoretical analysis of cross-border shopping. Finally, I provide policy background on local taxes and how they vary across states.

2.2.1 Comparison to the Existing Literature

2.2.1.1 Tax Competition and Distance

Previous literature on tax competition tries to explain asymmetries in tax rates among competing jurisdictions when each jurisdiction chooses a uniform tax rate within its boundaries. The approach to solving this problem has varied substantially. Kanbur and Keen (1993) develops a model with a single good in partial equilibrium where governments are revenue maximizers. Nielsen (2001) extends Kanbur and Keen (1993) to welfare maximizing governments, but relies on an additively separable relationship between consumer surplus and revenue. Hoyt (2001) studies the optimality of sales taxes within a federation.

The conclusions from these models highlight the way jurisdictions may reach an equilibrium with different tax rates. Haufler (1996) implies that a state like California may have higher tax rates than Oregon because of stronger preferences for public goods. This is in contrast to models such as Bucovetsky (1991), Kanbur and Keen (1993), Nielsen (2001), and Trandel (1994), which focus on country size or population as an explanation for variation in tax rates and find that larger jurisdictions set higher rates. This paper uses the result that tax differentials will exist at state borders as a starting point. The model proposed in my paper will extend Kanbur and Keen (1993) by incorporating multiple competitors and distance from the border as an explanation for tax differences within states. My model will help to explain responses of jurisdictions to tax notches, which according to Slemrod (2010) are widespread in any federalist tax system.

Recently, several studies have focused on the role of distance to a competing jurisdiction, which will be a key variable in this paper. Lovenheim (2008) studies how distance from the nearest lower-tax cigarette state relates to the demand elasticity of home state consumption. He finds that cigarette demand becomes more elastic

to the home state price the further individuals live from a lower price cigarette border because the cost to obtain a given amount of saving rises. This implies that cross-border shopping is most problematic for border localities.[1] In a similar light, Harding, Leibtag and Lovenheim (Forthcoming) find that the incidence of taxation varies depending on a firm's distance from the nearest lower-tax border. Both papers find that the nearest low-tax border is the only border which results in any geographic differentiation. Merriman (2010) also analyzes distance from the border and shows the likelihood of having an Indiana cigarette tax stamp (the low-tax neighboring state to Illinois) is decreasing in the distance from the Indiana border. The results from these papers clearly indicate that distance from the border shapes the responsiveness of individuals to cross-border shop.

2.2.1.2 Competition Among Localities

Many studies have analyzed how local sales taxes influence cross-border shopping with little emphasis on the geographic distribution of these local sales tax rates within a state. Empirically, a large literature attempts to quantify the price responsiveness of consumers to state border effects. The findings suggest that a 1% increase in a sales tax rate results in a 1% to 6% reduction in sales, although the geographic unit of analysis varies across different studies (Mikesell 1970; Fox 1986; Walsh and Jones 1988; Tosun and Skidmore 2007). This behavioral response is indicative of the degree of cross-border shopping and will play an important role in the following model.

The behavioral response is also important to determine with whom localities compete. Using data on local sales tax rates in Tennessee, Luna (2003) demonstrates that the sales tax rates in neighboring states influence the tax rate that a local government selects, both in the long run and in the short run. Similar inter-dependencies are found in Georgia (Sjoquist et al. 2007). However, Luna, Bruce and Hawkins (2007)

[1] Using a similar methodology, Lovenheim and Slemrod (2010) studies the effect of having a neighboring county with a lower minimum legal drinking age on the number of accident fatalities.

find that a jurisdiction does not decide to reach the maximum statutory rate based on its neighbors' decisions to do so.

2.2.1.3 Summary

Although the location of state borders are not chosen as a matter of policy, geographic borders between different states create a discontinuous tax policy. State borders are an example of a "line" in the tax system. The line creates a "notch" – or a discontinuous jump in the tax rate of the good based on the characteristics of the good – where the characteristic of the good is the location of purchase.

The existing literature indicates that competition over sales will naturally result in tax "notches" at state borders, but has ignored how localities will respond to these differentials. Given that these differentials arise, the empirical literature indicates cross-border shopping is highly elastic to tax rate changes. But, the responsiveness of cross-border shopping is empirically not uniform in a state, resulting in larger responses closer to the border. In light of this, local option taxes may provide a mechanism for smoothing tax differentials, which result in cross-border shopping and distorted firm location choices at state borders.

2.2.2 Background on the Local Option Sales Tax

Local option sales taxes (LOST) are widely used in the United States. Of the forty-five states that impose a sales tax ranging between 2.9% and 7%, thirty-six allow local or county governments to set LOST. Over 7,500 localities utilize this option. Among these towns, the local sales tax contributed anywhere from 1% to 52.2% of municipalities' revenues during fiscal year 2006 (Mikesell 2010). Sales taxes in the United States are levied *de facto* according to the origin principle.[2] This implies that

[2]When an individual cross-border shops, the sales tax is paid in the jurisdiction of purchase. However, the individual is legally responsible for filing a use tax – the excise tax on purchases of items for which the home state's sales tax was not previously paid – in the state of residence. The use tax is notoriously under-enforced. Because the use tax is often evaded, taxes are implicitly

the jurisdiction of sale rather than the jurisdiction of residence effectively determines the tax paid.

There is substantial variation in the way the local option sales tax works.[3] Fourteen states do not allow for LOST. Of the remaining states that allow for LOST, the locality's degree of autonomy varies greatly. For example, the smallest unit that is granted autonomy to assess a tax varies from the county level (example: Wyoming) to the town level (most states), to within-town jurisdictions such as fire, school or transportation districts (examples: Colorado or Georgia). Of states that allow municipalities to set a tax, some do not allow counties to assess an additional tax (example: South Dakota), although most do. In other states, a mandatory county rate is set uniformly across the state with the option to increase the rate (example: California). Some consumers face different tax rates street blocks away while others need to travel more than 50 miles before the tax rate changes.

States also vary in terms of how the tax base is defined. Lines are drawn on what goods are taxed under the retail sales tax. In most states, the definition of the tax base at the state level is the base that applies to LOST. Some exceptions exist. For example, in the state of Florida, only the first $5,000 of a purchase is taxable under LOST. Other states impose restrictions on the rate increases that localities can impose at any given time. For example, counties in Ohio can only select taxes in increments of 1/4 of a percentage point and the maximum rate a county can assess is capped (at a fairly high rate). On the other hand, the maximum LOST in Iowa is capped at 1 percentage point, so "maxing out" is common.

The method in which localities determine whether to implement LOST and the rate at which to set it also varies by state. In most states, only a city or town

paid based on the sales tax and, therefore, are based on the origin of the purchase rather than the destination of the sale.

[3] All of the descriptions in this section come from my own review of the Department of Revenue sites for each state. The *Guide to Sales and Use Taxes* (Research Institute of America 2010) also provides state-by-state information on state sales taxes.

government needs to pass LOST. In other states, such as Iowa, the process is more complicated. In Iowa, a referendum determines LOST. Voters determine the rate of the tax, the purpose of the tax, and the sunset provisions on the tax. North Carolina, on the other hand, requires approval of the state legislature for LOST rates. The method of collection also varies; businesses remit taxes directly to the state or the locality, depending on the state.

Finally, two states allow local jurisdictions to set implicitly negative tax rates. Within Urban Enterprise Zones in New Jersey and Empire Zones in New York, localities may set tax rates lower than the state tax rate at no revenue cost to the locality. In fact, some locations elect to implement the favorable rate.

2.3 Theory of Local Competition with Heterogeneous States

In this section, I develop a model to evaluate the equilibrium commodity tax rates when local jurisdictions can assess a local sales tax. The goal of this model is to show that the equilibrium local sales tax rates are not uniform within a state – even if the towns are identical in preferences and size.

My paper extends Kanbur and Keen (1993)'s two-state partial equilibrium model of cross-border shopping by allowing for multiple jurisdictions and multiple levels of government. Compared to standard models, I place jurisdictions on a Salop Circle rather than a Hotelling line so that all jurisdictions have two neighbors.

2.3.1 Setup of the Model

The model features three states located on a circle that are indexed by $j = H, M, L$ for high-, medium-, and low-tax states. Each state has three identical towns[4] indexed $i = A, B, BB$, where I arrange the towns along the circumference of a circle

[4]The towns could be equally sized sub-regions. The word "town" only need imply a governing entity.

as depicted in Figure 2.1. "A" denotes the towns "Away" from the border. "B" and "BB" denote the towns close to the "Border" in their respective states. Each town is of equal length and each town covers a distance of x units along the line segment.

Modeling jurisdictions on a circle rather than a line has important implications. On a line segment, the towns furthest away from the state border have only one neighbor rather than two by virtue of their position at the end of the line segment. The purpose of this model is to determine – from a town government's perspective – the pattern of geographic differentiation resulting from the notch at the state border. When considering such a problem, the local government will take into account the town's position along the line segment, which includes both the number of towns away from the state border and the number of borders with other towns. Therefore, if I use a line segment in the model, towns on the exterior of the line will possibly have different rates for two reasons – one, because they are far from the state border and two, because they can only undercut one town rather than two and therefore have a different elasticity. However, all the variation in the tax rates is a result of distance from the border and not from the number of neighbors present, when I place the states along the circumference of a circle.[5]

The model has three agents: producers, consumers, and governments. I assume firms providing one private consumption good locate exogenously at any point along the circle's circumference. Firms are perfectly competitive and set price equal to marginal cost. When purchases at a particular point along the circle are high, more stores enter. Although the distribution of stores need not be uniform, firms are simply responding to consumption decisions; firms do not manipulate cross-border shopping. The implication of perfect competition is that increases in the demand of a good in a particular town – resulting from cross-border shopping – will not alter the pre-tax

[5]Of course, the linear model may be more descriptive to some states. One example is Florida, where the northern border touches two states. The Florida peninsula is surrounded entirely by ocean water, which implies that there are many Florida towns that are likely to have "one" less neighbor.

price relative to the pre-tax price in another town. Nor will taxes alter the pre-tax price. Therefore, the pre-tax price is the same in all jurisdictions and is normalized to one.[6]

State governments levy a state sales tax rate, τ^j, on commodity purchases within the state. In this analysis, I assume that state tax rates are exogenous and known to all localities before localities compete over taxes.[7] Exogenously different preferences for a state public good will imply that the state tax rates will differ across the states.[8] State H sets the highest tax rate and State L sets the lowest tax rate. State M has a rate in between the other two rates such that $\tau^H > \tau^M > \tau^L$. Denote $S = \tau^M - \tau^L$, $R = \tau^H - \tau^M$, and $D = \tau^H - \tau^L$ so that it measures the size of the "notch" induced by the state tax differential. Note that because states are around a circle, it must be that $D = S + R$. Town governments i in state j levy local taxes on the consumption good at rate t_i^j.[9] Taxes are assessed under the origin principle so that the location of purchase defines the tax rate that the consumer pays. Denote the sum of the state and local tax rate in jurisdiction i of state j as T_i^j so $T_i^j = \tau^j + t_i^j$. Towns compete in a Nash game over the local sales tax rates. The objective of the local government is to maximize the tax revenue it raises from the local tax on the consumption goods purchased within its town borders.[10]

[6]The model assumes the incidence of the tax is fully passed forward to consumers as well.

[7]The explicit assumption requires that tax competition over local sales tax rates has no effect on equilibrium state sales tax rates. The intuition for this is that most localities in the United States are small. Therefore, small changes in the local tax rates should not induce a response in the state sales tax rates. Of course, changes in large city tax rates may violate this assumption. An alternative model would be one where states are Stackelberg leaders.

[8]All states are the same in size and population density. Haufler (1996) shows that differences in preferences for the public good are an alternative explanation for varying tax rates in equilibrium. I assume that these exogenously different preferences for state public goods result in different state tax rates, but make no further use of this assumption.

[9]The model assumes governments only have access to a sales tax. If other tax instruments were available, the results below would be similar so long as governments rely on sales taxes to some extent.

[10]An alternative to modeling revenue maximizing governments is to model welfare maximizing governments. If governments maximize welfare, the government will care not only about the revenue raised, but also the amount of the consumption good its residents can purchase. Revenue maximizing governments can be interpreted in several ways. The solution to the problem of revenue maximizing governments approximates the solution to welfare maximizing governments when consumers have

Consumers are distributed uniformly across each town and the populations are identical in all towns. Consumers cannot migrate. Demand is perfectly inelastic. Each consumer will purchase one unit of the consumption good, but will have a choice over the location of purchase and can travel along the circumference of the circle to do so. Transportation costs make purchasing goods in another jurisdiction less beneficial.

Let V denote the reservation value net of the producer price for each consumer. V is assumed to be large enough so that all consumers either purchase one unit of the good from their home town or elsewhere. If the individual decides to purchase in the home town, she simply goes to the store at the point of the circle corresponding to where she lives and does not incur any transportation costs. If the consumer elects to do this, she purchases the good for a price equal to the tax-inclusive marginal cost of production. The surplus she will receive from such a purchase is $V - T_i^j$.

Alternatively, each consumer can purchase the consumption good elsewhere.[11] If she elects to do so, she will drive along the circumference to the nearest town border and purchase the good from the store located exactly at the border. Let the distance to the nearest town border for any consumer be denoted s. The transportation cost of traveling to the border (and back) is denoted $\delta > 0$ per unit of travel. The surplus the consumer will receive from purchasing one unit of the private good abroad is $V - T_k^l - \delta s$, where $k \neq i$ indexes the tax rate in a foreign town of state l, which may

a high marginal value of the public good financed by the government. Second, so long as borders are open, the solution for revenue maximizing governments will approximate the welfare maximizing government's solution when the marginal utility of consumption is constant (which will be the case when the utility function of the individual is quasi-linear in consumption). Linearity in the consumption good seems to be a realistic assumption for cross-border shopping, which often results in many purchases of a few similar items. I use revenue maximizing governments in my model because the introduction of multiple towns and levels of government complicates the already complex problem of two competing jurisdictions and I need this simplifying assumption to make the model tractable.

[11] In reality, individuals also have the option to purchase goods on the Internet and completely evade taxes. As long as some people in each state still cross-border shop and the use of the internet does not vary by town, all the results of the model will carry through. The presence of an Internet with tax-free purchases will only affect the number of people cross-border shopping and the relative slopes of local tax rates, but the qualitative pattern will be the same.

be equal to j if she does not cross state lines.

A consumer will purchase the private good from the neighboring town if the surplus of purchasing the good is strictly greater than zero and if the surplus from buying the good elsewhere is strictly greater than buying the good at home. Comparing the consumer surplus from purchasing the good elsewhere with the surplus from purchasing the good from home, it is evident that a consumer will purchase the good elsewhere if:

$$\frac{T_i^j - T_k^l}{\delta} > s \text{ for } T_i^j > T_k^l. \tag{2.1}$$

The implication of Equation 2.1 is that all residents who live farther than $\frac{T_i^j - T_k^l}{\delta}$ units from a low-tax border will purchase the good at home, while all other consumers cross-border shop in the nearest low-tax jurisdiction. The model assumes that x is sufficiently large so that towns do not have incentives to target consumers multiple towns over.[12]

2.3.2 Benchmark Solutions

Given the assumptions outlined above, it is easy to see that each town government will extract all the surplus from its residents in equilibrium if all borders are closed or if a use tax could be perfectly enforced. Therefore, the equilibrium tax rates will be to set $t_i^j = V - \tau^j$ for $i = A, B, BB$ in $j = L, M, H$. If all states have equal state tax rates and borders are open, the equilibrium local tax rates would not be geographically differentiated and would be characterized by $t_i^j = \delta x$ for all i. This contrasts with a model where towns are located along a line segment. Along a line segment, the Nash equilibrium is for Town A to set local tax rates that are $4/5$ of the

[12]The assumption of shopping no more than one town over simplifies the nature of the problem, but would not change the qualitative results. Relaxing this assumption would require checking incentive constraints on the individual for all towns to which cross-border shopping is feasible. The assumption of a linear transport cost and a sufficiently large x rules out this possibility. Intuitively, relaxing this assumption would reduce the degree of geographic differentiation and the level of local tax rates as towns will compete over the tax rates more aggressively.

level of the tax rates at the extremity because of their ability to undercut two towns rather than one. If all state taxes are uniform or if all town borders are closed (e.g., use taxes are perfectly enforced), the equilibrium town tax rates are uniform within a state.[13]

2.3.3 Equilibrium with Three Heterogeneous States

It is essential to know the direction of cross-border shopping to solve the model. Recall that although the state rates are exogenous, the towns within the state are free to set whatever tax rate they desire. I consider all possible patterns of local option taxes along the circle; taxes increase from borders, decrease from borders, or taxes are random. I must also allow for the possibility that after local option taxes are assessed, the total tax rate on the low-tax side of the border may be higher than on the high-tax side.

The revenue for each town can be derived using Equation 2.1. The tax base is defined as the total number of consumers within a town i of state j minus those individuals in that town who shop elsewhere plus individuals from other towns who shop in i. Multiplying the tax base by t_i^j yields total revenue and the municipality maximizes total revenue by selecting t_i^j. The best response functions are linear and will be continuous when changing from a high- to a low-tax jurisdiction. This continuity as towns change regimes implies that one set of best response functions characterizes the equilibria for all possible cases – including non-symmetric cases.

The revenue functions for towns in State M are defined in Equation 2.2. The revenue functions for State H and L are omitted for simplicity, but are defined in a

[13]International borders are much more likely to be "closed" to cross-border shopping. Transportation costs, broadly defined, are much higher for crossing an international border. International borders usually take more time to cross, have some probability of vehicle search, and require individuals to exchange their currency before they can purchase goods abroad.

similar manner.

$$R_i^M = \begin{cases} t_A^M(x + \frac{t_{BB}^M - t_A^M}{\delta} + \frac{t_B^M - t_A^M}{\delta}) & \text{for Town A} \\ t_{BB}^M(x + \frac{t_B^H - t_{BB}^M + R}{\delta} + \frac{t_A^M - t_{BB}^M}{\delta}) & \text{for Town BB} \\ t_B^M(x + \frac{t_{BB}^L - t_B^M - S}{\delta} + \frac{t_A^M - t_B^M}{\delta}) & \text{for Town B.} \end{cases} \tag{2.2}$$

Notice that xt_i^j denotes the revenue in the absence of cross-border shopping. The second and third terms represent the in- and out-flows resulting from cross-border shopping with both neighbors. If these terms are positive, then cross-border shopping is inward. If they are negative, cross-border shopping is outward. If the neighboring state is a high-tax state, the discontinuity in tax rates enters positively, but if the neighboring state is a low-tax state, the differential in the state tax rates enters negatively.

Solving the problem by differentiating the revenue functions yields the following best response functions:

$$\begin{array}{lll} t_{BB}^H(\cdot) = \frac{1}{4}(\delta x - (R+S) + t_A^H + t_B^L) & t_A^H(\cdot) = \frac{1}{4}(\delta x + t_{BB}^H + t_B^H) & t_B^H(\cdot) = \frac{1}{4}(\delta x - R + t_A^H + t_{BB}^M) \\ t_{BB}^M(\cdot) = \frac{1}{4}(\delta x + R + t_B^H + t_A^M) & t_A^M(\cdot) = \frac{1}{4}(\delta x + t_{BB}^M + t_B^M) & t_B^M(\cdot) = \frac{1}{4}(\delta x - S + t_A^M + t_{BB}^L) \\ t_{BB}^L(\cdot) = \frac{1}{4}(\delta x + S + t_B^M + t_A^L) & t_A^L(\cdot) = \frac{1}{4}(\delta x + t_{BB}^L + t_B^L) & t_B^L(\cdot) = \frac{1}{4}(\delta x + (R+S) + t_{BB}^H + t_A^L), \end{array} \tag{2.3}$$

which imply that municipality i's neighboring local tax rates are strategic complements, but that t_i^j and τ^j are strategic substitutes.

Solving this system of equations for a Nash equilibrium yields the following results:

$$\begin{array}{lll} t_{BB}^H = \kappa(\omega - 12R - 11S) & t_A^H = \kappa(\omega - 6R - 3S) & t_B^H = \kappa(\omega - 12R - S) \\ t_{BB}^M = \kappa(\omega + 11R - S) & t_A^M = \kappa(\omega + 3R - 3S) & t_B^M = \kappa(\omega + R - 11S) \\ t_{BB}^L = \kappa(\omega + R + 12S) & t_A^L = \kappa(\omega + 3R + 6S) & t_B^L = \kappa(\omega + 11R + 12S), \end{array} \tag{2.4}$$

where $\kappa = 1/53$ and $\omega = 53\delta x/2$. The first section of the Appendix proves that the Nash equilibrium derived above is unique for a general number of states and towns.

Differencing the local tax rates on the high- and low-tax side of borders shows

that border towns in a low-tax state will set relatively higher tax rates than border towns within a high-tax state. The differential in tax rates at the state border shrinks to $\frac{30}{53}$ of the differential when there is no competition over local option taxes and this pattern holds for all three discontinuities.

To determine if the Nash equilibrium exists and its pattern, I need to verify three conditions. First, I need to verify that T_i^j is larger on the high-tax side of the border in any equilibrium and that all taxes are positive. Second, I need to determine the direction of the inequality between $t_A^j \gtreqless t_B^j$ and $t_A^j \gtreqless t_{BB}^j$ for all states. Finally, I need to verify whether the direction of these inequalities is consistent with the assumption that cross-border shopping occurs only one town over along the continuum. For a Nash equilibrium to exist, I must also verify that the number of residents of each town that cross-border shop is strictly less than the total population of the town.

From Equation 2.4, the Nash equilibrium will always have the following properties: $t_A^H > t_{BB}^H$, $t_{BB}^M > t_A^M > t_{BB}^M$, and $t_B^L > t_A^L$. The intuition of this result is explained below. However, in a Nash equilibrium, $t_A^H \gtreqless t_B^H$ and $t_A^L \gtreqless t_{BB}^L$. The direction of these inequalities depends on the relative sizes of the model's parameters – specifically the differentials in state tax rates. It can be shown that $t_A^H > t_B^H$ if $R > \frac{1}{4}D$ and $t_{BB}^L > t_A^L$ if $S > \frac{1}{4}D$. If the reverse is true then the pattern of those two tax rates will flip. It is also easy to verify that total taxes, T_i^j, on the high-tax side of the border are always greater than total taxes on the low-tax side of the border in equilibrium.

For a Nash equilibrium to exist, I must find the size of each town that guarantees that tax rates are strictly positive, all towns have some residents that shop at home, and that no one will shop more than one town away.[14] Denoting the value of x that satisfies all three of these conditions as x^*, a Nash equilibrium exists in pure strategies so long as $x > x^*$. Strictly positive tax rates can be guaranteed by setting $x > \kappa\frac{22D+2R}{\delta}$. Verifying that no one shops more than one town over and that some

[14] This is equivalent to finding a value of δ that is sufficiently large.

shoppers purchase the good at home requires $x > \kappa\frac{30D}{\delta} > \kappa\frac{22D+2R}{\delta}$. Therefore, this guarantees that a small deviation in the tax rate of a particular town cannot change revenues discontinuously. Thus, $x > \kappa\frac{30D}{\delta}$ guarantees these assumptions hold. If $x < \kappa\frac{30D}{\delta}$, then no Nash equilibrium exists in pure strategies. But, $x > \kappa\frac{30D}{\delta}$, combined with continuity and concavity of the best response functions in the strategies, guarantees the existence of the unique Nash equilibrium described above.

After defining two terms, I can state a proposition that characterizes the Nash equilibrium for a general model with many towns.

Definition 2.1. The *tax gradient* is defined as the slope of local option taxes away from the border. Define the distance of a town to its relevant state border as the length along the circumference from the center of the town to the relevant state border. Then, the tax gradient is increasing in distance from the border if local option taxes increase as towns are further from the relevant state border. The tax gradient is decreasing in distance from the border if local option taxes decrease as towns are further from the relevant state border.

In the presence of multiple borders, the gradient will likely switch slopes at some point within the state.

Definition 2.2. The *critical town* is defined as the town where the tax gradient changes sign from increasing to decreasing (or vice-versa) within a state. If the tax gradient is increasing for both state borders in a state, then the critical town corresponds to the town with the maximum tax rate. If the tax gradient is decreasing for both state borders in a state, the critical town corresponds to the town with the minimum tax rate. If the tax gradient is increasing for one border and decreasing for the other border, then no towns are critical.

Proposition 1 characterizes the Nash equilibrium for a model with multiple states and multiple towns. The general solution to an n town model, where states can have

more than three towns, is in the appendix .

Proposition 2.3. *If towns are sufficiently large in size and a state has two heterogeneous tax differentials at state borders, a Nash equilibrium for local option taxes exists in pure strategies and is characterized by the following statements:*

(1) A state with one high-tax neighbor and one low-tax neighbor will have a tax gradient that is decreasing from the high-tax border to the low-tax border.

(2) A state with two neighbors that set relatively low state tax rates of different magnitudes will have tax gradients that are increasing away from each border.

(3) A state with two neighbors that set relatively higher state tax rates of different magnitudes will have tax gradients that are decreasing away from each border.

(4) In the high- and low-tax states, the critical town will be located closer to the border with the smaller state tax differential. How close it is depends on the relative sizes of the state tax differentials.

Border towns in the low-tax state will set higher taxes than towns at the interior of a low-tax state; border towns in the high-tax state will set lower taxes than towns at the interior of a high-tax state. Combined with the result that tax differentials are smaller after local competition, the results of the theory suggest that local option taxes do help to reduce the size of the discontinuity at the state border and that this reduction occurs gradually. A tax gradient begins to emerge where in high-tax states, taxes rise as distance from the border increases, while in low-tax states, taxes fall as distance from the border increases. Additionally, the changes in local taxes within a state are largest for the towns located near the state border with the largest discontinuities. The model in the appendix adds one additional insight; the tax gradient will be steeper in a local region of the border than in regions far from the closest border.

For illustrative purposes, Figure 2.2 shows the Nash Equilibrium tax rates when the parameters of the model are set such that $\delta x = 4$, $D = 5$, $R = 3$, and $S = 2$.

Notice local taxes are highest in the lowest-tax state and that the tax gradient is steepest for the largest discontinuities. For the three state, nine town model, the solution can be characterized as follows. If $x < x^*$, then no Nash equilibrium exists in pure strategies. If $x > x^*$, the Nash equilibrium will have $t_A^H > t_{BB}^H$, $t_{BB}^M > t_A^M > t_{BB}^M$, and $t_B^L > t_A^L$ with the relationship between t_A^H, t_B^H and t_A^L, t_B^L determined by the relative sizes of the notches given above.

Consider the case of towns at the High-Low border when $D > 0$ and x is sufficiently large. The border town on the high-tax side can smooth the tax differential at the state border to reduce the number of cross-border shoppers while also attracting shoppers from interior towns. Any deviation that lowers the rates by Town A will result in a loss of revenue as the losses from lowering the tax rate will outweigh the gains from expanding the tax base. On the low-tax side, the border town will have a relatively high local sales tax rate. Here, the town seeks to export the tax burden to residents of the neighboring state who are already shopping within its borders. If the interior town deviates from the Nash equilibrium and increases the rate, the benefits of a higher tax rate will not be sufficient to outweigh the losses from contracting the tax base.

For a state that has a high- and a low-tax neighbor at its borders, local tax rates should always decline with the distance from the high-tax state and should increase with the distance from the low-tax border. These results are mutually consistent. Therefore, the gradients of taxes in State M are always declining as one moves away from the high-tax border. Of course, the size of the notches on the high- and low-tax side may affect the precise slope of the gradient, but they will not affect the general pattern (sign) of the gradient.

In this model, which town is critical depends on the magnitude of the differential. If the notch between the high- and low-tax state is especially large, then, additional towns past the center of the state will be pulled into its gradient. By definition, notch

D must be larger than notch R so that notch R can never pull additional towns into its gradient. A similar logic applies to the low-tax state as well. Another way of thinking of this problem is to imagine that $R = 0$, so that State H and State M set the same state tax rate. If this is the case, it is as if the high-tax states now have six towns. The local tax rates will be increasing away from the border. The problem becomes symmetric with respect to the two borders with non-zero differentials so that the tax rates will increase all the way to the state border between State H and M. As towns get further and further away from the larger notch, they are less likely to be pulled out of the gradient for the smaller notch closest to them. More generally, State H is bordered by two low-tax states. Starting from State L, there is upward pressure on taxes. Starting from State M, the pressure on the towns moves is downward. These two gradients are mutually inconsistent and they need to change at some point along the circumference.

These results differ from what comes out of Kanbur and Keen (1993). In Kanbur and Keen (1993), the "small" country, as defined by domestic population, always undercuts the large country in a Nash equilibrium. In the model I present above, towns set taxes following an inverse elasticity rule,[15] but what matters is the relative size of the "foreign" plus "domestic" market. For Town BB in State H, even if its local tax rate is zero, some residents will always shop abroad. Therefore, starting from a position where $t^H_{BB} = t^H_A > 0$, Town BB perceives a relatively small (in comparison to Town A) market of "foreign" plus "domestic" shoppers (it already has some of its residents shopping in the neighboring state, which reduces its market size). Town BB perceives the relatively larger elasticity (because its market is smaller) and undercuts Town A. On the low-tax side, starting from $t^L_B = t^L_A > 0$, Town B is already attracting residents from the neighboring states. Therefore, because of the

[15] Put differently, one needs to set the tax rate inversely proportional to the elasticity of demand to maximize revenue from a tax on sales. Here, the elasticity of demand is derived from cross-border shopping. When the elasticity of cross-border shopping is high, the town has an incentive to undercut the other town.

fact that B has already attracted some foreign residents, it perceives the "foreign" plus "domestic" market as larger than that of A. Town A perceives itself as small and undercuts Town B. Thus, the town with the the largest cumulative "foreign" plus "domestic" market will always set higher rates.

The theoretical model discussed above has several testable hypotheses for the empirical analysis to follow. First, municipal tax rates are lower on the high-tax side of the state border than on the low-tax side of the border. Second, local tax rates are decreasing away from the nearest high-tax border and are increasing away from the nearest low-tax border. Third, the distance of the municipality from the state border will determine how steep the tax gradient is. The treatment effect will be heterogeneous by distance, because the tax gradient will be steepest for the towns closest to the state border. Fourth, the size of the notch will determine how steep the tax gradient is. For larger discontinuities in state tax rates, the tax gradient will become steepest. This suggests that the treatment must be continuous (as a function of the size of the notch) rather than a simple binary treatment for high- or low-tax states. In addition, the theory implies that the critical town will depend on the relative size of the discontinuities at various borders. This suggests that it is important to account for the presence of multiple borders in the regression equations. Finally, the existence of a tax gradient may be tested in other contexts; the theory will also apply to county borders.

2.4 Data and Evidence on the Levels of Local Tax Rates

The goal of the analysis is twofold. One, do localities on the high-tax side of a border set lower local taxes than localities on the low-tax side? Two, is there a tax gradient that is a function of distance from the border?

2.4.1 Data

The data on tax rates come from Pro Sales Tax's national database.[16] The data contain state, county, municipal, and sub-municipal (district) tax rates for April 2010. Table 2.1 displays summary statistics for these tax rates by state. Because I am interested in combining the tax data with a measure of distance from the state border and with Census data, I restrict the sample to municipalities that are identified Census Places, which subsequently will be referred to as localities.[17] To do this, I merge Geo-coded data provided by the 2009 American Community Survey and the 2000 Census[18] SF3 file to the tax data set.[19]

A key variable in the analysis is a locality's distance from the nearest neighboring state. I draw on and substantially modify the method of Lovenheim (2008) to calculate distance from the border. Using 2000 Census geographic Tiger Line files and Arc-GIS software, I estimate the driving distance from each population weighted centroid of a locality to the closest intersection of a major road and a state border crossing. Unlike in Lovenheim (2008), I calculate distance from the nearest border rather than from the nearest low-price border. I also calculate distance from the population weighted centroid instead of the population weighted distance, which enables me to calculate driving distance

It is essential to have the most accurate measure of distance, because this paper

[16]The data is proprietary data, but was provided to me without charge. For a complete description of the data, see http://www.prosalestax.com/.

[17]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 relationship between a Census Place and governing authority is often different across states. Census Places contain some locations that may not have legal authority or jurisdiction to set sales taxes.

[18]SF3 files for the 2010 Census are not yet available.

[19]Merging the data requires name matching, which can introduce some error. However, the error of incorrectly matching a name is likely to be small because I name match the Places based on state, county, and locality names, where all three must match. Census Places may cross county lines and Places are matched to Counties using a procedure outlined in the appendix. Because of spelling errors, etc., it is possible that some matches that would be correct matches remain unmatched. I hand match these. It should also be noted that some jurisdictions are in one data set but not the other.

analyzes taxes away from the border. I estimate the driving[20] distance from the population center of each locality to the nearest intersection between a state border and a major road that minimizes travel time.[21] Relative to the "as the crow-flies" distance, this measure of distance more accurately captures the true commuting time from the nearest state border. Although the "crow-flies distance" is correlated with driving distance, it is not a very accurate measure of true commuting costs except in a local region of the border.

For a detailed description of how I calculate the distance from the border, please see the Appendix. Figure 2.3 shows that the distance calculation is accurate at the county level and gives information with regard to the range of distances.

Finally, Tribal Nations are treated as localities. Although the tax treatment of sales to non-tribal members often varies by state, most courts have ruled that sales to non-tribal members require tax collections.[22] I treat these reservations as being similar to localities within the states and thus do not consider a border with a tribal nation as being a state border. International borders are considered state borders even though crossing the border is more difficult and may restrict cross-border shopping. Canada assesses a 5% Goods and Services Tax (GST) but many provinces assess an additional provincial tax resulting in an implicit tax rate between 10 and 15.5%, depending on the province.[23] The Mexican Value Added Tax at the United States

[20]In some specifications, I derive the "as the crow-flies" distance. I do this in specifications where I account for the two-dimensional nature of borders by calculating the distance from the second closest border. Calculating driving distance from all neighboring state borders is feasible, but computationally time intensive. The crow-flies distance is a measure of the shortest distance between two points. It is not a perfect measure of the actual transportation distance because it does not account for the road system in place to move between two points nor does it make any adjustments to account for geographic impediments such as mountains or rivers.

[21]A major road is a Census classification including most non-residential roads. As pointed out by Lovenheim (2008), the exclusion of residential roads is "trivial because the vast majority of interstate travel does not occur on such roads" and it is unlikely that retail locations are on minor (residential) roads.

[22]This is the opposite of court rulings on excise taxes, where courts have ruled that tribal nations need not collect state excise taxes under most circumstances. For a discussion of tribal regulations see "Piecing Together the State-Tribal Tax Puzzle" by the National Conference of State Legislatures.

[23]One exception is Alberta, which assesses no provincial tax so that the total GST tax is 5%. Alberta shares a border with Montana, which has no state or local sales taxes. The territories are

border is 11%, which is higher than the state sales tax rate along any border state.[24]

2.4.2 Graphical Analysis and Summary Statistics

First, I present some graphical results in Figures 2.4, 2.5, and 2.6. Figure 2.4 depicts the state sales tax and the range of tax differentials at borders. Figure 2.5 presents the distribution of the county tax rate plus the average municipal tax rate with the county. I demonstrate the role of municipal taxes in the case of one state, Missouri, in Figure 2.6. Although it is difficult to discern an immediate border effect, some examples seem to be present. Figure 2.6 shows that urban areas set higher tax rates regardless of their proximity to neighboring states, implying the need to control for this factor. In addition, similar tax rates are clustered in particular regions. Figure 2.5 indicates that the largest amount of variance in county tax rates is in the central and southern states. Western states have some variance in their county tax rates, but counties are also significantly larger.

Table 4.1 presents summary statistics of all the variables used in this analysis. Statistics are presented for town (Census Place) level data. Many of the control variables appear to have relatively similar means on both the high- and low-tax sides of the border. The average tax differential at state borders is about 1.90 percentage points and the average local plus district tax is 0.71 percentage points.

Some average summary statistics can help analyze the pattern of tax rates. On each side of a state border, jurisdictions are sorted into towns that are 0-25 miles from the border, 25-50 miles from the border, 50-75 miles from the border, etc. I calculate the mean local and state tax rate in each of these bins. Table 2.3 presents the unweighted averages. Some convergence of average tax rates near the border indicates smoothing of the tax rates.

another exception, which face the federal rate only.

[24]Mexican goods and services are taxed at 16%, but within twenty kilometers of the border with the U.S., the preferential tax rate is 11%.

2.4.3 Methodology: The Tax Level Effect

How much higher are local tax rates on the low-tax side of the border relative to the high-tax side of the border? Recall that the theoretical model has one discontinuity in state tax rates of $D = \tau^H - \tau^L$. After localities assess local option taxes, the tax differential at the border becomes $\frac{30}{53}D$. This implies that local tax rates on the low-tax side of the border are higher than on the high-tax side of the border. To test this hypothesis, I conduct a regression discontinuity (RD) design following the local linear regression and bandwidth selection methodology of Imbens and Kalyanaraman (Forthcoming).[25] The results of this RD design can be interpreted as the effect of the border between states with different tax rates on the level of local option tax rates. The treatment is being on the border of a high- or low-tax state and is binary. Because many policies vary at state borders, the interpretation of the results would be causal only if there are no other state policies that are discontinuous at borders correlated with being a high- or low-tax state.

To implement the regression discontinuity design, I conduct local linear regression on the local tax rates with a triangle kernel. I include the same vector of controls and state fixed effects in the summary statistics table and allow them to vary locally. The bandwidth is selected optimally according to Imbens and Kalyanaraman (Forthcoming) and towns closest to a same-tax state border are dropped.

2.4.4 Result: Are Local Taxes Higher on One Side of the Border?

Figure 2.7 graphically presents the results of the regression discontinuity design without controls. Table 2.4 shows the regression results for various specifications. Specifications 1, 2, 3 and 4 present the results for local, local and district, county, and

[25] Spatial analysis using regression discontinuities at borders has been used extensively in models such as Holmes (1998) and Gopinath, Gourinchas and Hsieh (2011). A spatial fixed effects regression discontinuity model is used by Magruder (2011) to try to account for spatial proximity across localities.

total local rates at the state border. Specification 4 is most preferred; it demonstrates how much the tax differential at state borders falls after all local option taxes are assessed. Keeping in mind that many things change at state borders, columns 5 and 6 present the results at county borders that are not also state borders. Columns five and six are more likely to represent a causal effect, given that fewer policies vary discontinuously at county borders than at state borders.

Recall that the average tax differential at state borders is about 1.90 percentage points. The results from the RD design of the total local rate yield an estimate of -1.13: a town (located precisely near the border) on the high-tax side assesses a local sales tax rate that is 1.13 percentage points lower than an identical border town on the low-tax side. Thus, the average tax differential falls from 1.90 percentage points to 0.87 percentage points. This result is consistent with the theoretical prediction that local sales taxes reduce the state tax differential approximately by one-half. The results for local plus district tax rates are smaller, as expected, and imply that municipal taxes are about 0.352 percentage points lower on the high-tax side. With regard to county borders, the results remain negative suggesting that the discontinuity in county tax rates imply local taxes are higher on the low-tax side of the county. The results at county borders are smaller because tax differentials at county borders are much smaller than at state borders.

The RD results suggest that as the model developed in Section 3 predicts, local sales taxes reduce tax differentials at state borders.

2.5 Estimating the Tax Gradient

The previous section has shown that local taxes are lower on the high-tax side of the border than on the low-tax side of the border. What remains to be seen – and is the subject of this section – is if this reduction occurs gradually with distance from the border.

2.5.1 Methodology: The Tax Gradient Effect

My estimation strategy exploits the discontinuity in the state tax rates at borders to estimate how local tax rates depend on distance from the border. I use a global polynomial regression rather than the local linear regression above. A global polynomial regression is preferred to local linear regression because the treatment will vary with the size of the discontinuity. It is also preferred because I care about precisely estimating the marginal effects (the slope of the tax gradient) conditional on the size of the differential and on distance.

I will allow distance to have a different effect on local tax rates depending on the side of the border on which a town is located because the theory implies the geographic pattern is different on the low- and high-tax sides of the borders. The "treatment" is defined as the size of the difference in state tax rates at the border. As such, the state border will be allowed to have a different effect on local tax rates depending on the size of the notch at the border and depending on a locality's distance.

Let l index localities, c index counties and s index states. Then, X_{lc} denotes observable characteristics of locality l in county c. The matrix X includes variables from Census 2000 data at the Census Place level plus some other control variables generated using geographic files such as population, the median level of income, the fraction of the population with a college education, the fraction of seniors, the fraction of residents working in another state and county, the number of neighbors, town area and perimeter. To this matrix, I also add the vote share received by Obama in the 2008 Presidential election in case political affiliations determine a locality's choice of whether to use the sales tax, a dummy for proximity to international borders, and a dummy for proximity to oceans or water.

The variable t_{lc} will denote the local (town) plus sub-district tax rates.[26] Define

[26]To proceed, I must assume that local tax rates are in equilibrium in my data. Given the large number of observations that I have, this is a realistic assumption.

R_{lc} as the difference between the state tax rate of the high-tax state and the state tax rate of the low-tax state. For locality l, if the nearest neighbor is a relatively high-tax state, this will be a positive number. If the neighbor is a relatively low-tax state, the difference will be a negative number. Also, H_{lc} is a dummy variable that denotes whether locality l's state is a high-tax state relative to the nearest neighboring state of jurisdiction l. And S_{lc} is a dummy variable that is equal to one when locality l's state has the same state tax rate as its neighboring state.

Define the distance from a locality to the nearest state border as d_{lc} and note it is always positive, so that towns located fifty miles on either side of the border are identical with respect to distance. The role that distance plays may be non-linear and differ depending on the side of the border.[27] To do analysis on the tax gradient, I need to assume that the relationship between d and local taxes is sufficiently flexible, and I allow it to be a polynomial function of degree p. Denote this function in matrix form as $G(d_l)$, where each column represents a higher order term of the polynomial function $\sum_{k=1}^{p}(d_{lc})^k$. Note that the locality cannot manipulate d_l, so that it cannot select whether it is on the high- or low-tax side.

The reduced form equation designed to test whether the difference in state tax rates at the border influences local tax rates has the following specification:

$$\begin{aligned} t_{lc} = \beta_0 + \beta_1 H_{lc} + \beta_2 S_{lc} + \beta_3 R_{lc} + \beta_4 R_{lc} H_{lc} + X_{lc}\varphi + \\ G(d_{lc})\rho + R_{lc}G(d_{lc})\gamma + H_{lc}G(d_{lc})\delta + R_{lc}H_{lc}G(d_{lc})\alpha + \\ S_{lc}G(d_{lc})\theta + \lambda \mathrm{T}_{lc} + \zeta_s + \varepsilon_{lc}. \end{aligned} \tag{2.5}$$

In expression 2.5, ε_l are unobservable characteristics that are specific to a locality. ζ_s are state fixed effects and T_{lc} is the county tax rate for locality l. The variables γ, ρ, δ, α, and θ are the vectors of coefficients where each element of the vector denotes

[27] Identifying the slope of the tax gradient will rely on a functional form dependent method. Lovenheim and Slemrod (2010) use dummy variables based on distance to obtain a flexible function of distance. Lovenheim (2008) and Harding, Leibtag and Lovenheim (Forthcoming) impose $log(D)$ as the functional form because they do not have sufficient power to use a more flexible polynomial.

the coefficient on each term in the polynomial – linear, square, cubic, etc.

A standard RD would simply define the treatment as H_{lc}, as done in in Section 4.3. The above specification allows the treatment to also vary by the size of the state tax rate discontinuity R_{lc} and estimates separate distance functions for high-, low-, and same-tax state borders. The coefficient on $G(d_{lc})$ will pick up the average effect of distance on local tax rates – or put differently, the effect of distance from the border. The interaction of the distance term with R_{lc} in $R_{lc}G(d_l)$ allows the treatment to vary with the size of the discontinuity. Because R_{lc} is negative for towns in relatively low-tax states and positive for towns in relatively high-tax states, the treatment is different on opposite sides of the border, as the theory suggests. The interaction $H_{lc}G(d_{lc})$ will allow for the effect of being at a particular distance from the border on the high-tax side to be different than being at that same distance on the low-tax side. The coefficient on $R_{lc}H_{lc}G(d_{lc})$ allows for the treatment effect to vary depending on the side of the border. Lastly, $S_{lc}G(d_{lc})$ allows the estimated distance function to be different for towns where the nearest border has the same state tax rate on both sides. The term, $\beta_1 H_{lc} + \beta_2 S_{lc} + \beta_3 R_{lc} + \beta_4 R_{lc} H_{lc}$, represents the intercept of the polynomial function of distance. It captures the effect of the notch when $d = 0$ and allows for the polynomial to have a different intercept in relatively high- or low-tax states.[28]

The inclusion of state fixed effects controls for commonalities occurring within a state but not across borders and helps mitigate any geo-spatial correlation among towns in the same state. With state fixed effects, identification is coming from within state variation of local tax rates. The state dummies will help to control for variation in policies and unobservables that is constant within states, such as political climates or state business policies. Critically, the inclusion of the state fixed effects will also

[28]Note that, $\beta_1 H_{lc} + \beta_2 S_{lc}$ implies that when R_{lc} goes to zero, local taxes on the high-tax side may be different in level terms than the low-tax side. The theoretical model implies this is not possible. I include these terms in case there are unobservables correlated with the side of the border. However, imposing that the coefficients on these terms is zero does not change the findings.

control for the level of the state tax rate in state s.

Finally, Equation 2.5 uses town tax rates as the left-side variable. However, towns are embedded in counties, which are within states. While the state fixed effects control for the level of the state tax rate, they do not control for the level of the county tax rate. I cannot directly include the county tax rate, T_{lc}, on the right hand side because it is likely selected simultaneously with local rates and therefore is endogenous. As a result, I instrument for the county tax rate using a standard instrument in the tax competition literature – a subset of the X's at the county level. The justification of this set of instruments can be found in Brueckner (2003). Brülhart and Jametti (2006) apply the instrument to higher levels of government as well. Instead of using the entire subset of the X's as instruments, I will only use geographic variables (which are not often controls in previous studies) as instruments because demographic variables are likely to be correlated with other factors at the municipal level. Particularly, I use the county area, county perimeter, and county number of neighbors as instruments.

To justify these instruments, recall that the regression equation controls for town area, town perimeter, and the number of neighbors for the town. Then, the exclusion restriction requires that the instrument should have no partial effect on local taxes after controlling for these variables. The direct impact of the number of neighbors, area and perimeter of the county on local taxes is likely to be zero. County neighbors, area and perimeter affect the county's tax rates, but will have no direct impact on the locality's tax rate so long as there are multiple jurisdictions within a county and so long as counties are sufficiently large in size. County borders were likely to be historically drawn on latitudes and longitudes or broader geographic features. The number of neighbors, area, and perimeter depend on a county's characteristics such as whether along a body of water, broader geographic features, and how counties were divided historically. Because area and perimeter are historically drawn, the evolution

of time with these variables helps to make them exogenous. As a contrary point, the town's area, number of neighbors 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.[29] Furthermore, as an alternative to the instrumental variable approach, I present additional results where the dependent variable in the estimating equation is equal to the town plus county tax rate. Because county rates are now part of the variable of interest in these specifications, I avoid having to instrument for the county tax rate and can instead look for a gradient in the total effective tax rate. The sign and significance of the tax gradient remain the same.[30] The F-statistic for instrument strength from the first stage is 30.20 ($p = .00$). I can now proceed using the regression analysis in Equation (2.5) to estimate the effect within a state.

The coefficients within γ , ρ, δ, and α are not informative as stand-alone parameters because the marginal effect of distance is a non-linear function of d_{lc}. In the case of a p order polynomial, the marginal effects of distance on the local tax rate for the high-, low-, and same-tax side of the border are given by Equation 2.6:

$$\frac{\partial t_{lc}}{\partial d_{lc}} \equiv G'(d_l) = \begin{cases} \sum_{k=1}^{p} k[\rho_k + \delta_k + (\gamma_k + \alpha_k)R_{lc}](d_{lc})^{k-1} & \text{for } H_{lc} = 1 \,\&\, S_{lc} = 0 \\ \sum_{k=1}^{p} k[\rho_k + \gamma_k R_{lc}](d_{lc})^{k-1} & \text{for } H_{lc} = 0 \,\&\, S_{lc} = 0 \\ \sum_{k=1}^{p} k[\rho_k + \theta_k](d_{lc})^{k-1} & \text{for } H_{lc} = 0 \,\&\, S_{lc} = 1, \end{cases} \tag{2.6}$$

where the coefficients indexed by k indicate the coefficient on the kth order term of the polynomial.

[29] I conduct a Hansen J test of over-identification. The result of this test – $J = 2.28$ ($p = .32$), a failure to reject the null hypothesis that all the instruments are uncorrelated with the error – is suggestive that in the presence of one valid instrument, the other instrument is also valid.

[30] The results of these regressions are summarized in the appendix. This is not, however, the preferred specification because it will leave me unable to compare the tax gradient at state borders to the tax gradient at county borders – where the left side variable must be the local tax rate.

From this expression, I can calculate the mean marginal effect (or mean derivative) of distance on tax rates by calculating the sample mean of the estimated derivative conditional on being in a high-, low-, or same-tax state relative to the neighbor. Letting N_b denote the number of observations on a particular side of a border, the mean marginal effects represent the slope of the tax gradient away from the discontinuity after averaging across the conditional sample:

$$E\left[\frac{\partial t_{lc}}{\partial d_{lc}}\right] = \frac{1}{N_b}\sum_{l=1}^{N_b} G'(d_l). \tag{2.7}$$

The mean derivative of the full sample is of little interest given that the theory indicates it will be of opposite sign on different sides of the border. Therefore, the summation in Equation 4.18 is separately taken over all towns on the high-tax side, low-tax side, and same-tax sides of the border and N_b adjusts appropriately. Standard errors for mean derivatives are calculated using the Delta Method. I also calculate, but do not report, the marginal effect evaluated at the mean, or $G'(\overline{d}_{lc})$. The derivative evaluated at the mean is biased and inconsistent because of non-linearity in the derivative. The marginal effect given by Equation (4.18) is a consistent estimate of the mean derivative in the conditional population and is the preferred effect.

If the mean marginal effect is positive, tax rates are increasing as the distance from the nearest border increases. If the effect is negative, then towns further from the border are setting lower tax rates than those at the border. In general, because the marginal effect need not be identical for all values of d_{lc} or R_{lc}, I will also report the mean marginal effects evaluated at different possible values of d_{lc} and R_{lc}. The theory predicts that the mean marginal effect on the high-tax side of the border will be positive and on the low-tax side of the border will be negative. When state tax rates are the same, no effect should be evident.

The polynomial order is selected using "leave-one-out" cross-validation.[31] The

[31]The process of cross-validation estimates equation 2.5, omitting one observation from the data

implied root mean squared error from leave-one-out cross-validation is decreasing in the order of the polynomial until it reaches a minimum of .6625 at a polynomial order of five; thus, the quintic polynomial is the preferred specification. In addition, I ocularly compare the fit of the predicted values from a global polynomial regression to the calculations using the local linear regression described above. Figure 2.8 compares the fit of a quintic polynomial to the results of the local linear regression. The quintic polynomial fits the local linear regression almost identically – both with respect to the curvature and the levels of the tax rates. Conducting the same comparison for lower degree polynomials is less accurate.

2.5.2 Results: How Steep Is the Tax Gradient?

Table 2.5 presents the full set of coefficients on several specifications of Equation 2.5 along with the mean derivatives associated with them, scaled to represent a 100 mile change. Specification 1 includes only a quintic polynomial in distance. Column 2 adds local controls and column 3 adds state fixed effects. Column 4 instruments for the county tax rate, which is the preferred specification.

The coefficients on some control variables are worth highlighting. Having a higher fraction of people who work in-county and in-state implies the tax base is less mobile across state lines, and results in higher local tax rates. Higher income jurisdictions have lower tax rates, suggesting that they may be more able to use the property tax as an alternative revenue source. Towns near oceans have higher tax rates, consistent with a model located along a line-segment where the jurisdictions at the end of the line set higher rates. On average, jurisdictions near international borders set lower tax rates even though the international borders are all high-tax borders. This suggests

set. The process is repeated 16,781 times, omitting each observation from the data set once and calculating the root mean squared error each time. I then calculate the average root mean squared error for all the omitted observations. After repeating the cross-validation techniques for polynomials up to an order of eight, I select the polynomial that yields the smallest average root mean squared error.

that towns near the international border see no gain to exporting the tax to foreign consumers by raising the rates, but do not view the border as closed. The number of neighbors has a positive sign. Finally, the coefficient on the instrumented county tax rate is negative and significant – suggesting that higher county taxes reduce local sales taxes.[32]

In the preferred specification, the mean derivative on the low-tax side of the state border is significant and of the expected sign: -0.102. Thus, moving a town 100 miles away from its high-tax state neighbor – assuming constant marginal effects – will decrease its local tax rate by just over one tenth of a percentage point. Alternatively said, a one mile change has a -.102 basis point change on local rates. Interpreting this in the context of the average local tax rate, which is approximately 0.71 percent, implies that average local taxes are about 15% lower 100 miles away from the border. The gradient for towns with a same-tax state neighbor is insignificant. For a town on the high-tax side of the border, the marginal effects are slightly larger in absolute value (-.197) but are unexpectedly negative.

How robust are the results to reasonable alternative specifications? With regard to the marginal effects, Table 2.6 presents the marginal effects for a variety of specifications, with column 1 denoting the baseline results. In general, significantly negative results on the low-tax and high-tax side of the border are found. One specification eliminates towns near international borders, in case these borders are effectively closed to purchases. Another specification eliminates towns that are most proximate to the ocean.[33] One reason for doing this is that if the theoretical model were on a line segment instead of a circle, towns at the end of the line segment have incentive to raise their tax rates because they only have one neighbor. Column 3 indicates international

[32]Such a finding – the relative magnitude of the vertical and horizontal externalities in the presence of multiple federations (counties) – is interesting in its own right, and is the subject of Agrawal (2011*b*).

[33]As an alternative, I interact the dummy variables for proximity to international borders and water with the distance functions and find similar results.

borders may be different than state borders. Because the gradient becomes slightly steeper when excluding these jurisdictions (always located on the low-tax side), it suggests that towns near international borders are less likely to be able to charge a mark-up over their interior neighbors – perhaps because of differences in the pre-tax price. Column 4 also seems to indicate that the towns near the ocean set higher rates than their interior neighbors in low-tax states. Because the gradient becomes steeper when excluding these towns, this suggests that towns near the ocean and away from the border are setting higher rates than border towns, which is consistent with a Hotelling style model where towns at the end of the line segment set higher rates.

Specification 5 allows the reader to determine the marginal effect in hours instead of miles. The marginal effects with a quintic in driving time indicates that a one-hour drive from the border lowers tax rates .06 percentage points on the low side and by .10 percentage points on the high side. Comparing this to specification 1 indicates that the marginal effect of driving 100 miles is approximately equal to the marginal effect of a driving time of two hours. Columns 7 and 8 are important to decompose the results when neighboring states do not assess local option sales taxes (LOST). Recall that 16 states do not allow LOST. It may be expected that the nature of the competition at these borders is different than at a border where both states allow LOST. When the neighbor does not allow LOST, the slope of the gradient becomes much steeper on the low-tax side. Additionally, on the high-tax side, the slope of the gradient becomes positive – and matches the predicted results of the theory. This suggests that not having a local competitor on the opposite side of the border makes it more likely that localities will reduce discontinuities in state tax rates. In addition to those particular extensions, the decreasing tax gradient for low-tax states is highly robust.[34]

[34]Column 2 indicates the results are robust to a binary treatment. The results are also robust to excluding some states where local taxes are not extensively utilized (column 6), to using local taxes only as the dependent variable (column 9), and to using county plus local plus district taxes as the dependent variable, thus avoiding the need to instrument for county tax rates (column 10).

Table 2.7 shows that the results are robust to the order of the polynomial (columns 2 and 3). Columns 4 and 5 suggest that giving all states equal weight in the data and weighting by population preserves the expected results. The results weighted by population increase the steepness of the slope dramatically, suggesting that high population jurisdictions may be more likely to be located near the border, where the gradient is steepest. Specification 6 interacts all the state fixed effects and all the control variables with a polynomial in distance and allows the coefficients on these interactions to vary on the high-, low- and same-tax side of the border. Such a specification removes any possibility of coincidental correlation with the distance function and allows each state to have its own tax gradient. The results increase in magnitude and in significance. Column 8 includes state fixed effects and state-border-pair fixed effects, such that all the identifying variation comes from variation in local taxes within only one border region in a state. Such a specification controls for the level of all state controls and controls for the possibility that the east border of a state may be very different from the west border (i.e., the presence of mountains or rivers, etc.). The last three columns of Table 2.7 restrict the sample to towns within 150, 98, and 40 miles of the border. The restriction of 98 miles is the optimal bandwidth of the local linear regressions above. In these specifications I re-select the order of the polynomial for each restricted sample by the cross-validation method above. Notice when the sample is restricted to a local region around the border, the tax gradient becomes steepest and the gradient on the high-tax side becomes insignificant. From Table 2.7, it can be seen that as the estimating area becomes larger, the steepness of the tax gradient becomes smaller. This suggests that the mean derivatives are strongest in a local region of the border – where cross-border shopping is likely to be most salient.

Many of the coefficients for the high- and low-tax side are similar in magnitude and sign. Because of the similar estimates on both sides of the border, I must consider

the possibility that the coefficients identify something fundamental about borders, but do not identify the effect of tax discontinuities. While such a border effect cannot be dismissed with certainty, a variety of robustness checks suggest that the results are not driven by a pure border effect. I will show no significant effect at same-tax state borders. In the following sections, I will show that the effect varies with the size of the discontinuity and for both county and state borders. Thus, if the results are not driven by tax differentials at state borders, they must be driven by a policy that changes discontinuously at the border and is correlated with the tax differential and with distance at both state and county borders. In addition, the tax gradient would be the same on both sides of the border, if a pure border effect were operating. I can reject the null hypothesis that the slopes on the high- and the low-tax side of the border are identical in the baseline specification ($t = 3.90$ with a $p = .048$).

Previously, I have estimated a gradient effect that is common to all states. Table 2.8 displays the mean derivatives by border-type in every state that allows for LOST and highlights substantial variation in the gradients. Out of the twenty-eight states that have a high-tax neighbor and allow for LOST, twenty states have a negative gradient consistent with the theory. Out of the eight states with positive gradients, only three states – Colorado, Oklahoma, and West Virginia – have statistically significant gradients that imply local taxes increase away from the border. The negative gradient is steepest in Louisiana and Arkansas. Of the twenty-two states with a low-tax neighbor, eight states have positive gradients consistent with the theory. Of these states, only North Dakota and Oklahoma have statistically significant gradients that imply taxes increase away from the nearest low-tax neighbor.

2.5.3 Is the Tax Gradient Steepest Near the Border?

Consumers who live further away from the border have less incentive to cross state lines to avoid paying the sales tax, because of transportation costs. As a result,

jurisdictions in a local region of the border will adjust tax rates the most rapidly, either to attract cross-state shoppers to the interior of the state or to undercut their higher tax local neighbors. In fact, in the n town model in the Appendix, the tax gradient is steepest near the border. Using the unrestricted sample of the fixed effect IV model from Table 2.5, I evaluate the marginal effects conditional on one-mile intervals. With the quintic polynomial, Figure 2.9 indicates the marginal effects are steepest within a thirty mile radius of the neighboring high-tax state border. In this region, the marginal effects are almost five times as large as the average marginal effect for the full sample. These results are significant at the 95 percent confidence level. In a local region of the border on the low-tax side, towns rapidly adjust local tax rates downward the closer to the border. On the high-tax side of the border, the marginal effects are approximately zero for all distances, except in a vicinity of the border.

As a falsification test, it is useful to compare these results to towns where the neighboring state has the same state tax rate. Figure 2.10 plots the marginal effects conditional on various distances for these towns. As expected, if the state tax rates are uniform, the local tax rates are uniform but highly volatile, because of the small number of such towns in the sample.

Figure 4.2 presents the marginal effects conditional on distance, when the optimal bandwidth is selected to restrict the sample to towns within 98 miles of the border. I expect the marginal effects to increase in absolute value (note the different scale of the vertical axis), because the sample no longer contains towns at far distances, where the tax gradient is most likely to be flattest. Again, the figure indicates that, for low-tax states, the marginal effects are steepest in the vicinity of the border. On the high-tax side, in a very local region of the border, taxes increase with distance from the border in a manner consistent with the theory.

The tax gradient may be heterogeneous not only in distance from the border, but

also in the size of the tax discontinuity. Consumers may decide to cross-border shop both because they live closer to the border and because of a large difference in tax rates.

2.5.4 Is the Tax Gradient Steepest for Large Tax Differentials?

Figure 2.12 presents changes in the marginal effects conditional on both distance and the size of the discontinuity in tax rates. Standard errors are omitted for simplicity, but the standard errors unconditional on distance can be read from the columns of Table 2.9. The theory predicts that the tax gradient should be steepest for larger state tax rate differences. Figure 2.12 indicates that the theory is correct, especially in a local region of the border. In the region of zero to forty miles from the border, on the low-tax side, an increase in the size of the discontinuity from two to six percentage points increases the marginal effects by a factor of 1.5. On the high-tax side, the tax gradient slopes shift upward, implying that larger discontinuities create more incentives for localities to act as the theory predicts. Such a pattern begins to suggest the causal nature of the estimates; the existence of a discontinuous omitted variable at state borders that is also correlated with the difference in state tax rates and distance is less likely.

Figure 2.13 displays the predicted values of the regression equation, and thus converts the marginal effects into level terms conditional on the average size of the discontinuity and on various tax differentials. The level graphs have slopes that are consistent with the marginal effects above. The level of local option taxes on the low-tax side of the border shifts up as the discontinuity increases in absolute value. On the high-tax side of the border, the level of local taxes shifts down. In other words, as state tax rates increase, local tax rates decrease holding fixed the neighboring tax rate. This confirms the results of the RD design regarding the level of tax rates – and generalizes the results to a treatment that is heterogeneous.

Looking at Table 2.9, the average marginal effect of distance is almost perfectly monotonic in the size of the discontinuity. A significant decreasing gradient only emerges when the discontinuity is greater than 1.25 percentage points for towns in a low-tax state. Towns behave as if the neighboring state has assessed the same tax rate, when the discontinuity in state tax rates is small. On the high-tax side, the tax gradient is significant (but of the unexpected sign) over most relevant ranges of the discontinuity. However, an insignificant gradient begins to emerge if the size of the discontinuity is greater than 5 percentage points.

2.5.5 Extension: Second Closest State Borders and County Borders

Up until now, the empirical specifications have assumed a one-dimensional response: towns only respond to one border and one level of higher government. The existence of critical towns in the theory suggests that the process is actually multi-dimensional – and that towns respond based on their proximity to multiple borders. My theory does not account for a tax gradient near county borders, because I do not consider multiple levels of government.

To address the concern that towns may be responding to two state borders, I calculate the distance from the second closest intersection of a major road and a state border.[35] For computational feasibility, I use the "as the crow-flies" distance instead of driving distance. Column 1 of Table 2.10 is analogous to the baseline specification in Equation 2.5. The only difference is the use of the "crow-flies" distance instead of driving distance. Marginal effects are nearly identical, but are not always significant. In column 2, I add a polynomial in distance from the second closest border along with its interaction with the size of the difference in state tax rates at that border and dummies H and S. After controlling for multiple-state borders and their relevant discontinuities, the tax gradient increases in absolute value.

[35]The closest border intersection is always with respect to a contiguous state by definition. However, the second closest state border is often, but not always, a contiguous state.

The second concern is that towns can both reduce the tax differential at state borders through local option taxes, and reduce the tax differential at county borders through local sales taxes. To account for this, I calculate the driving distance from every population weighted centroid to the nearest intersection of a major road and a county border. I then regress town and district taxes (without county taxes) on a polynomial in distance from the county border, plus controls and interactions. Column 3 shows the marginal effects of distance from the state border while controlling for the second state border and the nearest county border. The results in column 3 remain unbiased estimates of the marginal effects of the distance from the state border because of the inclusion of state fixed effects. Note that the sign of the gradient on the low-tax side of the border flips, and becomes insignificant.

In columns 4 and 5, I present the marginal effects of towns with respect to county borders – as discontinuities at county borders are equivalent in spirit to discontinuities at state borders. The results are of the same sign but are much larger than the state effects. Intuitively, this arises because counties are smaller than states and localities have pressure to adjust their tax rates much more rapidly over shorter distances. Due to the exclusion of county fixed effects, the presence of institutional differences that vary systematically across counties within a state will bias the estimates. These results suggest that the addition of multiple levels of government to the model would not change the interpretation of the results, because accounting for multiple borders does not qualitatively alter my findings.

2.5.6 Extension: Who Adopts Local Taxes Within a State?

The emergence of the tax gradient in the data may be a result of towns always setting positive tax rates conditional on distance from the border. But, the tax gradient will also appear if towns near the border are either more or less likely to assess a non-zero tax rate. To test if distance is a determinant to setting a non-zero

tax rate, I run the following binary choice probit regression:

$$\begin{aligned} z_{lc} = \beta_0 + \beta_1 H_{lc} + \beta_2 S_{lc} + \beta_3 R_{lc} + \beta_4 R_{lc} H_{lc} + X_{lc}\varphi + \\ G(d_{lc})\rho + R_{lc}G(d_{lc})\gamma + H_{lc}G(d_{lc})\delta + R_{lc}H_{lc}G(d_{lc})\alpha + \\ S_{lc}G(d_{lc})\theta + \lambda \mathrm{T}_{lc} + \zeta_s + \varepsilon_{lc}, \end{aligned} \tag{2.8}$$

where all of the variables are defined as above with the exception of z_{lc}, which takes on the value of one if a locality assesses a non-zero town or district (sub-town) tax. The variable z_{lc} takes on a value of zero if both the town and district sales tax are zero.[36] I again instrument for the county tax rates. The last three columns of Table 2.9 present the marginal effects from the probit regression and Figure 2.14 presents the results conditional on distance.

The results indicate that towns near the high- and low-tax side of the border are more likely to assess a local option tax. For example, in low-tax states, the average marginal effect is -0.47. If this effect were constant, a town 100 miles from the border is 47 percent less likely to have a local option tax compared to an identical town on the border. For the high-tax state, the effect is 44 percent. The negative effect on the low-tax side is consistent with the tax gradient decreasing away from the border. Towns near the border realize they can export some of the town's tax burden to residents of the neighboring state and substitute to a local option sales tax – perhaps as a mechanism to reduce the property tax. On the high-tax side, towns near the border are more likely to have local option taxes as well, which is inconsistent with the theory and may explain why no tax gradient emerges in the tax rates. However, as the size of the discontinuity increases, the marginal effect on the probability of

[36]The probit specification above might be suggestive of the need for a two-tier hurdle model in the previous sections. In the previous sections, I assume that the choice of going from a 0% to 1% tax rate is equivalent to going from 1% to 2%. Given the background section on local option taxes in Section 2.2 this is likely to be the case. Each time a jurisdiction passes a sales tax, it requires a council vote or a town-wide referendum – each of which are equally likely to draw public opposition. Additionally, most local taxes are remitted to the state government and then forwarded to the town, suggesting the passage of a local option tax imposes few administrative costs (at start-up) on the local government. The results of the probit are also consistent with the results for Equation 2.5.

assessing a local option tax is decreasing.

2.5.7 Discussion

In summary, in a local region of the border, towns in low-tax states keep their local tax rates relatively high in order to export a large fraction of their tax burden to cross-border shoppers. This is especially salient when the size of the tax gap is largest – and when the town is most likely to attract cross-border shoppers who will bear some of the burden of the tax. As towns get further from the border, tax rates fall, and they fall the fastest the closest to the border – suggesting that towns undercut their neighbor in an effort to attract cross-border shoppers. However, towns farther away from the border need a much lower tax rate to attract residents from the low-tax state or from the border towns in their state. The empirical evidence here suggests that for towns beyond fifty miles from the border, tax rates are relatively constant with respect to distance and even slightly increasing.

On the high-tax side, the results do not line up with the theory, except in the case of very large state tax rate differences. An explanation for these effects is that the local tax rates on the high-tax side are much lower on average. Localities on the high-tax side of the border may not be constrained when assessing local option taxes because they do not anticipate further cross-border shopping problems. If the state tax rate is high, localities may anticipate that any remaining consumers have a low elasticity – leading the jurisdiction to raise the rate without fear of losing shoppers. An alternative explanation is suggested by Harding, Leibtag and Lovenheim (Forthcoming) who find that the incidence of the tax varies with distance from neighboring low-tax states. If firms near the border pass on less of the tax to consumers, the firms may be already smoothing the post-tax price in the same manner the localities would. If towns know that firms near the border bear more of the burden of the tax, the jurisdiction will have incentives to raise taxes near the border. These authors find

no evidence that the incidence varies with distance on the low-tax side of borders. If incidence effects do not vary with distance in low-tax states, then the towns will adjust the local tax rate as a function of distance because the local government cannot rely on firms to vary the incidence of the tax with distance. Another explanation is yardstick competition by municipalities. If jurisdictions engage in yardstick competition, governments observe the tax rate levels of nearby jurisdictions and mimic the tax rate. For border localities on the high-tax side of the border, some of the local neighbors are on the low-tax side of the border. Because these neighbors in the low-tax state set relatively high tax rates, border jurisdictions on the high-tax side may look to their neighbors in the opposing state and set a relatively high tax rate compared to a town interior to their state.

However, some of the above hypotheses are unsatisfactory in that they apply to towns on the high-tax side, but not to towns on the low-tax side of borders. An alternative model-based explanation that would apply uniformly to all towns may consider the political economy of tax setting behavior. The results estimated in this paper are consistent with a model where the median voter determines local sales tax rates. On the high-tax side of a border, the median voter is more likely to be a cross-border shopper if the town is closer to the border. Suppose the median voter's goal is to set taxes in a manner that reduces their own tax obligations. If the median voter is someone who engages in large amounts of cross-border shopping, they may seek to raise sales taxes in order to export the tax burden to residents who do not cross-border shop. In a world with multiple tax instruments, the higher sales tax rate may offset some of the tax burden for other taxes that the median voter pays. As towns are further from the border, the median voter is less likely to be a cross-border shopper or is less likely to purchase as many goods in the neighboring state. As a result, the median voter spends more on purchases domestically and would want lower tax rates. On the low-tax side of a state border, the median voter always shops in their home

state and the median voter will want to export the tax to non-domestic residents. The median voter in the town near the border can do this even with a relatively high local sales tax because the proximity to the border. For towns further away, the median voter needs to vote for lower taxes in order to attract residents from abroad. Future research should consider theoretical models – such as this simple voting-based model – which yields tax gradients consistent with the empirical findings.

To summarize, the level of local tax rates is substantially higher for border towns on the low-tax side of the border. The existence of an economically and statistically significant tax gradient implies that discontinuities at state borders continue to affect the elasticity of demand in local jurisdictions away from the border by inducing a unique form of tax competition where jurisdictions account for this discontinuity.

2.6 Conclusion

Identical local governments within a state will have incentives to differentiate their tax rates to smooth the discontinuity in the total tax rate at the border, even when the state government prefers a higher or lower state tax rate than its neighbors. Discontinuities at state borders make identical jurisdictions heterogeneous based on the jurisdiction's location within a state and its ability to attract cross-border shoppers.

The discontinuity in the tax system at state borders induces welfare distortions with respect to consumption, as some residents cross borders to purchase lower-tax goods. It is also likely to create horizontal inequities – where individuals with the same ability to pay actually pay different taxes – if some residents cross-border shop while others do not. A continuous tax system will reduce these welfare distortions, but will still have horizontal inequities if some individuals purchase goods a few neighboring jurisdictions away.

Moreover, as in Kleven and Slemrod (2009), "tax-driven product innovations" will distort the location of firms. Such innovations require no technological changes, but

distort the characteristics of the goods – where the characteristic is the location of sale – to avoid taxes. In the context of the retail sales tax, this innovation arises as firms distort the locational characteristics of the good to the favorable tax side of the border in order to capture cross-border shoppers.[37] Absent the tax differential, the firm may have decided to locate on the other side of the border. Instead, the locational characteristics become inefficient because of the firm's decision. I uncover a local tax gradient, which reduces these production and consumption distortions at state borders.

Even if the first-best tax system – as set by a global welfare maximizer who is constrained by varying state tax rates – is a perfectly continuous function of distance from the border, it is likely that this system is administratively infeasible. However, as in the model presented above, smoothing the size of the discontinuity through discrete steps in the tax system (that are a function of distance) is likely to happen as a result of tax competition over local tax rates – approximating a continuous policy. Although the discrete steps will induce additional notches within the tax system, the reduction in the larger distortion across the state will almost certainly be welfare improving. The new distortions at local borders are virtually guaranteed in any tax system given the infeasibility of an infinite number of differentiated tax rates.

Whether the tax gradient is optimal from a state government's perspective or if it is optimal for a state to allow localities to adopt LOST raise different questions than the global planner's problem. Agrawal (2011*a*) shows in a model where states maximize social welfare by selecting sales tax rates within the state, that the optimal pattern for the state depends critically on the elasticity of cross-border shopping, the inequality in consumption profiles, and the relative magnitudes of the tax base and exporting effects. If the elasticity of cross-border shopping is below a critical value, then it is advantageous to have higher tax rates – as empirically found in this paper –

[37] The model presented in this paper did not allow for firms to choose their location decisions. However, shopping patterns must be linked with retail locations.

in the border region of a low-tax state. For it to be optimal to set higher rates in the border region of a high-tax state, the within-state inequalities in consumption induced by cross-border shopping must be sufficiently large such that the social planner would want to eliminate these inequalities by raising the border region's rate. It may be welfare enhancing from the state government's perspective to allow localities to assess local sales taxes.

This paper shows that federalism is a partial solution to the administrative difficulty of centrally implementing a tax system where tax rates are a continuous function of distance from the border. If local jurisdictions can set additional retail sales taxes, the salience of the state border would directly influence localities close to that border. On the other hand, jurisdictions located far from the border worry less about the notches because these jurisdictions do not have in-flows or out-flows of cross-border shoppers.

Empirically, I show that conditional on a set of observables, distance is a significant factor in the pattern of local option taxes. Jurisdictions on the low-tax side of the border that are relatively close to the border have higher tax rates than identical jurisdictions farther away from the border. These jurisdictions have a smaller elasticity of demand and export the tax burden to cross-border shoppers from the neighboring state. Although local tax rates gradually offset border differences on the low-tax side, the results have an unexpected sign on the high-tax side. Identical towns on opposite sides of the border also differ in the level of local tax rates with local taxes substantially lower on the high-tax side of the border. Although a discrete tax gradient is not the first-best outcome of a perfectly continuous tax gradient, the empirical evidence suggests that local option taxes will gradually smooth tax differentials over longer distances on the low-tax side.

The methodologies I develop in the paper are broadly applicable to analyzing how local governments respond to any policies that vary discontinuously at each

border. For example, suppose states implement varying environmental restrictions on firms. At the margin, firms may choose to relocate to the side of the border with lower emission standards. Localities may adopt local regulations as a function of distance to the border in order to discourage firms from relocating and to reduce the regulatory differential at the border. More generally, "notches" – resulting from discontinuous policies – are widespread and often differ in magnitude at each border. The methodologies developed in this paper account for this heterogeneity. Researchers can apply the theoretical and empirical methodologies I introduce in this paper to test if "notches" induce jurisdictions to implement policies that vary with distance from the discontinuity.

2.7 Appendix

2.7.1 Proof of Uniqueness of the Equilibrium

In the paper, I derive conditions under which an equilibrium will exist. Existence follows from concavity and continuity of the best response functions, plus certain additional conditions outlined in the text. Once existence is shown, I prove below that any equilibrium in this model will be unique for the case of a multiple state, n town model. The solution to a three state model with n towns is characterized by the equation $\mathbf{At} = \mathbf{b}$. The proof generalizes to a multi-state model simply by changing

the dimensions of the vectors and the elements of $\mathbf{b}$. This system can be written as:

$$\begin{bmatrix} 1 & -\frac{1}{4} & 0 & & \cdots & & 0 & -\frac{1}{4} \\ -\frac{1}{4} & \ddots & \ddots & \ddots & & & & 0 \\ 0 & \ddots & \ddots & \ddots & & & & \\ & \ddots & \ddots & 1 & -\frac{1}{4} & & & \vdots \\ \vdots & & & -\frac{1}{4} & 1 & \ddots & \ddots & \\ & & & & \ddots & \ddots & \ddots & 0 \\ 0 & & & & \ddots & \ddots & \ddots & -\frac{1}{4} \\ -\frac{1}{4} & 0 & & \cdots & & 0 & -\frac{1}{4} & 1 \end{bmatrix} \begin{bmatrix} t_{BB}^{H} \\ \vdots \\ \vdots \\ t_{B}^{H} \\ t_{BB}^{M} \\ \vdots \\ \vdots \\ t_{B}^{M} \\ t_{BB}^{L} \\ \vdots \\ \vdots \\ t_{B}^{L} \end{bmatrix} = \frac{1}{4} \begin{bmatrix} \delta x - D \\ \delta x \\ \vdots \\ \delta x \\ \delta x - R \\ \delta x + R \\ \delta x \\ \vdots \\ \delta x \\ \delta x - S \\ \delta x + S \\ \delta x \\ \vdots \\ \delta x \\ \delta x + D \end{bmatrix}. \tag{2.9}$$

Proof. Matrix $\mathbf{A}$ is a strictly diagonally dominant matrix because the sum of the diagonal element in every row is greater than the sum of all the off-diagonal elements in absolute value. By the Levy–Desplanques theorem, a strictly diagonally dominant matrix is non-singular – has an inverse. For a given number of towns and parameters in the model, therefore, $\mathbf{A}^{-1}\mathbf{b}$ is unique. When a Nash equilibrium exists, it is guaranteed to be the unique Nash equilibrium and is characterized by $\mathbf{A}^{-1}\mathbf{b}$. □

2.7.2 Characterizing the Solution to an n Town Model

In this section, I generalize the solutions in the text for a model with three states and n total towns. Recall that the solution to an n town model – where n is defined

as the total number of towns in the model such that each state has $\frac{n}{3}$ towns – can be obtained by solving Equation 2.9, where the solution is given by $\mathbf{t}^{*}=\mathbf{A}^{-1}\mathbf{b}$.[38] Define $\mathbf{A}^{-1}$ as $\boldsymbol{\Xi}$. Then, letting $\ell = -\frac{1}{4}$, the determinant of $\boldsymbol{\Xi}$ is given by:

$$\det(\boldsymbol{\Xi}) = 1+\sum_{k=1}^{\lfloor\frac{n-1}{2}\rfloor}\frac{(-1)^k n(n-k-1)!\ell^{2k}}{(n-2k)!k!}+2\ell^n\mathbf{1}(n \text{ is odd})-4\ell^n\mathbf{1}(n \text{ is even} \,\&\, \frac{n}{2} \text{ is odd}), \tag{2.10}$$

where $\mathbf{1}(\bullet)$ is an indicator function that takes on the value one if $\bullet$ is true and takes on the value zero otherwise. Then, I can decompose $\boldsymbol{\Xi}$ such that $\boldsymbol{\Xi} = \frac{1}{\det(\boldsymbol{\Xi})}\boldsymbol{\Psi}$. Letting i index rows (the town) and j index columns (the town's neighbors' positions on the circle), the elements of $\boldsymbol{\Psi}$ are given by $\psi_{i,j}$. If n is even, the elements of the first row in $\boldsymbol{\Psi}$ can be solved for as:

$$\psi_{1,j} = \begin{cases} \phi_j & \text{for } j = 1, ..., \frac{n}{2} \\ -2\ell\phi_{\frac{n}{2}} & \text{for } j = \frac{n}{2}+1 \\ \phi_{2+n-j} & \text{for } j = \frac{n}{2}+2, ..., n, \end{cases} \tag{2.11}$$

while if n is odd, the elements are given by:

$$\psi_{1,j} = \begin{cases} \phi_j & \text{for } j = 1 \\ \phi_j + \xi_j & \text{for } j = 2, ..., \frac{n-1}{2}+1 \\ \phi_{2+n-j} + \xi_{2+n-j} & \text{for } j = \frac{n-1}{2}+2, ..., n, \end{cases} \tag{2.12}$$

where

$$\phi_j = \sum_{k=0}^{\lfloor\frac{n-j}{2}\rfloor}\frac{(-1)^k(n-j-k)!\ell^{2k+j-1}}{(n-j-2k)!k!}, \tag{2.13}$$

[38] Adding additional states would just change the position of the elements of $\mathbf{t}$ and $\mathbf{b}$. Similarly, making states heterogeneous with respect to the number of towns would change the position of the elements.

and

$$\xi_j = \sum_{k=0}^{\lfloor \frac{j-2}{2} \rfloor} \frac{(-1)^{j+k-2}(j-2+k)!\ell^{n-j+2k-1}}{(j-2)!k!}. \tag{2.14}$$

Note that as they are written above, equations 2.11 and 2.13 only yield the elements of the matrix's first row. Each of the subsequent rows will shift the elements of the previous row one place over, where the final element of the row will become the first element of the row. Therefore, for $i > 1$:

$$\psi_{i,j} = \psi_{1,r} \text{ where } r \equiv (j+n-r+1)(\operatorname{mod} n), \tag{2.15}$$

and $(j+n-r+1)(\operatorname{mod} n) = j+n-r+1-n\left\lfloor \frac{j+n-r+1}{n} \right\rfloor$.

The equilibrium tax rates can now be characterized as $\mathbf{t}^*=\mathbf{A}^{-1}\mathbf{b} = \mathbf{\Xi}\mathbf{b} = \frac{1}{\det(\mathbf{\Xi})}\mathbf{\Psi}\mathbf{b}$. It should be noted that

$$\sum_{j=1}^{n}(-\ell)\delta x \psi_{i,j} = \frac{1}{2}\delta x \ \forall i. \tag{2.16}$$

To complete the characterization of the solution, one must determine how the size of the discontinuities in $\mathbf{b}$ that multiply the inverted matrix enter the solution. Under the assumption that all states are uniform in size, the discontinuities $D, R,$ and S will additively enter into town i's equilibrium tax rates as $(-\ell)(\psi_{i,n}-\psi_{i,1})D$, $(-\ell)(\psi_{i,\frac{n}{3}+1}-\psi_{i,\frac{n}{3}})R$, and $(-\ell)(\psi_{i,\frac{2n}{3}+1}-\psi_{i,\frac{2n}{3}})S$. Recalling that $D = R+S$ because of the circular setup of the problem, if $x > x^*$, the equilibrium tax rates can be characterized as:

$$t_i^* = \frac{1}{2}\delta x - \frac{\ell(\psi_{i,n}-\psi_{i,1}+\psi_{i,\frac{n}{3}+1}-\psi_{i,\frac{n}{3}})R + \ell(\psi_{i,n}-\psi_{i,1}+\psi_{i,\frac{2n}{3}+1}-\psi_{i,\frac{2n}{3}})S}{\det(\mathbf{\Xi})} \ \forall i. \tag{2.17}$$

Notice that if the states were asymmetric with respect to the number of towns, the above expressions could be modified by picking the jth element of $\mathbf{\Psi}$ appropriately – where the number of towns are adjusted on the j index of ψ.

To get a better idea of what the pattern of geographic differentiation looks like,

Figure 2.15 presents numerical results when the number of towns is large. As is evidenced by the figure, the tax gradient is strongest in a local region of the border and the tax gradient is steepest for the largest tax differentials. At the interior of states, the gradient becomes relatively flat, which is due in part to the assumption that individuals are restricted from shopping one town over.

2.7.3 Methodology for Calculating Distance from the Border

In this section, I outline the methodology for calculating distance from the border. Arc-GIS is used to calculate this variable and all base map files necessary to calculate distance from the border are available on the Arc-GIS / ESRI map CD.[39] Figure 2.16 shows the methodology graphically.

I sometimes use the "as the crow-flies" distance from the population weighted average centroid of a place to the nearest intersection of a major road and a state border or foreign country to calculate the distance from the border. The District of Columbia is counted as a state, but Native American reservations are treated as localities. The justification for treating reservations as localities is that with some exceptions, purchases on Native American reservations by non-tribal members are subject to state sales taxes. Furthermore, reservations are often small and although they frequently sell cigarette purchases tax free, they do not have extensive shopping outlets for many larger items. Many reservations have also begun charging tribal tax rates on general sales.

To calculate distance from the border, I execute the following steps. When calculating distance, the projection system utilized in the map files is essential to guaranteeing that the distance measure is accurate for all latitudes and longitudes. This requires that the projection system selected preserves distance attributes and that it be the same on all maps before beginning any calculations. I select the North

[39]The section below utilizes jargon from mapping software, which may be unfamiliar to readers not familiar with Arc-GIS.

American Equidistant Conic Projection System. When the coordinate system is defined differently, I convert the coordinate system using the NAD 1983 to WGS 1984 _ 1 geographic transformation option. This transformation converts the coordinate system with an accuracy of plus or minus two meters.

First, in order to identify the tax rates at international crossings, I merge a detailed polygon file of the fifty states plus the District of Columbia with detailed files of Canada and Mexico. It is important to use a "detailed" file that precisely traces out the border. Smoothed files may be off several miles in many circumstances. I then convert the polygon file into a line file that explicitly identifies the geographic identification number of the "left" and "right" states. This identification will allow me to record the neighboring state's tax rate. Second, I overlay a detailed Census major roads file. Census major roads are Class 1, 2, and 3 roads, which include major highways and paved roads primarily used for transportation. These classes of roads exclude dirt roads and primarily residential roads. Then, I find the precise intersection of each state border line with a major road. This intersection is identified with a FID number, which can be used to identify the state border combination from the state line file. I drop all intersections that correspond to coastal areas or to major routes that are defined as ferry crossings.

Third, I identify the population weighted centroid as the point in which the place would balance on a scale if every person in that place were equal weight. To calculate this, I identify the population distribution within a place using the population of every Census block in the country.[40] Let b index each Census block point given by population P_b and has latitude ϕ_b and longitude λ_b. The population weighted center

[40] A Census block is the smallest unit of geography. In some cases, a block may be a large area with little or no population. In other areas, a Census block may contain an entire apartment complex or building and may have a population of several hundred.

of place i is the latitude $\bar{\phi}$ and the longitude $\bar{\lambda}$ given by:

$$\bar{\phi}_i = \frac{\Sigma P_b \phi_b}{\Sigma P_b} \qquad \bar{\lambda}_i = \frac{\Sigma P_b \lambda_b Cos(\phi_b(\frac{\pi}{180}))}{\Sigma P_b Cos(\phi_b(\frac{\pi}{180}))}.$$

Fourth, I run a "near" command on the 25,000 population weighted centroids and the several thousand intersections that I found above. This will calculate the nearest linear distance from the intersection of the major roads and the state borders. Fifth, I conduct a spatial join on the centroids with the level of geography I wish to analyze (call it a place polygon file). I define a centroid as being within a place polygon if its point is contained entirely within the polygon. This spatial join will attach the geographic identifier of the Census place or county to the centroid.

To calculate the second closest border crossing, I follow the method outlined above, but instead of executing a near command in ArcGIS, I use the near table command. This will calculate all of the nearest border crossings up to a particular threshold. I calculate 1000 of the nearest border crossings for each place centroid. This is a sufficient number for me to calculate the distance from the second closest border.

The data calculated above then can be merged based on geographic identification numbers to the Census data. However, the tax data does not contain geographic identifiers, so I must merge the data using name matching. In cases of merging by county, this is an easy process and I am able to obtain a 99.9% match rate. One county does not match because it is not in the tax data set. Census places are the closest to towns in the United States. Census places contain no county information. In some states, Census places (and towns) cross county lines. To deal with this issue, I intersect Census places and counties using a spatial join in ArcGIS. This matches each place to a county that it overlaps. I can then name match Census places to the tax data using place, county, and state names. Name matching to Census place data obtains nearly a 90% match rate.[41] I hand match any remaining observations.

[41] Recall some Census places are not towns and some towns are not Census places.

Inevitably, a better measure of distance is actual driving distance. I calculate driving distance using ArcGIS' network analyst toolbox. After following the first three steps above, I use ESRI's street file to calculate driving distance. The data in the street file contains all streets in the country, but note that the final destination I use will always be a major road as above. I convert the data to a network data set so that it has street driving speeds within it. To calculate driving distance, I locate the nearest minor street to a population weighted centroid and to the major road crossings by searching within a fifty-mile radius. After doing this, I need to specify how ArcGIS will calculate driving distance. Using the centroids as origins and the border crossings as destinations, I use a time criterion to calculate distance – that is I have GIS minimize the driving time to the nearest location.[42]

In addition, I need to make assumptions on how the individual drives to the border. I assume that individuals follow a "hierarchical" method of driving – that is whenever possible, I have ArcGIS route their travel via larger roads. I also require that individuals must obey one-way streets or turn restrictions onto roads. However, I do not impose any other restrictions – that is I do not restrict individuals from using alleys, four-wheel drive roads, or ferry crossings.[43] Using the network analyst, ArcGIS returns the driving distance (in miles) and time (in minutes) for the shortest time path from each population weighted centroid to the nearest intersection of a major road and state border.

2.7.4 An Alternative Empirical Specification

One alternative to the regression in the text is to run the same specification with the dependent variable being the total effective local option tax: the local plus district plus county tax rate. The difference is that the county tax rate is now included in the dependent variable. Because the county tax rate is included in the dependent

[42]The alternative would be to use a distance criterion.
[43]I impose these restrictions and find the driving distances are almost perfectly correlated.

variable, I need not include it on the right side of the equation and avoid the need of instrumenting for it. The estimating equation becomes:

$$\begin{aligned} t_{lc} + \mathrm{T}_{lc} = \beta_0 + \beta_1 H_{lc} + \beta_2 S_{lc} + \beta_3 R_{lc} + \beta_4 R_{lc} H_{lc} + X_{lc}\varphi + \\ G(d_{lc})\rho + R_{lc}G(d_{lc})\gamma + H_{lc}G(d_{lc})\delta + R_{lc}H_{lc}G(d_{lc})\alpha + \\ S_{lc}G(d_{lc})\theta + F_{lc}\iota + \zeta_s + \varepsilon_{lc} \end{aligned} \tag{2.18}$$

where F_{lc} is a matrix of county level demographic and geographic controls. The inclusion of F_{lc} is necessitated by the fact that the regression equation must now control for predictors of both the local and county tax rates. With this change, the interpretation of the marginal effects becomes the gradient the total local option sales taxes rather than the slope of only the municipal tax rates.

The probit specification in the text also needs to be modified. In this specification, the left-side variable takes on the value zero if a locality does not have a local, district or county tax rate and takes on a value one otherwise. It should be noted that because the sample includes only states that allow for local option sales taxes, most municipalities have at least one of these taxes being greater than zero and, in fact, some outcomes are perfectly predicted.

The following paragraphs summarize the results of the alternative specification presented above. Tables and figures of these results are excluded for simplicity and because the results under this alternative specification are extremely similar to the preferred results in the text.

Across almost all columns and specifications, the marginal effects presented in Tables 2.6 and 2.7 are matched in sign and significance. One difference in the results is that the marginal effects are – on average – smaller in absolute value when the county tax rate is included in the dependent variable. One reason for this is that counties are much larger and base their decisions on their distance to the border rather than the municipality's distance to the border. In the baseline specification, a

one hundred mile increase in distance from the border reduces local option taxes by .077 percentage points on the low-side of the border. The same change in distance reduces the total local rate on the high-side of the border by .18 percentage points.

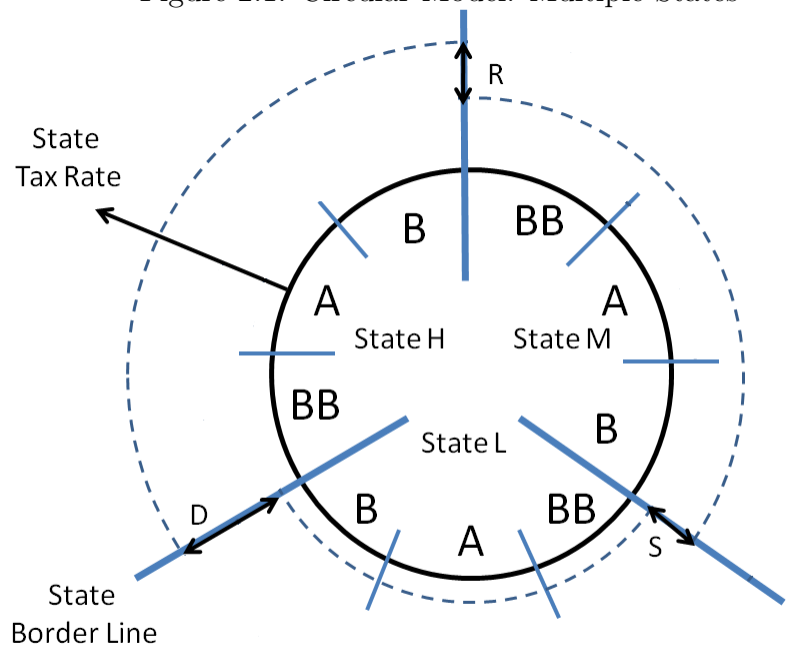
When looking at the results conditional on the size of the discontinuity in state tax rates in Table 2.9, the results are also similar in spirit. On the low-tax side, the gradient becomes steeper as the notch size increases. On the high-tax side, the gradient is closer to a positive slope as the discontinuity increases. One difference is that the low-tax side results are less significant, but it is likely this is driven by the fact that the point estimates become smaller after including the county tax rate. With regard to the probability that a jurisdiction has a local option tax, the results are insignificant in this specification. With regard to the probit results in the table, the probit regression has a dependent variable that equals one if the total local rate is greater than zero. This can occur if either the local rate, the district rate, or the county rate is greater than zero. Because the sample only includes states allowing for LOST, most jurisdictions in the sample meet at least one of this criteria. In addition, in some states, every county has a county tax rate, which implies some observations are perfectly predicted. As a result, the probit results in the text are more reliable and answer the more policy relevant question – what determines if a jurisdiction assesses a tax? The probit results to the modified equation instead answer: does distance determine if a jurisdiction has a local option tax or is in a county that has such a tax?

Lastly, the results conditional on various distances are also similar to the results in the text. Figures 2.9 and 2.10 match the alternative specification, except that the marginal effects are smaller in absolute value for the total effective rate. Results conditional on both the size of the discontinuity and distance in Figure 2.12 are also similar under the alternative specification. One difference here is that the results for the high-tax side of the border are more likely to be positive compared to the results in

the paper. In Figure 2.12, the largest discontinuities implied marginal effects close to zero. When considering the total effective rate, the marginal effects become positive at some distances for discontinuities of four and six percentage points.

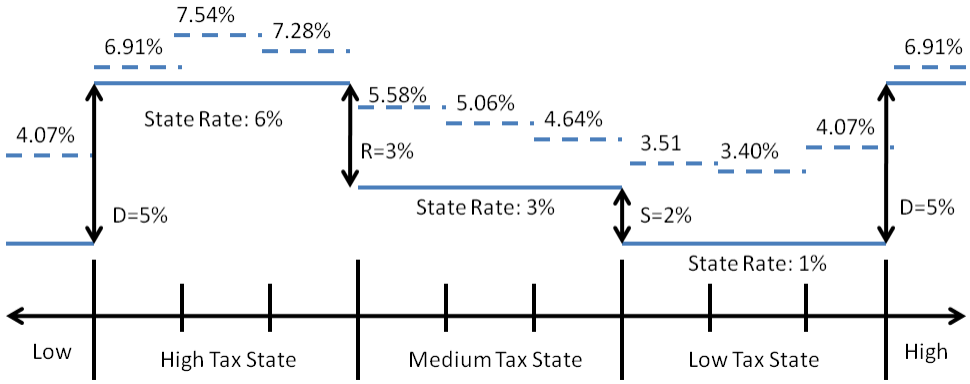
In conclusion, this appendix develops an empirical specification that avoids having to instrument for county tax rates in the estimating equation. Such a specification is derived by an analyzing the steepness of the tax gradient in the "total effective local tax rate," which is the county tax rate plus all local taxes in a municipality. This modification allows me to move the county tax rate to the left-side of the estimating equation – eliminating any need to control or instrument for it. After estimating all the equations in the text with this modification, the signs and significance of most results are unaffected. The main difference in the results of this specification is that the slope of the gradient becomes smaller in absolute value.

Figure 2.1: Circular Model: Multiple States



The dotted lines around the circle represent the level of the state tax rate. The large solid lines are state borders and the smaller solid lines are town borders.

Figure 2.2: Example of Equilibrium Tax Rates



The figure above flattens the circular model of Figure 2.1. The solid horizontal lines are the state tax rates. The dotted lines are the municipal plus state tax rates in the Nash Equilibrium when the parameters of the model are set such that $\delta x = 4$, $D = 5$, $R = 3$, and $S = 2$.

Figure 2.3: Driving Distance from the Nearest Border

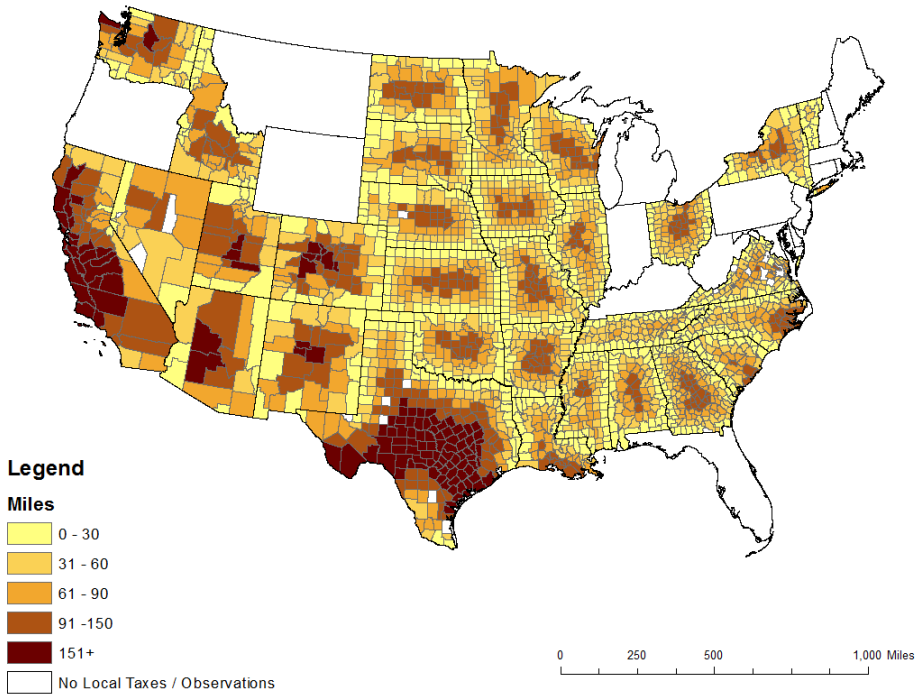


Figure 2.4: State Sales Tax Rates

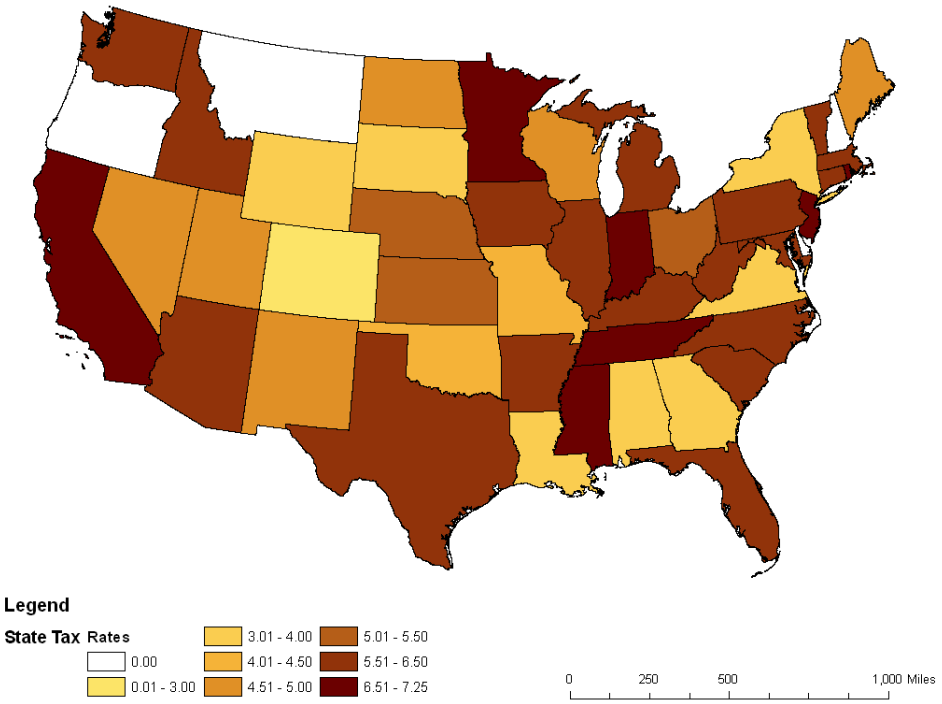


Figure 2.5: County Tax Rates by Nearest Border Type

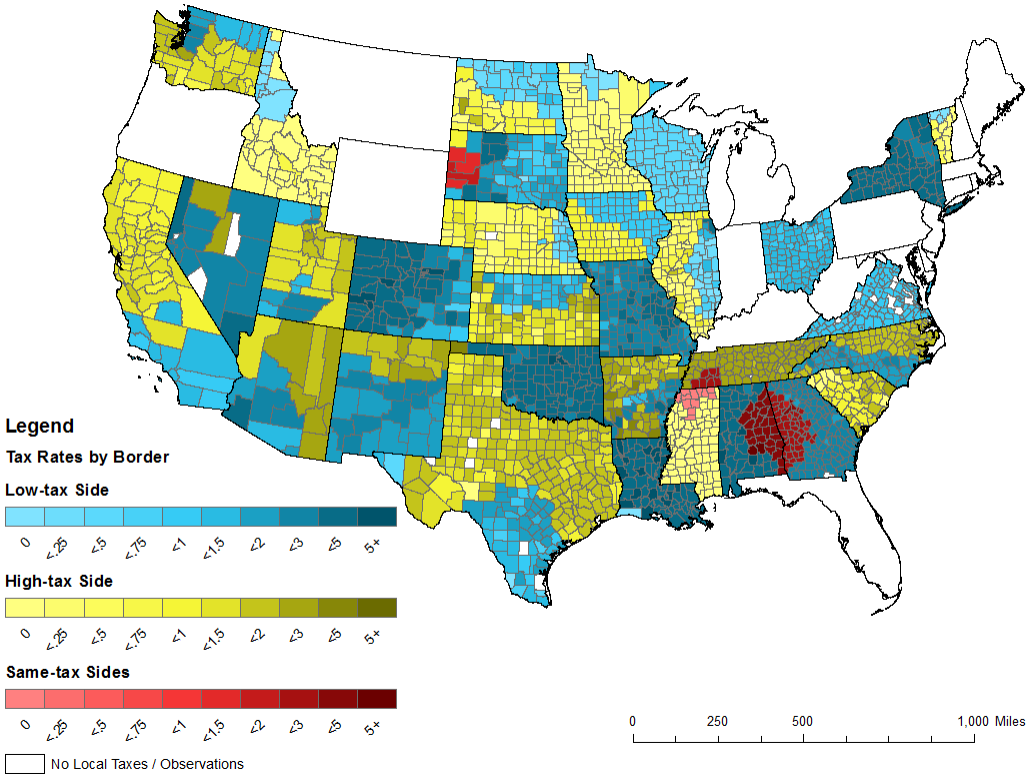


Figure 2.6: Missouri Example of Local Option Taxes

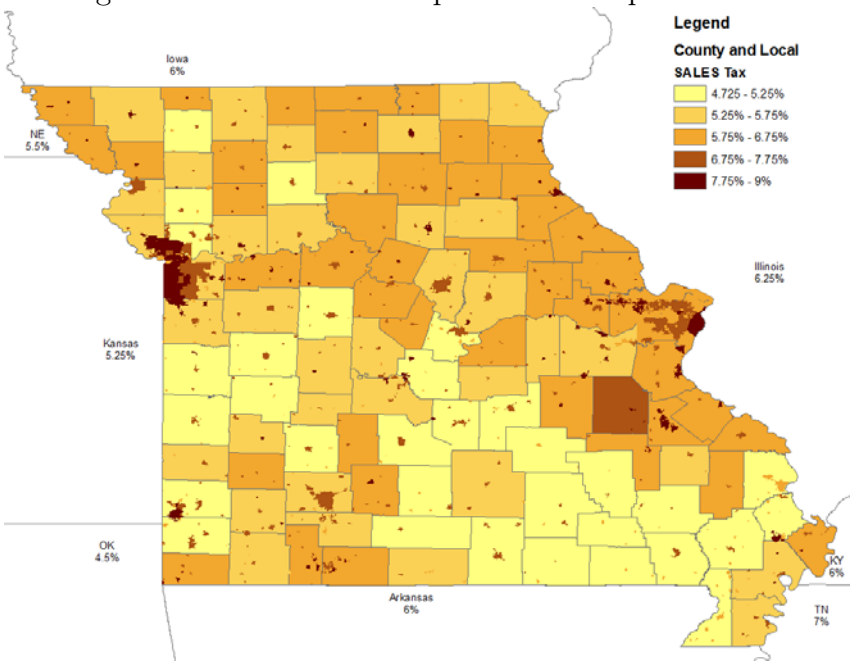
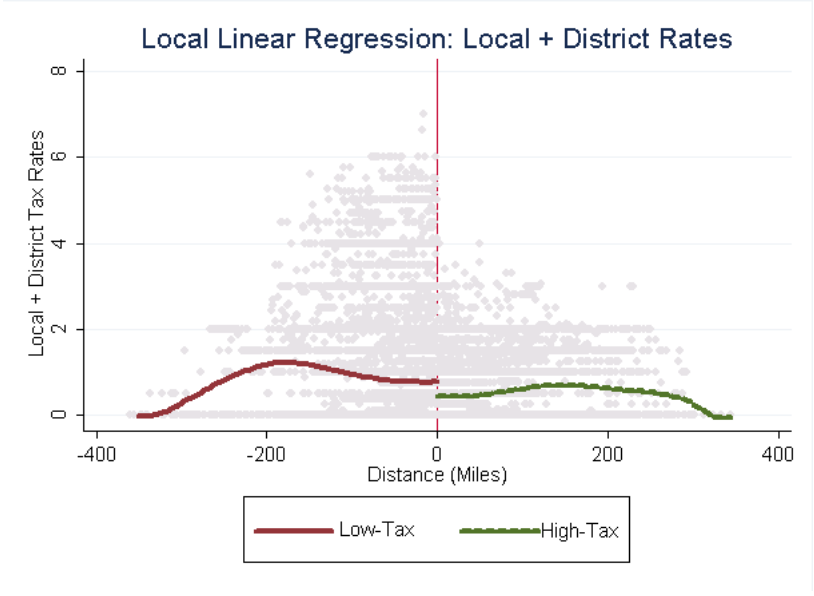
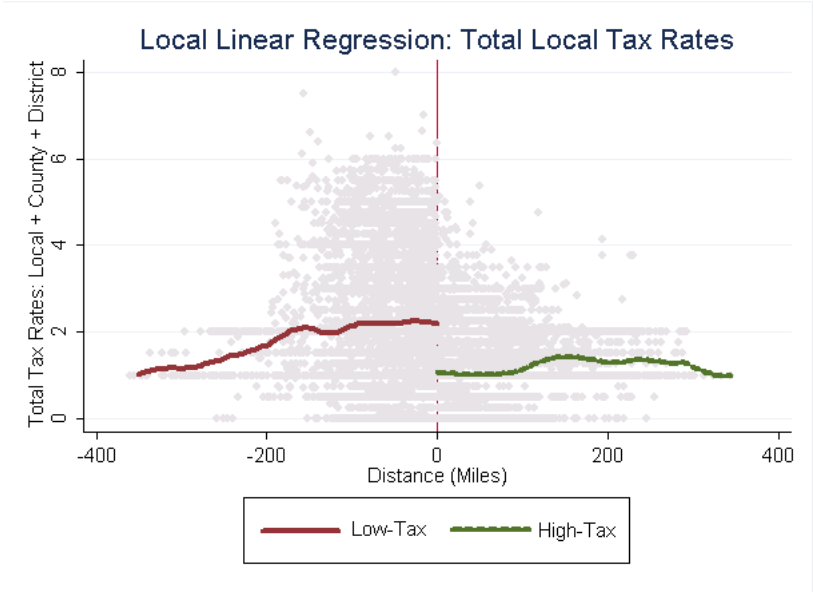


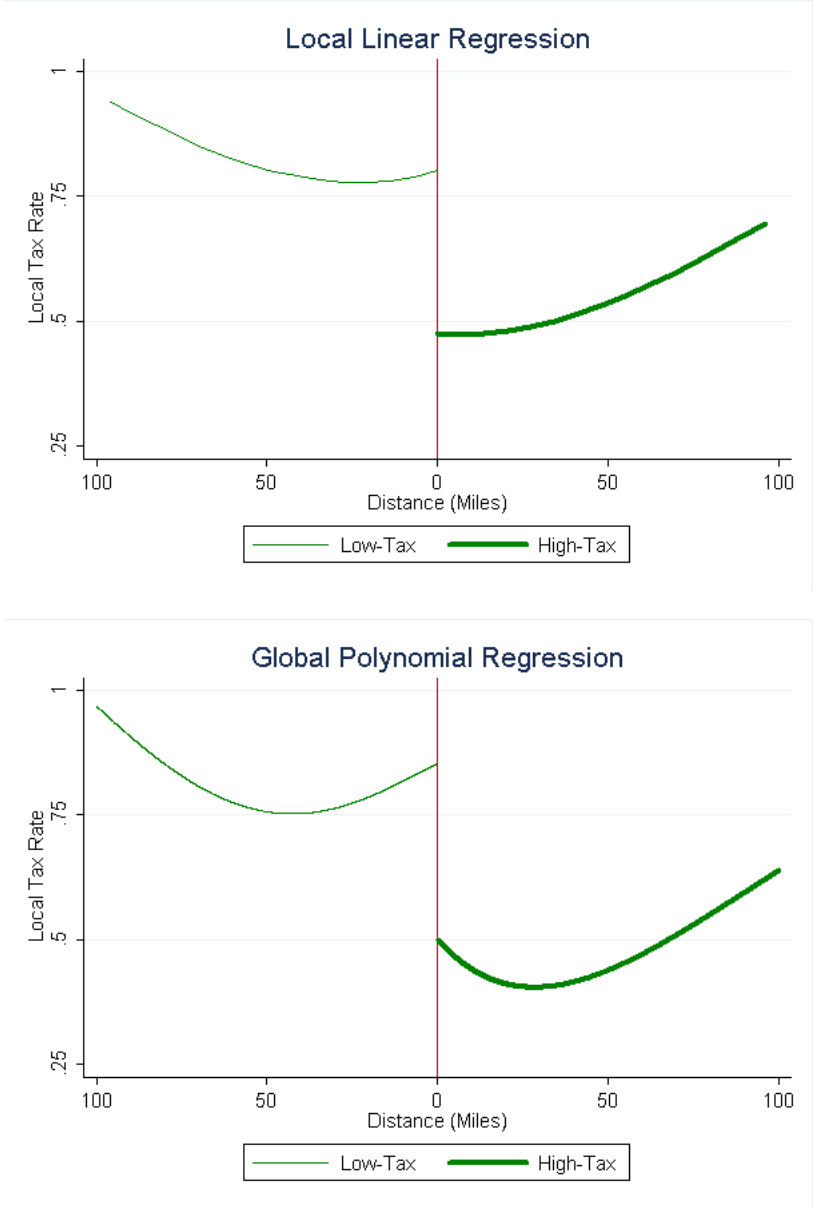
Figure 2.7: Regression Discontinuity Results





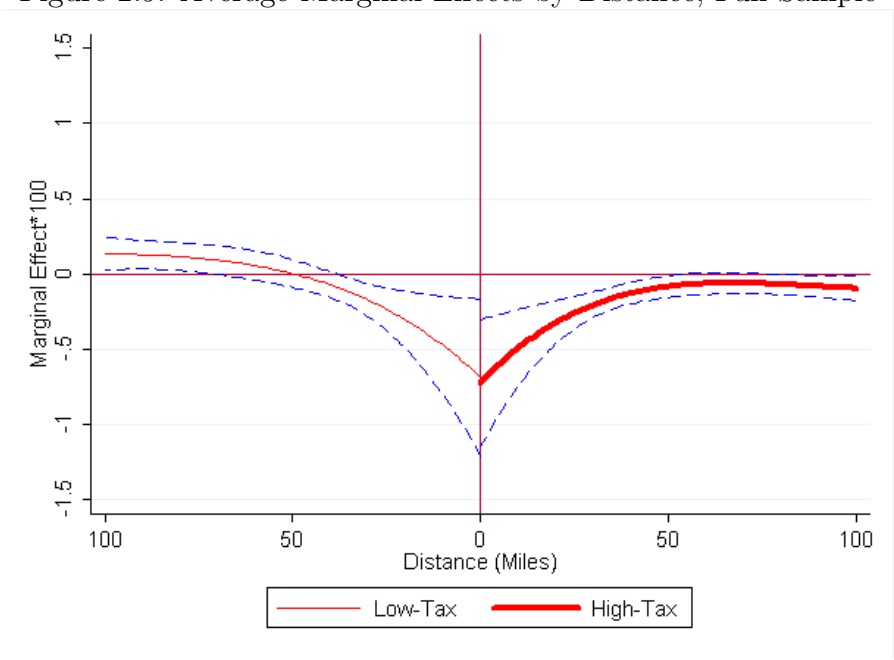
The top graph uses municipal tax rates only, while the second graph depicts the RD for the total effective local tax rate.
RD methodology follows Imbens and Kalyanaraman (Forthcoming) with a triangle kernel. The graphs above do not include additional covariates.

Figure 2.8: Local Linear Regression Versus Global Polynomial Regression



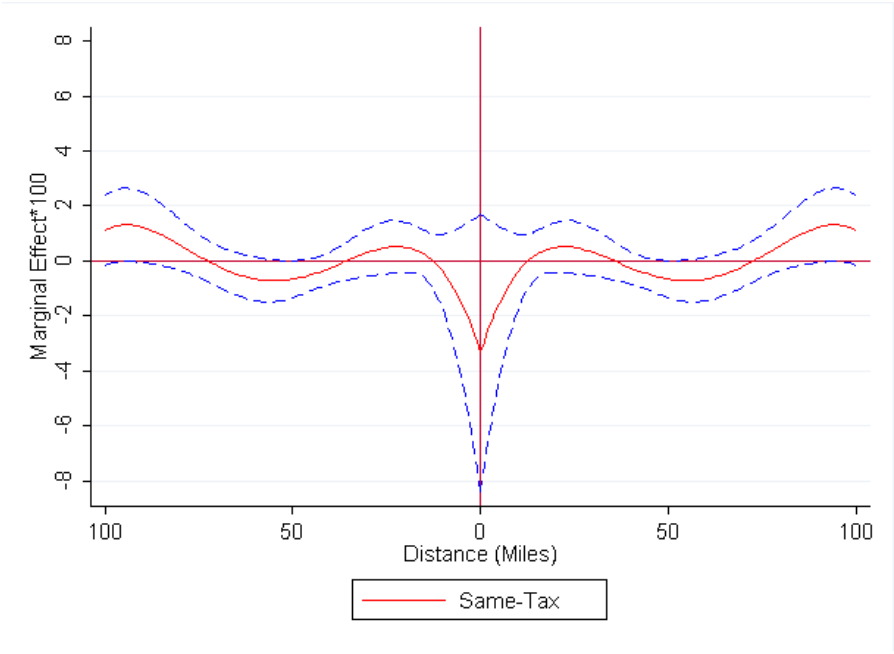
The top graph zooms in on the top graph in Figure 2.7. The bottom graph plots the predicted value from a fifth degree global polynomial evaluated at the average discontinuity size with no covariates.

Figure 2.9: Average Marginal Effects by Distance, Full Sample



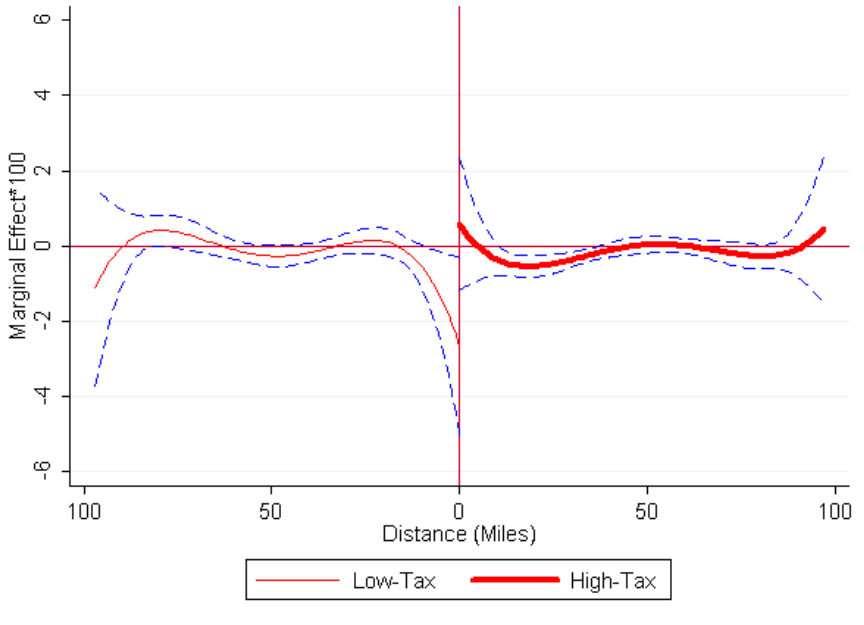
Marginal effects represent the effect of moving **away** from the border in both directions. To derive the per mile marginal effects, divide by 100. Confidence intervals are 95%.

Figure 2.10: Average Marginal Effects by Distance, States with Same Tax Rate



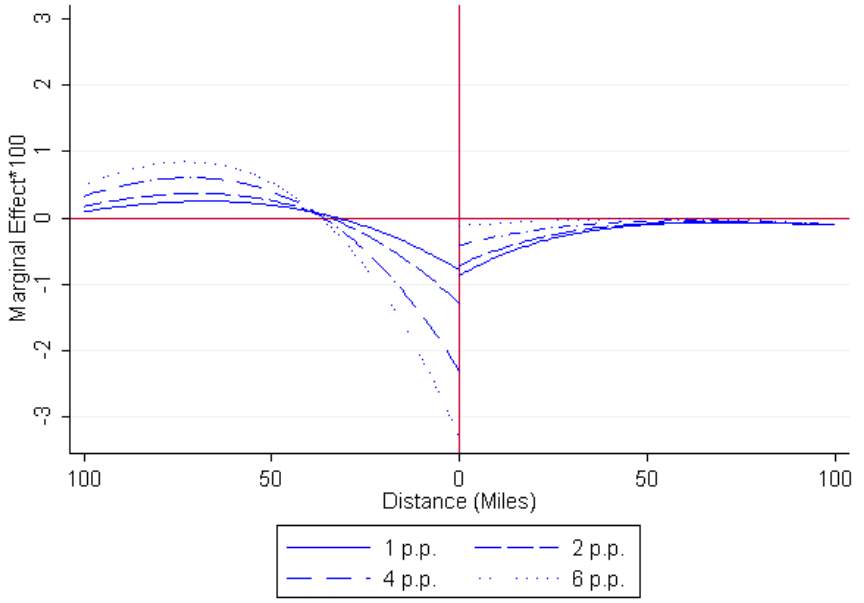
Marginal effects represent the effect of moving **away** from the border in both directions. To derive the per mile marginal effects, divide by 100. Confidence intervals are 95%.

Figure 2.11: Average Marginal Effects by Distance with 98 Mile Sample Restriction



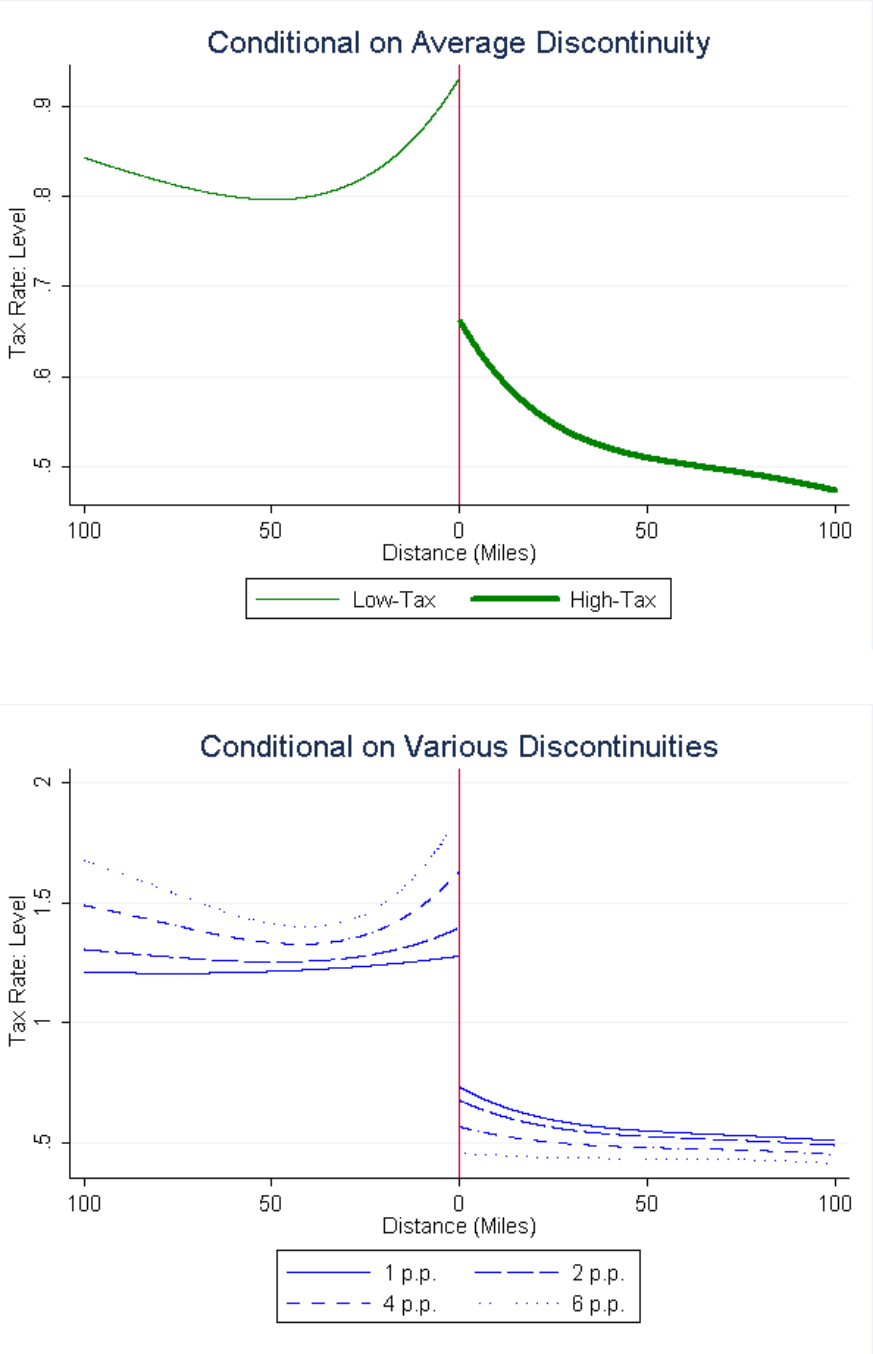
Marginal effects represent the effect of moving **away** from the border in both directions.
Confidence intervals are 95%.

Figure 2.12: Average Marginal Effects by Distance and Size of the Discontinuity



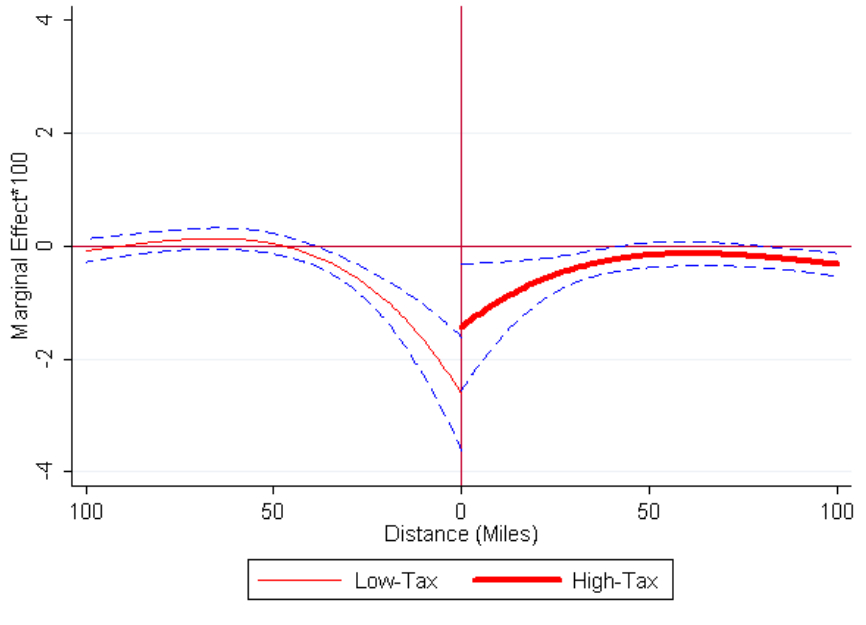
Left side is the relatively low-tax state. Right side is the relatively high-tax state.
Marginal effects represent the effect of moving **away** from the border in both directions.

Figure 2.13: Predicted Values of Local Rates from a Global Polynomial Regression



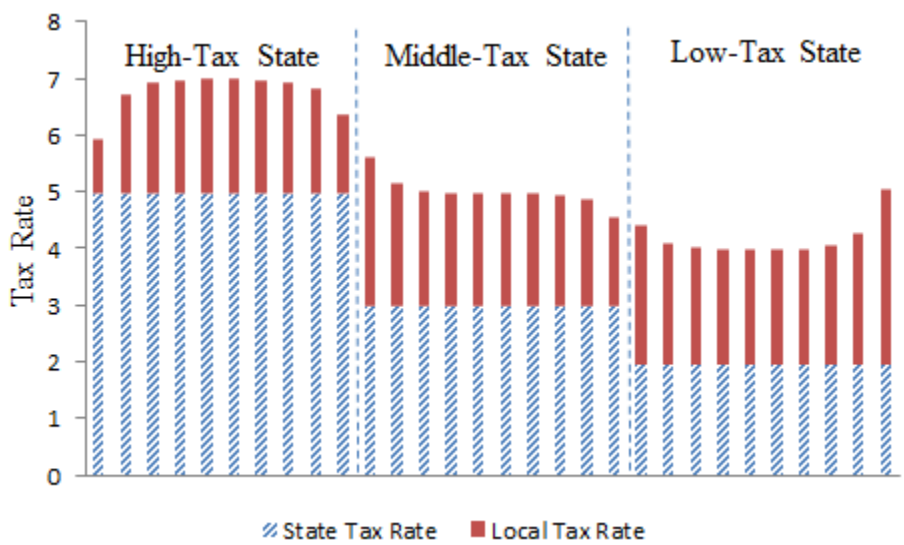
These graphs plot the predicted values of local tax rates from a global polynomial regression including state fixed effects and covariates. The top graph evaluates the predicted value conditional on the average discontinuity and the set of covariates. The bottom graph conditions on various discontinuity sizes. The marginal effects in Figures 2.9 and 2.12 correspond to the slopes of the lines above, respectively.

Figure 2.14: Average Marginal Effects by Distance of the Probit Regression



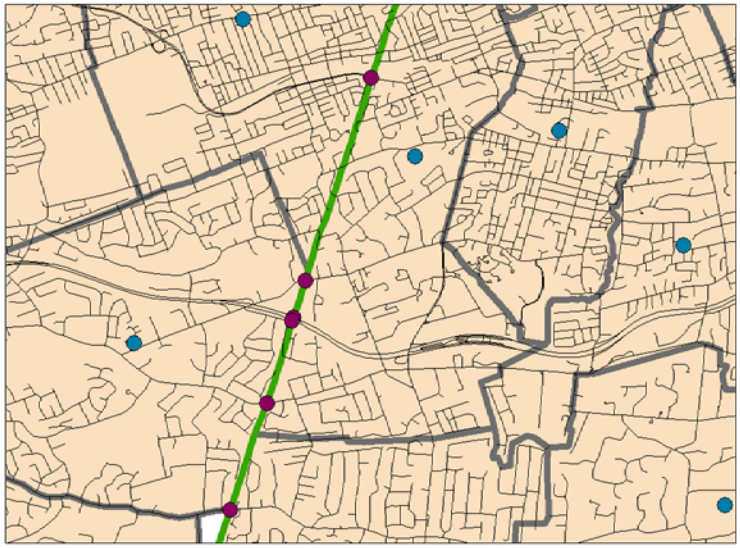
Left side is the relatively low-tax state. Right side is the relatively high-tax state. Marginal effects represent the effect of moving **away** from the border in both directions. To derive the per mile marginal effects, divide by 100.

Figure 2.15: Simulation for a Multiple Town Model



The figure above shows the Nash Equilibrium when the model has a large number of towns (thirty) and the parameters of the model are set such that $\delta x = 4$, $D = 5$, $R = 3$, and $S = 2$.

Figure 2.16: Methodology for Calculating Distance



To calculate driving distances: (1) Find the population weighted centroid. These are the dots at the center of the polygons in the file above. (2) Calculate major road crossings at state borders. These are the dots along the straight line. (3) Plot a street network data set. Allow GIS to optimize over the shortest route. Notice some routes require significant side road travel, while others can follow one main road to the border. This yields a better approximation of travel distance relative to the "as the crow-flies" distance.

Table 2.1: Sales Tax Summary Statistics by State (April 2010)

Mean Rates Listed as a Percent

	State Rate†	County Rate‡	Local Rate	District Taxes?	Total Rate	N	Neighboring States
Alabama	4.00	1.74	2.20	Yes	8.05	483	FL, GA, MS, TN
Alaska	-	.43	1.07	No	1.50	330	CAN
Arizona	5.60	1.09	1.17	No	7.27	179	CA, MEX, NM, NV, UT
Arkansas	6.00	1.39	0.68	No	8.07	560	LA, MO, MS, OK, TN, TX
California	7.25	1.00	0.04	Yes	8.82	897	AZ, MEX, NV, OR
Colorado	2.90	0.99	2.05	Yes	6.20	308	KS, NE, NM, OK, UT, WY
Connecticut	6.00	-	-	No	6.00	172	MA, NY, RI
Delaware	-	-	-	No	0.00	58	MD, NJ, PA
D.C.	6.00	-	-	No	6.00	1	MD, VA
Florida	6.00	0.74	-	No	6.74	562	AL, GA
Georgia	4.00	2.94	0.001	No	6.94	571	AL, FL, NC, SC, TN
Hawaii	4.00	0.13	-	No	4.13	66	-
Idaho	6.00	0.02	0.02	No	6.04	199	MT, NV, OR, UT, WA, WY
Illinois	6.25	0.34	0.13	Yes	6.90	1528	IN, IA, KY, MO, WI
Indiana	7.00	-	-	No	7.00	651	IL, KY, MI, OH
Iowa	6.00	0.93	0.01	No	6.94	980	IL, MN, MO, NE, SD, WI
Kansas	5.30	0.97	0.31	Yes	6.58	656	CO, MO, NE, OK
Kentucky	6.00	-	-	No	6.00	486	IL, IN, MO, OH, TN, VA, WV
Louisiana	4.00	1.71	2.66	Yes	8.66	367	AR, MS, TX
Maine	5.00	-	-	No	5.00	426	CAN, NH
Maryland	6.00	-	-	No	6.00	297	DC, DE, PA, VA, WV
Massachusetts	6.25	-	-	No	6.25	397	CT, NH, NY, RI, VT
Michigan	6.00	-	-	No	6.00	708	CAN, IN, OH, WI
Minnesota	6.875	0.01	0.01	Yes	6.93	1048	CAN, IA, ND, SD, WI
Mississippi	7.00	-	0.01	No	7.00	309	AL, AR, LA, TN
Missouri	4.225	1.40	0.97	Yes	6.81	977	AR, IA, IL, KS, KY, NE, OK, TN
Montana	-	-	-	No	0.00	236	CAN, ID, ND, SD, WY
Nebraska	5.50	0.00	0.26	No	5.77	629	CO, IA, KS, MO, SD, WY
Nevada	4.60	2.72	-	Yes	7.36	74	AZ, CA, ID, OR, UT
New Hampshire	-	-	-	No	0.00	221	CAN, MA, ME, VT
New Jersey	7.00	-	-	No	7.00	528	DE, NY, PA
New Mexico	4.85	0.62	1.29	No	6.77	167	AZ, CO, MEX, OK, TX
New York	4.00	3.93	0.05	Yes	8.10	1310	CAN, CT, MA, NJ, PA, VT
North Carolina	5.75	2.03	-	Yes	7.78	632	SC, TN, VA
North Dakota	5.00	0.04	0.28	No	5.32	421	CAN, MN, MT, SD
Ohio	5.50	1.28	-	Yes	6.78	1043	IN, KY, MI, PA, WV
Oklahoma	4.50	1.08	2.33	No	7.91	544	AR, CO, KS, MO, NM, TX
Oregon	-	-	-	No	0.00	306	CA, ID, NV, WA
Pennsylvania	6.00	0.06	-	No	6.06	1418	DE, MD, NJ, NY, OH, WV
Rhode Island	7.00	-	-	No	7.00	46	CT, MA
South Carolina	6.00	0.13	0.43	Yes	7.11	344	GA, NC
South Dakota	4.00	-	1.16	Yes	5.16	359	IA, MN, MT, ND, NE, WY
Tennessee	7.00	2.47	0.03	No	9.50	426	AL, AR, GA, KY, MO, MS, NC, VA
Texas	6.25	0.23	1.17	Yes	7.75	1386	AR, LA, MEX, NM, OK
Utah	4.70	1.14	0.25	Yes	6.37	253	AZ, CO, ID, NM, NV, WY
Vermont	6.00	-	0.04	No	6.04	252	CAN, MA, NH, NY
Virginia	4.00	0.93	0.07	No	5.00	335	DC, KY, MD, NC, WV
Washington	6.50	0.31	1.33	Yes	8.25	367	CAN, ID, OR
West Virginia	6.00	-	-	No	6.00	270	KY, MD, OH, PA, VA
Wisconsin	5.00	0.43	-	Yes	5.45	788	IA, IL, MI, MN
Wyoming	4.00	1.10	-	No	5.10	150	CO, ID, MT, NE, SD, UT
United States	5.31	0.81	0.38	Yes	6.61	25,721	CAN, MEX

†Denotes the state statutory rate. Excludes Indian reservations. U.S. values are the mean weighted by number of localities in the state.

‡Weighted by the number of jurisdictions in the county.

Table 2.2: Summary Statistics

Averages with Standard Deviations in ( )

Place Level Data – Full Sample

Variable	Full	Low-Side	High-Side	Same-Tax
Differential in State Tax Rate (R)	-.26	-1.90	1.93	0
	(2.42)	(1.65)	(1.43)	(-)
Differential in County Tax Rate	-.01	-.08	.10	-.09
	(1.03)	(1.13)	(.88)	(.87)
Driving Distance from State Border	60.21	56.33	66.15	50.44
(miles)	(50.44)	(46.55)	(55.71)	(31.20)
Travel Time from State Border (min.)	77.88	72.89	85.57	64.33
	(62.62)	(59.19)	(67.59)	(36.52)
Driving Distance from County Border	7.29	7.30	7.41	5.62
	(6.97)	(7.28)	(6.70)	(3.97)
Crow-Fly Distance from State Border	45.59	43.18	49.36	38.41
	(37.51)	(36.03)	(39.88)	(23.59)
Second Closest State Crow-Fly Distance	92.83	85.34	104.23	74.94
	(59.96)	(55.35)	(65.37)	(31.19)
Number of Neighbors	1.86	1.96	1.71	2.00
	(2.09)	(2.25)	(1.83)	(2.37)
Town Area	6.35	6.22	6.26	9.93
	(23.04)	(24.44)	(21.36)	(33.98)
Town Perimeter	15.41	14.70	15.47	27.32
	(28.66)	(27.37)	(29.05)	(40.80)
Population	10,244	11,264	8,919	9.642
	(90,367)	(109,828)	(58,781)	(33,976)
Senior	15.83	15.75	16.04	14.55
	(7.63)	(7.55)	(7.85)	(5.67)
College	21.97	22.40	21.57	19.55
	(13.93)	(14.44)	(13.31)	(12.24)
Income	37,312	37,920	36,780	33,501
	(18,048)	(18,833)	(17,250)	(12,571)
Work in County (%)	67.62	67.69	68.12	59.69
	(20.43)	(20.60)	(20.06)	(20.66)
Work in State (%)	96.20	96.26	96.12	95.57
	(8.68)	(8.52)	(20.06)	(10.26)
Obama Vote Share	43.99	45.22	42.62	40.22
	(13.66)	(13.16)	(13.90)	(16.44)
Local Rate	0.62	0.73	0.44	1.20
	(1.08)	(1.24)	(0.70)	(1.53)
Local + District Rate	0.71	0.83	0.51	1.27
	(1.13)	(1.30)	(0.75)	(1.64)
Local + District + County Rate	1.76	2.19	1.09	3.14
	(1.45)	(1.55)	(0.91)	(1.44)
Sample Size	16,799	9331	6952	516

High-side means that the nearest state to the location is a low-tax state.

Table 2.3: Average Tax Rates by Driving Distance

	Low-Tax Side					High-Tax Side				
Miles to Nearest State	100+	75-100	50-75	25-50	0-25	0-25	25-50	50-75	75-100	100+
Total Tax	7.56	6.77	6.33	6.26	6.25	7.04	6.97	7.11	7.23	8.01
	(2.42)	(2.71)	(2.36)	(2.28)	(2.25)	(1.31)	(1.24)	(1.40)	(1.57)	(1.61)
State Tax	5.64	4.88	4.90	4.85	4.79	6.20	6.24	6.21	6.17	6.68
	(1.55)	(1.08)	(.97)	(.97)	(.92)	(.60)	(.61)	(.63)	(.72)	(1.11)
County Tax	.91	1.28	1.39	1.49	1.44	.59	.60	.55	.40	.48
	(.97)	(1.30)	(1.30)	(1.32)	(1.19)	(.81)	(.84)	(.78)	(.60)	(.55)
Local Tax	.91	.91	.68	.64	.66	.38	.33	.41	.49	.64
	(1.29)	(1.44)	(1.27)	(1.17)	(1.15)	(.67)	(.62)	(.65)	(.72)	(.80)
District Tax	.10	.08	.07	.12	.13	.08	.07	.08	.09	.05
	(.36)	(.36)	(.30)	(.41)	(.38)	(.26)	(.24)	(.26)	(.28)	(.21)
Local + District	1.01	.99	.75	.76	.79	.46	.38	.49	.57	.68
	(1.35)	(1.52)	(1.29)	(1.23)	(1.21)	(1.05)	(.66)	(.73)	(.78)	(.83)
County + Local + District	1.95	2.27	2.14	2.32	2.23	1.06	1.00	1.04	1.01	1.32
	(1.40)	(1.70)	(1.60)	(1.59)	(1.48)	(1.01)	(.96)	(.93)	(.87)	(.69)
Sample Size	1368	1250	1675	2334	2704	1793	1614	1191	868	1486

Sample only includes states that allow for Local Option Taxes.

Table 2.4: RD Estimates

	(1) Local Tax Rates	(2) Local + District Tax Rates	(3) County Tax Rates	(4) Total Local Rates (L+D+C)	(5) Local Tax Rates	(6) Local + District Tax Rates
No Controls	-.531***	-.352***	-.521***	-1.130***	-.176**	-.043
	(.055)	(.034)	(.124)	(.068)	(0.075)	(.092)
Controls	-.505***	-.328***	-.634***	-1.242***	-.205***	-.123
	(.055)	(.034)	(.179)	(.068)	(.074)	(.090)
Controls and State Dummies	-.434***	-.324***	-.650***	-1.247***	-.154**	-.105
	(.055)	(.034)	(.178)	(.066)	(.073)	(.087)
Optimal Bandwidth	46.12	97.75	65.30	38.90	12.21	7.26
Type of Border	State	State	State	State	County	County
Unit of Observation	Locality	Locality	County	Locality	Locality	Locality
Border Counties Included?	Yes	Yes	Yes	Yes	No	No

Results represent the effect on the level of local tax rates in a border town from the high-tax side relative to an identical border town on the low-tax side. Columns (1) - (4) estimate the level effect at state borders for various local tax rates – local only, local plus district, county only, and local plus district plus county rates. The running variable is the driving distance from the locality to the state border with the exception of column (3), where the running variable is the driving distance from the county centroid to the state border. Columns (5) and (6) estimate the level effect of local tax rates and local plus district tax rates at county borders, where the county borders specifications exclude counties along state borders. ***99%, **95%, *90%

Table 2.5: Regression Coefficients for the Estimating Equation

Variable	(1)	(2)	(3)	(4)
Same (S)	.560	.640*	.403**	.350**
	(.403)	(.375)	(.162)	(.145)
High (H)	.066	.240***	.226***	.209***
	(.095)	(.089)	(.060)	(.062)
Notch (R)	-.174***	-.347***	-.095***	-.118***
	(.039)	(.038)	(.025)	(.025)
$H \cdot R$	.129***	.302***	.027	.063**
	(.045)	(.043)	(.030)	(.030)
Distance (d)	1.346**	1.394**	-.227	.279
	(.643)	(.628)	(.384)	(.414)
$S \cdot d$	-.8168	-8.445	-4.541	-3.626
	(-7.033)	(6.497)	(2.878)	(2.599)
$H \cdot d$	-1.436*	-.578	-.810	-1.294**
	(.843)	(.800)	(.532)	(.554)
$R \cdot d$	.888***	.766***	.443***	.508***
	(.295)	(.285)	(.168)	(.166)
$H \cdot R \cdot d$	-1.246***	-1.283***	-.344	.358*
	(.355)	(.343)	(.215)	(.211)
d^2	-4.137**	-3.051*	.228	-.705
	(1.709)	(1.698)	(1.040)	(1.087)
$S \cdot d^2$	61.003	62.433*	28.903*	22.476
	(37.526)	(34.504)	(15.979)	(14.447)
$H \cdot d^2$	2.853	-.081	1.277	2.592*
	(-2.298)	(2.197)	(1.455)	(1.501)
$R \cdot d^2$	-2.077***	-1.457**	-.895**	-.993**
	(.719)	(.715)	(.429)	(.425)
$H \cdot R \cdot d^2$	3.733***	3.383***	.991*	.711
	(.891)	(.870)	(.549)	(.542)
d^3	5.295***	3.079*	-.082	.404
	(1.811)	(1.843)	(1.131)	(1.152)
$S \cdot d^3$	-158.815*	-158.969**	-66.955*	-54.118*
	(81.579)	(74.585)	(35.64)	(32.082)
$H \cdot d^3$	-3.493	.165	-1.09	-2.193
	(2.492)	(2.395)	(1.582)	(1.611)
$R \cdot d^3$	2.059***	.993	.679	.652
	(.726)	(.742)	(.454)	(.449)
$H \cdot R \cdot d^3$	-3.747***	-2.837***	-.870	-.371
	(.913)	(.898)	(.573)	(.572)
d^4	-2.491***	-1.170	.055	.010
	(.802)	(.844)	(.517)	(.519)
$S \cdot d^4$	168.433**	164.557**	64.788*	54.247*
	(76.777)	(69.842)	(34.107)	(30.510)
$H \cdot d^4$	1.923*	.053	.387	.769
	(1.132)	(1.096)	(.722)	(.730)
$R \cdot d^4$	-.850***	-.273	-.203	-.147
	(.315)	.332	(.205)	(.203)
$H \cdot R \cdot d^4$	1.425***	.910	.281	.012
	(.400)	(.398)	(.255)	(.257)

Table 2.5 Continued

Variable	(1)	(2)	(3)	(4)
d^5	.366***	.134	-.016	-.028
	(.125)	(.136)	(.083)	(.083)
$S \cdot d^5$	-61.463**	-58.759**	-22.133*	-18.966*
	(26.006)	(23.557)	(11.668)	(10.362)
$H \cdot d^5$	-.331*	-.017	-.046	-.093
	(.180)	(.176)	(.116)	(.117)
$R \cdot d^5$	.117**	.021	.020	.007
	(.048)	(.053)	(.033)	(.032)
$H \cdot R \cdot d^5$	.178***	-.093	-.029	.017
	(.062)	(.063)	(.040)	(.041)
Number of		.070***	.034***	.054***
Neighbors		(.008)	(.005)	(.006)
Area		-.540***	-.446***	-.493***
		(.115)	(.077)	(.063)
Perimeter		.724***	.479**	.413***
		(.093)	(.057)	(.053)
Population		-.031	.014	-.004
		(.036)	(.019)	(.011)
Senior		-.012***	-.001	.001
		(.001)	(.001)	(.001)
College		.006***	.007***	.006***
		(.001)	(.001)	(.001)
Income		-.783***	-.215***	-.286***
		(.077)	(.054)	(.058)
International		-1.450***	-.429***	-.567***
		(.068)	(.087)	(.090)
Ocean		-1.370***	.163***	.220***
		(.029)	(.019)	(.022)
County Worker		.005***	.001***	.001***
		(.001)	(.000)	(.000)
State Worker		-.004***	.005***	.005***
		(.001)	(.001)	(.001)
Obama Vote		-.015***	.004***	.007***
		(.001)	(.001)	(.001)
County Tax (IV)				-.562***
				(.098)
Constant		1.127***	1.470***	2.345***
		(.123)	(.106)	(.179)
Observations	16,799	16,781	16,781	16,781
R^2	.054	.170	.625	.659
State FE	N	N	Y	Y
Border FE	N	N	N	N
IV?	N	N	N	Y
Marginal Effect:	.088	.291***	-.153***	-.102**
Low Side	(.059)	(.060)	(.043)	(.044)
Marginal Effect:	.046	.125***	-.209***	-.197***
High Side	(.035)	(.042)	(.032)	(.032)
Marginal Effect:	.559	.589	-.171	-.199
Same Side	(.491)	(.449)	(.195)	(.165)

The dependent variable is the local plus district tax rate. Standard errors are robust. ***99%, **95%, *90%

Table 2.6: Mean Derivatives for Several Specifications

Mean Derivative	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Low-Tax State	-.103**	-.116***	-.133**	-.240***	-.055***	-.094**	-.086	-.286***	-.135**	-.077*
	(.044)	(.042)	(.058)	(.058)	(.015)	(.049)	(.070)	(.054)	(.064)	(.041)
High-Tax State	-.197***	-.181***	-.198***	-.211***	-.102***	-.209***	-.239***	+.097*	-.113***	-.179***
	(.032)	(.030)	(.033)	(.039)	(.014)	(.035)	(.041)	(.058)	(.034)	(.032)
Same-Tax State	-.199	-.227	-.198	-.186	-.045	-.185	-.204	-.126	-.221	-.107
	(.165)	(.162)	(.164)	(.167)	(.069)	(.174)	(.210)	(.123)	(.198)	(.144)
Variable†	L+D	L+D	L+D	L+D	L+D	L+D	L+D	L+D	L	L+D+C
Bandwidth	N	N	N	N	N	N	N	N	N	N
Restriction	N	Binary Treat-ment	No Intern.	No Ocean	Time	MS, NC, NV, WI	Nbr. LOST	Nbr. No LOST	N	N
Observations	16,781	16,781	15,642	14,255	16,781	15243	12,516	3126	14,563	16,781

The marginal effects represent a **per 100 mile** change (i.e., per mile, the units are basis points), except for (5), which is per hour.

(1) is derived from column 4 of Table 2.5. (2) uses a binary treatment. (3) eliminates towns where the closest border is an international one. (4) eliminates towns proximate to the ocean. (5) uses driving time instead of driving distance. (6) drops states where the primary local taxes are district taxes. (7) utilizes only towns where the nearest neighboring state allows for LOST. (8) includes observations where the nearest neighbor is a state that does not allow LOST. (9) uses local taxes as the left side variable. (10) uses city plus district plus county taxes as the left-side variable and controls for the county X's. †L is local taxes, D is district taxes, and C is county taxes.

Standard errors are robust and calculated using the Delta Method. ***99%, **95%, *90%.

Table 2.7: Mean Derivatives for Several Specifications

Mean Derivative	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Low-Tax State	-.103**	-.049	-.083**	-.051	-.704***	-.203***	-.203***	-.118**	-.180**	-.695*
	(.044)	(.030)	(.046)	(.045)	(.152)	(.072)	(.070)	(.054)	(.094)	(.398)
High-Tax State	-.197***	-.150***	-.181***	-.169***	-.230***	-.129**	-.490***	-.190***	-.194***	-.490
	(.032)	(.024)	(.031)	(.031)	(.061)	(.058)	(.149)	(.041)	(.072)	(.330)
Same-Tax State	-.199	-.012	-.212	-.204	-.513	-.486*	-.380**	-.199	-.395	-.051
	(.165)	(.128)	(.149)	(.161)	(.460)	(.251)	(.179)	(.167)	(.342)	(2.616)
Variable†	L+D	L+D	L+D	L+D	L+D	L+D	L+D	L+D	L+D	L+D
Bandwidth	N	N	N	N	N	N	N	152	97.75	40
Restrict?	No	Degree 3	Degree 7	State Weights	Pop. Weights	Interact All	Border & St. FE	N	N	N
Observations	16,781	16,781	16,781	16,781	16,781	16,781	16,781	16,124	13,739	7236

The marginal effects represent a **per 100 mile** change (i.e., per mile, the units are basis points).

(1) is derived from column 4 of Table 2.5. (2) uses a cubic polynomial. (3) uses an order seven polynomial. (4) weights each state equally in the regression. (5) weights by the population of the locality. (6) interacts the control variables and state fixed effects with a polynomial in distance along with a full set of interactions with the high- and same-side dummies. (7) includes state fixed effects and state-border-pair fixed effects. (8) restricts the sample to observations within 150 miles of the nearest border. (9) restricts the sample to observations within the optimal bandwidth from the local linear regression. (10) restricts the sample to observations within forty miles of the nearest border. †L is local taxes, D is district taxes, and C is county taxes.

Standard errors are robust and calculated using the Delta Method. ***99%, **95%, *90%

Table 2.8: State by State Marginal Effects

State	Low Side	High Side	Same Side
Alabama	.001		-.056
	(.001)		(.366)
Arizona	-.563	-.014	
	(.366)	(.436)	
Arkansas	-.981***	-.544**	
	(.303)	(.249)	
California	-.016	.002	
	(.027)	(.019)	
Colorado	1.358***		
	(.210)		
Georgia	.006		-.084
	(.008)		(.162)
Idaho	.025	.058	
	(.017)	(.074)	
Illinois	-.434***	-.517***	
	(.154)	(.093)	
Iowa	-.042**	-.001	
	(.019)	(.038)	
Kansas	.038	-.555***	
	(.136)	(.136)	
Louisiana	-1.189***		
	(.382)		
Minnesota	-.199***	-.041**	
	(.057)	(.021)	
Mississippi		.012	-.131
		(.010)	(.224)
Missouri	.051		
	(.190)		
Nebraska	-.880***	-.200	
	(.312)	(.164)	
Nevada	-.008***	-.108***	
	(.002)	(.016)	
New Mexico	-.012	-1.244*	
	(.391)	(.676)	
New York	-.378***		
	(.081)		
North Carolina	-.139***	-.171***	
	(.032)	(.020)	
North Dakota	-.636**	.564*	
	(.278)	(.308)	
Ohio	-.026		
	(.024)		
Oklahoma	.905***	7.936*	
	(.183)	(4.198)	
South Carolina		-.085	
		(.218)	
South Dakota	-.528	-.265	-.591
	(.325)	(1.342)	(.513)
Tennessee		.027	-.003
		(.026)	(.069)
Texas	-.103	.027	
	(.086)	(.036)	
Utah	-.310	-.137	
	(.369)	(.106)	
Vermont	-.017	1.235**	
	(.055)	(.561)	
Virginia	-.117		
	(.122)		
West Virginia	.311***	.114	
	(.103)	(.085)	
Wisconsin	-.095***		
	(.018)		

The marginal effects represent a **per 100 mile** change (i.e., per mile, the units are basis points). The regression specification allows for state fixed effects to be interacted with the distance function such that the gradient is allowed to vary by state.

Standard errors are robust and calculated using the Delta Method. ***99%, **95%, *90%

Table 2.9: Average Marginal Effects by Size of the Treatment/Discontinuity

Specification	IV Regression of Rates			Probit Binary Variable		
	Low-Tax Side	High-Tax Side	Same-Tax	Low-Tax Side	High-Tax Side	Same-Tax
Not Conditioned on Notch	-.103**	-.197***	-.199	-.469***	-.435***	-.135
(R)	(.044)	(.032)	(.165)	(.087)	(.092)	(.467)
5th Percentile of R	-.029	-.234***		-.348***	-.663***	
$R_{Low} = -.25$; $R_{High} = .25$	(.059)	(.049)		(.126)	(.119)	
10th Percentile of R	-.041	-.234***		-.365***	-.663***	
$R_{Low} = -.50$; $R_{High} = .25$	(.055)	(.049)		(.118)	(.119)	
20th Percentile of R	-.054	-.217***		-.382***	-.571***	
$R_{Low} = -.75$; $R_{High} = .875$	(.051)	(.041)		(.110)	(.102)	
30th Percentile of R	-.060	-.206***		-.391***	-.516***	
$R_{Low} = -.875$; $R_{High} = 1.25$	(.050)	(.037)		(.106)	(.096)	
40th Percentile of R	-.079*	-.192***		-.416***	-.443***	
$R_{Low} = -1.25$; $R_{High} = 1.75$	(.046)	(.033)		(.097)	(.095)	
50th Percentile of R	-.105**	-.192***		-.450***	-.443***	
$R_{Low} = -1.75$; $R_{High} = 1.75$	(.044)	(.033)		(.090)	(.095)	
60th Percentile R	-.117***	-.189***		-.467***	-.424***	
$R_{Low} = -2$; $R_{High} = 1.875$	(.045)	(.032)		(.088)	(.096)	
70th Percentile of R	-.117***	-.185***		-.467***	-.402***	
$R_{Low} = -2$; $R_{High} = 2.025$	(.045)	(.032)		(.088)	(.097)	
80th Percentile of R	-.155***	-.178***		-.518***	-.369***	
$R_{Low} = -2.75$; $R_{High} = 2.25$	(.052)	(.032)		(.095)	(.101)	
90th Percentile of R	-.168***	-.139***		-.534***	-.163	
$R_{Low} = -3$; $R_{High} = 3.65$	(.056)	(.040)		(.101)	(.146)	
95th Percentile of R	-.275***	-.073		-.678***	.183	
$R_{Low} = -5.125$; $R_{High} = 6$	(.102)	(.073)		(.179)	(.256)	

The marginal effects represent a **per 100 mile** change (i.e., per mile, the units are basis points).

Standard Errors are robust and calculated using the Delta Method. ***99%, **95%, *90%

Table 2.10: Mean Derivatives for Multiple Borders

Mean Derivative	(1)	(2)	(3)	(4)	(5)
Low-Tax State	-.069	-.127**	.029	-.461	-.396
	(.065)	(.065)	(.067)	(.458)	(.462)
High-Tax State	-.181***	-.189***	-.129**	-1.060**	-1.206***
	(.043)	(.044)	(.061)	(.417)	(.410)
Same-Tax State	-.111	-.162	.010	-1.003***	-.935***
	(0.192)	(.193)	(.192)	(.174)	(.173)
Marginal Effects	State	State	State	County	County
1st State	Y	Y	Y	N	Y
2nd State	N	Y	Y	N	Y
County Border	N	N	Y	Y	Y
Border Counties?	Y	Y	Y	N	Y

The marginal effects represent a **per 100 mile** change (i.e., per mile, the units are basis points).

(1) is derived from column (4) of Table 2.5 but uses the "as the crow-flies" distance. (2) adds a polynomial in distance from the second border plus interactions. (3) uses polynomials in distance from the closest state border, the second closest state border and the closest county border plus the appropriate interactions. (4) includes a polynomial to the closest county border and drops state border counties. (5) is the same as (3) but calculates marginal effects from the county border.

Standard errors are robust and calculated using the Delta Method. ***99%, **95%, *90%

CHAPTER III

Games Within Borders: Are Geographically Differentiated Taxes Optimal?

3.1 Introduction

"Jeux Sans Frontières" – the title of a classic commodity tax competition paper (Kanbur and Keen 1993) – translates from French to "Games Without Borders." What happens if countries play games *within* borders by setting geographically differentiated tax rates?

Geographic borders between different states and countries create a discontinuous tax policy. Forty-five American states impose a sales tax ranging between 2.9% and 7.25%. Among European Union member states, the Value Added Tax (VAT) ranges from 15% to 27%. Given these disparities in tax rates, consumers have large incentives to engage in cross-border shopping. As cross-border shopping is distortionary, it is important to know if a uniform tax policy is socially optimal compared to a policy that differentiates tax rates based on geographic location.

The Mexican VAT features a differentiated tax rate depending on proximity to the United States border. The standard tax rate in Mexico is 16%, but goods purchased within twenty kilometers of the lower tax United States border are assessed a rate of 11%. Davis (2011) finds that there is also a modest, but statistically significant distortion that encourages Mexicans living in the high-tax zone to shop or locate in the preferred tax zone. If the size of the distortion induced within Mexico is small

relative to the decrease in the size of the distortion at the international border, then such a policy may be optimal for Mexico. In addition, the Netherlands differentiated gasoline taxes near the German border in the 1990s and Italy differentiated gas taxes near the Slovenian border in the early 2000s (Kessing 2008). Duty free shopping at airports is the most visible representation of reduced tax border-zones – where the border zone is sufficiently small and restricted to cross-border shoppers.

In the tax competition literature, differences in tax rates arise because of differences in size or public good preferences. However, in all these models, the state competes over only *one* sales tax rate. In this paper, acknowledging that tax competition will imply differentials in tax rates at borders, I consider what a state's optimal policy is when the state can select *two* sales tax rates within its borders – a rate near the state border and a rate away from the border.[1] The evidence suggests that in addition to country size and population, the spatial composition of regions within a state is an essential factor to consider when setting tax rates.

In the context of a Haufler (1996) model that combines elements from both Mintz and Tulkens (1986) and Kanbur and Keen (1993), I demonstrate that from a welfare maximizing state planner's perspective, the optimal tax system in the presence of borders is almost always a geographically differentiated tax. The model that I will outline in this paper is a partial equilibrium model where one state selects an optimal tax policy taking as given an exogenous tax rate in the neighboring state. If regions within a state are homogeneous, then tax differentiation within a state is most likely to arise where there are differences between states. Thus, solving for a Nash equilibrium when states each choose two tax rates would not be interesting if states were identical and is uninformatively complicated when states are different. This simplification has the advantage of allowing for welfare maximization.

Geographic differentiation of tax rates is not surprising, but the specific pattern

[1] The word "optimal" in this paper will always refer to what is "optimal from the prospective of a state planner" rather than what is "globally optimal" unless specifically noted otherwise.

of geographic differentiation is novel. A shocking result is that a tax rat in the the border region that is higher away from the border is sometimes optimal – a result not attainable with Leviathan governments. For a low-tax state, if the elasticity of cross-border shopping is sufficiently small, the social planner will want to raise taxes in the region closest to the border in an effort to export the tax burden to foreign consumers. In high-tax states, this possibility follows from a desire to equalize consumption across the two regions that outweighs the desire to expand the tax base. Inequality in consumption is a concern for the social planner in a high-tax state where residents near the border have the opportunity to increase consumption through cross-border shopping – but would play no role if the government cared solely about revenue. In these circumstances, the state planner may want to raise taxes at the border.

Policymakers often discuss inequality, expanding the tax base, and exporting the tax burden to "foreigners" in the context of the tax system. As is noted by Slemrod (2007), tax evasion is not uniform across the entire population and this can create serious inequities as the people who evade taxes realize more income. Applied to the commodity taxes, cross-border shopping is primarily done by residents near borders and results in higher consumption for those who evade. Examples of the desire to export taxes and to expand the tax base are common. For example, municipalities on the low-tax side of a border set larger local option sales tax rates than municipalities on the high-tax sides of borders. In doing so, municipalities on the low-tax side of the border can export some of the sales tax burden to cross-border shoppers. Additionally, high hotel taxes in cities tourists frequent are another mechanism to export the tax burden. This exporting effect contrasts with the desire of municipalities to expand the tax base. In 2009, the Arkansas state legislature passed a law that requires the cigarette tax in border towns be equal to the cigarette tax in the neighboring town outside of the state. The law only applies if the resulting rate is less than the statutory Arkansas state rate and if the border town in Arkansas and the neighboring state have

a population of more than 5000 people. The law also applies to retailers less than 300 feet from the border.[2] The legislation implemented in Arkansas is a clear example of a state trying to reduce cross-border shopping in order to expand the tax base.

3.2 Tax Competition and Cross-Border Shopping

The literature on tax competition has developed with an emphasis on trying to explain asymmetries in tax rates among competing jurisdictions. Haufler (1996) implies that a country may have higher tax rates than the neighbor because of a stronger preference for public good provision. This is in contrast to models such as Kanbur and Keen (1993), Nielsen (2001) and Trandel (1994), which focus on country size or population as an explanation for variation in equilibrium tax rates. These models only allow for one tax rate within a state.

Several studies have focused precisely on the role of distance to a competing jurisdiction as a key variable of interest. Lovenheim (2008), Merriman (2010), and Harding, Leibtag and Lovenheim (Forthcoming) show that the elasticity of demand, tax evasion, and the tax incidence are a function of distance from the border. These papers clearly indicate that geographic space shapes the responsiveness of individuals. The results, together with the theoretical results that tax differentials will exist at state borders, provide a powerful argument for geographic differentiation of tax rates.

Although discriminatory taxation has been studied in the context of multiple tax bases (Janeba and Peters 1999; Keen 2001), relatively few papers have studied the geographic differentiation of taxes. Two known exceptions are Kessing (2008) and Nielsen (2010), which study geographically differentiated taxes in the context of a Kanbur and Keen (1993) model. Kessing (2008) and Nielsen (2010) differ from this paper by considering the problem in the context of revenue maximizing governments

[2]The law providing for this provision is Act 180 of the Regular Session of the 87th General Assembly.

and only for high-tax states.

This paper builds upon the model in Haufler (1996) because differences in public goods provision is a natural starting point for explaining how differences in tax rates arise across borders and, more importantly, because a model of welfare maximizing governments (without a separable utility function) will allow for a rich treatment of the problem.

3.3 Model

3.3.1 Setup of the Model

The model features two states indexed by $k = H, N$. The home state (H) has two identical regions[3] indexed $i = A, B$, denoting the away region (A) and border region (B). The neighboring state (N) is identical in terms of the number of regions, but will never set differentiated tax rates. The geographic setting is depicted in Figure 1.

3.3.2 Consumers and Firms

A representative consumer lives at the center of each region. Each consumer is endowed with M dollars of income. The consumer has the choice to purchase a quantity of the consumption good, c, in her home region or in a neighboring region or state. Let *superscripts* on c index where the person lives and *subscripts* index the region that the good is purchased in. Thus, c^i_j denotes the quantity of the consumption good that the resident of jurisdiction i purchases in region j. Let c^i denote the total consumption of a resident of region i. Then, $c^i = \sum_j c^i_j$ for all i and that $c^i = c^i_i$ if $c^i_j = 0$ for all j.

Firms providing the private good are assumed to be located exogenously at any

[3]The regions could be two equally sized jurisdictions. The word "region" need not imply a non-governing entity. As an alternative, the two regions can be viewed as two towns – with one town being the preferred tax region – as in the case of Arkansas.

point where consumers shop. Firms are perfectly competitive and set price equal to the marginal cost of production. Increases in demand in a particular region resulting from cross-border shopping will not alter the pre-tax price. The pre-tax price is normalized to one.

Transportation costs constrain consumers from purchasing goods in another jurisdiction. If the individual decides to purchase in her own region, she simply goes to the store at the point of the line corresponding to where she lives and no transportation costs are incurred. If shopping in a neighboring region, the individual faces a transportation cost function denoted $D_i(c_j^i)$ for j not equal to i.[4] I assume that the marginal cost of purchasing the first unit abroad is zero for residents of each region. This will guarantee that if taxes are not equal, the individual in the high-tax region will immediately begin to purchase some of the goods in the low-tax region. In addition, $D_i'(c_j^i) > 0$ and $D_i''(c_j^i) > 0$ for $c_j^i > 0$. Additionally, $D_i(0) = 0$, $D_i'(0) = 0$ and $D_i''(0) > 0$. These assumptions on the transportation cost function guarantee that as the tax differential between the regions increase, the individual will purchase more of his consumption from the low-tax region. For notational convenience, let S_i denote the transportation cost function for crossing a state border. I assume $S_i = D_i$. Individuals are willing to travel, at most, one jurisdiction over to make cross-border purchases.

A convex transportation cost function can be justified by a composite consumption good that represents consumption goods that are heterogeneous in terms of the ease of their transportability. More importantly, a convex transportation cost makes the representative agent problem behave like a Hotelling-style model where consumers are along a continuum. Convex transportation cost functions implicitly underlie models of cross-border shopping where agents live along a continuum and are heterogeneous in their distance to the border. The convex transportation cost function is a modeling

[4]The transportation cost is independent of distance. Each consumer lives an identical amount of distance from the nearest region or state border.

technique for mimicking a Kanbur and Keen (1993) model, while also allowing for welfare maximizing governments with a representative agent. Thus, the representative agent will proxy for heterogeneous consumers located at different distances from the border.

Consumers have preferences over aggregate consumption and the publicly provided good. The functional form of the utility function is identical across all individuals. Preferences are given by the utility function $U^i(c^i, G)$, which is strictly quasi-concave. Individuals can only choose how much of the consumption good to purchase at home and abroad.

3.3.3 The State Planner's Problem

The home state sets taxes, t_i, on goods purchased in region i to fund a public good. Taxes are levied according to the origin principle, which implies that if an individual crosses a border, he will pay the tax to the jurisdiction of purchase.[5] The paper assumes that states provide a state-level public good, meaning that all the revenue raised within the state is aggregated to provide a public good that is uniform across all regions.[6] The assumption of a state public good merits some discussion. The model considers whether a state should implement discriminatory taxation in geographic space. The model does not consider what are the optimal geographically differentiated tax rates when towns are allowed to set tax rates and keep revenue earned in the town. As discussed, examples of border-zones include the Mexican Value Added Tax, the Netherlands' and Italy's historical gas taxes, duty-free shops, and the Arkansas cigarette tax. In these examples, taxes are differentiated within the state, but the provision of public goods is not or was not related to the revenue

[5]The sales tax in the United States is levied *de facto* according to the origin principle. The use tax is notoriously under-enforced. Because the use tax is often evaded, taxes are implicitly paid based on the location of purchase rather than the destination of the sale. For a discussion of how enforcement policies and closed borders changes the optimal tax policy, please see the appendix.

[6]When the public good is provided at the state level, this means that the size of the tax base in one region is independent of the level of public good provision that the region receives.

intake from each region within the state. Furthermore, the models of Kessing (2008) and Nielsen (2010) make the same assumption. For these reasons, I assume a state provided public good common to both regions. The appendix discusses a model with local provision of public goods.

Letting R_i denote the total revenue raised in jurisdiction i and letting $0 \leq \rho \leq 1$ denote the rate of transformation between revenues and expenditures, $G = \sum_{i=A,B} \rho R_i(t_i)$. I assume that the production technology is unity, so that $\rho = 1$. The home state and the neighboring state differ in their preferences for public goods.

The home state selects the optimal tax rates for its two localities by maximizing the social welfare of its two residents. The government of State H chooses tax rates for Regions A and B taking as given an exogenously fixed tax rate of $\bar{t}$ in the neighboring state. Subject to the government and individual budget constraints, State H sets tax rates by maximizing:

$$W = \sum_{i=A,B} U^i(c^i, G). \tag{3.1}$$

Because the utility function is concave, the social welfare function captures equity concerns. The social welfare function must be continuous and strictly quasi-concave in the strategies. The first order conditions of the government's constrained maximization problem must be continuous when a region switches from being a high-tax jurisdiction to a low-tax jurisdiction. This will be the case if the transportation cost function, described above, is identical across all regions. I assume weak conditions hold such that the first order conditions are concave in the strategies. Continuous and concave first order conditions and quasi-concavity of the objective function will guarantee an interior solution.

Because producer prices are equal in all jurisdictions and the incidence of the tax is fully passed through to the consumer, a resident of a high-tax region will cross-border shop until the marginal benefit (tax savings) of doing so is equal to the marginal cost (transport). As in Haufler (1996), in any equilibrium, the following

consumer arbitrage condition holds:

$$D_i'(c_j^i) = \begin{cases} t_i - t_j & \text{for } t_i > t_j \\ 0 & \text{for } t_i \leq t_j. \end{cases} \tag{3.2}$$

The function in Equation (3.2) implicitly defines the level of cross-border shopping c_j^i:

$$c_j^i(t_i - t_j) = \begin{cases} (D_i')^{-1} & \text{for } t_i > t_j \\ 0 & \text{for } t_i \leq t_j. \end{cases} \tag{3.3}$$

For a given pattern of tax rates, the individual consumer's budget constraint will imply that income is equal to consumption purchases in the location of residence plus consumption purchases in neighboring low-tax region and/or state. Differentiating equation (3.3) using the inverse function theorem and totally differentiating the budget constraints yields the following derivatives for the high-tax region $(t_i \geq t_j)$:

$$\begin{array}{lll} \frac{\partial c_j^i}{\partial t_i} = \frac{1}{D_i''} > 0 & \frac{\partial c_i^i}{\partial t_i} = -\frac{c_i^i}{1+t_i} - \frac{1}{D_i''} < 0 & \frac{\partial c^i}{\partial t_i} = -\frac{c_i^i}{1+t_i} < 0 \\ \frac{\partial c_j^i}{\partial t_j} = -\frac{1}{D_i''} < 0 & \frac{\partial c_i^i}{\partial t_j} = -\frac{c_j^i}{1+t_i} + \frac{1}{D_i''} \gtreqless 0 & \frac{\partial c^i}{\partial t_j} = -\frac{c_j^i}{1+t_i} < 0 \end{array} \tag{3.4}$$

and for the low-tax region $(t_i < t_j)$:

$$\frac{\partial c^i}{\partial t_i} = -\frac{c^i}{1+t_i} < 0 \quad \frac{\partial c^i}{\partial t_j} = 0. \tag{3.5}$$

3.4 Results

3.4.1 Solving the Model

I derive the optimal tax rates assuming that State N sets a fixed and uniform tax rate of $\bar{t}$ and that this state does not respond competitively to geographic differentiation in the home state. Therefore, the equations that follow characterize the

home *state's optimal response* to a fixed tax rate rather than a *Nash equilibrium*. Tax competition, especially with geographically differentiated rates on both sides of the border would add additional complexities and cases. Introducing tax competition would require additional simplifications – such as revenue maximization – that will eliminate the interaction of the effects presented.[7]

To solve this problem, it is important that I specify the direction of cross border shopping in several cases. The first case is when State H has a low preference for public goods and wants its tax rate at the border strictly less than State N. The second case is when State H has a high preference for public goods and wants its tax rate at the border strictly greater than State N. Within each case, there are two sub-cases to consider: the border region of State H sets higher rates than the away region and vice-versa. Sub-case 1 will denote where State H's away region sets a higher rate than the border region, while Sub-case 2 will denote the reverse of this. The four possible scenarios are presented in Figure 2.

In all of the cases, State H selects t_A and t_B taking as given and fixed $\bar{t}$ by maximizing equation (3.1) subject to the constraints below.

Case Low – The Case of a State with Low Public Good Preferences: $t_B \leq \bar{t}$

In this scenario, the neighboring state is a high-tax state and the home state has a preference for lower levels of public goods. The problem facing the low-tax state's social planner is that she wishes to determine the optimal combination of regional tax rates given that the neighbor has selected a high fixed rate of $\bar{t} \geq t_B$.

Sub-Case 1 – Preferential Tax Rates at the Border: $t_A \geq t_B$.

Accounting for cross-border shopping that is inward to region B from both sides, the individual budget constraints for Regions A and B and the government budget

[7] An alternative way to proceed would be to maintain welfare maximization, start from a symmetric equilibrium where all tax rates are identical and then shock the public good preferences in one state. With only one tax rate, these results are extremely complex.

constraint are as follows:

$$
\begin{aligned}
(1+t_A)c_A^A + (1+t_B)c_B^A + D_A(c_B^A) &= M \\
(1+t_B)c^B &= M \\
G = t_A c_A^A + t_B c_B^A + t_B c^B + t_B c_B^N.
\end{aligned}
\tag{3.6}
$$

I can use Equations (3.4) and (3.5) to solve this problem. As shown in the appendix, the two first order conditions can be solved for the marginal benefit of the public good ($U_G^A + U_G^B$) for all i, which can then be equated. The following condition must hold at an optimum:

$$
\frac{U_C^A \frac{c_A^A}{1+t_A}}{MR_A} = \frac{U_C^B \frac{c^B}{1+t_B}}{MR_B} + \frac{U_C^A \frac{c_B^A}{1+t_A}}{MR_B}
\tag{3.7}
$$

where the marginal revenue from a change in a tax rate is denoted:

$$
\begin{aligned}
MR_A &= \tfrac{c_A^A}{1+t_A} - \tfrac{t_A - t_B}{D_A''} \\
MR_B &= \tfrac{c^B}{1+t_B} + \tfrac{c_B^A}{1+t_A} - \tfrac{t_B - t_A}{D_A''} - \tfrac{t_B}{S_N''} + c_B^N.
\end{aligned}
\tag{3.8}
$$

Sub-Case 2 – Higher Taxes at the Border: $t_B \geq t_A$.

Accounting for cross-region shopping from Region B to A and cross-state shopping from the neighboring state, the individual budget constraints for Regions A and B and the government budget constraint are as follows:

$$
\begin{aligned}
(1+t_A)c^A &= M \\
(1+t_A)c_A^B + (1+t_B)c_B^B + D_B(c_A^B) &= M \\
G = t_A c^A + t_A c_A^B + t_B c_B^B + t_B c_B^N.
\end{aligned}
\tag{3.9}
$$

Solving the first order conditions for the marginal benefit of the public good,

implies:

$$\frac{U_C^A \frac{c^A}{1+t_A}}{MR_A} + \frac{U_C^B \frac{c_A^B}{1+t_B}}{MR_A} = \frac{U_C^B \frac{c_B^B}{1+t_B}}{MR_B} \tag{3.10}$$

where the marginal revenue from a change in a tax rate is denoted:

$$\begin{aligned} MR_A &= \tfrac{c^A}{1+t_A} + \tfrac{c_A^B}{1+t_A} - \tfrac{t_A - t_B}{D_B''} \\ MR_B &= \tfrac{c_B^B}{1+t_B} - \tfrac{t_B - t_A}{D_B''} - \tfrac{t_B}{S_N''} + c_B^N. \end{aligned} \tag{3.11}$$

Case High – The Case of a State with High Public Good Preferences:$t_B \geq \bar{t}$

In this scenario, the neighboring state is a low-tax state, while the home state has a stronger preference for public goods. The problem facing high-tax state's social planner is that she wishes to determine the optimal combination of regional tax rates given that the neighbor has selected a low fixed rate of $\bar{t} \leq t_B$.

Sub-Case 1 – Preferential Tax Rates at the Border: $t_A \geq t_B$

Cross-border shopping occurs from region A to B and from region B to the neighboring state. The individual budget and government constraints for this model are as follows:

$$\begin{aligned} (1+t_A)c_A^A + (1+t_B)c_B^A + D_A(c_B^A) &= M \\ (1+t_B)c_B^B + (1+t_N)c_N^B + S_B(c_N^B) &= M \\ G = t_A c_A^A + t_B c_B^A + t_B c_B^B. \end{aligned} \tag{3.12}$$

The first order conditions are rewritten as:

$$\frac{U_C^A \frac{c_A^A}{1+t_A}}{MR_A} = \frac{U_C^B \frac{c_B^B}{1+t_B}}{MR_B} + \frac{U_C^A \frac{c_B^A}{1+t_A}}{MR_B} \tag{3.13}$$

where the marginal revenue from a change in a tax rate is denoted:

$$\begin{aligned} MR_A &= \tfrac{c_A^A}{1+t_A} - \tfrac{t_A - t_B}{D_A''} \\ MR_B &= \tfrac{c_B^B}{1+t_B} + \tfrac{c_B^A}{1+t_A} - \tfrac{t_B - t_A}{D_A''} - \tfrac{t_B}{S_B''}. \end{aligned} \tag{3.14}$$

Sub-Case 2 – Higher Taxes at the Border: $t_B \geq t_A$

Accounting for cross-border shopping that is outward from region B in two directions, the individual budget and government constraints for this model are as follows:

$$\begin{gathered}(1+t_A)c^A = M \\ (1+t_A)c_A^B + (1+t_B)c_B^B + (1+t_N)c_N^B + D_B(c_A^B) + S_B(c_N^B) = M \\ G = t_A c^A + t_A c_A^B + t_B c_B^B\end{gathered} \tag{3.15}$$

The first order conditions are rewritten as:[8]

$$\frac{U_C^A \frac{c^A}{1+t_A}}{MR_A} + \frac{U_C^B \frac{c_A^B}{1+t_B}}{MR_A} = \frac{U_C^B \frac{c_B^B}{1+t_B}}{MR_B} \tag{3.16}$$

where the marginal revenue from a change in a tax rate is denoted:

$$\begin{aligned}MR_A &= \tfrac{c^A}{1+t_A} + \tfrac{c_A^B}{1+t_A} - \tfrac{t_A - t_B}{D_B''} \\ MR_B &= \tfrac{c_B^B}{1+t_B} - \tfrac{t_B - t_A}{D_B''} - \tfrac{t_B}{S_B''}.\end{aligned} \tag{3.17}$$

3.4.2 The Marginal Cost of Funds in a Federation

Equations 3.7, 3.10, 3.13, and 3.16 equate the marginal cost of funds (or MCF) across jurisdictions within a state such that $MCF_A = MCF_B$.[9] As in Dahlby and Wilson (1994), in a federation, the marginal cost of funds must be equal across all jurisdictions at an optimum.[10]

The MCF can be decomposed into two parts, which I call the "within cost of funds" (WCF) and the "private consumption externality" (PCE). The WCF would

[8] Note I need to add one more derivative to Equation (3.4) to solve this problem. This derivative can be obtained by totally differentiating Region B's individual budget constraint.

[9] A large literature on the marginal cost of funds has emerged, including Slemrod and Yitzhaki (1996) and Slemrod and Yitzhaki (2001).

[10] The intuition for this can be seen in an example. If the MCF is 1.5 in Region A but is 1.1 in Region B, then raising an additional dollar of revenue in Region B is less costly than raising an additional dollar in Region A. Raising additional revenue from Region B is welfare enhancing for the state even though it is not a Pareto improvement for Region B.

be the MCF without any externalities and it specifically measures the direct cost in jurisdiction i of changing tax rate i. The WCF is present on both sides of the equalities in equations 3.7, 3.10, 3.13, and 3.16. On one side of each equality, a second term is also present. This term is the PCE and is only present in the marginal cost of funds for the low-tax region of each sub-case.[11] This term captures how changes in the low-tax region's tax rate distort the consumption decisions of individuals in the high-tax region. As the tax differential between the two regions becomes larger, the high-tax region will have a larger fraction of its goods purchased abroad – resulting in relative distortions to the consumption profile. The social planner accounts for this effect on consumption when choosing the tax rate in the low-tax region. The choice of the high-tax region's rate never affects consumption of the low-tax region because the residents of the low-tax region always purchase commodities at home.

Marginal revenue is a key component of the MCF. Suppose i denotes the relatively high-tax jurisdiction and j denotes the relatively low-tax jurisdiction. Then, the government can possibly expand the base by attracting additional cross-border shoppers with a lower rate. The terms MR_i and MR_j in the sub-cases above contain a $-\frac{t_i}{D_i''}$ and $-\frac{t_j}{D_i''}$ respectively. I will refer to this as the "tax base effect" or TBE.[12] This is the change in revenue in the jurisdiction resulting from changes in the amount of cross-border shopping due to a tax rate change. Increases in the tax rate decrease the amount of consumption purchased within the region via reductions in cross border shopping. Alternatively, the government can increase the tax rate, thereby increasing the revenue raised from those shoppers who continue to shop within the jurisdiction. In the low-tax regions for the sub-cases above, MR_j contains a c_j^i term. I will call this the "tax exporting effect" or TEE.[13] As more residents of i cross the border to shop in j, the larger are the incentives of j to raise its tax rate to extract addi-

[11] The equations are arranged so that it is always the second additive term on its side of the equality.

[12] In the language of Mintz and Tulkens (1986) these are "public consumption effects."

[13] Mintz and Tulkens (1986) refer to this as a "private consumption effect."

tional revenue from non-residents. This effect moves in the opposite direction of the "tax base effect." The "tax exporting effect" is only present in the relatively low-tax jurisdictions.

3.4.3 Characterizing a Pattern to the Optimal Tax Rates

At an optimum, it must be the case that the marginal cost of funds is equal in all jurisdictions within a federation. Therefore, the optimal tax rate will be geographically differentiated if $MCF_A \neq MCF_B$ when the tax rates are equal in Region A and B. The exact conditions for geographic differentiation will be specified below.

Proposition 3.1. *For a low-tax state with open borders, the optimal tax system depends on the relative size of the tax base effect and tax exporting effect. If, at the state border, the tax base effect is larger in absolute value than the tax exporting effect, then $t_A > t_B$ is the optimal response to a neighboring state's high tax rate. If, at the state border, the tax exporting effect is larger in absolute value than the tax base effect, then $t_A < t_B$ is optimal. Only if the tax base effect exactly offsets the tax exporting effect will uniform taxation be optimal.*

Proof. See appendix. □

Intuitively, because the regions are located in the relatively low-tax state, the residents do not have any opportunity to shop in the other state. Starting from a position where tax rates are equal across regions, the consumption profiles are equal in both regions and there is no within state cross-border shopping. Differences in the marginal cost of funds will be determined solely by the relative magnitudes from any efficiency losses to the public good, which are driven entirely by the differences in marginal revenue across regions. Differences in marginal revenue result from the tax base and tax exporting effect in the border region.

Define the absolute value of the elasticity of cross-state shopping with respect to the border region's tax rate as:

$$\epsilon = \left| \frac{\partial c_B^N}{\partial t_B} \frac{t_B}{c_B^N} \right| = \frac{1}{S_N''} \frac{t_B}{c_B^N} \tag{3.18}$$

Corollary 3.2. *The tax exporting effect will exactly offset the tax rate effect if the elasticity of cross-state shopping with respect to the border region's tax rate is unit elastic. If $\epsilon > 1$, the tax base effect will dominate the tax exporting effect and $t_A > t_B$ is optimal. The tax exporting effect dominates if $\epsilon < 1$ in absolute value and $t_A < t_B$ is optimal.*

Proof. See appendix. □

If cross-state shopping is very price responsive, then small adjustments down in the border region's tax rate will result in large quantities of cross-border shopping and additional revenue gains from the neighboring state. On the other hand, if cross-state shopping is not very responsive with respect to the neighboring jurisdiction's tax rate, then the government can increase revenue by raising the tax rate and exporting the tax to foreign residents.

Therefore, on the low-tax side, the optimal tax is geographically differentiated if cross-border shopping of the other state's resident's is not unit elastic. In the following discussion I will subsequently return to discuss what the magnitude of this elasticity is likely to be.

Now I consider a state with high preferences for public goods. On the high-tax side, cross-border shopping allows border-zone residents to obtain additional consumption. This implies that the proof defining the conditions for geographic differentiation must account for the benefit some consumers receive from cross-border shopping.

Proposition 3.3. *For a high-tax state with open borders, the optimal tax system depends on the relative size of the tax base effect and differences in the marginal*

utility of consumption across the two regions. If the tax base effect is sufficiently large in absolute value and differences in the marginal utility of consumption are sufficiently small ($U_C^A \approx U_C^B$), then $t_A > t_B$ is the optimal response to a neighboring state's low tax rate. If the tax base effect is sufficiently small in absolute value and differences in the marginal utility of consumption are sufficiently large ($U_C^A \ggg U_C^B$), then $t_A < t_B$ is optimal. If the tax base effect exactly offsets the differences in the marginal utility of consumption, then uniform taxation is optimal.

Proof. See appendix. □

Intuitively, when a state desires higher taxes than its neighbors, the consumer in the border region will purchase some of his consumption abroad. This is bad for the state because of revenue leakage. But, when the tax rate in the two regions within a state are equal, residents of the border region may have more consumption than residents of the away region because of the arbitrage opportunity that exists. The state planner may want lower taxes in the border region to decrease the revenue leakage from cross-border shopping. However, this will also raise consumption in the border region relative to the away region, furthering inequality. For this reason, when $t_B = t_A$, the direction of the inequality of:

$$\frac{t_B}{S_B''} \gtreqless \frac{(U_C^A - U_C^B)}{U_C^A} \frac{c_B^B}{1 + t_B} \tag{3.19}$$

will determine the relative pattern of geographic differentiation.[14]

Lowering taxes in the border region will reduce the revenue leakage. This will be optimal if the right side of (3.19) is smaller than the left side. If the tax base effect is large in absolute value, it is more likely to be optimal to lower taxes in the border region because individuals are very responsive. The tax base effect will be larger if

[14]Both sides of the equation in the text are positive because $U_C^A \geq U_C^B$. The left side illustrates the revenue leakage through the tax base effect. The right hand side illustrates the social planner's concern for equalizing consumption across all residents. The derivation of this equation is in the appendix.

the elasticity of cross-state shoppers in the border-zone are highly responsive to the border-zone's own tax rate. Inequalities in the marginal utilities of consumption – the right side of (3.19) – may be small depending on how concave the utility function is, as well as how much of the gains to consumption are offset by the total transportation cost.

When tax rates are equal in a state, if the tax base effect is significantly responsive and the consumption profiles in the two regions are similar or if the utility function is not very concave so that even large differences in consumption do not concern the social planner, then a preferential tax rate near the border will be optimal. However, if the resident of Region B has much more total consumption than the resident of Region A or if the utility function is very concave so that even small differences in consumption concern the social planner, it may be possible for tax rates to be higher in the border region. This will be especially true if the tax base effect is sufficiently small in absolute value. Intuitively, this scenario arises because the social planner cares about the equality of consumption between the two regions due to the concavity of the social welfare function. Thus, if the tax base effect is sufficiently small, but the discrepancies in the marginal utility of consumption are large, the social planner will want to raise taxes in the border region.

Corollary 3.4. *If the utility function is linear in consumption, i.e. has a marginal utility that is constant, and tax rates are non-zero, then $t_A \leq t_B$ is never optimal. Preferential tax rates in the border region will always be optimal.*

Proof. See appendix. □

If the individual has a utility function that is linear in consumption (for example utility that is quasi-linear with respect to the consumption good), then the social planner does not care about how consumption is allocated across all individuals and the problem is similar to revenue maximizing governments. Because utility is no

longer concave with respect to consumption, equity concerns vanish from the planner's problem. With no equity concerns, raising the tax rate near the border never is optimal because it will only increase the revenue leakage across state lines, while also inducing a distortion within the state.

It is reasonable that utility is linear in consumption if cross-border shopping is a matter of a few easily carried items purchased in bulk. Differences in the marginal utilities across these goods seem unlikely to be large, unlikely to have significant complementarity or substitutability with public goods, and unlikely to be a major policy concern.

3.5 What Determines the Outcome?

From the above analysis, it is clear that a strong taste for public goods will lead to Case High and that a weak taste for public goods will lead to Case Low.

3.5.1 A State with Low Public Good Preferences

If state H were acting in autarky and state H were restricted to choosing a single tax rate to maximize domestic welfare, the optimal commodity tax is simply a head tax. At an optimum, the home state will choose the tax rate $\tilde{t}$ such that $2\frac{U_G}{U_C} = 1$ (the Samuelson rule holds), which corresponds to private consumption of $\widetilde{c}$. Recalling that the country has two agents, in autarky, social welfare is $\widetilde{W} = 2U(\widetilde{c}, 2\tilde{t}\widetilde{c})$.

Now consider the model where the home state sets an optimal tax rate in response to the neighbor's (exogenous) tax rate t_N. Suppose the state N has a higher preference for public goods such that $t_N > \tilde{t}$. How will the optimal tax rate for state H compare to the optimal tax rate under autarky? Note that if the state chooses a tax rate of $\tilde{t}$, there will be inward cross-border shopping to the state. As a result, the added revenue from cross-border shopping will guarantee that public goods will be over-provided relative to the case of closed borders. Private consumption, on the other

hand, will remain the same as in the case of autarky. Starting from $\tilde{t}$, social welfare is now $W = 2U(\widetilde{c},\, 2\tilde{t}\widetilde{c} + \tilde{t}c_B^N)$. Consider the incentive that the home state has to either raise or lower its tax rate. The change in welfare from a small tax change will be proportional to:

$$\frac{U_G}{U_C}[2\frac{\widetilde{c}}{1+\tilde{t}} + (1-\varepsilon)c_B^N] - \frac{\widetilde{c}}{1+\tilde{t}}. \tag{3.20}$$

The over-provision of public goods at $\tilde{t}$ implies that $\frac{U_G}{U_C} < \frac{1}{2}$. When the public good is over-provided, lower taxes will help to re-allocate domestic consumption toward the more valuable private good. Depending on the value of ϵ, a lower tax could raise more or less revenue from foreigners. Because ϵ is determined endogenously, equation 3.20 has certain predictions about whether the elasticity of cross-border shopping is larger or smaller than unity. Rearranging equation 3.20 yields:

$$(1-\epsilon)c_B^N = (\frac{U_C}{U_G} - 2)\frac{\widetilde{c}}{1+\tilde{t}}. \tag{3.21}$$

Because the public good is over-provided, the right side of equation 3.21 is unambiguously positive. Therefore, the first order conditions imply that ϵ must be less than 1 for an optimizing low-tax state.[15] This need not be the case for a high-tax state.

The implication of equation 3.21 is that the solution to a low-tax country's optimal tax rate given an exogenous neighboring tax rate must be a tax rate at which $\epsilon < 1$. I must determine, however, if ϵ can be less than one. Totally differentiating equation 3.18 yields that $\frac{\partial \epsilon}{\partial \tilde{t}} > 0$ if $S_N''' \geq 0$. Noting that as $\tilde{t} \to 0$, $\varepsilon \to 0$, it must be that there exists a critical tax rate, $\hat{t}$, such that $\epsilon < 1$ for all t in the interval $[0, \hat{t})$ and $\epsilon > 1$ for all t in $(\hat{t}, t^N)$. And $\hat{t}$ may be greater than t^N. Note that $\epsilon = 1$, implies that

[15] In this model, the only decision is over where to shop. Thus, commodity taxation in a closed economy will amount to a head tax. In an open economy with a high-tax neighbor, commodity taxation is non-distortionary and the marginal cost of funds is less than 1. Of course, this would not necessarily hold if other distortions existed in the economy, such as an endogenous labor-leisure choice.

$\hat{t} = c_B^N S_N''$. Thus, if c_B^N or S_N'' are sufficiently large and $S_N''' \geq 0$, then ϵ must be less than 1.

Proposition 1 above is proven by comparing the marginal cost of public funds when the tax rates are equal in both regions. Without loss of generality, the proofs can compare the MCF across regions for equality of any tax rate – including $\tilde{t}$. All that matters is that the preference for public goods in state H is low relative to the neighboring state and thus that $\tilde{t} < \bar{t}$. Therefore, under the assumption that $S_N''' \geq 0$ and $\hat{t}$ is sufficiently large, the optimal tax must be a tax rate where ϵ must be less than 1, which implies $t_A < t_B$.

A higher tax rate at the border can only arise in the low-tax state if the the elasticity of *cross-border shopping* with respect to the *border-zone's* tax rate is less than unity in absolute value. Whether or not the assumptions above guaranteeing that $\epsilon < 1$ hold is an empirical question. The literature has focused on estimating the elasticity of *total* consumption with respect to the *own*-jurisdiction tax rate. Estimates of this elasticity vary from -5.9 (Walsh and Jones 1988) to -1.38 (Tosun and Skidmore 2007). Fox (1986) shows that the elasticity in response to a one percent increase in the tax rate varies by location suggesting the size and density of the border region may influence the relevant elasticity for a social planner. In part because of data limitations, the literature has not focused on estimating the elasticity of cross-border shopping with respect to the neighboring jurisdiction's tax rate. However, the heterogeneity of responses across jurisdictions in Fox (1986) suggests that the size of ϵ will depend critically on the characteristics of the neighboring jurisdiction. For a very dense region, empirically speaking, the elasticity is likely to be much smaller than for a sparse area.

To summarize, if a state has a preference for low public good provision, the optimal response of the state depends on the relative size of the tax base and tax exporting effects at the border. Starting from $t_A = t_B$, considering the following policy change:

the low-tax state raises the tax rate at the border and lowers it away from the border in a manner that keeps tax revenues from domestic consumers constant. This policy will raise welfare if and only if the tax change will raise tax collections from foreigners. Starting from $t_A = t_B$, the policy change will have no first order effect on domestic revenue because $c^A = c^B$ initially. If $\epsilon < 1$, the policy will raise tax collections from foreigners. If If $\epsilon > 1$, tax revenue from foreigners will fall. Thus, if $\epsilon < 1$ when $t_A = t_B$, the government can "export" a large portion of its revenue raising capabilities to non-residents by increasing the tax rate at the expense of losing relatively little cross-border shopping. Without loss of generality, when $t_A = t_B = \tilde{t} < t_N$, for a low-tax state to be maximizing welfare, I have demonstrated that ϵ must be less than 1 under relatively mild assumptions. This will not be the case for the high-tax state.

3.5.2 A State with High Public Good Preferences

Up to now, I have assumed that the social welfare function, $W = \sum_{i=A,B} U^i(c^i, G)$, gives equal weight to consumers in both regions. Proposition 2 implies that the concavity of the social welfare function will help to determine whether taxes are higher or lower in the border-zone. This proposition suggests that the weight that the social planner will give to each agent is also important – because if one consumer obtains excessively large amounts of consumption, but the social planner gives this region additional weight, the outcome will be very different than if the planner gives this region zero weight. One reason the planner may want to give extra weight to one region is if population in the two regions are different – but that all people in a region are identical and equidistant from the region border.

Consider the monotonic concave transformations, $\Omega_i[U^i(c^i, G)]$, of the utility functions where $\Omega_i'(\cdot) > 0$ and $\Omega_i''(\cdot) \leq 0$ for all i. The social welfare function now becomes $\hat{W} = \sum_{i=A,B} \Omega_i[U^i(c^i, G)]$. If $\Omega_A = \Omega_B$, maximizing $\hat{W}$ is equivalent to maximizing W. If $\Omega_A \neq \Omega_B$, then the magnitude of $\frac{\Omega_A'}{\Omega_B'}$ is essential to determining the pattern of

geographic differentiation. If $\frac{\Omega'_A}{\Omega'_B} > 1$, the planner is giving additional weight to the resident of region A. It is as if region A has a larger population.

As $\frac{\Omega'_A}{\Omega'_B}$ become smaller and less than 1, the $MCF_A(t_A = t_B)$ grows larger relative to $MCF_B(t_A = t_B)$. This suggests that even if the consumption differences between the residents of Region A and Region B are very large, if $\frac{\Omega'_A}{\Omega'_B}$ is sufficiently small to overcome these inequality differences, the social planner will always want to lower taxes in the border region. As $\frac{\Omega'_A}{\Omega'_B}$ become larger and greater than one, the $MCF_A(t_A = t_B)$ becomes smaller relative to $MCF_B(t_A = t_B)$. As a result, the social planner will more likely want higher taxes in the border region even if consumption differences are small between the two regions. Therefore, the monotonic concave transformation of the utility function suggests that in addition to the sizes of the tax base effect and the inequality in consumption across the two regions, the social planner's weight to each region is also important. The transformation is useful to help clarify when the unintuitive case that municipal tax rates should rise near a low-tax border is possible. Higher taxes in the border-zone become more likely when the border region is sparsely populated relative to the away region.[16] With this transformation, geographic differentiation may arise even if the marginal utility of consumption is equal across agents.

To summarize, if a state has a preference for high taxes, it can never capture cross-border shoppers from the neighboring low-tax state. Starting from $t_A = t_B$, considering the following policy change: the high-tax state lowers the tax rate at the border and raises it away from the border in a manner that keeps tax revenues raised from domestic consumers constant. By reducing the tax gap with the neighboring state, there is a reduction in cross-state shopping. It must be the case that the policy change will increase total private consumption $(c^A + c^B)$ in the state. Starting at

[16]When interpreting this result, it is important to remember that all agents in each region are identical and located the same distance from borders. The interpretation of the results would change if the population of the region were distributed heterogeneously with respect to distance to borders.

$t_A = t_B$, the reduction in cross-state shopping is large relative to the added cross-border shopping within the state. As a result, of the increase in consumption, if the marginal utility of private consumption were the same in both regions and each region received equal weight in the welfare function, the policy change must increase overall welfare. However, at $t_A = t_B$, residents of the border region have $c^A > c^B$. Thus, when the social welfare function is concave, the social utility of private consumption is higher for residents of the region away from the border. If the welfare measure is sufficiently "redistributive", then the distributional effect will create incentives for the planner to reduce t_A relative to t_B. Which of these two offsetting effects dominates is also dependent on how much weight the social planner gives to each region – and therefore, a function of the relative populations.

3.6 Extensions and Discussion

3.6.1 Extensions

Many states utilize multiple tax instruments such as income taxes and sales taxes. This raises the question as to whether geographically differentiated sales taxes are optimal if the state has two tax instruments. Intuitively, a state with a high preference for public goods could equalize its sales tax rate to the rate of the neighboring state and then assess a higher income tax. In the absence of migration and large labor supply distortions, such a solution would eliminate any possibility of cross-border shopping while obtaining the desired level of public services. However, uniform sales tax rates across the states is rare. One reason for this is that sales tax revenues and income tax revenues are imperfect substitutes as states often seek to rely on multiple taxes to different degrees. Taking as given that these differences in the sales tax rate will exist and assuming that the state cannot geographically differentiate its income tax rate, then all of the above results will remain applicable.

Second, I assume that producers fully pass forward the tax to the consumer. This may not be the case. Furthermore, firms may adjust their prices depending on how far they are located from the border. Harding, Leibtag and Lovenheim (Forthcoming) find that the incidence of taxation varies depending on a firm's distance to the nearest low-tax border. As such, firms near low-tax borders pass through less of the tax to consumers. So long as the individual firms do not change prices in a manner that completely reduces the after-tax price at borders, then geographic differentiation of tax rates will still be optimal. The degree of geographic differentiation in the tax rates will be different than if the incidence of the tax is entirely on consumers because the firms are making some of the price adjustments that the social planner would do through the geographic differentiation.

Finally, geographic differentiation of tax rates presents issues relating to horizontal equity. Although consumers in both regions have equal incomes, the tax burden that they face will vary depending on their residence. Such a violation of horizontal equity is a concern, however, it would also be present if the tax system were uniform within a state. Under uniform taxation, residents with equal incomes have heterogeneous opportunities for cross-border shopping in the neighboring state. The implication is that the uniform tax system will be horizontally unequal on the basis of some residents cross-border shopping while other residents will be unable to cross-border shop because of how far they reside from the border.

3.6.2 Comparison to the Existing Literature

Although the above results are derived under welfare maximizing governments, a large portion of the tax competition literature has relied on revenue maximizing governments. In the context of this model, revenue maximizing governments are characterized by the marginal utility of consumption being zero. Therefore, a model with governments as revenue maximizers is nested within the model presented above.

The results for the high-tax side of the border would simplify such that the state government will always want an unambiguously lower tax rate near the border. As noted above, revenue maximizing governments is the subject of two working papers: Kessing (2008) and Nielsen (2010). Revenue maximization allows for the characterization of an equilibrium – when it does exist.

In Kessing (2008), the equilibrium policy is to raise tax rates in the region away from the border if the size of the border-region is sufficiently large. Consider the case where the home country is a high-tax country and where the border-zone is sufficiently large. In this case, the neighboring country will set the same tax rate as in the non-preferential case. The home country will set the same tax rate in the border-region as it would select in the non-preferential case, but will add an additional mark-up to the tax rate in the region away from the border. In Nielsen (2010), no Nash equilibrium will exist in pure strategies when the government can select tax rate and the size of the border-region. Selecting the size of the border-region will make it so that the the away region is composed of some "swing-shoppers." These shoppers are willing to cross the border-zone and purchase the good in the foreign country. The implication is that the neighboring country can lower tax rates to obtain a discrete jump in revenues. However, in a Stackelberg equilibrium with the neighboring country as a leader, an equilibrium will exist with preferential tax rates at the border. These results suggest that depending on the nature of the strategic interaction and whether the size of the border-region is exogenous, tax competition may or may not be intensified. Whether tax competition occurs in response to border-zones remains an open question.

Unlike Kessing (2008) and Nielsen (2010), I assume that tax competition will not occur in response to preferential border-zones. The results of this paper differ from the revenue maximizing models discussed above in that tax rates may be higher in the border-region – suggesting that preferential tax rates need not result in a race to the bottom. In addition, I do not restrict preferential tax rates to only high-tax states.

As I suggest in this paper, preferential or discriminatory rates may be desirable for low-tax jurisdictions as well. Whether or not introducing tax competition into a model of welfare maximizing governments with discriminatory tax rates will induce a race to the bottom will inevitably depend on the structure of the game and the size of the border-region.

3.7 Conclusion

The model presented here suggests that when the tax system is characterized by a line resulting from geographic borders, uniform within-state taxation is not an optimal policy under most conditions. The border-line encourages cross-border shopping. This distortion is indicative of deadweight loss, but benefits the low-tax state. When states differ in preferences for public goods, it is not optimal from a state planner's perspective to levy a single rate.

The model above presents an administratively feasible tax system,[17] where discrete changes in the tax system based on location is likely to be welfare enhancing. The discrete steps will induce additional discontinuities within the tax system, but the state planner can utilize these differences to increase revenue or smooth consumption. Shopping in a different region within the state incurs inefficiencies in transportation, it does not create revenue loss within the state, while also having the possibility of obtaining additional revenues from residents outside of the state. The result is a first order gain with a second order loss.

The results presented in this paper suggest that as states seek ways to increase revenue and as the European Union continues its process of integration, additional research on geographically differentiated taxes should be an important part of future

[17] It is worth recognizing that there are significant administrative difficulties of implementing border-zones under a Value Added Tax system. For example, it raises the question of how inter-region purchases of intermediate goods are treated. Although similar issues arise with sales and excise taxes, such concerns are not as significant. These administrative costs play no role in this paper.

research. In the presence of tax differentials, a state can improve the social welfare of residents through geographic differentiation of the tax rate if neighboring states do not respond strategically. The principle of geographic differentiation within a state is likely to apply to other types of non-tax policies where similar distortions result from policy differentials at the border.

3.8 Appendix

3.8.1 Derivation of MCF Terms

The derivations of the marginal cost of public funds occur repeatedly. I present the algebraic derivations of Equation 3.7 below. The derivations of all the other marginal cost of fund equations follow a similar process.

Recall that for this case, $t_A > t_B$ and the state is a low-tax state. The government then maximizes $U^A(c^A, t_A c_A^A + t_B c_B^A + t_B c^B + t_B c_B^N) + U^B(c^B, t_A c_A^A + t_B c_B^A + t_B c^B + t_B c_B^N)$. Differentiating with respect to t_A and t_B yields:

$$t_A: \; U_C^A[\tfrac{\partial c^A}{\partial t_A}] + (U_G^A + U_G^B)[c_A^A + t_A \tfrac{\partial c_A^A}{\partial t_A} + t_B \tfrac{\partial c_B^A}{\partial t_A} + t_B \tfrac{\partial c^B}{\partial t_A}] = 0$$

$$t_B: \; U_C^A[\tfrac{\partial c^A}{\partial t_B}] + U_C^B[\tfrac{\partial c^B}{\partial t_B}] + (U_G^A + U_G^B)[t_A \tfrac{\partial c_A^A}{\partial t_B} + c_B^A + t_B \tfrac{\partial c_B^A}{\partial t_B} + c^B + t_B \tfrac{\partial c^B}{\partial t_B} + c_B^N + t_B \tfrac{\partial c_B^N}{\partial t_B}] = 0$$

Using equation 3.4 and 3.5 yields:

$$t_A: \; -U_C^A[\tfrac{c_A^A}{1+t_A}] + (U_G^A + U_G^B)[c_A^A - t_A \tfrac{c_A^A}{1+t_A} - t_A \tfrac{1}{D_A''} + t_B \tfrac{1}{D_A''}] = 0$$

$$\begin{aligned} t_B: \; & U_C^A[-\tfrac{c_B^A}{1+t_A}] - U_C^B[\tfrac{c^B}{1+t_B}] \\ & +(U_G^A + U_G^B)[-t_A \tfrac{c_B^A}{1+t_A} + t_A \tfrac{1}{D_A''} + c_B^A - t_B \tfrac{1}{D_A''} + c^B - t_B \tfrac{c^B}{1+t_B} + c_B^N - \tfrac{t_B}{S_N''}] = 0 \end{aligned}$$

Rearranging terms yields:

$$t_A: \; U_G^A + U_G^B = \frac{U_C^A \frac{c_A^A}{1+t_A}}{\frac{c_A^A}{1+t_A} - \frac{t_A - t_B}{D_A''}}$$

$$t_B: \; U_G^A + U_G^B = \frac{U_C^B \frac{c^B}{1+t_B}}{\frac{c^B}{1+t_B} + \frac{c_B^A}{1+t_A} - \frac{t_B - t_A}{D_A''} - \frac{t_B}{S_N''} + c_B^N} + \frac{U_C^A \frac{c_B^A}{1+t_A}}{\frac{c^B}{1+t_B} + \frac{c_B^A}{1+t_A} - \frac{t_B - t_A}{D_A''} - \frac{t_B}{S_N''} + c_B^N}.$$

Setting these two equation equal to each other yields Equation 3.7.

3.8.2 Derivation of Equation 3.19

To derive this equation, I start with the equality of the marginal cost of funds in a high-tax state, given by equations 3.13 and 3.16:

$$\frac{U_C^A \frac{c_A^A}{1+t_A}}{\frac{c_A^A}{1+t_A} - \frac{t_A - t_B}{D_A''}} = \frac{U_C^B \frac{c_B^B}{1+t_B}}{\frac{c_B^B}{1+t_B} + \frac{c_B^A}{1+t_A} - \frac{t_B - t_A}{D_A''} - \frac{t_B}{S_B''}} + \frac{U_C^A \frac{c_B^A}{1+t_A}}{\frac{c_B^B}{1+t_B} + \frac{c_B^A}{1+t_A} - \frac{t_B - t_A}{D_A''} - \frac{t_B}{S_B''}}$$

$$\frac{U_C^A \frac{c^A}{1+t_A}}{\frac{c^A}{1+t_A} + \frac{c_A^B}{1+t_A} - \frac{t_A - t_B}{D_B''}} + \frac{U_C^B \frac{c_A^B}{1+t_B}}{\frac{c_B^B}{1+t_B} - \frac{t_B - t_A}{D_B''} - \frac{t_B}{S_B''}} = \frac{U_C^B \frac{c_B^B}{1+t_B}}{\frac{c_B^B}{1+t_B} - \frac{t_B - t_A}{D_B''} - \frac{t_B}{S_B''}}$$

To proceed, I need to compare these expressions in the limit where taxes in this state are equal, but still less than taxes in the neighboring state (i.e. where $t_A \to t = t_B < t_N$). The limit for both of the equations above are the same. Without loss of generality, work with the second equation. In the limit, there can be no within state cross-border shopping implying that $c_A^B = 0$. Because $c^B = c_A^B + c_A^A + c_A^N = c_A^A + c_A^N$ and $c^A = c_A^B + c_A^A = c_A^A$, it is easy to see that the ability of consumers in B to arbitrage and cross the state border implies that in the limit $c^B \geq c^A$ and therefore, $U_C^B \leq U_C^A$. Substituting into the above equation and letting tax rates converge to $t = t_B$, yields

$$\frac{U_C^A \frac{c^A}{1+t}}{\frac{c^A}{1+t} - \frac{t-t}{D_B''}} = \frac{U_C^B \frac{c_B^B}{1+t}}{\frac{c_B^B}{1+t} - \frac{t-t}{D_B''} - \frac{t}{S_B''}}$$

which reduces to

$$\frac{U_C^A \frac{c^A}{1+t}}{\frac{c^A}{1+t}} = \frac{U_C^B \frac{c_B^B}{1+t}}{\frac{c_B^B}{1+t} - \frac{t}{S_B''}}$$

and simplifying further implies

$$U_C^A \left(\frac{c_B^B}{1+t} - \frac{t}{S_B''}\right) = U_C^B \frac{c_B^B}{1+t} \ ,$$

which then yields the equation in the text:

$$\frac{t_B}{S_B''} \gtreqless \frac{(U_C^A - U_C^B)}{U_C^A} \frac{c_B^B}{1+t_B}.$$

3.8.3 Proof of Proposition 1

Proof. It must be the case that equation (3.7) or (3.10) holds at an optimum. The limit of the equations as $t_A \to t_B$ implies that the first order conditions are continuous when a state changes regimes from sub-case 1 to sub-case 2. Start from a point where all jurisdictions in State H have equal tax rates ($t_A = t_B = t$). Because the tax rates are equal at this point, $c_j^i = 0$ for all $i \neq j$. Because no one from Region B crosses the state border, this immediately implies that $c^A = c^B = c_A^A = c_B^B = c$ and the marginal utilities of consumption are equal. Residents of State N will cross-border shop so that $c_B^N > 0$. Using the simplifications above, equation (3.7) implies $MCF_A(t_A = t_B) - MCF_B(t_A = t_B) = \frac{\frac{c}{1+t}}{(\frac{c}{1+t})} - \frac{\frac{c}{1+t}}{(\frac{c}{1+t} - \frac{t}{S''_N} + c_B^N)}$. A little algebra will show that this expression is equal to zero if $c_B^N = \frac{t}{S''_N}$, implying equal tax rates. The tax system will be differentiated if $MCF_A(t_A = t_B) \neq MCF_B(t_A = t_B)$. Using the limit of equation (3.7) and the above substitutions, $MCF_A(t_A = t_B) - MCF_B(t_A = t_B) < 0$ if $c_B^N < \frac{t}{S''_N}$. $MCF_A < MCF_B$ implies the government must lower the tax rate in Region B relative to Region A to equalize the marginal cost of funds. Therefore, $c_B^N < \frac{t}{S''_N}$ guarantees that $t_A > t_B$ is welfare improving. Similarly, it can be shown that $MCF_A > MCF_B$ if $c_B^N > \frac{t}{S''_N}$, which implies $t_A < t_B$ is optimal. □

3.8.4 Proof of Corollary 1

Proof. The term of interest in the marginal revenue terms is $-\frac{t_B}{S''_N} + c_B^N$. This can be rewritten as $t_B \frac{\partial c_B^N}{\partial t_B} \frac{c_B^N}{c_B^N} + c_B^N = (1 - \epsilon) c_B^N$, where ϵ is the elasticity of c_B^N with respect to t_B. ϵ is positive. □

3.8.5 Proof of Proposition 2

Proof. It must be the case that equation (3.13) or (3.16) holds at an optimum. The limit of the equations as $t_A \to t_B$ implies that the first order conditions are continuous when a state changes regimes from sub-case 1 to sub-case 2. Start from a point where

all jurisdictions in State H have equal tax rates ($t_A = t_B = t$). Because the tax rates are equal, $c_j^i = 0$ for all $i \neq j$ in State H. However, some consumption from Region B is purchased across the state border so that $c_N^B > 0$. This implies that $c^A \leq c^B$, $c_A^A \geq c_B^B$ and $U_C^A \geq U_C^B$. Substituting into equation (3.13), $t_A = t_B$ is optimal if $MCF_A(t_A = t_B) - MCF_B(t_A = t_B) = U_C^A - \frac{U_C^B \frac{c_B^B}{1+t}}{(\frac{c_B^B}{1+t} - \frac{t}{S_B''})} = 0$. The tax system will be differentiated if $MCF_A(t_A = t_B) \neq MCF_B(t_A = t_B)$. $MCF_A(t_A = t_B) - MCF_B(t_A = t_B) < 0$ if $U_C^A < \frac{U_C^B \frac{c_B^B}{1+t}}{(\frac{c_B^B}{1+t} - \frac{t}{S_B''})}$ and the government must lower the tax rate in Region B relative to Region A to equalize the marginal cost of funds. Similarly, it can be shown that $U_C^A > \frac{U_C^B \frac{c_B^B}{1+t}}{(\frac{c_B^B}{1+t} - \frac{t}{S_B''})}$ implies $t_A < t_B$. □

3.8.6 Proof of Corollary 2

Proof. If the marginal utility is constant, then $U_C^A = U_C^B$ because the utility functions are the same. To have $t_A \leq t_B$, then $MCF_A(t_A = t_B) \geq MCF_B(t_A = t_B)$ implies $1 \geq \frac{\frac{c_B^B}{1+t}}{(\frac{c_B^B}{1+t} - \frac{t}{S_B''})}$. But, this requires $-\frac{t}{S_B''} \geq 0$, which is never true. However, $MCF_A(t_A = t_B) < MCF_B(t_A = t_B)$ is always true because $-\frac{t}{S_B''} < 0$. This immediately implies $t_A > t_B$ is optimal. □

3.8.7 The Characteristics of Lines and Optimal Tax Policy

Beyond tax rate considerations, optimal tax policy should consider where to draw lines in the tax system. In the case of borders, where to draw the lines are not a matter of policy. However, because borders have different characteristics, optimal tax policy must also consider the characteristics of the line such as how fluid the "drawn" line is. Some lines differentiating types of goods are likely to easily result in distortions, while others will not – for example, the characteristics of goods can be altered more easily in some circumstances. I demonstrate this for the case of border-lines: where some borders are more fluid than other borders. As such, the

optimal tax system will be characterized by the tax rates outlined in the main text, but also by the characteristics of the lines – which are partially dependent on tax enforcement and partially dependent on the transportation cost function. I make this point for two extreme cases – where the enforcement system is perfect and where the transportation cost function prohibits all cross-border shopping.

3.8.7.1 The Case of Closed Borders: Perfect Enforcement

Consider a state that chooses tax rates for its regions but does not allow for residents to cross-border shop – either across the state border or the region border. In this scenario, it is as if the borders are effectively closed to cross-border shopping because the the tax system is levied (and effectively enforced) under the destination principle.

Noting that because the border is closed, $c^i = c_i^i$, the individual budget constraints are given by the following equation for all i:

$$(1 + t_i)c^i = M. \tag{3.22}$$

The government budget constraint for the state provided public good is given by:

$$G = \sum_{i=A,B} t_i c^i. \tag{3.23}$$

The state government selects t_A and t_B to maximize (3.1) subject to the two constraints above. The first order conditions for this problem immediately imply that for all i:

$$\sum_{i=A,B} U_G^i = U_C^i. \tag{3.24}$$

Proposition 3.5. *If taxes are credibly levied according to the destination principle or if all borders are closed, the optimal tax system features uniform tax rates ($t_A = t_B$)*

within a state.

Proof. Equation 3.24 will hold for i equal to A and B. The equality of $U_C^A = U_C^B$ across both regions, implies that total consumption is equal in both regions. Individuals in both regions are identical in terms of income, so it immediately follows that the tax rates must be identical across all jurisdictions to equalize consumption. □

Intuitively, under effective enforcement of the use tax, borders are closed and there is no reason for the government to differentiate the tax rate if all regions are identical. From a consumer choice perspective, the level of the public good is the same within the state. The individual's only choice is over how much of the consumption good to purchase at home. The social welfare planner wants identical consumers to choose identical consumption bundles, and therefore, the welfare maximizing choice is to equate the marginal utilities of consumption across all individuals.

3.8.7.2 The Case of Open Borders Within a State

Now, I consider the case where individuals can cross-border shop within the state but are prohibited from crossing the state border. Because the state border is closed, the optimal tax rates will be symmetric in both a low- and a high-tax state.

This case is applicable to several states. First, international borders are much harder to cross than state borders. For example, crossing the United States-Mexico border comes with added time costs along with some probability that customs agents will search for items being brought across the border. Exchange rates are often uncertain and conversion fees must be paid. The added costs and uncertainty may make crossing the border much more costly, perhaps effectively closing it to cross-border shopping. Second, even within a federation, this scenario is highly relevant because some state borders are effectively closed as a result of geographic barriers along borders.[18] If this is case, the optimal tax system at international borders may

[18]Many state borders are delineated by the presence of rivers or mountains, which effectively limit

be different compared to the solution when the border is open.

Solving this problem requires specifying possible directions of cross-border shopping within a state. Because both regions within the state are identical, I can consider the case of cross-border shopping in one direction without loss of generality. For this purpose, assume $t_A \geq t_B$.

The individual budget constraints are given by:

$$\begin{aligned} &\text{Region A:} \quad (1+t_A)c_A^A + (1+t_B)c_B^A + D_A(c_B^A) = M \\ &\text{Region B:} \quad (1+t_B)c^B = M. \end{aligned} \tag{3.25}$$

Recalling that I have assumed that Region A is the high-tax region, the state budget constraint is given by:

$$G = t_A c_A^A + t_B c_B^A + t_B c^B. \tag{3.26}$$

The first order conditions of this problem imply

$$\frac{U_C^A \frac{c_A^A}{1+t_A}}{MR_A} = \frac{U_C^B \frac{c^B}{1+t_B}}{MR_B} + \frac{U_C^A \frac{c_B^A}{1+t_A}}{MR_B}, \tag{3.27}$$

where $MR_A = \frac{c_A^A}{1+t_A} - \frac{t_A}{D_A''} + \frac{t_B}{D_A''}$ and $MR_B = \frac{c^B}{1+t_B} + \frac{t_A c_B^A}{1+t_A} - \frac{t_B}{D_A''} + \frac{t_A}{D_A''} + c_B^A$ denote the marginal revenue from a change in the tax rate.

Proposition 3.6. *If the state border is closed or if the transportation cost function effectively closes the border, equal tax rates in Region A and Region B ($t_A = t_B$) is the socially optimal solution.*

Proof. The first order conditions imply that the MCF must be equal at an optimum. If $t_A = t_B$, $c_B^A = 0$ because no cross-border shopping will occur. This implies that $c_A^A = c^A = c^B$ because incomes are identical. Given identical preferences, $U_C^A = U_C^B$.

border crossings to areas where bridges or major roadways exist. For policymakers in these states, the regions close to such a border essentially face a closed state border.

After substituting these equalities in the first order conditions, it is easy to see the MCF above are equal if $t_A = t_B$. □

In the case of closed state borders, the optimal tax rate from a state planner's perspective is characterized by equal tax rates. Intuitively, the regions are identical and the state cares equally about both within-state regions. Differentiating the tax rate within a state will incur only wasted resources through the realized transportation cost.

3.8.7.3 How Do Line Characteristics Matter?

The comparison of the results with open and closed borders suggests that geographic differentiation need not be the only policy remedy to tax differentials at borders. Effectively enforcing the tax system on the basis of the destination principle within a federation would be another way of eliminating the inefficiency from tax differentials at state borders. Furthermore, inefficiencies from tax differentials at borders may naturally be eliminated at international or state borders if border enforcement, exchange rate uncertainty and added time costs effectively close the border. Therefore, if country borders are effectively closed to cross-border shopping, uniform taxation will be optimal. For this reason, the optimal tax system may depend not only on the existence of tax differentials, but also on the precise nature of the border – and the ability of the planner to enforce taxes on the basis of the destination principle.

The implications for optimal tax policies are more general than border-lines. Lines in the commodity tax system arise naturally. Policymakers have a finite number of tax rates. Traditional Ramsey Rules suggest that the optimal tax rate on consumption goods be set inversely proportional to the good's elasticities. In a world with an infinite number of goods and characteristics of goods, policymakers must decide on where to draw the line when differentiating tax rates. Optimal line drawing is the

subject of Kleven and Slemrod (2009). The results in this paper suggest that optimal line drawing must also consider that some lines are more fluid and distortionary than other lines. Lines that create distortions easily – such as open borders between states – merit closing the differences in tax rates. If lines can be drawn between goods without inducing any distortion – as in the case of borders delineated by rivers – then tax rates need not be set in a manner to reduce the differences in tax rates. Thus, when considering where to draw lines in characteristic space, the degree of differentiation in the rates will be highly sensitive to how easily goods can be distorted in their characteristics across the line.

3.8.8 Geographic Differentiation with Local Public Goods

If the public good is provided at the local level, the revenue raised in town i funds a public good for town i and has no benefit for people in other jurisdictions.[19] If the revenue raised in one town is higher than the other, the state will provide that town with more of the public good. Letting g_i denote the local public good and R_i denote the total revenue raised in jurisdiction i, it will be that $g_i = R_i(t_i)$. If this is the case, the social welfare function will replace G with g_i. When the public good is provided at the local level, the solution describes the state optimum to a decentralized problem (as if the state planner were choosing the rates and public good that localities would choose).

[19] I now refer to the two "regions" in the models as "towns" to illustrate that the case of local public goods characterizes the optimal tax system when municipal governments have taxing authorities in which the municipality can keep all of the revenue from the sales tax. Therefore, this robustness check sheds light on whether the geographic pattern of differentiation in Agrawal (2011*c*) is optimal from a state's perspective.

3.8.8.1 Closed Borders

If the public good is provided locally, 3.22 will still hold. However, the government budget constraint must now obey

$$g_i = t_i c^i \tag{3.28}$$

for all towns within the state. State H will solve

$$\max \sum_{i=A,B} U^i(c^i, g_i). \tag{3.29}$$

The first order conditions for the maximization problem will imply

$$\frac{U_g^i}{U_C^i} = 1, \tag{3.30}$$

which is no different than the standard Samuelson Rule for the optimal provision of public goods in a locality. Individuals in both towns are identical; it immediately follows that the tax rates should be identical across all jurisdictions.

3.8.8.2 Closed State Border, but Open Internal Borders

If the public good is provided locally, 3.25 will still hold. However, the government budget constraint must now obey

$$\begin{aligned} g_i &= t_i c_i^i \qquad \text{if } t_i \geq t_j \\ g_i &= t_i c^i + t_i c_j^i \quad \text{if } t_i < t_j \end{aligned} \tag{3.31}$$

for each respective town within the state. The state will solve 3.29 and the first order conditions for the maximization problem will imply

$$t_A : \ \frac{U_g^A}{U_C^A} = \frac{\frac{c_A^A}{1+t_A}}{MR_A} - \frac{U_g^B(\frac{t_B}{D_A''})}{U_C^A MR_A} \tag{3.32}$$

$$t_B: \ \frac{U_g^B}{U_C^B} = \frac{\frac{c^B}{1+t_B}}{MR_B} - \frac{U_g^A(\frac{t_A}{D_A''} - \frac{t_A c_B^A}{1+t_A})}{U_C^B MR_B} + \frac{U_C^A \frac{c_B^A}{1+t_A}}{U_C^B MR_B}, \tag{3.33}$$

where $MR_A = \frac{c_A^A}{1+t_A} - \frac{t_A}{D_A''}$ and $MR_B = \frac{c^B}{1+t_B} - \frac{t_B}{D_A''} + c_B^A$ denote the marginal revenue from a change in the tax rate. The only difference of these equations from the state public good FOCs is that the second term in both of the equations above suggest the existence of differential effects on the public good provision in the neighboring town. Equal tax rates will still satisfy the conditions for an optimum.

3.8.8.3 All Open Borders

Case Low – A State with Low Public Good Preferences: $t_B \leq \bar{t}$

The individual budget constraints are unchanged, but the government now faces a government budget constraint for each town in the optimization problem:

$$\begin{aligned} g_A &= tc_A^A + t_A c_A^B \\ g_B &= t_B c_B^A + t_B c_B^B + t_B c_B^N \end{aligned} \tag{3.34}$$

where $c_A^B = 0$ for Sub-Case 1 and $c_B^A = 0$ for Sub-Case 2. State H selects t_A and t_B taking as given and fixed $\bar{t}$ by maximizing (3.23) subject to the the individual budget constraints in the previous section and the government budget constraints above. The first order conditions are rewritten such that they imply modified Samuelson rules:

$$\begin{aligned} t_A: \ \frac{U_G^A}{U_C^A} &= \frac{\frac{c^A}{1+t_A}}{MR_A} - \frac{U_G^B(\frac{t_B}{S_N''} - \frac{t_B c_A^B}{1+t_B})}{U_C^A MR_A} + \frac{U_C^B \frac{c_A^B}{1+t_B}}{U_C^A MR_A} \\ t_B: \ \frac{U_G^B}{U_C^B} &= \frac{\frac{c^B}{1+t_B}}{MR_B} - \frac{U_G^A(\frac{t_A}{D_A''} - \frac{t_A c_B^A}{1+t_A})}{U_C^B MR_B} + \frac{U_C^A \frac{c_B^A}{1+t_A}}{U_C^B MR_B} \end{aligned} \tag{3.35}$$

where the marginal revenues from a change in the tax rate is denoted $MR_A = \frac{c^A}{1+t_A} - \frac{t_A}{D_B''} + c_A^B$ and $MR_B = \frac{c^B}{1+t_B} - \frac{t_B}{D_A''} - \frac{t_B}{S_N''} + c_B^N + c_B^A$ and $c_A^B = 0$ for Sub-Case 1 and $c_B^A = 0$ for Sub-Case 2 in the above first order conditions.

Case High – A State with High Public Good Preferences: $t_B \geq \bar{t}$

The individual budget constraints are unchanged from the uniform public good case, but the government now faces two constraints in each optimization problem:

$$\begin{aligned} g_A &= t_A c_A^A + t_A c_A^B \\ g_B &= t_B c_B^A + t_B c_B^B \end{aligned} \tag{3.36}$$

where $c_A^B = 0$ for Sub-Case 1 and $c_B^A = 0$ for Sub-Case 2. Again, State H selects t_A and t_B taking as given and fixed $\bar{t}$ by maximizing (3.23) subject to the above constraints and the individual budget constraints. The first order conditions are rewritten such that they implied modified Samuelson rules:

$$\begin{aligned} t_A: \ \frac{U_G^A}{U_C^A} &= \frac{\frac{c^A}{1+t_A}}{MR_A} - \frac{U_G^B(\frac{t_B}{S_N''} - \frac{t_B c_A^B}{1+t_B})}{U_C^A MR_A} + \frac{U_C^B \frac{c_A^B}{1+t_B}}{U_C^A MR_A} \\ t_B: \ \frac{U_G^B}{U_C^B} &= \frac{\frac{c_B^B}{1+t_B}}{MR_B} - \frac{U_G^A(\frac{t_A}{D_A''} - \frac{t_A c_B^A}{1+t_A})}{U_C^B MR_B} + \frac{U_C^A \frac{c_B^A}{1+t_A}}{U_C^B MR_B} \end{aligned} \tag{3.37}$$

where the marginal revenues from a change in the tax rate is denoted $MR_A = \frac{c^A}{1+t_A} - \frac{t_A}{D_B''} + c_A^B$and $MR_B = \frac{c_B^B}{1+t_B} - \frac{t_B}{D_A''} - \frac{t_B}{S_B''} + c_B^A$ and $c_A^B = 0$ for Sub-Case 1 and $c_B^A = 0$ for Sub-Case 2 in the above first order conditions.

3.8.8.4 Discussion of Local Public Goods

When the public good is financed as a state public good, taxes do not impose an externality on any domestic region's provision of the public good because a change in the tax rate of one town affects the public good provision of the other town in a uniform manner. However, if the public good were a local public good, the effect on the revenue of the two towns will be different. The second additive terms of Equations 3.35 and 3.37 capture the social welfare planner's concern for how changes in the tax rate of jurisdiction i affects the revenue raising capacity of town j. The numerator represents the gain to jurisdiction j from a change in the public good funded by taxes, while the denominator represents the cost to what town i is giving up. When this term has a negative sign in front, it implies an overall benefit to society, thus reducing

the cost of raising funds in the jurisdiction. I refer to this term as the "public good effect" or PGE because it demonstrates the planner's need to consider changes in the public good provision resulting from tax changes of within-state neighbors.

With respect to the PGE in 3.35 and 3.37, changes in the high-tax town's rate will directly affect the low-tax town's ability to raise revenue through the tax base effect because the level of cross-border shopping responds to the difference between the two tax rates. There is no tax exporting effect from a change in the high-tax rate in the numerator. With respect to the PGE for the high-tax town, the social planner accounts for how changes in the low-tax town's rate influences revenue in the high-tax town. Here both the tax exporting and tax base effects influence how much the high-tax town's revenue changes. Depending on the relative magnitudes, the PGE may either raise or lower the MCF.

The main difference compared to the uniform public good case is that the PGE's must now be considered when the state government equalizes the marginal cost of funds. The numerator of each PGE (the second additive term in both first order conditions) represents the marginal benefit from the additional revenue that the town j receives from a change in the tax rate of town i. The denominator represents the marginal revenue for the town whose tax rate is changing. In the case of a state public good, the effect of a tax change has the same effect on the revenue of both towns within the state because the level of public good provision was common to both states. However, when the public goods are different, the PGE represents how much benefit is provided to town j when the state planner wants town i to raise additional revenue. Therefore, when the state social planner solves the decentralized problem, the solution must consider the relative magnitude of the public good externality in addition to the sizes of the tax base and tax exporting effects. The explicit characterization of the optimum is left to future research.

Figure 3.1: Geographic Layout of the Model

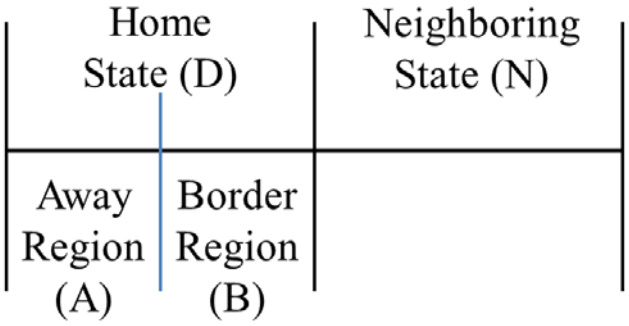
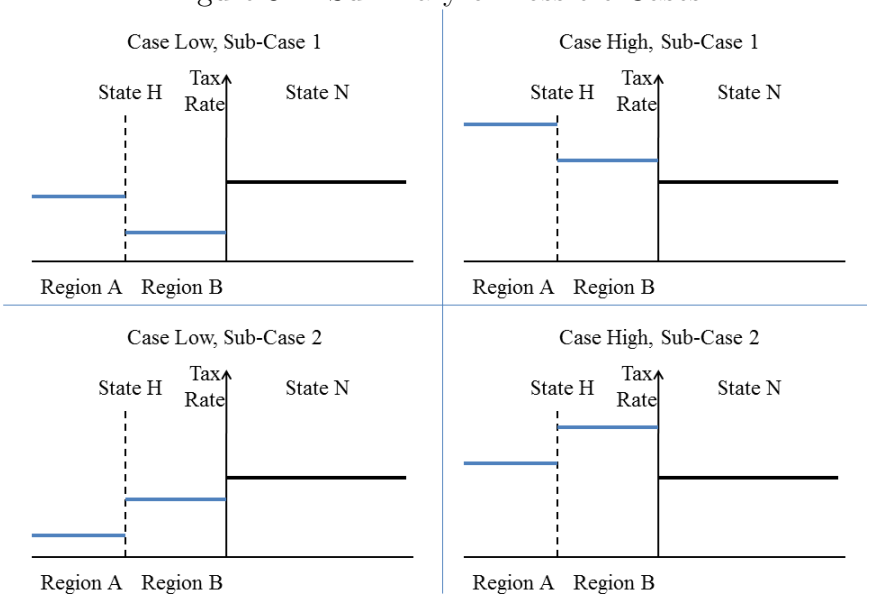


Figure 3.2: Summary of Possible Cases



Note: Equality of tax rates is a knife's edge case and is included in all of the cases above. Tax rates show the magnitude of the rates relative to the neighboring jurisdiction only.

CHAPTER IV

Inter-Federation Competition: Sales Tax Externalities with Multiple Federations

4.1 Introduction

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

Vertical externalities arise between different levels of government that share the same tax base. The state of New York and New York City both realize tax revenue from consumers in New York City. When New York City raises its tax rate, a vertical externality arises because total consumption of goods in New York City declines. The

state's tax revenue falls, even though the state tax rate remains unchanged. Although state revenues fall, city only internalizes the effect of the rate change on its own tax base. Vertical externalities result in governments underestimating the social marginal cost of public funds. Thus, taxes are too high in equilibrium.

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

In such a context, how does introducing inter-federation competition into a model of sales taxation change the strategic interaction between governments of different levels? In a world where the federal government has no horizontal competitors of its own (other counties), federal taxes will (under most assumptions) be too high as the federal government only faces a vertical externality. Inter-federation competition inevitably constrains the federal government by inducing horizontal externalities at the federal level, while also triggering additional competition at the sub-federal level across federation borders. As a result of the discontinuity induced by the federation's border, this cross-federation competition is inevitably different than horizontal competition within the federation. These differences have strong implications for

estimating the strategic reaction functions in a federation.

Inter-federation competition is traditionally used to mean competition among nations. However, I use the term inter-federation competition to highlight competition among the higher levels of government. As such, counties are federal to towns just as the national government is federal to the state governments. Throughout this chapter, inter-federation competition will refer to competition among multiple counties that are composed of multiple towns, but the theoretical models in the paper could similarly apply to the national-state reltionship as well.

I argue that inter-federation competition is essential to fully understand horizontal and vertical externalities. Modeling and estimating vertical externalities with only one "federal" government ignores horizontal competition among neighboring federal governments – which will put downward pressure on rates. An assumption of a unitary federation may be valid when the federal government is the national government – as it realizes any leakage out of the United States boundaries is relatively small. However, when the "federal" governments are state or county governments, such an assumption is no longer valid. The more decentralized the "higher" level of government, the more likely it will have significant horizontal competitors of its own and this will constrain its taxes from being too high. Empirically analyzing the reaction functions of governments has been mostly restricted to how states respond to national taxes. However, there is no reason to believe that the nature of the strategic reaction function of states is the same as that of localities – especially given that more decentralized levels of government have more horizontal competitors – not to mention institutional differences.

Introducing multiple federations will also allow me to develop a more realistic model where sub-federal governments have multiple borders. In traditional single federation models, there are usually only two-sub federal governments. With only two sub-federal governments, both governments have only one border (one competitor).

In such a model, all members of the federation are peripheral (they are located at the federation's borders). In reality, not all members of a federation are located at the border. This paper shows that vertical externalities affect members of the federation that are located at its periphery in a manner differently than they affect members of the federation that are located internal to the federation. I will also show that "diagonal externalities" will arise. A diagonal externality is the effect of a county's tax rate on a municipality in the neighboring county. For a peripheral jurisdiction, it is identical in the nature of the strategic response to a horizontal externality. The results are derived in a model where consumers have a downward sloping demand function.

In addition to these theoretical considerations, the estimation strategy for determining the strategic response to vertical and horizontal externalities becomes more complex in the presence of inter-federation competition. The empirical methodology presented in this paper indicates, even if the federal government has no horizontal competitors of its own, it is essential to consider the interaction of horizontal and vertical externalities. Additionally, when the federal government has horizontal competitors, externalities induced by the federal government on its competitor must also be considered. The existing literature estimates a strategic reaction function for the average jurisdiction assuming that all externalities are of the same magnitude – no matter the spatial composition of the federation. The empirical results suggest that, with multiple federations, in order to obtain consistent estimates the researcher must empirically allow for strategic reactions to vary based on distance to the federation's borders. The paper also uses decentralized data within federations to determine if the strategic reaction at lower levels of government differs from existing estimates of the state-national interactions mainly estimated in the literature.

This paper will focus on two dimensions of "spatial economics" – "spatial interdependence or contagion" and "spatial location." One, spatial interdependence is the

process by which one jurisdiction has a contagion effect on another (perhaps neighboring) jurisdiction's tax rate. For example, when a jurisdiction sets a tax rate, it maximizes an objective function that aggregates the welfare of residents within the jurisdiction, but does so while competing with neighboring jurisdictions for a mobile tax base. This competitive process will influence the tax setting behavior of other geographically close and possibly overlapping jurisdictions. Two, spatial location is the process by which distance from or proximity to a particular geographic feature influences tax setting behavior. An example of spatial location is that the slope of a jurisdiction's reaction function may be a function of proximity to a border or to an amenity. This paper will show that the strength of spatial interdependence can be heterogeneous with respect to a jurisdiction's spatial location. Adding elements of spatial location will complicate the analysis, but will provide me with a unique and convincing identification strategy to identify tax competition where I rely on demonstrating that strategic interaction is heterogeneous with respect to spatial location in a manner predicted by theory.

The baseline empirical results show that a one percentage point increase in county sales tax rates reduces municipal sales tax rates by .640 percentage points. Alternatively, a one percentage point increase in the average neighboring tax rate will increase a municipality's local sales tax by .399 percentage points. The results are similar in spirit but smaller in magnitude when the definition of neighborliness is made more restrictive. Relative to a specification accounts for horizontal and vertical interaction effects, diagonal externalities, and distance based effects, the estimates from the specification currently estimated in the literature are approximately 30% too large in absolute value. The implications of the estimating strategy are threefold. (1) Omitting interaction effects of horizontal and vertical externalities will induce a substantial bias (between 30% and 60% depending on the definition of neighborliness) that overestimates the magnitude of the true strategic vertical and horizontal interactions.

This bias will arise even if the federal government has no horizontal competitors. (2) If a "federal" government has horizontal competitors as is the case of states and counties, then the neighboring federation's tax rate and the proximity to the neighboring federation must be accounted for in order to obtain an unbiased estimate of the true strategic reactions. (3) The strategic nature of local governments is vastly different than that the strategic reaction of states. The existing state-based empirical literature indicates that most U.S. states consider federal tax rates as strategic complements (although the regression estimates are often not significantly different from zero). The results in this paper suggest that municipalities consider county sales tax rates as strong strategic substitutes. Because the theoretical model in this paper suggests that the direction of the vertical externality depends on the relative magnitudes of the elasticity of demand and the elasticity of cross-border shopping, the empirical results also shed light on the relative magnitudes of these elasticities. The empirical finding that county and town tax rates are strategic substitutes is consistent with municipalities having a much larger elasticity of cross-border shopping than states. Such a conclusion is not unreasonable given that the elasticity of cross-border shopping is often decreasing as state size increases.

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

4.2 Sales Tax Competition in Federations

4.2.1 Theoretical Models of Fiscal Federalism

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

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

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

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

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

4.2.2 Empirically Estimating Reaction Functions

In the first serious attempt to estimate vertical externalities, Besley and Rosen (1998) uses panel data and regresses the state cigarette tax rate on the federal cigarette tax rate. Such a regression fails to account for horizontal externalities and the simultaneity in the tax rates. Esteller-Moré and Solé-Ollé (2001) corrected for these problems by instrumenting for federal tax rates and including (instrumented) average neighboring tax rates on the right-hand side of the equation. However, because the federal tax rate does not vary in any cross-section, such a procedure does not allow for the inclusion of time effects in any panel. Other authors have followed approaches different to the instrumental variable approach of Esteller-Moré and Solé-Ollé (2001). Fore example, Hayashi and Boadway (2001) claim to avoid the endogeneity problem altogether by assuming that the interaction between different levels of government occurs with a time lag so that the values of the federal tax rates are no longer simultaneous. Revelli (2001) argues that estimating the reaction function in first differences and instrumenting for the differenced county tax rates with the lagged level of the tax rates eliminates the endogeneity problem.[2] In addition to these issues, Fredriksson and Mamun (2008) point out that the assumption of putting the sub-national tax rate on the left-side of the equation is completely arbitrary. It is conceivable that the vertical externality works in the opposite direction – that state tax rates cause the federal government to respond.

Brülhart and Jametti (2006) tries to estimate the vertical and horizontal externalities without identifying the effects off of the slopes of reaction functions. Brülhart and Jametti (2006) theoretically shows that when horizontal externalities dominate, an increase in the number of sub-national governments will lower equilibrium tax rates. If vertical externalities dominate, an increase in the number of sub-national

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

governments will raise equilibrium tax rates. Using a cross-section of Swiss cantons, the paper regresses local tax rates on canton tax rates and the number of localities in a canton. By doing so, they implicitly assume that there is no horizontal competition between localities of different cantons – even if these localities are contiguous. Under this assumption, the coefficient on the number of localities is positive, which suggests that vertical externalities dominate horizontal externalities.

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

4.2.3 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 the industrial organization and tax competition literature would benefit from more familiarity with each other.

4.3 Model

The following model expands Devereux, Lockwood and Redoano (2007). The geographic setup of the model features sub-federal jurisdictions (towns) located on a (possibly) infinite length horizontal line segment. Federations (counties) are indexed $j = 1, 2, ...M$ and are composed of $m+j$ towns each such that each sequential county has one more town than the previous; county 1 will have $m+1$ towns, but county M will have $m + M$ towns. The ordering of counties in this manner is not important, but it will allow me to characterize the model's solution using a simpler notation than if counties were organized arbitrarily along a line segment. Towns are indexed $i = 1, 2, ...M(m + \frac{1+M}{2})$. The M federations in the model compete with each other and with the towns in the model. All of the federations are within a common union (state). Towns may differ in size (both population and length). Each town has n_i residents and is l_i units long on the line segment. The relationship between length and population is that $n_i = \phi_i l_i$ so that ϕ_i denotes the population density. Consumers and producers are located at every point along the continuous line segment.

Consumers have preferences over a consumption good and another untaxed good (i.e., dollars or leisure). I assume that the producer price of the consumption good is fixed and constant at p_i across all towns and normalize it to one. This assumption follows from having a continuum of producers in the model. Producers cannot manipulate cross-border shopping by changing the pre-tax price. Rather, they are perfectly competitive and set prices equal to marginal cost no matter the location of the firm along the line segment. Firms enter or exit to meet additional or reduced demand.

Demand for the taxed good is denoted x. The untaxed good is assumed to be the numéraire and is used as payment for purchase of the consumption good. Every consumer has a utility function $u(x, \bullet)$ that is strictly increasing and concave with respect to x. Utility is linear in the untaxed numéraire good.

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

Individuals face a choice regarding how much of the consumption good to purchase, as well as whether to purchase the good at home or abroad. When purchasing the good in the home town of residence, the consumer goes to the store located at the same point of the line segment that she lives on. If this is the case, no transportation cost is incurred and the resident pays $q_i x_i$. Alternatively, the shopper can purchase the good in a neighboring town. If the resident of town i shops in jurisdiction $k \neq i$, the individual will pay $q_k x_k$ plus any transportation cost of traveling to the border. The transportation cost function $C_i(d)$ is assumed to be linear in distance to the border, d, such that $C_i(d) = c_i d$. Note c_i denotes a constant per unit of distance cost for traveling to the border and is independent of x, so that the amount of the good

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

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

purchased does not change the transportation cost of the buyer. All cross-border shoppers will purchase the good from the first store in the neighboring jurisdiction and are constrained from shopping multiple towns over.[5]

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

$$d \leq \frac{v(q_k) - v(q_i)}{c_i}, \tag{4.1}$$

from the border of town k.

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

$$R_i = t_i B_i \tag{4.2}$$

where B_i is the tax base defined below.[7] The tax base for a town is influenced by tax

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

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

[7]Maximizing revenue is a simplifying assumption that allows for explicit characterization of the

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 tax base is defined as the sum of residents who shop at home plus the individuals that cross-border shop, which are multiplied by the demand function $x(q)$ to account for elastic demand. In order to define the tax base, the direction of cross-border shopping needs to be specified. Because I am introducing asymmetric federations in the model, county tax rates will be different in equilibrium because counties differ in size. With an infinite line segment, there are an infinite number of possible cases to consider, so simplifying assumptions need to be placed on the problem. Recall that counties are indexed $j = 1, 2, ...M$ and contain $m + j$ towns. I assume that the length of a town is identical for all towns in the model. Because each county is ordered such that it has one more identical town than the previous county along the line, the length of each county increases as j increases, which implies that n is also increasing in j. Kanbur and Keen (1993) and Nielsen (2001) show that tax rates are increasing as the size (population or geographic size) increases. Using the intuition from Kanbur and Keen (1993) and Nielsen (2001) that the perceived elasticity of cross-border shopping in a big county (one with more identical towns) is inelastic relative to a small county, I can conclude that the Nash equilibrium of county tax rates will follow the following pattern: $\tau_1 < \tau_2 < ...\tau_M$. This assumption places no restriction on the pattern of local taxes within a county. Agrawal (2011*c*) shows that revenue maximizing towns of identical size will set higher rates the closer to a high-tax county neighbor and lower rates closer to a low-tax neighbor. I assume this pattern holds when evaluating the reaction functions below because the models differ only in whether demand is perfectly inelastic or not.[8] Under this assumption,

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.

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

the tax rate in town $i-1$ will always be lower than the tax rate in town i. Therefore, residents on the west portion of town shop abroad, while additional entry occurs on the eastern side of the town.

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

$$B_i = x(q_i)[n_i + \rho_{i+1}(v(q_i) - v(q_{i+1})) - \rho_i(v(q_{i-1}) - v(q_i))] \tag{4.3}$$

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

The goal of the subsequent exercise is to derive the slope of the reaction functions with respect to the tax rates of a higher level of government. The key is to see how the tax rate of a competing jurisdiction affects marginal revenue in the region of marginal revenue equal 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 4.1 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.

pattern in the perceived elasticities.

4.3.1 Nash Equilibrium

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

$$\frac{\partial R_i}{\partial t_i} = B_i + t_i \frac{\partial B_i}{\partial q_i} = x_i s_i + t_i s_i x_i' - t_i x_i^2 (\rho_i + \rho_{i+1}) = 0, \tag{4.4}$$

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

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

$$\frac{t_i}{q_i} = \frac{1}{-\frac{q_i}{B_i}\frac{\partial B_i}{\partial q_i}} = \frac{1}{\varepsilon_i + \theta_i}, \tag{4.5}$$

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

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

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

4.3.2 Strategic Interaction

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

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

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

4.3.2.1 Slopes With Respect to County Tax Rates

A town is called "interior" if it borders two towns within the same county. A town is called "peripheral" if it borders one town in another county.

Proposition 4.1. *In the neighborhood of a Nash Equilibrium, the slope of a towns' reaction functions with respect to the county tax rate is larger for interior towns than for periphery towns.*

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

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

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

Consider the case of a homogeneous ρ for all towns. Because the slope of the reaction function measures how responsive localities are to county tax rates, a larger slope implies the vertical externality is more likely to generate upward sloping reaction functions. Consider a town at the interior of the county. This town borders two other towns that fall under the jurisdiction of the same county rate. Changes in the county tax rate will directly affect the tax base of this town via three channels. It will change the demand function for individuals because x is a function of q. Changes in the county tax rate will also distort the number of individuals living in i and purchasing

goods in $i-1$.[10] It will also distort the inflow in cross border shopping from town $i+1$. However, because the price rises by the same amount in both locations, it mitigates the change in cross-border shopping relative to the case if prices rose in only one town. On the other hand, for a town at the county border, the change in the county tax rate will directly distort the demand function for individuals and the number of individuals cross-border shopping in one direction. However, the post-tax price remains unchanged on one side of the town's border. Therefore, for a town neighboring a high-tax county, an increase in the county rate will substantially reduce inflows. For a town neighboring a low-tax county, an increase in the county tax rate will substantially increase outflows because the rate changes on only one side of the border. Either way, the tax base becomes smaller relative to an interior town and the elasticity becomes larger. Therefore, the peripheral town has more incentive to lower its rate in response to a higher county rate relative to the interior neighbor.

Another way to say this is that being located on the periphery makes it more likely that the reaction function is downward sloping. The intuition is that periphery towns want to capture additional cross-border shoppers from the neighboring county if they are in a low-tax county or they want to discourage their residents from leaving the county if they are in a high-tax county. Unlike interior towns, this leakage or capture is exaggerated when the county tax rate increases. For interior towns, the leakage changes somewhat, but remains relatively similar.

Mathematically, define the following cross-price elasticity

$$\theta_{i,k} = \frac{q_k}{B_i}\frac{\partial B_i}{\partial q_k}\frac{q_i}{q_k} = \begin{cases} \frac{\rho_k x_k q_k}{s_i}\frac{q_i}{q_k} & \text{for } i < k = i+1 \\ \frac{\rho_i x_k q_k}{s_i}\frac{q_i}{q_k} & \text{for } i > k = i-1 \end{cases} \tag{4.6}$$

Recall $\theta_i = \frac{q_i x_i(\rho_i+\rho_{i+1})}{s_i}$ is the response of town i's number of shoppers with respect

[10] 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.

to changes in the town's own price, q_i. The interpretation of $\theta_{i,k}$ is the elasticity of the number of shoppers purchasing goods in town i with respect to a change in the neighboring price. For ease of interpretation, the elasticity is scaled by the price ratios to account for the fact that the equilibrium is asymmetric. Note that in a symmetric equilibrium, θ_i is approximately two times $\theta_{i,k}$ because the town's own price influences two borders.

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

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

$$[\varepsilon_i(\theta_i - \varepsilon_i - \eta_i) + \theta_i(\theta_{i,k} - \theta_i)]/D_i \tag{4.7}$$

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

$$[\varepsilon_i(\theta_i - \varepsilon_i - \eta_i) + \theta_i(\theta_{i,i+1} + \theta_{i,i-1} - \theta_i)]/D_i \tag{4.8}$$

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

neighbor's tax rate will cause less leakage out of your town or more entry in to your town, but not both. How large this distortion is will depend on ρ and increases as density increases or if costs of shopping decline. More generally, Proposition 1 can be presented as:

Corollary 4.2. *In the neighborhood of a Nash Equilibrium, the slope of a town's reaction function with respect to county tax rates is increasing in the number of within county neighbors if all towns have the same number of total neighbors. Keeping in mind that this slope may be positive or negative, the more within county neighbors a town has, the more likely the reaction function is upward sloping.*

4.3.2.2 The Diagonal Externality

Towns at the periphery of a county are also affected by the county tax rate in the neighboring county. This occurs because the peripheral town shares a border with the neighboring county. When the neighboring county changes its tax rate, towns bordering it will react to this change and adjust their tax rates accordingly. Such an externality is horizontal because it involves a neighbor's tax rate, but it is vertical because it is with respect to another level of government. Call this externality a diagonal externality.

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

Proposition 4.3. *In the neighborhood of a Nash equilibrium, the slope of a peripheral town's reaction function with respect to the neighboring county tax rate is identical to the slope of the reaction function with respect to the neighboring town's tax rate.*

The implication of this proposition is that the diagonal externality is exactly equivalent to a horizontal externality when individuals are restricted to shopping

one town over.[11] 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 literature has ignored this externality by focusing on models of tax competition that are restricted to a single federation. Failing to account for this externality in any regression will yield to biased estimates of the horizontal externality if the researcher specifically desires to estimate the same level of government horizontal effect.

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

$$[\theta_i \theta_{i,k}]/D_i \tag{4.9}$$

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

Notice that the slope of the town's 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

[11] If the assumption of shopping one town over were relaxed, diagonal and horizontal externalities would not likely be identically equivalent, but would probably be of the same sign.

with respect to the neighboring county rate. Also recall that the elasticity θ increases with the intensity of horizontal tax competition – it increases when the 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 regression.

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

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

Proposition 4.4. *The closer the tax rate of an interior town i is to the tax rate of its high-tax neighbor (relative to the low-tax neighbor), the more likely the reaction function of town i will be upward sloping with respect to the county tax rate.*

Recall $i-1$ is the low-tax neighbor and $i+1$ is the high-tax neighbor. When a town's rate is close to the high-tax neighboring town and the county rate increases in both jurisdictions, the change in $v(q_i)$ and $v(q_{i+1})$ are relatively similar – because the prices are similar in both jurisdictions. However, when the county tax rate changes, the change in $v(q_{i-1})$ is much larger than the change of $v(q_i)$ – because the indirect utility function is downward sloping and convex with respect to prices.[12] The

[12]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

implication of this is that the elasticity of the tax base with respect to the town's own price θ_i is small relative to the cumulative cross price elasticities. As a result, $(\theta_{i,i+1} + \theta_{i,i-1} - \theta_i) > 0$.

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

Consider Figure 4.3.2.3. In the figure, the prices start at q_i, q_{i-1} and q_{i+1}. Assume that town i is internal to the county. Then, if the county rate increases, all prices rise by a constant amount to the bolder lines on the graph. The amount of cross-border shopping is proportional to the difference between the indirect utilities. In the top graph, the neighbor has a higher price so cross-border shopping is inward to i. In the second graph, the neighbor has a lower price, so cross-border shopping is outward from i. Note that $v(q_i) - v(q_{i+1})$ becomes smaller after the tax increase and $v(q_{i-1}) - v(q_i)$ also becomes smaller – but it falls by a much larger amount because of the convexity of the indirect utility function. The closer q_i is to its high-tax neighbor, the smaller the change in $v(q_{i-1}) - v(q_i)$ and vice-versa. This mitigates the change in outward cross-border shopping and amplifies the change in inward cross-border shopping.

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

Intuitively, the transportation cost function does not depend on the tax rate. Therefore, no matter the tax rate, an individual must pay cd to cross-border shop. When the price increases because the county tax rate increases, demand will decrease.

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.

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.

4.3.2.4 Interaction With State Tax Rates

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

4.4 Implications of the Theory for the Empirical Analysis

I now turn to estimating the strength of the strategic reaction empirically. The empirical component of this paper is important for three reasons. First, the theoretical results imply that many of the parameters – including the sign of the vertical reaction – are ambiguous. Thus, whether county and town tax rates are strategic complements or strategic substitutes remains an empirical question. Determining whether the slope of the reaction function is positive or negative will also provide information regarding the relative magnitudes of the elasticity of cross-border shopping and the elasticity of demand. Second, even in cases where the theoretical predic-

tions are unambiguous – such as the case of horizontal interactions – the theory is not informative of the intensity of the strategic interaction. The empirical results below will indicate how intense the vertical and horizontal relationships are and will be informative of whether taxes may be too high or too low in equilibrium. Finally, the theoretical model above demonstrates that the effect of vertical externalities is heterogeneous with respect to spatial location. This source of heterogeneity makes identification of the strategic interaction convincing; any omitted variable that may be a threat to identification must now also be correlated with distance to the federation's borders. Before explaining the empirical methodology, I summarize the theoretical implications for estimating horizontal, vertical, and diagonal reactions.

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

$$t_{ij} = \alpha_0 + \alpha_1 t_{-i} + \alpha_2 \tau_j + \delta X_{ij} + \gamma F_j + \varepsilon_{ij} \tag{4.10}$$

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

$$t_{-i} = \sum_{k \neq i} w_{ik} t_k. \tag{4.11}$$

Denote w_{ik} as exogenous weights normalized such that $\Sigma_{k \neq i} w_{ik} = 1$.[13] Tax rates on the right-hand side are usually instrumented for with the weighted average of the neighbors' X. The specification above is the one implemented in Devereux, Lockwood and Redoano (2007). The coefficient on t_{-i} identifies the strategic reaction in response to a jurisdiction's neighboring tax rates. The coefficient on τ_j identifies the strategic reaction to the vertical externality or $\frac{\partial t}{\partial \tau}$.

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

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

4.4.1 Why Interaction Effects Are Essential for Estimation

The standard literature has estimated vertical externalities using only the level of the local tax rate and the federation tax rate. From the slopes of the reaction functions, it is evident that θ_i captures the strength of the horizontal externality. Recall that θ_i represents the response in the number of cross-border shoppers and that it is a function of how costly cross-border shopping is and the density at the border of the town, which is endogenous to local tax rates. Because the slopes of the reaction functions depend on θ_i, the reaction to the vertical externality becomes more intense as the horizontal externality is increased. Devereux, Lockwood and Redoano (2007) recognize "...there is an interaction between vertical and horizontal tax competition. ...an increase in horizontal tax competition makes it more likely that the vertical slope

[14] 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.

is positive" but in their empirical analysis do not include an interaction of neighboring local rates with the federation rate. Any specification omitting this interaction will suffer an omitted variable bias with a sign that is discussed below.

To see this empirically, it is useful to consider a multi-level model[15] of tax competition.[16] 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 (4.10) – on neighboring tax rates

$$t_{ij} = \alpha_{0j} + \alpha_{1j} t_{-i} + \varepsilon_{ij}, \tag{4.12}$$

but where it is also known that each i jurisdiction is within a j jurisdiction. As a result of having multiple levels of government, it is known that county tax rates τ_j, which only vary across the j level and not within the j level of the model, affect t_{ij} with some error. The following equation includes this effect to (4.12):

$$\alpha_{0j} = \gamma_{00} + \gamma_{01}\tau_j + u_{0j}. \tag{4.13}$$

It is clear that substituting (4.13) into (4.12) will yield (4.10) sans controls. This is where the literature on tax competition within federations stops. However, the theoretical results in this paper and in Devereux, Lockwood and Redoano (2007) indicate that the empirical specification is further complicated by an interaction effect which also determines t_{ij}. From the theory, the effect of t_{-i} depends on τ_j and vice

[15] For a critical review of multi-level modeling, see Franzese (2005).

[16] 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.

versa. Assuming that this interaction also occurs with error, this implies

$$\alpha_{1j} = \gamma_{10} + \gamma_{11}\tau_j + u_{1j}. \tag{4.14}$$

Substituting (4.13) and (4.14) into (4.12) yields

$$t_{ij} = \gamma_{00} + \gamma_{01}\tau_j + (\gamma_{10} + \gamma_{11}\tau_j)t_{-i} + u_{1j}t_{-i} + u_{0j} + \varepsilon_{ij}, \tag{4.15}$$

which is not the same as equation (4.10). The implication is that the existing literature has estimated the strategic reaction as $\frac{\partial t}{\partial \tau} = \gamma_{01}$ despite the true reaction being $\frac{\partial t}{\partial \tau} = \gamma_{01} + \gamma_{11}t_{-i}$ and where the estimates of γ_{01} are different across the two specifications because they are derived from different models. Failure to account for this interaction effect will yield biased estimates of γ_{01} where the bias is given by $\hat{\gamma_{11}}\frac{cov(t_{-i},t_{-i}\tau_j)}{var(t_{-i})}$ with $\hat{\gamma_{11}}$ expected to be positive and the covariance is of an indeterminate sign.

4.5 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.

4.5.1 Data

I make us of 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 created the

[17]The data is proprietary but was provided to me without charge. For a complete description of the data see http://www.prosalestax.com/

following variables: the driving distance 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 the length of each jurisdiction's perimeter and area.

Driving distances are measured from the population weighted centroid of each Census Place to the nearest intersection of a major road and a state or county border crossing.[19] Population weighted centroids are calculated as the balance point at which an imaginary flat surface of the Census place would balance given the population distribution of Census blocks within the place. The driving distance calculated is the distance that minimizes the time to drive to the closest border. For a detailed description of this calculation, see Agrawal (2011*c*).

Contiguity is often not a satisfactory measure of the strength of horizontal competition because many places have zero contiguous neighbors if completely surrounded by unincorporated places. To calculate broader measures of neighborliness, I define a jurisdiction as neighbors if they are within a twenty-five or fifty mile buffer of each other.[20] To measure the diagonal externality, I define the neighboring county as the county that is closest to the town based on the distance criteria above.

Control variables are from the 2000 United States Census plus geographic and political controls that I have calculated myself. The set of possible control variables

[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 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.

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

include a municipality's area, perimeter, number of contiguous neighbors, population, the fraction of individuals who work in state or in county, average income, education, the fraction of seniors, whether the town is near the ocean or an international border,[21] and the Obama vote share from 2008. Table 4.1 lists all of the control variables used in the regression equation, along with summary statistics. In addition, Agrawal (2011*c*) 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 distance to the nearest state border and the complete set of interactions of these variables.[22] Including these terms will help to control for diagonal externalities across state lines that are not the focus of this paper.

4.5.2 Estimation in a Cross-Section

I will focus on how towns react to the county level rates, neighboring county rates, and neighboring town rates. Thus, the specification I estimate is

$$\begin{aligned} t_{ij} = \quad & \alpha_0 \underbrace{+\alpha_1 \tau_{i,j}}_{Vertical} \underbrace{+\alpha_2 t_{-i}}_{Horizontal} \underbrace{\alpha_3 t_{-i}\tau_j}_{Interaction} \underbrace{+\alpha_4 \tau_{i,-j}}_{Diagonal} \\ & \underbrace{+\alpha_5 \tau_{i,j} d_i}_{Vertical-Distance} \underbrace{+\alpha_6 \tau_{i,-j} d_i}_{Diagonal-Distance} \\ & +X_{ij}\beta + S_s + \varepsilon_{ij} \end{aligned} \tag{4.16}$$

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

[22] These are the same terms in Agrawal (2011*c*) 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*c*) 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.

where t_{ij} is municipal plus sub-municipal local option taxes and t_{-i} is defined in equation 4.11. Specifically, if town k is within fifty miles of town i, the weights in the main specification are equal to one divided by the number of jurisdictions within fifty miles of town i and zero otherwise.[23] Thus, the interpretation of t_{-i} is the average tax rate of town i's neighbors. Defining N_i as the set of towns within a fifty mile region of town i, then

$$w_{ik} = \begin{cases} \frac{1}{n_i} & \text{if } k \in N_i \\ 0 & \text{if } k \notin N_i \end{cases} \tag{4.17}$$

where n_i is the number of towns in N_i.

The matrix X_{ij} are the local controls listed above and S_s are state fixed effects that control for the level of the state tax rate in a state along with other within-state policies. Then, $\tau_{i,j}$ is defined as the county tax rate that town i is located in[24] 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 multi-dimensional problem into a single dimension. Finally, in a similar spirit, d_i is a measure of distance from the town centroid to the nearest county border and is either linear in distance to the county border, D_i, or the log of it, $\log(D_i)$. The inclusion of this distance function is driven by the theory and the use of the log specification implies that the effect of distance will be close to zero for very large distances. Proposition 1 states that the strategic reaction to one's own county tax

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

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

rate is more likely to be positive for towns closer to the border.

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

$$
\begin{aligned}
E[\tfrac{\partial t}{\partial \tau_j}] &= \frac{1}{M}\sum_{i=1}^{M}(\alpha_1 + \alpha_3 t_{-i} + \alpha_5 d_i) \\
E[\tfrac{\partial t}{\partial t_{-i}}] &= \frac{1}{M}\sum_{i=1}^{M}(\alpha_2 + \alpha_3 \tau_{i,j}) \\
E[\tfrac{\partial t}{\partial \tau_{-j}}] &= \frac{1}{M}\sum_{i=1}^{M}\alpha_4 + \alpha_6 d_i,
\end{aligned}
\tag{4.18}
$$

where M is the total number of observations in the estimating sample. Standard errors for mean derivatives are calculated using the Delta Method. The marginal effects given by Equation (4.18) are a consistent estimate of the mean derivative in the conditional population.

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

$$\varepsilon_{ij} = \rho w_{ik}\varepsilon + \upsilon_{ij} \tag{4.19}$$

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

The presence of the spatial lags of tax rates on the right hand side of the equation implies that standard ordinary least squares results will be biased because neighboring town and county tax rates are endogenous; tax rates are determined simultaneously. Two possible solutions exist – maximum likelihood estimation (Case, Hines and Rosen 1993) and instrumental variables estimation (Figlio, Koplin and Reid 1999). The presence of both horizontal, vertical, and diagonal externalities in the estimating equation along with spatial auto-correlation in the error term make use of the maximum likelihood method extremely difficult. Instrumental variables via general-

ized method of moments has the advantage of generating consistent estimates of the strategic reactions, even in the presence of the spatial error dependence (Kelejian and Prucha 1998).

The standard instruments for neighboring tax rates t_{-i} are the weighted average of several control variables (i.e., the entire set of $\Sigma_{k \in N_i} w_{ik} x_{kj}$, where x_{kj} is a column in X.) Although this instrument is used mostly for state level tax competition, it has also been applied to local level tax competition. I, however, am less inclined to believe that the neighbors' X's (for example, population) have no direct effect on town i's tax rate. In fact, standard theoretical models imply that this is exactly the case. Although population directly influences town k's tax rate, which then indirectly changes town i's tax rate, town k's population – and density – directly determine the elasticity of demand that town i realizes, which is inversely proportional to the optimal tax rate.

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

In order to justify the instruments, recall that the regression equation controls for town area and town perimeter. Then, the exclusion restriction requires that the county-level and neighboring jurisdiction instruments should have no partial effect on local taxes *after* controlling for these individual town-level variables. Absent any non-linear relationships between county variables and local variables, this is likely to

be the case. The direct impact of county area and county perimeter on local taxes is likely to be zero. County area and perimeter affect the county's tax rates, but will have no direct impact on the locality's tax rate so long as there are multiple jurisdictions within a county and so long as counties are sufficiently large in size.

County borders were likely to be historically drawn on latitudes and longitudes or broader geographic features. The area and perimeter of a county depend on a county's characteristics such as whether along a body of water, broader geographic features, and how counties were divided previously. Because area and perimeter are historically drawn, the evolution of time with these variables helps to make them exogenous. As a contrary point, the town's area and perimeter often depend on how municipalities were formed within the county and the characteristics within the county when the town borders were drawn – which in most cases were not at the same time county lines were delineated.[25]

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

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

4.6 Empirical Results

4.6.1 Hypotheses to Test

Before presenting the results, recall that the theoretical model provides the following testable hypotheses regarding the sign on the coefficients from equation 4.16. I relate the predictions to the precise magnitudes predicted by the theoretical model above.

- The coefficient representing the horizontal interaction, t_{-i}, is expected to be positive because towns mimic their neighbors. The magnitude is equal to $[\theta_i \theta_{i,k}]/D_i$.

- 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 magnitude is equal to $[\theta_i \theta_{i,k}]/D_i$ if individuals shop only one town over; otherwise, the effect may be muted.

- The vertical externality,τ_j, is ambiguous as the relative magnitudes of θ, ε, and the cross-price elasticities determine the sign of the vertical externality. The magnitude is equal to $[\varepsilon_i(\theta_i - \varepsilon_i - \eta_i) + \theta_i(\theta_{i,k} - \theta_i)]/D_i$ for a peripheral town and $[\varepsilon_i(\theta_i - \varepsilon_i - \eta_i) + \theta_i(\theta_{i,i+1} + \theta_{i,i-1} - \theta_i)]/D_i$ for an interior town.

- The interaction effect $t_{-i}\tau_j$ may also be ambiguous. If $\varepsilon > \theta$, then horizontal externalities will fuel vertical externalities and the coefficient should be positive. For this reason, the sign on this term will be informative of the relative magnitude of the elasticity of demand and the elasticity of cross-border shopping.

- The vertical externality will be affected by distance $d_{ij}\tau_j$ in a positive manner as interior towns are more likely to mimic the federation's tax rate because the leakage from a county government increase is mitigated by being far from the border. Internal towns have reaction functions that are approximately $(\theta_{i,i+1} + \theta_{i,i-1} - \theta_{i,k})/D_i$ larger than peripheral towns.

- Finally, the diagonal externality will be affected by distance through $d_{ij}\tau_{i,-j}$ and the effect is expected to be negative; interior towns are less likely to mimic the neighboring federation's tax rate because they are far from the border.

4.6.2 Main Results

Table 4.2 presents the baseline results using GMM-IV estimation, where the measure of the horizontal interaction is the average tax rates within a fifty mile region (i.e., a buffer) of the town of interest. All of the regression equations in the table contain state fixed effects and the control variables discussed above. The coefficients on the endogenous regressors from the second stage are reported in addition to the mean derivatives, which are the slopes of the reaction functions after accounting for interaction effects and distance. Although previous papers have focused on the coefficients directly, these estimates are likely to be misspecified if the interactions are not included as regressors. The mean derivative is the most important measure of strategic interaction and, therefore, will be the focus of the following discussion. Each of the columns from 1 to 7 build sequentially on what the current literature has estimated, with column 7 being the specification that the theory suggests is accurately specified.

Column 1 presents the results where the vertical reaction is estimated alone as in the baseline specification of Besley and Rosen (1998). Although the theory predicts the sign of this reaction may be ambiguous, Besley and Rosen (1998) find that the coefficient is positive for state cigarette and gasoline taxes with respect to the level of the federal tax. I find a significant and largely negative result that is consistent with the regression discontinuity results in Agrawal (2011*c*). Several explanations exist for this finding. Towns may react in a different manner to county rates than states will react to the federal government as the institutional structure of lower level governments is different. Alternatively, municipalities face a much more mobile

tax base relative to states. The increased mobility of the tax base implies that the elasticity of cross-border shopping, θ, is more likely to be larger for local governments than for state governments – and the theoretical model implies the larger θ is relative to ε, the more likely that county and local rates will be strategic substitutes. Column 2 estimates the horizontal reaction function and finds a significant positive relationship, perhaps indicative of yardstick competition. This positive relationship is predicted unambiguously by the theory and has been found in many studies of horizontal tax competition.

Column 3 presents a similar specification to Devereux, Lockwood and Redoano (2007), where the vertical and horizontal reactions are estimated jointly, but without any interaction effects. In general, Devereux, Lockwood and Redoano (2007) find positive coefficients on the sign of the horizontal interaction and positive but insignificant results for the vertical interaction. Again, the results in Column 3 indicate large and positive effects for neighboring local tax rates and a large negative effect with respect to the county rate. The estimates in this equation suggest that a 1 percentage point increase in county tax rates lowers municipal tax rates by .836 percentage points. Contrarily, a 1 percentage point increase in the average of the neighbors tax rate, increases a municipality's local tax rate by .456 percentage points. Keeping in mind that this is the most robust specification currently estimated in the literature, it is useful to compare the future results to these benchmark numbers.

Before proceeding, whether the instruments are valid and are strong is important for identifying consistent estimates of the coefficients. The first stage R^2, the magnitude and the precision of the instrumental variables – area of the county, perimeter of the county, and the averages of neighboring areas and perimeters – indicate that the instruments are able to explain variation in the endogenous regressors. In the case of two endogenous regressors, instrument weakness does not appear to be a concern. In the table, I report the robust Kleibergen-Paap Wald rk F statistic and the Stock

and Yogo (2005) critical values for tests of 10 percent maximal bias induced by weak instruments. When the critical value falls below the test statistic, the bias from weak instruments is less than 10%.[26] In most every specification where critical values are tabulated, the bias is less than 5%. Rejecting instrument weakness using this test is comparable to rejecting the instrument validity with an F statistic less than 10 in cases of a single endogenous regressor. The tables also report the p-values for a Hansen J test of over-identification. Failure to reject the null hypothesis suggests that if one instrument is valid, the other instrument is also valid.

Specification 4 adds the interaction of the vertical and horizontal externalities and specification 5 accounts for the diagonal externality. Looking at the mean derivatives, the slopes of the reaction function fall by almost .15 percentage points. Although the diagonal externality is of an unexpected sign, the result is not significant at traditional levels in most specifications. Specification 6 and 7 (the preferred specification) add in the interaction of the vertical and diagonal externalities with a distance function. Notice that the slope of the reaction function with respect to the county tax rate falls to -.640 and the slope of the horizontal reaction function falls to .399. The estimates in column 3 are 30% and 15% larger than the vertical and horizontal reactions in the complete specification of column 7. Columns 8 and 9 include only a sub-set of the endogenous regressors in column 7 in order to demonstrate what the reaction functions would look like if only the vertical and diagonal elements were included.

Finally, column 10 provides an important verification that the instruments in column 7 are not weak, which is important because the Stock and Yogo (2005) critical values are not tabulated for the case of six endogenous regressors. With many variables in need of instrumenting, the concern of weak instruments become more worrisome. As an alternative, I estimate specification 7 by limited information maximum

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

likelihood (LIML). If the estimating equation is correctly specified and the distributional assumptions hold, LIML will produce unbiased coefficient estimates even if weak instruments are present. Because the point estimates in column 10 are similar to column 7 and because the mean derivatives are still smaller than specification 3, it suggests that the absence of weak instruments in the GMM estimates.

Before proceeding to the robustness checks, it is worth summarizing the implications of the results above. Possible misspecifications of the identifying equation by failing to account for diagonal externalities, the interaction of horizontal and vertical externalities, and the distance to federation borders will yield the researcher to believe that the slopes of reaction functions are larger in absolute value than they actually are. Incorrectly estimating these parameters has policy implications. Whether taxes are too high or too low in equilibrium depends on the relative magnitudes of the vertical and horizontal externalities. While the slopes of the reaction functions are not directly informative of the relative magnitudes, a larger (more positive) vertical strategic interaction will imply that local and federal taxes are more intensely strategic complements. As a result, a larger estimate will imply that taxes are less likely to be too high in equilibrium. Therefore, when estimating vertical and horizontal reaction functions it is essential to include these terms in the regression equation. Even if the federation has no horizontal competitors of its own, the interaction effects of neighboring jurisdictions and higher levels of jurisdictions are essential variables. Lastly, if the federation has multiple competitors of its own, diagonal reactions and reactions contingent on distance to the nearest federation should be included in the estimating equation if possible.

Finally, the magnitudes of the strategic interactions in this paper are different and, for the vertical externalities, opposite in sign to the traditional literature. This suggests that using comprehensive local data will produce different strategic interactions than using state level data. Again, the intuition is that the elasticity of cross-border

shopping is much larger for smaller jurisdictions, which from the theory suggests that the vertical reaction is more likely to be negative. Relatively little research has analyzed local reaction functions and those papers (Revelli 2001; Büttner 2003) that have analyzed local data have often used data from only one state within a country. The evidence in this paper suggests that it would be incorrect to assume that the slope of the reaction function for a state government looks anything like the reaction function for a local government.

4.6.3 How Robust Are The Results?

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

4.6.3.1 Alternative Definitions of Neighborliness

In the results presented above, I assumed that jurisdictions competed with neighbors in a fifty mile region of the town borders. As an alternative, I use the average tax rate of jurisdictions within 25 miles as a measure of neighborliness.[27] Table 4.3 presents the results. Columns 1 - 7 are identical in all respects to Table 4.2 except the measure of neighborliness is restricted to jurisdictions in a smaller vicinity of the town. Column 8 reproduces the previous column using LIML instead of GMM. Generally, the intensity of horizontal competition is more intense when neighborliness is defined in a more restrictive manner. This is consistent with a municipality reacting more intensely to closer jurisdictions.

Again, column 3 is the specification currently estimated in the literature. The slope of the vertical reaction function implies a 1 percentage point increase in the county tax rate lowers municipal tax rates by .705 percentage points. A 1 percentage

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

point increase in the average neighbor tax rate raises municipal tax rates by .387 percentage points. The sign and significance of these results is similar to Table 4.2, but the magnitude of the marginal derivatives is slightly smaller. In the preferred specification (including interaction effects, distance effects, and diagonal externalities), the slope of the vertical reaction function is -.439 and the slope of the horizontal reaction function is .374. With regard to the estimate of the reaction function with respect to the county rate, the slope is smaller than before. However, the results in column 3 are now 60% larger than the results from the complete specification in column 7, suggesting that the incomplete specification of column 3 potentially will induce a larger bias when the definition of one's neighbors is more restrictive.

4.6.3.2 Alternative Distance Functions

In all of the previous specifications, I have assumed that the distance function to county borders is $\log(D)$. Such a parametrization imposes that as distances become very large, the effect of distance converges to zero; thus, it estimates how an increase of distance by one percent influences the strategic interaction. Although some counties are very large in size, most counties are small in size. In fact, the average town is located approximately 8 miles from the nearest county border. Because distances to county borders are relatively small, the effect of distance on the strategic interactions may actually be linear in distance. In Table 4.4, I report the results when the distance function is linear. Columns 1, 2, and 3 correspond to the results in Columns 6, 7 and 10 from Table 4.2. Columns 4, 5, and 6 of Table 4.4 correspond to the results from columns 6-8 in Table 4.3.

When the distance function is linear, the point estimates change slightly but the interpretation of the results is similar in spirit. In some specifications, the interaction of the linear distance function and the county tax rate become significant at the 90% confidence level. For example, in Column 2, $\tau_{i,j} d_i$ has a positive and significant coeffi-

cient of .012. This implies that holding the county tax rate fixed, a one mile increase in distance from the nearest county border increases the slope of the reaction function by .012 percentage points. Such a result is consistent with the theory – interior jurisdictions are more likely to have upward sloping reaction function. However, the small magnitude of this coefficient suggests that the strategic reaction to the county rate will remain significantly negative for all reasonable distances.

4.6.3.3 Various Model Specifications and Sample Restrictions

Table 4.5 reports robustness checks for various sample restrictions and weighting schemes. Column 1 presents the baseline results. Column 2 drops jurisdictions for which the nearest county border is also a state border. This specification reduces the possibility that the neighboring state is also producing a diagonal externality on towns proximate to its borders and allows me to cleanly identify the diagonal externality resulting from neighboring counties. The measure of the diagonal interaction becomes larger in absolute value and the measures of the horizontal and vertical interactions become slightly smaller in absolute value. As a second robustness check, column 3 drops jurisdictions where the nearest state border would be an ocean or a major body of water. One reason for this is that theoretical models of tax competition using a Hotelling line segment where some jurisdictions have one neighbor and other jurisdictions have two neighbors produce very different results compared to if all jurisdictions have the same number of neighbors. Specifically, jurisdictions at the end of the line segment – e.g., proximate to the ocean – have incentives to raise tax rates because the elasticity of cross-border shopping is perceived as smaller. Dropping jurisdictions near water amplifies the strategic reaction of both the vertical and horizontal interactions. Such a result is suggestive that the elasticity of cross-border shopping is, in fact, larger for jurisdictions not near the ocean, which is consistent

with the theoretical model.[28]

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

Column 5 weights each observation in the sample by population such that cities are given more weight than small towns in the data set. For the weighted sample, the interaction effects of the vertical and horizontal interaction are strongest. However, the mean derivative with respect to the county tax remains negative and significant. Although the theory suggests that horizontal interactions are unambiguously positive, the weighted results suggest that the mean derivative with respect to neighboring tax rates is negative. This results fits with the existing empirical evidence that cities set much higher sales tax rates than their local neighbors, perhaps because the elasticity of demand is smaller in cities because the tax base is relatively large in cities. As such, the benefit of raising city sales taxes may outweigh the loss of cross-border shopping to the suburbs. The weighted results is also consistent with survey results in Janeba and Osterloh (2011), which finds that large cities view other large cities rather than small nearby jurisdictions as competitors. An implication of the survey in Janeba and Osterloh (2011) is that cities should only have a weak strategic interaction with

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

neighboring jurisdictions.

Finally, columns 7-9 alter the exogenous weights that determine t_{-i}. Define N_i as the set of towns within a fifty mile radius of town i, s_{ik} as the distance between town i and town k, and ϕ_k as the population of town k. Recall that $t_{-i} = \Sigma_{k \neq i} w_{ik} t_k$. Then column 7 specifies exogenous weights that are normalized to sum to 1 given by Equation 4.20:

$$w_{ik} = \begin{cases} \frac{1/s_{ik}}{\sum_{k \in N_i} (1/s_{ik})} & \text{if } k \in N_i \\ 0 & \text{if } k \notin N_i \end{cases}, \tag{4.20}$$

which can be interpreted as inverse distance weights. In this specification towns closer to town i are given more weight than towns far away. Column 8 estimates the equation using the weights from Equation 4.21:

$$w_{ik} = \begin{cases} \frac{\phi_k}{\sum_{k \in N_i} \phi_k} & \text{if } k \in N_i \\ 0 & \text{if } k \notin N_i \end{cases}, \tag{4.21}$$

such that neighboring towns within the fifty mile radius of town i are given more weight if they have a larger population. Lastly, Equation 4.22 specifies the weights used in Column 9:

$$w_{ik} = \begin{cases} \frac{\phi_k/s_{ik}}{\sum_{k \in N_i} \phi_k/s_{ik}} & \text{if } k \in N_i \\ 0 & \text{if } k \notin N_i \end{cases}, \tag{4.22}$$

which gives the most weight to highly populated towns that are closer to town i and the least weight to towns with small populations that are far from town i. The weights above and the weights given by equation 4.17 are the most common weights

in the literature. The weights used in equation 4.17 are most likely to be interpreted as spatial weights, while the weights given by the three equations above are more indicative of economic flows.

When using distance based weights, the strategic interaction with the county tax rate remains approximately the same magnitude as the baseline specification, but the horizontal interaction shrinks in absolute value. Population based weighting schemes shrinks both the slopes of the vertical and horizontal reaction functions in absolute value. Finally, the population-distance weights shrink the vertical interaction but increase the horizontal interaction relative to the baseline specification. Although I present these results as standard weighting schemes in the literature on tax competition, I believe they are less appropriate 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 only neighbors (e.g. by population) because this implies that municipalities would not account for large jurisdictions that are far away. Most jurisdictions in America are relatively small. Distance based weights would be more reasonable for small municipalities, but less reasonable for large cities. For this reason and because of its ease in interpretation, I prefer using the unweighted average of tax rates.

4.7 Conclusion

Introducing inter-federation competition into a model of sales tax competition that combines the vertical elements of Keen (1998) and the horizontal elements of Kanbur and Keen (1993) and Nielsen (2001) indicates that the spatial composition of towns within a federation is essential to determine the strategic nature of the

tax competition. First, this paper argues that the geo-spatial nearness to borders of sub-federal governments in a federation – particularly the spatial proximity to discontinuous changes in the tax rate resulting from the federation's borders – makes it less likely that a peripheral local government will mimic the federal government. Second, inter-federation competition results in a diagonal externality – an externality induced by a different level of government that does not share the same tax base – that has similar consequences as a horizontal externality.

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

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

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

With respect the nature of the strategic interaction, local sales taxes are strategic complements with neighboring sales taxes – a finding consistently found in the literature and unambiguous in theory. With respect to vertical interactions, the theoretical literature suggests the interaction may be positive or negative. I find that a one percentage point increase in county tax rates lowers municipal tax rates by about .60 percentage points; this suggests that county and municipal taxes are strategic substitutes. The result is inconsistent with the positive and small effects found for state level governments in response to the federal government. The results in this paper suggest it would be inappropriate to generalize results from state level studies to municipal level interactions. But, as relatively few studies of tax competition have exploited comprehensive local tax data, the use of local data will provide a continued avenue for future research within federations with multiple levels of government.

Figure 4.1: Reaction Functions

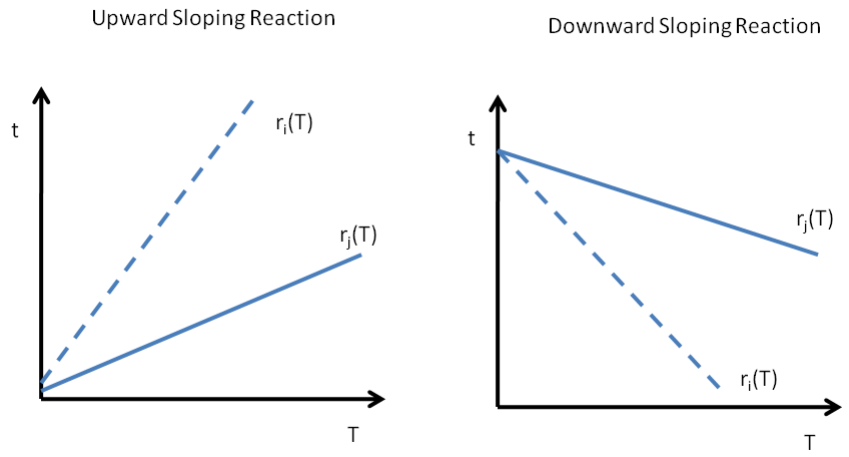


Figure 4.2: Change in Cross-Border Shopping due to a Change in the County Rate

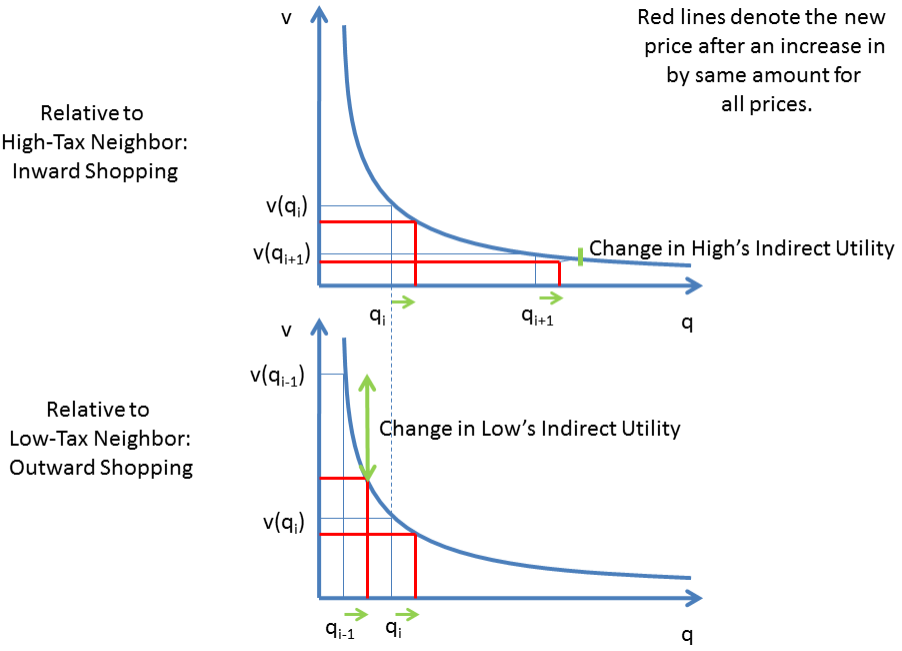


Table 4.1: Summary Statistics
Averages of Variables by Type
Standard Deviations in ( )

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

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

Table 4.2: Slopes of Reaction Functions – Neighbors Within 50 Miles

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

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

All of the specifications above include state fixed effects and the local control variables outlined in the text. The instruments are always the area and perimeter of the respective jurisdictions. In the case of horizontal externalities, the instruments are the average area and perimeter in the buffer zone of fifty miles around each town. The instrument for interactions are the interaction of the instruments. When the regression includes a tax term interacted with distance, the distance variable also enters the regression equation as a stand-alone variable.

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

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

Table 4.3: Slopes of Reaction Functions – Neighbors Within 25 Miles

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

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

The estimating equations in this table are identical to Table 4.2 except that horizontal neighborliness is defined within a 25 mile buffer region of the town rather than a 50 mile region.

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

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

Table 4.4: Slopes of Reaction Functions – When Distance Is Linear

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

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

The estimating equations in this table are identical to Tables 4.2 and 4.3 except the interaction with the distance function uses a linear distance term, rather than the log of distance.

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

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

Table 4.5: Slopes of Reaction Functions – Robustness Checks

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

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

All the specifications include the same set of controls and fixed effects as the previous tables. (1) presents the baseline results from Table 4.2. (2) drops towns where the nearest county border is also a state border. (3) drops towns where the nearest state border would be an ocean or a major lake. (4) and (5) restrict the estimating sample to towns within ten and five miles of a county border, respectively. (6) weights each observation in the regression equation by the population of the jurisdiction. (7) weights each town in t_{-i} by the inverse distance to the neighbor. (8) weights each town in t_{-i} by the population of the neighbor. (9) weights each town in t_{-i} by the population of the neighbor neighbor and then by the inverse of distance to the neighbor.

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

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

CHAPTER V

Conclusion

Chapter I presents three fundamental questions in public finance. With the evidence of this dissertation, I am now in the position to discuss some parts of these questions. In many circumstances, the discussion below suggests the continued need for research to answer the three questions.

Does strategic competition among localities and states result in a race to the bottom? As is evident, two offsetting fiscal externalities arise – one positive and one negative. The evidence in Chapter II and Chapter IV indicates that being located within a high-tax state or county results in lower municipal taxes compared to being in a low-tax state or county. As a result, if a higher level of government cuts its tax rate, the lower level of government is likely to respond by raising its tax rate. Although this reaction is not one-for-one, the possibility still exists that the increase in municipal taxes helps to mute a downward spiral in tax rates from horizontal competition among counties. Of course, this effect is counter-acted by the positive strategic reactions with respect to neighboring municipal jurisdictions. This positive reaction function implies that when the home town's neighbors reduce their tax rates, the home jurisdictions also has incentives to reduce tax rates and taxes may be too low in equilibrium. However, if the county government undercuts its neighbors, then the municipal government will raise its tax rate – and in turn, its neighbors will also

raise tax rates. Whether tax competition is reduced will inevitably depend on the form of the game, the number of players, and the slopes of the reaction functions.

This dissertation indicates that the negative strategic response of local governments to county governments is substantial in magnitude. This is in contrast to traditional state-federal interactions which finds small (and oftentimes zero) strategic reaction. Therefore, accounting for fiscal federalism at lower levels of government has the possibility to mitigate the race to the bottom induced by horizontal competition. In other words, although a state or county government may undercut its neighbors, this has the potential to allow for municipal governments to raise their local sales tax rates. As the municipal governments do this, they reduce the incentives of the neighboring state or county government to also undercut. In addition, Chapter III suggests that with welfare maximization, higher tax rates are possibly optimal in the border-region. This is not possible in a model of revenue maximization and implies that using more comprehensive objective functions for governments may mitigate the possibility of heightened tax competition.

Is coordination or harmonization of tax rates desirable? Policymakers, in the European union especially, have been concerned about fierce tax competition reducing the revenue streams of nations. Tax harmonization has been proposed as a solution to this dilemma. Tax harmonization inevitably requires member states to agree on tax setting behavior; this is no easy task. In Canada, complete harmonization of the Goods and Services Tax failed after one province objected to harmonization. Suppose that this political difficulty were overcome and state tax rates were harmonized, but that municipal rates were allowed to vary. Such a proposal would certainly make some municipalities – particularly, the municipalities previously located in states with high rates – better off in terms of revenue raising capacity. Of course, if county governments and local governments were still allowed to select different rates, then harmonization of state rates would result in municipalities in high-tax counties being

less able to raise sales tax rates than those in low-tax counties. An alternative is to harmonize municipal tax rates, which is already practiced in some states. However, harmonization of municipal rates ignores the advantages of decentralizing the tax system in the first place – that states and municipalities can select higher or lower taxes because of different preferences for public goods, preferences for equity of the tax system, or preferences for smoothing volatility in revenue streams over time. Chapter II and III start from the premise that different preferences for public goods result in tax differences across states. Harmonization of local tax rates eliminates these benefits.

If the United States were to adopt a Value Added Tax, complete harmonization may become necessary. Preferential tax rates under a Value Added Tax regime would come with much higher administrative costs than under a retail sales tax regime. However, in the context of the existing retail sales tax, harmonization of local option sales tax rates within a state would mitigate distortions across localities within the state but may accentuate distortions at borders. How states would respond to harmonization remains unknown. Instead of harmonization other possibilities exist to mitigate tax competition. For example, developing new technologies to enforce taxation on the basis of destination would reduce tax competition. Alternatively, policymakers may simply embrace decentralization as a means of reducing first-order distortions in the economy and replacing them with much smaller second-order distortions, while also providing for heterogeneity in public services.

How decentralized should the tax system be and how many levels of government should have taxing authority? More decentralization will inevitably result in more competition. Taking as given that states setting different tax rates is optimal because of differing preferences, the optimal municipal tax rates may potentially be a continuous function in distance to borders. Excluding such a tax system as being infeasible because of a finite number of tax rates, the next best solution is to discretize states

into small jurisdictions that have taxing authority. The evidence in this dissertation suggests that doing so will reduce distortions at borders. This is a first-order reduction of deadweight loss. However, there is no guarantee that the decentralized tax equilibrium needs to correspond with the state, let alone, the federal planner's optimal tax system. Although the United States tax system is highly decentralized, looking at the total effective tax rate in every jurisdiction makes it seems as if tax rates across the United States are much more similar than a comparison of the tax rates of one level of government. Looking at states sales tax rates suggests substantial heterogeneity in taxes, but looking at the population-weighted state plus local rates makes states look much more similar. Much of this is a result of the within state variation of tax rates in response to the higher levels of government.

There is no need to believe that the marginal cost of public funds need be constant within a jurisdiction. If it is not constant, then optimal taxation would predict that taxes should be set inversely proportional to the marginal costs of funds. Ignoring concerns of tax competition, the social planner would like to do this across jurisdictions that are as small as possible to best approximate the continuous tax treatment of sales. But, positive horizontal reaction functions imply that smaller units of decentralization result in more fiscal competition. Another possibility is to eliminate decentralized choice over tax rates, but instead to have the state government set local tax rates depending on their beliefs of the elasticity of demand and the marginal cost of public funds across localities. Some states already adopt this policy and have approved local option sales tax rates for particular cities or regions within the state. However, in other states where local option taxes are determined centrally, the state government has simply imposed a uniform state-wide local tax – suggesting that the state planner may be motivated by simplicity concerns.

Decentralization comes with benefits and with costs, but it is a reality of our tax system. For this reason, researchers must incorporate fiscal federalism into existing

models of tax competition and should embrace the variation arising from differences in local tax rates as a means of continuing to make progress on answering these three questions. The preceding chapters apply various theoretical and empirical approaches to study lines and notches in the tax system. The aim of these chapters has been to unify three distinct bodies of literature on tax competition, fiscal federalism, and geo-spatial economics.

This dissertation develops rigorous, yet intuitive, theoretical models that predict patterns of tax rate differentiation and then develops corresponding empirical techniques that help to identify strategic reaction in a robust and convincing manner. This dissertation has also utilized unique data sets on local taxes and geo-spatial variables. It is my continued hope that these types of local and geographic databases are more frequently utilized in public finance. Local tax data provides researchers with a rich, yet underexploited, amount of variation to identify fundamental parameters of interest. Utilizing local data allows the researcher to exploit geo-spatial data. Geographic data should be viewed as an advantage of studying local taxes; the evidence in this dissertation suggests that geo-spatial relations abound when studying local taxes. While this provides added complexities to the researcher, it also provides for ample identification strategies and novel instrumental variables.

The preceding chapters of the dissertation have focused on "notches," which refers to a discontinuity in tax liability after crossing a "line" in the tax base where the lines analyzed are geographic borders. While geographic borders are a particular type of a line – one that is historically drawn and no longer a matter of tax policy – other types of lines are ubiquitous in the tax system. For example, legislatures must determine where to draw the line between taxed and tax-free goods; changes in the tax year is a line in time space; and environmental policymakers need to create lines with respect to regulations. The presence of lines results in a notch, yet notches remain an understudied aspect of tax systems. The theoretical and empirical methods developed

are a starting point to analyzing lines and notches in the tax system. The previous chapters have focused on towns as the unit of analysis. One key aspect of towns is that they cannot move their position to the favorable side of a border-line. This is not the case with firms or with taxable income. This unique feature of geographic borders has helped to simplify the theoretical and empirical models used to study lines and notches, but my hope is that the methods introduced in this situation can be modified and tweaked to study more complex forms of lines.

BIBLIOGRAPHY

BIBLIOGRAPHY

Agrawal, David R. 2011*a*. "Games Within Borders: Are Geographically Differentiated Taxes Optimal?" Working Paper, University of Michigan.

Agrawal, David R. 2011*b*. "Inter-federation Competition: Sales Tax Externalities with Multiple Federations." WorkingPaper, University of Michigan.

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

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

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

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

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

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

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

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

Dahlby, Bev, and Leonard S. Wilson. 1994. "Fiscal Capacity, Tax Effort, and Optimal Equalization Grants." *Canadian Journal of Economics*, 27: 657–672.

Davis, Lucas. 2011. "The Effects of Preferential VAT Rates Near International Borders: Evidence from Mexico." *National Tax Journal*, 64(1): 85–104.

Devereux, Michael P., Ben Lockwood, and Michela Redoano. 2007. "Horizontal and Vertical Indirect Tax Competition: Theory and Some Evidence from the USA." *Journal of Public Economics*, 91: 451–479.

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

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

Fox, William F. 1986. "Tax Structure and the Location of Economic Activity Along State Borders." *National Tax Journal*, 39(4): 387–401.

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

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

Gopinath, Gita, Pierre-Olivier Gourinchas, and Chang-Tai Hsieh. 2011. "International Prices, Costs and Markup Differences." *American Economic Review*, 101(6): 2450–2486.

Harding, Matthew, Ephraim Leibtag, and Michael Lovenheim. Forthcoming. "The Heterogeneous Geographic and Socioeconomic Incidence of Cigarette Taxes: Evidence from Nielsen Homescan Data." *American Economic Journal: Economic Policy*.

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

Haufler, Andreas. 1996. "Tax Coordination with Different Preferences for Public Goods: Conflict or Harmony of Interest." *International Tax and Public Finance*, 3(1): 5–28.

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

Holmes, Thomas J. 1998. "The Effect of State Policy on the Location of Manufacturing: Evidence from State Borders." *The Journal of Political Economy*, 106(4): 667–705.

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

Imbens, Guido, and Karthik Kalyanaraman. Forthcoming. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." *Review of Economic Studies*.

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

Janeba, Eckhard, and Wolfgang Peters. 1999. "Tax Evasion, Tax Competition and the Gains From Nondiscrimination: the Case of Interest Taxation in Europe." *The Economic Journal*, 109(452): 93–101.

Kanbur, Ravi, and Michael Keen. 1993. "Jeux Sans Frontières: Tax Competition and Tax Coordination When Countries Differ in Size." *American Economic Review*, 83(4): 877–892.

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

Keen, Michael. 2001. "Preferential Regimes Can Make Tax Competition Less Harmful." *National Tax Journal*, 54(4): 757–762.

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

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

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

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

Kleven, Henrik Jacobsen, and Joel Slemrod. 2009. "A Characteristics Approach to Optimal Taxation and Tax-Driven Product Innovation." Working Paper, London School of Economics and Political Science.

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

Lovenheim, Michael F., and Joel Slemrod. 2010. "Fatal Toll of Driving to Drink: The Effect of Minimum Legal Drinking Age Evasion on Traffic Fatalities." *Journal of Health Economics*, 29(1): 62–77.

Luna, LeAnn. 2003. "Local Sales Tax Competition and the Effect on County Governments' Tax Rates and Tax Bases." *The Journal of the American Taxation Association*, 26(1): 43–67.

Luna, LeAnn, Donald J. Bruce, and Richard R. Hawkins. 2007. "Maxing Out: An Analysis of Local Option Sales Tax Rate Increases." *National Tax Journal*, 60(1): 45–63.

Magruder, Jeremy R. 2011. "High Unemployment Yet Few Small Firms: The Role of Centralized Bargaining in South Africa." Working Paper, University of California, Berkeley.

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

Merriman, David. 2010. "The Micro-geography of Tax Avoidance: Evidence from Littered Cigarette Packs in Chicago." *American Economic Journal: Economic Policy*, 2(2): 61–84.

Mikesell, John L. 1970. "Central Cities and Sales Tax Rate Differentials: The Border City Problem." *National Tax Journal*, 23(2): 206–213.

Mikesell, John L. 2010. "The Contribution of Local Sales and Income Taxes to Fiscal Autonomy." Working Paper.

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

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

Nielsen, Søren Bo. 2010. "Reduced Border-Zone Commodity Tax?" Working Paper.

Research Institute of America. 2010. *Guide to Sales and Use Taxes.* New York, NY:Thomson Publisher.

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

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

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

Sjoquist, David L., William Smith, Sally Wallace, and Mary Beth Walker. 2007. "An Analysis of the Time to Adoption of Local Sales Taxes: A Duration Model Approach." *Public Budgeting and Finance*, 27(1): 20–40.

Slemrod, Joel. 2007. "Cheating Ourselves: The Economics of Tax Evasion." *The Journal of Economic Perspectives*, 21(1): 25–48.

Slemrod, Joel. 2010. "Buenas Notches: Lines and Notches in Tax System Design." University of Michigan Working Paper.

Slemrod, Joel, and Shlomo Yitzhaki. 1996. "The Cost of Taxation and the Marginal Efficiency Cost of Funds." *IMF Staff Papers*, 43(1): 172–198.

Slemrod, Joel, and Shlomo Yitzhaki. 2001. "Integrating Expenditure and Tax Decisions: The Marginal Cost of Funds and the Marginal Benefit of Projects." *National Tax Journal*, 54(2): 189–202.

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

Tosun, Mehmet S., and Mark L. Skidmore. 2007. "Cross-Border Shopping and the Sales Tax: An Examination of Food Purchases in West Virginia." *B.E. Journal of Economic Analysis and Policy*, 7(1): 1–18.

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

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

Walsh, Michael, and Jonathan Jones. 1988. "More Evidence on the Border Tax Effect: The Case of West Virginia, 1979-84." *National Tax Journal*, 41(2): 261–265.