THE INTERNET AS A TAX HAVEN?: THE EFFECT OF THE INTERNET ON TAX COMPETITION

David R. Agrawal[*]
University of Georgia

This Version: August 2014
First Version: September 2013

Abstract

I study the effect of broadband availability on local sales tax rates. High Internet penetration puts downward pressure on tax rates as jurisdictions seek to reduce revenue leakage; but, taxable online sales will put upward pressure on tax rates because the Internet helps enforce sales tax collection. A one standard deviation increase in Internet penetration lowers local tax rates by 9%. Exploiting state borders, I find that that an increase in Internet penetration induces large municipalities on the low-tax side of state borders to lower their local tax rates by more than municipalities on the high-tax side.

JEL: H25, H71, H73, L81, R50

Keywords: Commodity Taxation, Fiscal Competition, e-Commerce, Tax Havens, Tax Evasion

---

[*] Contact information: University of Georgia Department of Economics, 527 Brooks Hall, Athens, GA 30602. Email: dagrawal@uga.edu. Thanks to Spencer Banzhaf, Christopher Bollinger, Sebastien Bradley, Leah Brooks, Donald Bruce, Bill Fox (discussant), Andreas Haufler, Jeffrey Hoopes, Olga Malkova, Giorgia Maffini (discussant), Tuukka Saarimaa (discussant), Georg Schaur, Ian Schmutte, Joel Slemrod, Gregory Trandel, Christian Traxler, John Turner, Ronald Warren, David Wildasin, and Jonathan Williams for helpful comments. Special thanks to Bill Fox and Jeffrey Hoopes for sharing data with me. Seminar and conference participants at Clemson University, DIW Berlin, the Federal Reserve Bank of Boston, the Institut d'Economia de Barcelona, the Lincoln Institute of Land Policy, Ludwig Maximilian University of Munich, the University of Georgia, the University of Kentucky, the University of Tennessee, the Workshop on Federalism and Regional Policy at the Universität Siegen, and Zentrum für Europäische Wirtschaftsforschung (ZEW) improved the paper. Thanks to Pro Sales Tax for providing the sales tax data. Any remaining errors are my own.

This paper studies the tax setting behavior of jurisdictions in the modern "new economy." Does the Internet cause cities and towns to raise or lower their sales tax rates? If online transactions are untaxed, the Internet may act as a tax haven and put downward pressure on local tax rates. On the other hand, if e-commerce is taxed, the Internet may act as an anti-haven, allowing cities and towns to collect taxes on remote transactions that previously went untaxed.

The Marketplace Fairness Act of 2013 was recently introduced in Congress, which if passed, will allow states to require large firms – including online vendors – to remit state and local sales taxes. Current law requires only firms with nexus within a state to remit the retail sales tax; a firm only has nexus in a state if it has a physical presence in the state from which it profits. The United States sales tax system is decentralized as most municipalities and counties are allowed to levy local sales tax rates. Once a firm has nexus within a state, it must remit sales taxes on online purchases for every municipality in the state.

In 2011, consumers spent approximately 200 billion dollars (5% of all retail sales) on online purchases.[1] Of these purchases, approximately, 11% occurred on eBay and approximately 13-19% occurred on Amazon (Einav et al. 2014); many of these were tax free.[2] Although the total amount of business to consumer online sales is relatively small, online sales have grown rapidly in recent years and policymakers expect these transactions to grow moreso in the coming years. States have reacted to this growth in online sales by negotiating deals with large online retailers and by adopting laws that require more online firms to remit retail sales taxes. This paper studies whether governments also may adjust (or stop raising) their tax rates in response to e-commerce.

Tax competition arises when governments compete with each other in order to attract a mobile tax base – in the case of sales taxes, the mobile tax base is composed of consumers who have choice on from where they want to purchase a good. Traditional studies of tax competition focus on how the tax base is mobile

---

[1] Although the focus of the theory in this paper is on consumers, firms make online transactions. This substantially increases the total volume of online purchases mentioned above.

[2] Over the last ten years, online sales have grown exponentially. News stories, council meetings, and state legislature debates indicate that lawmakers are extremely concerned about online sales. In fact, many states, such as Georgia have tried to redefine what it means to have nexus in the state in order to capture online sales tax revenue. See Appendix A.1 for a discussion of the salience of these revenue losses from the perspective of a town Mayor.

between different players (governments) of the game. But, what happens if the tax base is like a "leaky bucket" and some shoppers can go to a "jurisdiction" that is not a participant in the game, for example the Internet? At the same time that the tax base is a leaky bucket, can the Internet also expand the tax base for local jurisdictions and how does this effect influence tax setting behavior? The paper emphasizes the effect of the Internet on the strategic behavior of governments given that we know the Internet induces large changes in consumer behavior; these elasticities are important determinants of tax rates in a competitive game.[3]

I highlight the two roles of the Internet in the current regime: (1) the Internet creates a leaky bucket where jurisdictions see declines in tax revenue as a result of online vendors without nexus and (2) the Internet allows local jurisdictions to effectively enforce the sales tax on the basis of the consumer's location for online transactions originating from firms with nexus. The first effect suggests that the Internet may have similar effects on tax competition as parasitic tax havens and will result in inefficiently low tax rates. Tax havens can intensify tax competition and are thus responsible for revenue and welfare losses (Slemrod and Wilson 2009), but they may also be beneficial as tax havens can make it less attractive to compete for the mobile factor (Johannesen 2010).[4] The second point regarding the Internet as a means of tax enforcement is often ignored when thinking about online sales. Firms such as Amazon.com and eBay.com represent only a fraction of online sales. Many other online transactions occur on sites such as Walmart.com and Lowes.com – for which these firms have nexus within

---

[3] Goolsbee (2000) studies the effect of local sales taxes on whether individuals buy online or not and finds big effects in high-tax localities. Ballard and Lee (2007) studies the probability of shopping on the Internet and finds that it is highest in high-tax counties and that consumers who live adjacent to counties with low taxes are less likely to shop on the Internet. Goolsbee, Lovenheim and Slemrod (2010) merges data on Internet access from the CPS and smoking rates (by state) and show that there has been substantial increase in the sensitivity of taxable cigarette sales. Einav et al. (2014) finds that a 1 percentage point increase in sales tax rates reduces purchases on eBay from sellers that are in-state by 4% as individuals substitute to out of state sellers. In response to Amazon Laws, Baugh, Ben-David and Park (2014) find a 9.5% decline in Amazon expenditures following the passage of such a law. Ellison and Ellison (2009) also show a large response of online sales to state tax rates. The revenue implications of these responses are discussed in Bruce and Fox (2000), Alm and Melnik (2010), and Alm and Melnik (2012). Studies such as Brown and Goolsbee (2002) have shown the Internet lowers prices, but no studies have analyzed the effect on taxes.

[4] Hines (2010) describes the effects of tax havens more generally.

most every state in the United States. For many towns with no retail sales stores within the town, the presence of Walmart.com and Lowes.com elsewhere in their state actually allows for the local jurisdictions with no firms of its own to collect some taxes. This lowers the marginal cost of raising tax rates on sales as the jurisdiction can now collect taxes from online purchases; the tax rate will rise as jurisdictions shift away from other tax instruments. More generally, some consumers who would have engaged in cross-border shopping now buy taxable goods online, which reduces the competitive pressure between jurisdictions.

Thus, these two offsetting channels make the effect of the Internet on tax rates an empirical question: the pressures from having a tax "haven" for online shopping must be traded off with the "anti-haven" effect whereby the Internet helps jurisdictions easily and effectively collect taxes from remote firms with nexus. Because I can determine which of these two effects dominates, understanding the pressures arising from taxable online sales versus tax-free online transactions will shed light on how tax competition affects the current regime as well as how the Marketplace Fairness Act of 2013 would affect tax competition.

To preview the results, the model modifies the Kanbur and Keen (1993)-Nielsen (2001) model of tax competition by introducing revenue leakages from the Internet and by allowing for taxable e-commerce.[5] Consider the case of two jurisdictions with uniform population and size, but where one town is the border-town in a high-tax state and the other town is a border town in a low-tax state. If the Internet is tax-free, then an increase in access to the Internet in both jurisdictions will lower local tax rates in both jurisdictions. Critically, local taxes will fall in the low-tax state by more than in the high-tax state because the revenue lost to online sales will be lower for towns in the high-tax state given these towns set low local tax rates without e-commerce. This result intuitively arises because, even without a tax-free Internet, border towns in high-tax states are already constrained by a fear of cross-border shopping when setting their local tax rates relative to towns in low-tax states. The asymmetry of the effect of the Internet naturally yields a border design to empirically test for the effect of the Internet on tax rates. The model then shows that if the Internet is not tax-free and if some consumers have a preference to shop online (added product variety,

---

[5] Models include Mintz and Tulkens (1986), Haufler (1996), and Agrawal (forthcoming).

convenience, etc.), the presence of more Internet shoppers will place upward pressure on tax rates as some previous cross-border shoppers, now purchase taxable goods.

The paper then tests the theoretical predictions empirically using a border discontinuity design (Holmes 1998).[6] The results indicate that going from no Internet access to complete access lowers local sales taxes by between .15 and .20 percentage points, which is about 19% to 25% of the average municipal rate in 2011. Put differently, a one standard deviation increase in Internet penetration lowers local tax rates by 9% of the average local rate; Internet usage today in the average town corresponds to an approximately two standard deviation increase in penetration. I demonstrate that for very small jurisdictions, the marginal effect of additional Internet penetration has a small effect on local sales tax rates. However, for the large jurisdictions containing more than 95% of the population, an increase in Internet penetration has strong negative effects on local tax rates. Internet penetration places strong downward pressure on local tax rates on the low-tax sides of borders; the effect is insignificant on the high-tax side of the border where municipal jurisdictions were already forced to lower their tax rates in order to reduce cross-border shopping. Instrumenting for Internet usage with lightning flash density as suggested in Andersen et al. (2012) increase the magnitudes.

The results in this paper suggest that many towns perceive the Internet as a tax haven; this story is consistent with news headlines featuring politicians asking Congress to pass legislation requiring firms to remit taxes on e-commerce.

1 Local Sales Taxes and Nexus

The United States system of commodity taxation is decentralized. Sales taxes are levied at the state, county, municipal and sub-municipal level. Over thirty states allow for some sort of local taxation within the state. Within states that allow

---

[6] Border-based approaches were first used in Holmes (1998) and have been modified to study tax competition by tax avoidance by Lovenheim (2008), Merriman (2010), and Engel et al. (2013). Eugster and Parchet (2013) and Agrawal (forthcoming) modify border approaches to study tax competition. Empirical studies on the taxation of commodities include Besley and Rosen (1998) and Devereux, Lockwood and Redoano (2007). The study of local sales taxes in a competitive environment has been studied on a state by state basis in (Luna 2003; Sjoquist et al. 2007; Burge and Piper 2012). For a survey, see Brueckner (2003).

municipalities to set local sales tax rates, the statutory local rates add anywhere between zero percentage points to six percentage points on top of the state sales tax rate. Sales taxes are the second largest own-source of municipal revenue, comprising an average of 12% of local revenue.

Commodities in the United States are taxed using two separate but highly-related taxes – sales and use taxes. Based on the Supreme Court of the United States ruling *Quill Corp. v. North Dakota*, online firms are only required to remit sales taxes from consumers living in a state where the firm has nexus – a physical presence from which it profits.[7] This creates two types of online firms – those required to remit sales taxes (Walmart) and those not required to remit sales taxes (eBay in most states). Consumers of goods are then required to pay use taxes on any goods purchased in neighboring jurisdictions or from online vendors not required to remit sales taxes. The use tax is notoriously under-enforced and as a result, consumers are (*de facto*) taxed on the basis of the origin of the sale. Because the use tax is under-enforced, consumers are able to avoid (or evade) their tax liability from online purchases from firms without nexus.

2 Tax Competition, the Internet and Borders

In this section, I modify the Nielsen (2001) model of tax competition; the modification will allow for online shopping and state borders.

The model features two border towns that are both one unit long and with population density of unity. Each town indexed $i = H$, L is located in a different state; states exogenously set different state sales tax rates that apply to the locality in their state: $T_H \geq T_L$ where $b \equiv T_H - T_L \in [0, 1]$.[8] Like consumers, firms are everywhere. Consumers have inelastic demand and must purchase one unit of a consumption good with a pre-tax price normalized to one. The consumer may purchase the good at home, in which case no transportation cost is incurred or abroad, in which case the consumer pays a transportation cost, d, that is linear in the distance traveled, s. The two towns set local tax rates in order to maximize

---

[7] Issues of nexus also arise with the corporate income tax – discussed in Becker and Fuest (2012). More firms with nexus in a region may also mean stricter tax enforcement, which is related to the literature on fiscal competition via tax enforcement (Stöwhase and Traxler 2005).

[8] Regarding the exogeneity of state tax rates, state tax rates will not depend on any one municipal choice if towns are small relative to the state tax base.

local tax revenue while competing in a Nash game. The town in the high-tax *state* sets a rate t_H and the town in the low-tax *state* sets t_L.

Letting V be the reservation price of the good and assuming that V is sufficiently high, an individual in the high-tax state will purchase the good in the other jurisdiction if $V - 1 - t_L - T_L - ds > V - 1 - t_H - T_H$ which implies that consumers that are within $\frac{b+t_H-t_L}{d}$ units of the border will engage in cross border-shopping if $t_H + T_H > t_L + T_L$; otherwise, no one will cross-border shop. A similar cutoff rule can be derived for the neighboring town.

2.1 No Online Shopping

As a benchmark, consider the case where no one has access to the Internet for online purchases. When no shoppers have access to online transactions, local tax revenue in the two jurisdictions can be constructed as $R_H = t_H(1 - \frac{b+t_H-t_L}{d})$ and $R_L = t_L(1 + \frac{b+t_H-t_L}{d})$. Noting that the revenue functions are continuous when the direction of the inequality $t_H + T_H \gtreqless t_L + T_L$ flips, I can differentiate the revenue functions to solve for the best responses and the Nash equilibrium. Using superscripts to denote the equilibrium tax rates, the local tax differential between the two jurisdictions is given by:

$$t_L^N - t_H^N = \frac{2b}{3}, \tag{1}$$

and consistent with Agrawal (forthcoming), the town in the low-tax state sets a higher local option sales tax. For a Nash equilibrium to exist it must be that d is sufficiently large; I maintain this condition throughout. I proceed in two extreme cases: one where the Internet acts as a conduit for the purchase of tax-free goods and second where online firms remit taxes.

2.2 The Internet as a Tax Haven

Consider the case where a revenue leakage exists. Tax-free online sales are an example of a possible revenue leakage. In the presence of tax-free online sales, some individuals purchase goods online and tax revenue declines in the two jurisdictions. The purpose of this section is not to formally model the consumer incentives, but rather to show that when e-commerce is tax free, the effect of these sales on tax rates will be different depending on whether the locality is in a high- or low-tax state. Individuals previously shopping at home may be the source of the revenue leakage; equally possible is that the leakage in the low-tax jurisdiction arises from less cross-border shoppers.

To model revenue leakages to tax-free sources, assume that the tax revenue lost is given by the function $f(t_i, \theta, T_i)$ where $f(0, \theta, T_i) = 0$, $f_1 > 0$, $f_{11} > 0$, $f_2 > 0$ and $f_3 > 0$. Denote $\theta \in [0, 1]$ as a parameter that captures the fraction of consumers with access to the Internet.[9] Tax evasion is increasing at an increasing rate with respect to the local tax and some revenue is lost to the Internet as soon as the local tax rate is non-zero. Modeling tax evasion of the sales tax is likely to be different from standard models of income tax evasion (Slemrod 1994) where the cost of evasion is increasing in the amount evaded. In this model, demand is inelastic. However, even if demand were not fixed, the cost of making more online purchases is not likely going to trigger additional audits. To obtain a closed form solution assume that the revenue leakage function is a quadratic function in the local tax rate: $\theta t_i(t_i + T_i)$.[10] A quadratic cost of taxation has been used previously (for example, see Bolton and Roland (1996)).

The revenue functions are now $R_H = t_H(1 - \frac{b+t_H-t_L}{d}) - \theta t_H(t_H + T_H)$ and $R_L = t_L(1 + \frac{b+t_H-t_L}{d}) - \theta t_L(t_L + T_L)$ and the Nash equilibrium in the presence of the Internet as a tax haven (derived in Appendix A.2) is characterized by:

$$t_L^I - t_H^I = \frac{(2 + d\theta)b}{3 + 2d\theta}. \tag{2}$$

In the appendix, I show that $\frac{\partial t_H^I}{\partial \theta} < 0$ and $\frac{\partial t_L^I}{\partial \theta} < 0$ such that $t_i^I < t_i^N$ for $\theta > 0$, $\forall i$. Intuitively, more consumers with access to an avoidance technology will place

---

[9] I restrict θ to be the same in both jurisdictions. For spatially proximate jurisdictions, it is unlikely that Internet penetration rates vary; this is verified in the empirical section to follow.

[10] Note that the function $\theta(t_i + T_i)^2 = \theta(t_i^2 + T_i^2 + 2t_iT_i)$ does not make sense in this context because local revenue leakages increase in the square of the state tax rate and would be positive even if the local government sets $t_i = 0$. Given that tax revenue is given by $t_iB(t_i, T_i)$ where B is the tax base, the state tax rate only affects the tax base. The local tax rate affects revenue through two channels: changes in the base and the rate. The equation in the text imposes that the *base* shrinks linearly in both the state and local rate, but that *revenue* leakages are decreasing in the square of the local tax rate because t_i pre-multiplies that base.

downward pressure on tax rates. Important to the empirical design is that

$$\frac{\partial(t_L^I - t_H^I)}{\partial\theta} = \frac{-db}{(3+d\theta)^2} \leq 0, \tag{3}$$

which indicates that the tax rate in the high-tax state must fall by less than the tax rate in the low-tax state in response to an increase in Internet access in both jurisdictions. The effect of a change in Internet penetration will only be symmetric when $b = 0$. Thus, the tax haven effect of online shopping will be strongest in low-tax states. Intuitively, this arises because residents of the high-tax state already have access to low-tax goods through cross-border shopping. As a result, the higher state tax rate already constrains municipal tax rates in high-tax states. Thus, an increase in access to online goods in the high-tax state has a muted effect on the equilibrium tax rate.

Proposition 1. *When online sales are not taxable, revenue leakages from e-commerce will place downward pressures on tax rates. An increase in Internet penetration in both jurisdictions will, however, lower local tax rates in low-tax states more than in high-tax states.*

Proof. See Appendix A.3. □

Notice as well in equation 3, the asymmetries between towns in high- and low-tax states are increasing in b. That is to say, that the larger the state tax differential at the state border, the more likely that Internet penetration will have an asymmetric effect across the jurisdictions. As $b \to 0$, the effect of the Internet should be similar in both states.

Because state tax rate differentials at state borders average two percentage points, the incentives to engage in cross-border shopping are potentially quite large.[11] The most recent estimate of the elasticity of cross-border shopping places it at unity (Tosun and Skidmore 2007). To build intuition of proposition 1, consider the extreme example of Massachusetts (6% tax) and New Hampshire (0% tax). For a Massachusetts town right on the border, residents could avoid the Massachusetts sales tax by crossing over the border and buying tax free goods

---

[11] Agrawal (forthcoming) shows that towns in high-tax states have local tax rates that are on average 1.25 percentage points lower than towns in low-tax states independent of e-commerce.

in the pre-Internet era. In fact, this arbitrage opportunity is made easier by the fact that as you drive across the border, most of the retail stores are located on the New Hampshire side. Thus, the rise of a tax-free Internet is not likely to place much added pressure on the MA town to lower its tax rate from the already low pre-Internet rate t_H^N. Because the MA municipalities already lower their tax rate in the pre-Internet era, the presence of a tax-free Internet has a small effect. For a town on the low-tax side of the border, the incentives of the jurisdiction are the opposite given that in the pre-Internet era, they set relative high tax rates t_L^N.

2.3 The Internet as an Anti-Haven

What if Internet purchases are taxed? To model this, I assume that some individuals have a preference to shop online and pay sales taxes to their jurisdiction even if they could have purchased the good from a lower-tax jurisdiction. Individuals may engage in such transactions if they derive utility from online shopping or if the cost of purchasing online (for example, the breadth of goods available on walmart.com may be larger than from a brick-and-mortar store; consumers may have access to online reviews that allow them to make a more informed decision) is sufficiently low relative to the opportunity cost of driving to a brick-and-mortar store. Assume that the number of individuals with a preference to shop online is increasing one-for-one in θ.

When online sales are taxable, some residents previously buying from brick-and-mortar stores within their home jurisdiction now shop online; this does not affect the revenue function of the jurisdiction. However, residents with preferences to buy goods online and who previously purchased their goods abroad will now buy online and revenue affected. This reduces the number of cross-border shoppers by $1-\theta$ times; the new revenue functions are given by $R_H = t_H(1-(1-\theta)\frac{b+t_H-t_L}{d})$ and $R_L = t_L(1+(1-\theta)\frac{b+t_H-t_L}{d})$. The Nash equilibrium in the presence of this anti-haven is characterized by:

$$t_L^A - t_H^A = \frac{2b}{3}. \tag{4}$$

and as shown in the appendix, $\frac{\partial t_H^A}{\partial \theta} = \frac{\partial t_L^A}{\partial \theta} > 0$ such that $t_i^A > t_i^N$ for $\theta > 0$, $\forall i$.

Proposition 2. *When online sales are taxable and some consumers have a pref-*

erence to buy online, an increase in Internet penetration in both jurisdictions will raise local tax rates in both jurisdictions.

Proof. See Appendix A.3. □

Intuitively, more individuals with a preference to shop online will reduce the Nash competitive pressures and will place upward pressure on the tax rates. Because consumers in the high-tax state who now shop online are potential losses to the low-tax jurisdictions, the upward pressure is the same in both the high- and low-tax state. By exploiting people's preferences to shop online, the Internet lowers the marginal cost of raising revenue through the sales tax; jurisdictions can worry less about lost use tax revenues from cross-border shopping.

2.4 The Role of Size

In the model above, both jurisdiction have the same population. Intuitively, the tax haven effect is larger in low-tax states because border towns in low-tax states have higher tax rates and a larger tax base. The literature (Kanbur and Keen 1993, Nielsen 2001) indicates that jurisdictions with large populations set higher tax rates, which suggests that the tax-free Internet may affect larger jurisdictions differently than small jurisdictions.

In Appendix A.4, I introduce online shopping into the standard Nielsen (2001) model of tax competition where the localities differ in their population. All of the assumptions of the above model are maintained except I assume both jurisdictions have the same state tax rate in order to isolate the role of population size.

Proposition 3. *When online sales are not taxable and jurisdictions differ in size, an increase in Internet penetration in both jurisdictions will lower local tax rates in large jurisdictions more than in small jurisdictions.*

Proof. See Appendix A.4. □

The intuition is similar to the baseline model. In the pre-Internet era, the large jurisdiction realizes a size advantage; following an inverse elasticity rule, the large jurisdiction marks up its tax rate. When tax-free sales are possible, the large jurisdiction is more concerned about revenue leakage both because its tax base is big and because its tax rate is high. The small jurisdiction, with the already low tax rate and small tax base, will lower its tax rate by less. Thus, in very small jurisdictions, theory suggests that the Internet may have little effect.

3 Data and Baseline Results

The question arises as to whether the effects in proposition 1 or 2 dominate. Allowing the theoretical model to shape the empirical research design, I now test this empirically. Proposition 3 implies that a split sample analysis based on jurisdiction size or an interaction with population is appropriate. The theory leads me to expect that I should find small effects in high-tax states and in small towns, but I should find large effects in high-tax states and large towns.

3.1 Data

I have data on sales tax rates for every town, county, state and sub-municipal district from Pro Sales Tax's national database. Given the analysis will occur at the town level, states without municipal sales taxes are excluded from the analysis; towns that set a tax rate of zero are included in the analysis if they are in a state allowing for municipal taxes.[12] The tax data contain all local tax rates for December 2011, where the total tax rate in a jurisdiction is the sum of the rates. For each town in the tax data set, I calculate the minimum driving time from the population weighted centroid to the nearest state border major road intersection. If the state tax rate in the nearest neighboring state is lower than the state tax rate in the municipality's own state, the municipality is referred to as being in a "high-tax state." The data is described in in Appendix A.5.

I merge the tax data with data on Internet penetration from the National Broadband Map, which is collected by the National Telecommunications and Information Administration (NTIA) in conjunction with the Federal Communication Commission. The data on Internet penetration is from July 2011; earlier panel data is not available. The NTIA matches provider service maps to Census block maps. In doing so, they are able to calculate the fraction of people within a place that have access to Internet service providers. I know the percent of households with access to various types of services and the percent with access

---

[12] A state's decision to allow for LOST does not depend on Internet use. Only one state has adopted local option taxes in the last ten years and no states have repealed this authority.

to various download and upload speeds. I also know the percent of individuals who have access to more than one provider of Internet services, more than two providers, more than three providers, etc.

3.2 A Baseline Specification and Internet Penetration

A baseline cross-sectional regression specification that can be used to determine how the Internet influences tax competition takes the form:

$$\tau_i = \alpha + \beta I_i^* + \beta_1 I_i^* log(pop_i) + \zeta + \sum_m X_{im}\gamma_m + \epsilon_i, \tag{5}$$

where I_i^* is a measure of Internet usage in town i, ζ are state fixed effects, and X_{im} denotes the m control variables. Usage and population are interacted as suggested by proposition 3.[13] Depending on the specification, τ_i is either the total local tax rate (county plus town) or the town tax rate in town i. Within the set of controls are all of the geographic, political, and demographic variables listed in Table Appx.2 including (log) population. I also include a dummy variable for whether the jurisdiction is proximate to an international border or to a major body of water; the latitude and longitude of the town's population weighted centroid are controls designed to capture geographic or topographical features that change from east to west or north to south within a state. State tax rates are accounted for with the state fixed effects along with other state-by-state institutions.

Local policymakers likely set tax rates on the basis of the fraction of people who are able (and willing) to use the Internet.[14] This would suggest that a measure of I_i^* would be the fraction of people who *use* the Internet at home or the number of online purchases per person. Defining I_i as a measure of *penetration* from the National Broadband Map, if

$$I_i^* = \lambda + \delta I_i + \nu_i \tag{6}$$

---

[13] Such an interaction comes with some imposition of functional form assumptions. Therefore, I will also show the results are robust to a split sample analysis.

[14] Alternatively, one may argue that policymakers know the penetration rate, I_i – because the data are publicly available – but do not know the Internet usage rate I_i^* for their municipality. If this is the case, local tax policy would be expected to respond to the observable variable. Then, equation 7 will be the true policy reaction function.

and if δ is significantly different for zero, then Internet penetration can be used as a proxy for Internet usage. Simply plugging in the penetration measure for I^* in equation 5 will create attenuation bias on the true coefficient of interest because usage is measured with error if I estimate:

$$\tau_i = \hat{\alpha}_0 + \hat{\beta}_0 I_i + \hat{\beta}_1 I_i log(pop_i) + \zeta + \sum_m X_{im} \hat{\gamma_m} + \hat{\epsilon}_i. \tag{7}$$

3.2.1 The Best Measure of Internet Usage

The National Broadband Map has several potential candidates for I_i but Internet usage is not measured at the town level. I select a single proxy variable from the Broadband Map as the variable that maximizes the R^2 of the univariate regression in equation 6 where I_i^* is defined as state-level Internet usage from the Consumer Population Survey (CPS). I assume that the variable that maximizes the R^2 at the state level would also do so in the cross-section of lower levels of governments. The most recent supplementary data on "Computer and Internet Use" is from October 2010. In this supplement, respondents are asked "At home, do you or any member of this household access the Internet?", "At home, do you access the Internet?" and "Do you or any member of this household access the Internet at any location outside the home?". Using this data, I am able to match state level data on Internet penetration from the NTIA with Internet access statistics from the CPS. In case the reader is worried that Internet usage at home does not translate into online shopping, I alternatively define I_i^* as the per capita number of eBay purchases in each state as created by Einav et al. (2014). Assuming that the number of eBay transactions within a state are representative of online shopping more generally, eBay purchases allow me to identify the best proxy for Internet usage for e-commerce purposes.

Table 1 summarizes the results of this exercise. Each row of column 1 reports the estimated value of δ and the R^2 of equation 6 with no other controls. Column 3 uses the per capita number of eBay purchases as calculated by Einav et al. (2014) as I^*. The R^2 results are even stronger when using eBay purchases.

Table 1 indicates that the fraction of individuals with access to any type of Internet service is not a strong proxy for Internet usage; this variable has little variation. Access to three or more providers demonstrates a strong (positive) relationship between penetration and usage. This strong correlation is remains for having access to four or more providers and five or more providers. The fraction of people with access to four or more providers yields the strongest R^2 value and for this reason it is the preferred metric that I use throughout the paper. The predictive power of this variable remains strong, especially with respect to the eBay usage variable, even after controlling for state level characteristics. Figure 2 summarizes these results.

Why does access to more providers perform much better than broader measures of Internet access? The existing literature provides some theoretical evidence that increased competition by broadband companies will increase take-up of the Internet. For example, Faulhaber and Hogendorn (2000) shows that "the subgame equilibrium capacity and price strategies depend only on the number of networks to which a household has access." Thus, the number of providers serving an area (the outcome of the first stage) is, from a theoretical perspective, the most important determinant of price in this industry. Second. as shown in Distaso, Lupi and Manenti (2006) and verified empirically, inter-platform competition such as DSL versus cable technologies (rather than intra-platform competition), increase Internet usage. Prieger and Hu (2008) also show empirically that competition in broadband markets is an important contributing factor of the Digital Divide that exists across races even though prices do not vary substantially across various markets; the authors argue that more intense competition increases Internet usage because companies compete more intensely on installation, service fees, and other charges. All of this evidence taken together suggests that markets with more intense competition will have higher Internet usage rates and that penetration is also correlated with online purchases, which should then feed back into the tax setting behavior of the jurisdictions.[15]

3.2.2 Summary Statistics

I next present maps showing the distribution of local tax rates in Figure Appx.2. Although much of the within state tax variation is swamped out by the cross-state variation, the map still demonstrates that variation within a state is noticeable.

---

[15] For general discussions of the demand and supply of broadband, see Nevo, Turner and Williams (2013), Prieger (2007), and Varian (2001).

The average town tax rate is 0.77 percentage points.

The national numbers from the CPS indicate that 65% of people use the Internet at home. In the average town, 71% of the population have access to four or more providers. Figure Appx.3 shows the percentage of residents with access to four or more providers; there is ample variation of this variable within a state.

For visual evidence, I present correlations of state tax rates and measures of Internet usage from the CPS. The results of this exercise are summarized in Figure 3. Having access to the Internet outside of the home has no relationship with the state sales tax rate. This result is expected because it is not likely that individuals are using cafe cyberstations or work computers to do the bulk of their shopping. The question asking whether individuals use the Internet at home shows a negative relationship with sales taxes, as does the fraction of people living in a household where one person has access to the Internet.

Table Appx.1 shows summary statistics for changes in tax rates at the town level from 2003 to 2011. The average change is 0.11 percentage points, which suggests that taxes have been rising over this period – a trend consistent with the shift from the property tax to the sales tax over the last decade. On average, taxes have risen by more in low-penetration towns than in high-penetration towns and this effect is most pronounced on the low-tax sides of state borders. This provides some initial first order evidence that towns may be responding to Internet usage.

3.3 Cross-sectional Results

Table 2 shows the results of equation 7 at the town level using both the total local rate (town plus county) and the town rate as the dependent variable; the town tax rate is preferred given that is the unit of analysis, but the total rate is informative of the effect on both county and town taxes. In all tables, standard errors are clustered at the county level to account for the fact that when Internet providers service one town they may also service nearby towns. Column 2 – including all the controls listed previously and state fixed effects – shows a significant effect of an increase in Internet penetration on the town tax rates. The mean derivative is -0.132, which indicates that a 100 percentage point increase in the Internet penetration rate (going from a world with no Internet access, to complete access) decreases the local tax rate by .132 percentage points.[16] Internet penetration rates have not increased 100 percentage points in recent years; the average town has a penetration rate of about 70% (or 1.65 standard deviations of I_i). A one standard deviation increase in Internet penetration lowers local tax rates by 6% of the average rate. The results can also be benchmarked in magnitudes relative to the average local tax change from the last decade in Table Appx.1. Compared to the average change, the effect of the Internet is large.

One econometric concern is that tax rates are constrained to be above zero, but estimation using ordinary least squares (OLS) places no restrictions on the predicted values. In fact, OLS may result in predicted values that are well outside this interval; furthermore, the linear form may miss important non-linearities that arise in the fractional case. As is well know, estimation of censored data using OLS will yield biased estimates. I implement a fractional response model (Papke and Wooldridge 1996 and Wooldridge, fc) to address this.[17] In the fractional response model, the tax variable in equation 7 is allowed to take on values $0 \leq \tau_i \leq 1$ where the extreme values can – but need not – occur with positive probability. I then estimates a general nonlinear model under the assumption that

$$E(\tau_i|\mathbf{x}_i) = G(\mathbf{x}_i\beta) = G(\alpha_0 + \beta_0 I_i + \beta_1 I_i log(pop_i) + \zeta + \sum_m X_{im}\gamma_{0m}), \tag{8}$$

where $\mathbf{x}$ are the right hand side variables from equation 7 and $G(\cdot)$ is some function that will result in predicted values between zero and one. To mechanically implement this, I apply quasi-maximum likelihood estimation (QMLE) using a Bernoulli log likelihood function where I use the probit functional form for the mean response to bound the predicted tax rates. Because the model is non-linear, I always present marginal effects (mean derivatives) defined as $\partial E(\tau_i|\mathbf{x}_i)/\partial I_i$. Comparing the fractional response marginal effects with the OLS results in table 2 yields similar results – this holds true for more complex specifications. A one standard deviation increase in Internet penetration lowers local tax rate by .07 (.43 × .159) percentage points or 9% of the average rate. This indicates that the extra structure imposed by the fractional response model is not driving the

---

[16] If the mean local tax rate is .77 with no Internet access, it will be .64 with complete access.

[17] The fractional response model is more appropriate than a Tobit model in this context.

results; as a test of the structure, I verify the results are similar for a Tobit model as well as various G functions including the Cauchy distribution, which is more robust to outliers. All subsequent results use fractional response models.

3.3.1 Sensitivity to Controls and County Fixed Effects

The obvious threat to identification is an omitted variable that is correlated with both Internet penetration and tax rates. Some obvious candidates include the degree of urbanization, amenities, and agglomeration. All three of these variables are likely to be correlated positively with both tax rates and Internet penetration implying that my cross-sectional estimates, which are negative, are biased toward zero. This combined with measurement error suggests these results are a lower bound. To formally address the omitted variable problem, I will rely on a border discontinuity design grounded in the theoretical model.

I also address this empirically in table 3. The first robustness exercise is to rely on within county variation by including county fixed effects. Most counties within the United States are relatively small in size and commuting zones (with common labor markets) often times cover three to four counties at a time. After including county fixed effects, the marginal effects shrink slightly to -.108, but remain strongly significant. This provides powerful evidence that the results are not driven by omitted factors that are common to a local region within a state. The remainder of the table shows that the results are not sensitive to removing or adding control variables including important factors such as density and income.

3.4 Heterogeneity In Responses: Jurisdiction Size

The theoretical model suggest size may also play a role. The average town in the sample has just under 10,000 people and most likely does not have a single big-box brick-and-mortar store. Resident may need to drive long distances to find the closest mall or major retail shopping center. As a result, such a jurisdiction has little or no tax base in a world where no one has access to the Internet and the use tax is not enforced. Proposition 3 suggests the tax haven effect is largest in big jurisdictions because they set high rates and have broad tax bases.

To test for a heterogeneous response and to eliminate the interaction with the (log) population, I split the sample between jurisdictions that are above the median population size (1189) and jurisdictions below the median size. When splitting the sample at the median, the sub-sample of “large” towns contains *97% of the population from the full sample*; this suggests this subsample analysis is still economically meaningful and applicable. Table 4 shows that the response of small towns is statistically indistinguishable from zero, but that the response in larger towns is 7 times as large and significantly negative – suggesting that town tax rates are 10% (.075/.77) lower as a result of a change from zero to complete Internet penetration. This implies that the tax haven effect is stronger in large jurisdictions than small jurisdictions.[18] If splitting the sample at the mean, Internet penetration lowers large jurisdiction municipal tax rates by almost four times the effect in large jurisdictions when splitting at the median.

A split sample approach is preferable to interacting I with the log of population – and therefore, I proceed with a split sample approach.

4 Exploiting Discontinuities at State Borders

Modifying Holmes (1998), to find the causal effect of Internet access, I exploit the fact that borders result in discontinuous changes (“notches”) in state sales tax rates. This borders-based approach provides evidence that the channel through which the Internet influences tax rates is competitive rather than from some spatially correlated variable. The threat to identification is a spatially correlated variable that has an asymmetric effect on local tax rates depending on the state the municipality is in. To exploit this identification strategy, I implement a modified version of equation 8:

$$\begin{aligned} E(\tau_i|\cdot) = \quad & G(\alpha_0 + \beta_0 I_i + \beta_1 H_i + \beta_2 I_i H_i + \sum_{k=1}^{K} \delta_k (d_i)^k + \sum_{k=1}^{K} \varphi_k H_i (d_i)^k + \textstyle\sum_{k=1}^{K} \rho_k I_i (d_i)^k + \sum_{k=1}^{K} \lambda_k I_i H_i (d_i)^k \\ & + \zeta + \sum_{m} X_{im} \gamma_m), \end{aligned} \tag{9}$$

where X_{im} denotes the m control variables including latitude and longitude of the jurisdiction to control for position along the border,[19] ζ are state fixed effects, H_i is a dummy variable equal to one if the town is located in a high-tax state relative

---

[18] Jurisdictions with very small populations may also be less sophisticated in their tax setting behavior than large jurisdictions. If small jurisdictions are less sophisticated (and perhaps less strategic as a result), these small jurisdictions may see revenue declines over time, and instead of reducing their tax rate to mute the effects of the Internet, they may raise the tax rate in an attempt to satisfy a revenue requirement.

[19] Population enters as a control, but is not interacted with I_i. Instead, I proceed with a full sample and then the split sample approach from the prior section.

to its nearest neighbor, and d_i is the shortest driving time to the nearest state border from the population centroid of the town.[20] The interaction terms of I_i and H_i allows Internet penetration to have a different effect on the high-tax side of the border compared to the low-tax side of the border as $d \to 0$. In the above equation, I also include several flexible distance functions of polynomial order K.[21] The first two polynomials in distance $\sum_{k=1}^{K} \delta_k (d_i)^k$ and $\sum_{k=1}^{K} \varphi_k H_i (d_i)^k$ capture the effect of distance from the border on local tax rates independent of the Internet and allow it to vary on the high- and low-tax side. The term $\sum_{k=1}^{K} \rho_k I_i (d_i)^k$ allows for Internet penetration to have a heterogeneous effect on taxes depending on proximity to the border. The term $\sum_{k=1}^{K} \lambda_k I_i H_i (d_i)^k$ allows for this effect to be different on the high-tax side of the border compared to the low-tax side of the border. The polynomial order K is selected using leave-one-out cross-validation, which yields a cubic polynomial.

I am interested in how this effect differs as distance to the border approaches zero $d_i \to 0$ although the effects discussed in the theory should persist away from the border. In the limit, I can compare bordering towns in high-tax states with bordering towns in low-tax states with the idea being that towns on the other side of the border are a good counter-factual. I maintain the split sample analysis given that the previous section demonstrated that there is no significant response to the Internet in small towns; the results for the larger towns will be more appropriate to characterize how the Internet influences tax rates for the average person in America. In addition, I can interact the Internet terms in the regression specification above with the state tax differential at the border to determine where the response is largest. Recall from the theory that the asymmetry across borders is smallest as $b \to 0$. In all tables to follow, given the non-linear specification and the polynomial in distance, I report the mean derivatives (marginal effects) conditional on the side of the border that the town

---

[20] I exclude observations for which the nearest neighboring state sets the same tax rate. Alternatively, I could include a separate dummy variable. The results are robust to this, but requires adding multiple terms. Relatively few towns (3% at three borders) are proximate to a border where the state tax differential is the same.

[21] I will restrict the sample to towns within one hundred minutes of driving time to the nearest state border. The restriction of one hundred minutes reduces the sensitivity to far out jurisdictions. In Appendix A.9, I show that the results are robust to changing this restriction and including all jurisdictions.

is located within.

4.0.1 Testing Assumptions of the Border Design

A threat to identification using a border-based approach would arise if Internet penetration changes discontinuously at state borders conditional on observables because neighboring towns would then not be a good counter-factual. I test whether this is the case by regressing the measure of Internet penetration on a dummy variable for the high-tax side of the border, a distance polynomial, the same set of control variables used above, and fixed effects. Coefficients on the high-tax state dummy variable are insignificant and close to zero. The results are reported in Appendix A.6.

A second test to identification is to show that no other variables change discontinuously at state borders that are correlated with being in a high-versus a low-tax state. Of course, I cannot test this for unobservables, but I am able to run a regression discontinuity design for observable variables to determine if borders are balanced. Appendix A.6 shows that almost no observable control variables change discontinuously at state borders.

4.1 Border Discontinuity Design Results

Table 5 presents results for the average town on the high-tax and low-tax side at any distance. The results in column 2 for the full sample indicate that when I account for border asymmetries, the marginal effect of the Internet is most negative for jurisdictions on the low-tax side of borders (.06 percentage points larger at any d on average and .16 percentage points larger as $d \to 0$) . The effect independent of distance in low-tax states remains significant throughout specification – a result consistent with local taxes being higher throughout low-tax states. When focusing on large towns, the response of towns in low-tax states is much larger than towns in high-tax states; the magnitudes are statistically different from each other. This is consistent with a theory where jurisdictions on the high-tax side of borders were previously constrained by high state tax rates, which mutes the effect of the Internet on taxes. Column (5) includes border pair fixed effects and shows the results are unchanged. This controls for the specific border being analyzed such that all identifying variation comes from within a specific border pair; the previous results accounted for this using only latitude and longitude controls.

Table 6 shows the heterogeneity as a function of distance to the border. As $d \to 0$, the asymmetric effects are amplified – consistent with state sales tax differentials being most salient near borders. In the limit, as distance approaches zero, towns on the other side of the border become better counter-factuals for the opposite side and cross-border concerns become most salient. The effect of the Internet is 2.5 times as large on the low-tax side when $d \to 0$. On the low-tax side of borders, a one standard deviation increase in Internet penetration lowers local tax rates by 25% of the average. The effect of Internet penetration decreases over the first 20 minutes of distance in low-tax states at which point it levels off, but remains stronger than the effect for towns of a similar distance in the high-tax state. On the high-tax side, the results are generally insignificant from zero, which may lead the reader to wonder why the Internet does not place downward pressure on tax rates for interior towns in high-tax states. The intuition of this result is that even for towns 90 minutes away from the border, the fact that the state sales tax rate is relatively high still constrains the local rate. This result is consistent with large level effects – independent of distance – of state tax rates on local rates and relatively small tax gradient effects in high-tax states (such that towns near the border and away from the border are relatively similar in their local tax rates). Independent of the Internet, towns at the interior of high-tax states still set lower rates than towns in low-tax states.

The theory also predicts that municipalities facing the largest tax differentials (b) should have the most asymmetric responses. To test this, I interact (and include as stand alone variables) the magnitude of the tax differential of the home state minus the neighboring state with the terms including I_i. I allow a one percentage point increase in the differential to have different effects depending on the side of the border. Panel B of table 6 shows that municipalities have much larger responses in low-tax states relative to their high-tax state neighbors as the tax differential increases. In fact, comparing towns with 5 percentage point and 1 percentage point differentials implies that the difference is 5 times as large. As $b \to 0$, the effects converge as predicted by theory. The result is consistent with equation 3 from the theory, which implies the asymmetric response increases as

b increases.

Appendix A.7 shows that the border design results are robust to using other proxy variables with a high R^2 from equation 6 and Appendix A.8 shows that state based Internet regulations provides little additional information beyond the penetration variables. In Appendix A.9, I show and discuss that the results are robust to a variety of robustness checks including: changing the order of the polynomial in distance, using towns more than 100 minutes away from a border, various weighting schemes, focusing only on linear state borders likely to be in a featureless plane, clustering standard errors conservatively at the state level, and including a spatial lag of tax rates.

4.1.1 County Borders, Placebo Tests, and MSAs that Cross Borders

In states that allow for both county and town sales taxes, the asymmetric effects of the theoretical model should arise at county borders: towns on the low-tax side of a county border should realize larger effects of the Internet. County borders provide two critical checks. First, confounding policies that are correlated with county tax rates are less likely at county borders than at state borders. Second, many county borders have the same county tax rate on both sides of the border; these borders can be used as a placebo test.

Table 7 – using county borders with tax differentials – shows the Internet acts as a tax haven in low-tax counties and that the anti-haven effects are strongest in high-tax counties. The results are consistent with the results at state borders and similar in magnitude, except the results in high-tax counties show stronger positive effects than at state borders.

As a placebo exercise, I use the county borders where county tax rates are the same (either because the counties set the same rate or because the counties are prohibited from setting county sales taxes) to test whether the effect of the Internet is driven by a border effect. To do this, I take this subset of same-tax borders and randomly assign one side of the border to be the “high-tax” side and the other side to be the “low-tax” side. I repeat this randomization 600 times and estimate equation 9 for each random assignment. Figure 4 plots the distribution of the difference in the marginal effect of the Internet on the placebo “low-tax” side of the border minus the effect on the “high-tax” side of the border. As is evident, the distribution is centered on zero which suggests that an increase in Internet penetration has a similar effect on town tax rates for both sides of county borders where county tax rates are the same. When using borders where county tax rates differ, the effects are much more negative on the low-tax side of borders; the difference calculated from table 7 lies well to the left of the distribution from same-tax borders. This suggests that I am not identifying a border effect, but rather the effect that the Internet lowers local tax rates by more on the low-tax side of borders.

As a check to control for amenities and agglomeration, I identify all towns that are located in metropolitan statistical areas (MSA) that cross state borders. MSAs are defined based on commuting patterns such that they have a common labor market and have similar amenities. Approximately fifty MSAs with seventy-five million people cross state lines. As a result, towns in these MSAs may have different state tax rates but share a common economic market. Identifying towns in these MSAs allows me to use a different source of variation – within an MSA – rather than within a state. I can then estimate a version of equation 9 that contains MSA fixed effects instead of state fixed effects – where a town is in the high-tax state if its side of the MSA has a higher state tax rate. The identification strategy is discussed in Appendix A.10 and the results of this exercise confirm the results using state borders with state fixed effects. A one standard deviation increase in Internet penetration lowers municipal tax rates by approximately 35% of the average municipal rate in the low-tax side of the MSA but has very small effects on the high-tax side.

4.2 Heterogeneity In Responses: The Role of Nexus

From a theoretical perspective, the number of firms and sales by firms with nexus should be an important factor for determining whether the tax haven or anti-haven effect dominates. Once a firm has nexus within a state, the firm is required to remit all local option sales taxes. This is appealing because the decision to establish nexus in a particular state is independent of the local tax rate of any one town in the state. States that have a disproportionally high number of firms with nexus – all else equal – should have a higher fraction of online transactions that are taxable simply because the number of tax-free firms in the state is limited because most firms have nexus in the state.

Data on nexus is not readily available, however, Bruce, Fox and Luna (2014) construct nexus data by hand. In order to construct this data, Bruce, Fox and Luna (2014) use sample of approximately 200 large e-tail firms and visit the websites of these companies; the authors then attempt to place a transaction from different states. The authors record whether the firm assesses sales tax in each state upon checkout. A company is defined as having nexus in the state if sales taxes are due at checkout. Appendix A.11 provides a detailed explanation of how the authors construct this data. Using data collected for Bruce, Fox and Luna (2014), I know whether each state has an above average number of firms with nexus. I also know the fraction of sales volume from firms with nexus in the state is above or below average.[22] Using this information, I can include additional interactions of all the Internet terms in the regression with a dummy variable for whether the state has an above average number of firms with nexus.

The results of this exercise, presented in Table 8, indicate the anti-haven effect emerges for towns with an above average number of e-tail firms with nexus. In states with relatively few e-tail firms with nexus, the tax haven effect of the Internet is strongest. This is true both on the high-tax and low-tax side of state borders, but the effects are still strongest in low-tax states. For example, for border towns in low-tax states that are also in the top quartile with respect to the Bruce, Fox and Luna (2014) nexus measure, an increase in Internet penetration raises municipal tax rates by 0.68 percentage points. However, for border towns in low-tax states that are not in the top quartile of states based on the nexus measure, the effect is the opposite in sign. This asymmetric pattern also emerges in high-tax states and even for small towns in high-nexus states. This result is consistent with the anti-haven effect being strongest when more firms remit sales taxes on online transactions – e.g., when more firms have nexus in a state. Similarly, when relatively few e-tail firms have nexus in a state, then it is highly likely that regardless of the number of brick-and-mortar firms in the municipality, fewer online transactions will be subject to tax collection – making the tax haven

---

[22] As a robustness check, I show in Appendix A.12 that the results are robust to using a much more restrictive proxy for nexus – the count of the number of retail firms that have headquarters in a particular state from Compustat data.

effect of online shopping dominant. This is additional evidence – still using the border discontinuity design – that Internet penetration is affecting tax rates through the mechanisms outlined theoretically.

4.3 Measurement Error and Multiple Proxies

One downside of the proxy variable based approach is that it introduces concerns relating to measurement error in the Internet variable. This measurement error is likely to attenuate the coefficient estimates. In all of the above regressions, I am using a single proxy variable that was selected using an R^2 criterion. In reality, I have access to multiple proxy variables – the fraction of the population with access to any number of providers or greater, the fraction of the population with access to various speeds, and the fraction of the population with access to various types of technologies. Lubotsky and Wittenberg (2006) develop a procedure (for linear regressions) that estimates the coefficient of interest if all proxy variables are included in the regression simultaneously.[23] In large samples, the estimator is superior to methods constructing an ad hoc index or to using only a single variable and will minimize measurement error.

Appendix A.13 describes how I implement the Lubotsky and Wittenberg (2006) procedure. The results from Table 9 can be compared to the previous results. As expected, the effects increase in absolute value across all of the specifications; in some specifications they are 1.5 times as large. For example, in column 2, the result with the index implies going from no to complete penetration lowers tax rates .189 percentage points. This result can be compared to column 2 of table 2, which was -.132. The border discontinuity results also increase in absolute value. This suggests that – as expected – the proxy variable approach used previously biases me against finding results.

Given the results increase in absolute value (and that they are normalized using the same proxy variable utilized in table 2), use of a single proxy may attenuate my results by at least 30%. Given the sign on the coefficients is the same and the results significant even in the presence of this attenuation bias, this suggests that the results using single proxy variable represent a lower bound.

---

[23] All of the results in this section use OLS models rather than the fractional response model given that the procedure was developed for linear regressions. I have also applied the procedure to the non-linear models above and the effects only marginally differ.

5 Lightning as an IV for IT

As a final robustness check, I instrument for Internet usage. Andersen et al. (2012) show that lightning strikes are a powerful predictor of IT usage at the state level in the United States during the period from 1996 to 2006. Andersen et al. (2012) argue that in places with high lightning density, more power disturbances occur. These power disturbances increase the cost of investing in IT, which then lowers IT investment and Internet usage. Figure 5 shows the strong negative relationship of IT usage and lightning. This result suggests a possible instrument.

As an instrumental variable, I construct a measure of the flash density using the National Oceanic and Atmospheric Administration's Severe Weather Database. I use data on the annual number of ground strikes from 1996 to 2011 to construct the per year average number of strikes per square mile. Define the flash density of lightning in a state

$$lightning_j = \frac{(\sum_{t=1996}^{2011} strikes_{j,t})/16}{area_j}, \tag{10}$$

where $strikes_{jt}$ is the number of lightning strikes in state j in year t and where $area_j$ is the area of the state. Table 10 presents state level regressions where I instrument for Internet usage from the CPS with the natural logarithm of $lightning_j$.[24] I follow the control function approach of Wooldridge (fc) to estimate the fractional response model with endogenous regressors. With respect to the exclusion restriction, it is unlikely that lightning strikes directly influence local sales tax rates. Lightning may be correlated with weather conditions or topographical features, which are unlikely to be correlated with tax rates unless weather related amenities influence public good preferences. Lighting is strongest in the southern states. To address this possibility, I include Census region dummies in the regression specifications such that identification comes from variation across lightning strikes *within a Census region.*

Because these are state level regressions, in some specifications I use the state

---

[24] Given I am using a direct measure of Internet usage rather than a proxy variable, measurement error concerns are also limited in this specification. Weather related instruments have been previously used in the tax setting. See Rincke and Traxler (2011).

tax rate only, while in other specifications I use the state plus population weighted average municipal tax rate. Using the average local tax rate in the state allows me to study the effect of the Internet on local tax setting behavior. The results of this exercise are simply meant as a robustness exercise given that I am constrained to studying only the forty-eight contiguous states in this analysis; the border design remains the preferred method given that it is grounded in the theory.

In all specifications, as in Andersen et al. (2012)'s IV regressions, the instrument is strong and negatively correlated with Internet usage at home as measured by the CPS; the F-test statistic is large and the R^2 high. The results of the IV exercise confirm the same sign of Internet usage on tax rates as in the large towns case. The results indicate that Internet usage has a larger effect on the (population weighted) average local sales tax rate in the state than the state tax rate itself. This is consistent with municipalities being able to respond to the Internet only through adjusting tax rates; states may have more dimensions on which they can respond such as redefining the tax base or changing nexus laws. This suggests that endogeneity concerns are limited. The IV results taken together with the border design, which is grounded theoretically, are strong evidence that the channel through which the Internet influences local tax rates is through downward pressure that a tax-free virtual "haven" places on taxes.

6 Conclusion

E-commerce represents a non-trivial fraction of commodity purchases. Purchases from remote vendors without nexus are sales tax-free while purchases from online vendors with nexus in a state are subject to state and local sales tax remittance. Governments seeking to maximize revenue in a competitive game, will account for any revenue leakage to the Internet when setting tax rates. I modify a standard model of tax competition to account for the Internet. If the Internet acts as a tax-free source, towns in low-tax states will be most adversely affect by e-commerce. Jurisdictions in high-tax states will have a muted response to the tax haven pressures of e-commerce. If the Internet is taxed, then Internet sales are an effective way of collecting state and local use tax revenues; this mutes competitive pressures between jurisdictions as some cross-border shoppers now buy online. I demonstrate theoretically that this will place upward pressure on tax rates through an anti-haven effect.

Numerous studies (Ballard and Lee 2007; Goolsbee 2000; Goolsbee, Lovenheim and Slemrod 2010; Einav et al. 2014) show that consumers are responsive to tax differentials resulting from online shopping. Because the Internet influences the elasticity of demand for goods and because the elasticity of demand is an important determinant of tax rates in the context of fiscal competition, the natural step taken in this paper is to determine how differences in Internet penetration distort tax rates. This paper represents the first attempt at doing so.

Descriptive evidence shows an upward trend in municipal sales taxes from 2003 to 2011 (as documented in Agrawal 2013). Are the results in this paper in conflict with this trend? First, although municipal tax rates have been rising, in the most recent four years, the average monthly growth rate in LOST has substantially slowed from 0.17% prior to 2008 to 0.04%. This may be a response to increases in e-commerce. Second, the results in this paper are consistent with high-Internet penetration municipalities lowering their sales tax rate. The results are also consistent with these municipalities raising their tax rates – or not changing their tax rate at all – at a *relatively* slower rate than municipalities with low-penetration rates. Although the theoretical model predicts that tax rates should strictly fall in response to the introduction of the Internet, all else has not been held constant over the last decade. Given that jurisdictions have been shifting away from the property tax due to its unpopularity, the empirical results in this paper are consistent with the shift towards the sales tax being larger in towns with low Internet penetration rates.

The results of this study are both statistically and economically significant. If the Marketplace Fairness Act of 2013 passes, the results of this study indicate that the collection of online sales taxes may place upward pressure on tax rates if residents have a strong preference for online shopping. Certainly, as more and more states expand their definitions of nexus, the results of this paper suggest that tax rates will rise as the effect of the Internet as a tax haven is muted. Beyond that channel, online sales tax remittance by firms has the potential to lower the marginal cost of collecting tax revenue using the sales tax relative to other tax instruments; as this happens, the anti-haven effects discussed theoretically will likely arise placing further upward pressure on tax rates.

References

Agrawal, David R. 2013. “LOST in America: Evidence on Local Sales Taxes from National Panel Data.” University of Georgia Working Paper.

Agrawal, David R. forthcoming. “The Tax Gradient: Spatial Aspects of Fiscal Competition.” *American Economic Journal: Economic Policy.*

Alm, James, and Mikhail I. Melnik. 2010. “Do eBay Sellers Comply with State Sales Taxes?” *National Tax Journal*, 63(2): 215–236.

Alm, James, and Mikhail I. Melnik. 2012. “Cross-border Shopping and State Use Tax Liabilities: Evidence from eBay Transactions.” *Public Budgeting & Finance*, 32(1): 5–35.

Andersen, Thomas Barnebeck, Jeanet Bentzen, Carl-Johan Dalgaard, and Pablo Selaya. 2012. “Lightning, IT Diffusion, and Economic Growth Across U.S. States.” *Review of Economics and Statistics*, 94(4): 903–924.

Ballard, Charles L., and Jaimin Lee. 2007. “Internet Purchases, Cross-Border Shopping, and Sales Taxes.” *National Tax Journal*, 55(4): 711–725.

Baugh, Brian, Itzhak Ben-David, and Hoonsuk Park. 2014. “The ”Amazon Tax”: Empirical Evidence from Amazon and Main Street Retailers.” NBER Working Paper.

Becker, Johannes, and Clemens Fuest. 2012. “The Nexus of Corporate Income Taxation and Multinational Activity.” *FinanzArchiv / Public Finance Analysis*, 68(2): 1–21.

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.

Bollinger, Christopher R. 2003. “Measurement Error in Human Capital and the Black-White Wage Gap.” *The Review of Economics and Statistics*, 85(3): 578–585.

Bolton, Patrick, and Gerald Roland. 1996. “Distributional Conflicts, Factor Mobility, and Political Integration.” *American Economic Review*, 86(2): 99–104.

Brown, Jeffrey R., and Austan D. Goolsbee. 2002. “Does the Internet Make Markets More Competitive? Evidence from the Life Insurance Industry.” *Journal of Political Economy*, 110(3): 481–507.

Bruce, Donald, and William F. Fox. 2000. “E-Commerce in the Context of Declining State Sales Tax Bases.” *National Tax Journal*, 53(4): 1373–1388.

Bruce, Donald, William F. Fox, and LeAnn Luna. 2014. "The Effect of State Tax Policies on Where E-Tailers Collect Sales Tax." University of Tennessee Working Paper.

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

Burge, Gregory S., and Brian Piper. 2012. "Strategic Fiscal Interdependence: County and Municipal Adoptions of Local Option Sales Taxes." *National Tax Journal*, 65(2): 387–416.

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.

Distaso, Walter, Paolo Lupi, and Fabio M. Manenti. 2006. "Platform Competition and Broadband Uptake: Theory and Empirical Evidence from the European Union." *Information Economics and Policy*, 18(1): 87–106.

Einav, Liran, Dan Knoepfle, Jonathan Levin, and Neel Sundaresan. 2014. "Sales Taxes and Internet Commerce." *American Economic Review*, 104(1): 1–26.

Ellison, Glenn, and Sara Fisher Ellison. 2009. "Tax Sensitivity and Home State Preferences in Internet Purchasing." *American Economic Journal: Economic Policy*, 1(2): 53–71.

Engel, Melissa, Gerlinde Fellner, Rupert Sausgruber, and Christian Traxler. 2013. "Higher Taxes, More Evasion? Evidence from Border Notches in TV License Fees." Working Paper.

Eugster, Beatrix, and Raphäel Parchet. 2013. "Culture and Taxes: Toward Identifying Tax Competition." Discussion Paper no. 2013-39.

Faulhaber, Gerald R., and Christiaan Hogendorn. 2000. "The Market Structure of Broadband Telecommunications." *The Journal of Industrial Economics*, 48(3): 305–329.

Goolsbee, Austan D. 2000. "In a World Without Borders: The Impact of Taxes on Internet Commerce." *Quarterly Journal of Economics*, 115(2): 561–576.

Goolsbee, Austan, Michael Lovenheim, and Joel Slemrod. 2010. "Playing With Fire: Cigarettes, Taxes and Competition From the Internet." *American Economic Journal: Economic Policy*, 2(1): 131–154.

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.

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

Holmes, Thomas J. 1998. “The Effect of State Policy on the Location of Manufacturing: Evidence from State Borders.” *The Journal of Political Economy*, 106(4): 667–705.

Johannesen, Niels. 2010. “Imperfect Tax Competition for Profits, Asymmetric Equilibirum and Beneficial Tax Havens.” *Journal of International Economics*, 81: 253–264.

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.

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.

Lubotsky, Darren, and Martin Wittenberg. 2006. “Interpretation of Regressions with Multiple Proxies.” *The Review of Economics and Statistics*, 88(3): 549–569.

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.

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.

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.

Nevo, Aviv, John L. Turner, and Jonathan W. Williams. 2013. “Usage-Based Pricing and Demand for Residential Broadband.” Working Paper.

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

Papke, Leslie E., and Jeffrey M. Wooldridge. 1996. “Econometric Methods for Fractional Response Variables with an Application to 401(K) Plan Participation Rates.” *Journal of Applied Econometrics*, 11: 619–632.

Prieger, James E. 2007. "The Supply Side of the Digital Divide: Is There Equal Availability in the Broadband Internet Access Market?" *Economic Inquiry*, 41(2): 346–363.

Prieger, James E., and Wei-Min Hu. 2008. "The Broadband Digital Divide and the Nexus of Race, Competition, and Quality." *Information Economics & Policy*, 20: 150–167.

Rincke, Johannes, and Christian Traxler. 2011. "Enforcement Spillovers." *Review of Economics and Statistics*, 93(4): 1224–1234.

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. 1994. "Fixing the Leak in Okun's Bucket: Optimal Tax Progressivity When Avoidance Can Be Controlled." *Journal of Public Economics*, 55(1): 41–51.

Slemrod, Joel, and John D. Wilson. 2009. "Tax Competition with Parasitic Tax Havens." *Journal of Public Economics*, 93(11-12): 1261–1270.

Stöwhase, Sven, and Christian Traxler. 2005. "Tax Evasion and Auditing in a Federal Economy." *International Tax and Public Finance*, 12(4): 515–531.

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.

Varian, Hal. 2001. "The Demand for Bandwidth: Evidence from the INDEX Project." Brookings Institution.

Wooldridge, Jeffrey M. fc. "Quasi-Maximum Likelihood Estimation and Testing for Nonlinear Models with Endogenous Explanatory Variables." *Journal of Econometrics*.

Figure 1: Geography

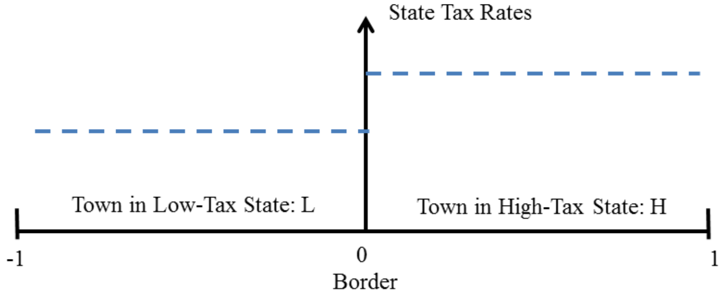
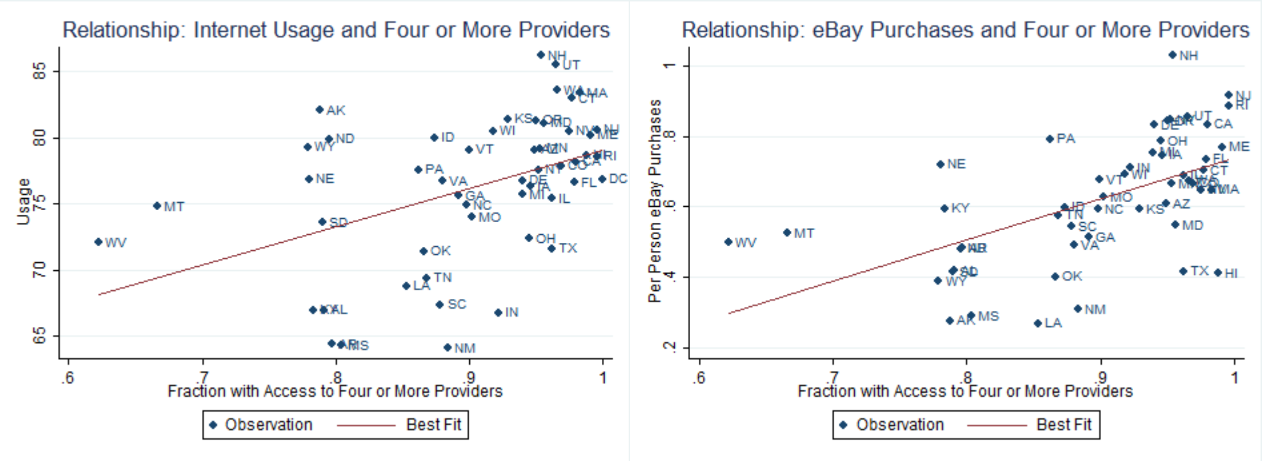
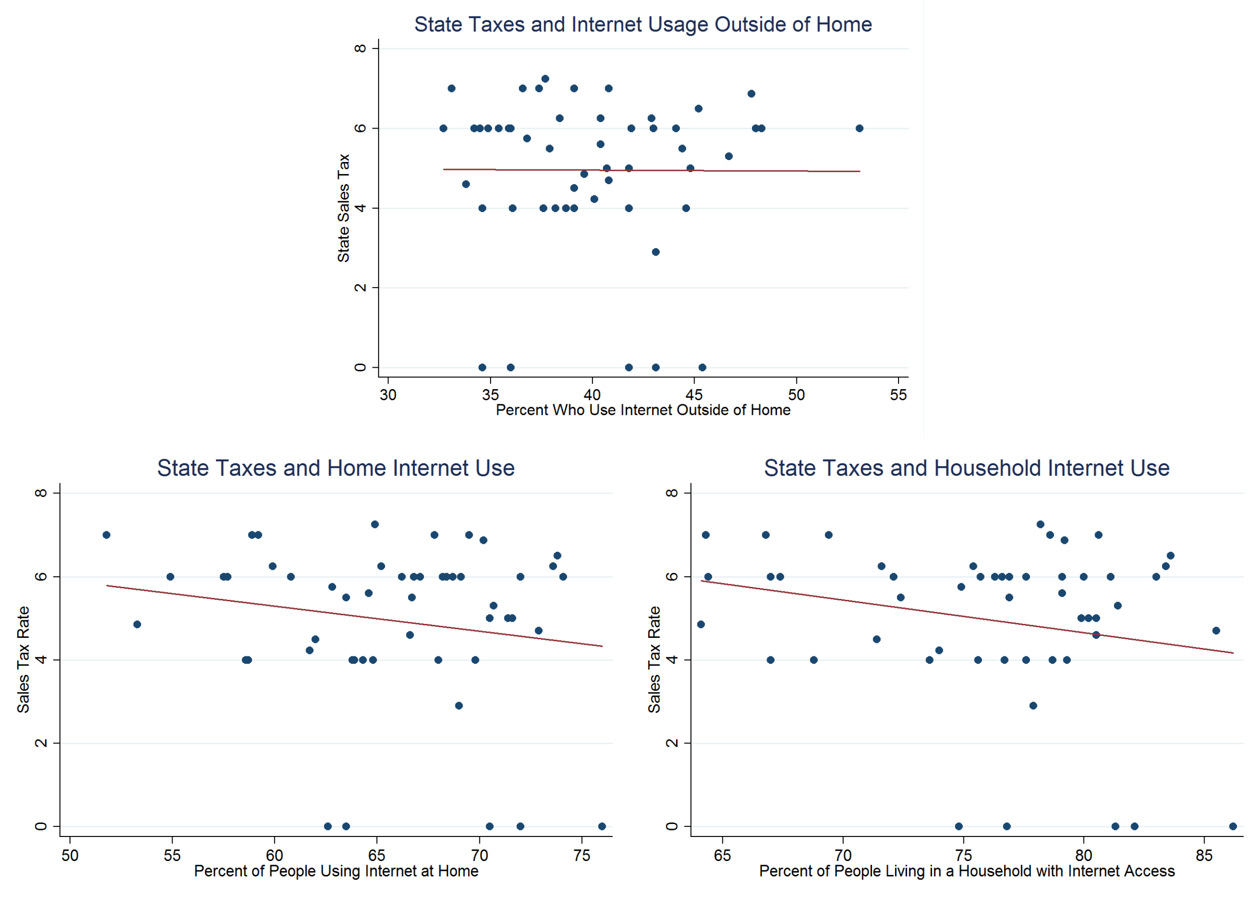


Figure 2: Relationship of Proxy Variable with Usage



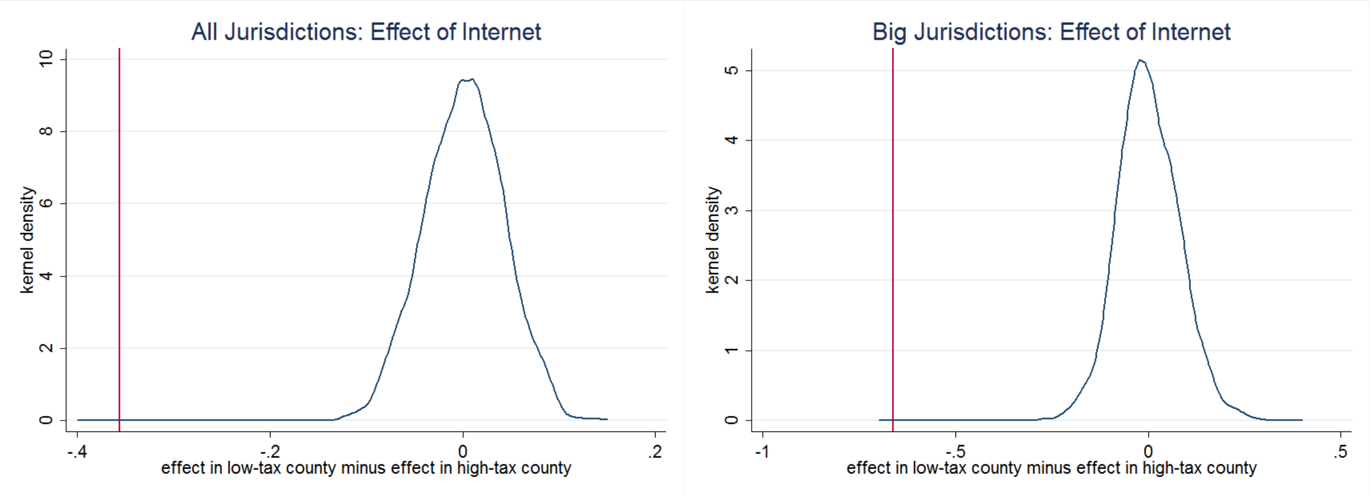
The first graph shows the relationship between usage as measured by the CPS ("At home, do you or any member of this household access the Internet?") and the fraction of the population with access to four or more Internet providers. The second graph shows the relationship between the average number of eBay purchases per person as released in Einav et al. (2014) and the fraction of the population with access to four or more providers.

Figure 3: State Internet Use and Tax Rate Correlations



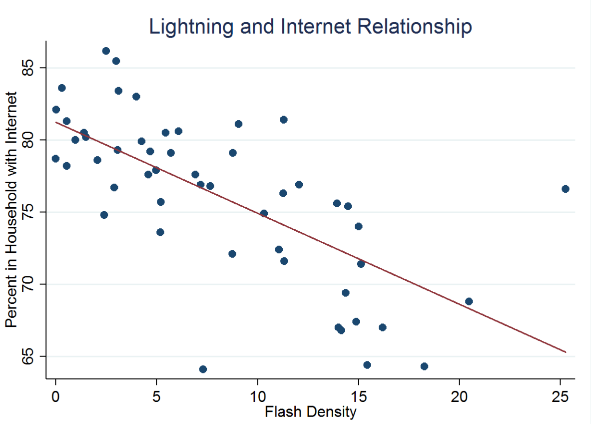
Each graph corresponds to one of the questions: "Do you or any member of this household access the Internet at any location outside the home?", "At home, do you access the Internet?", and "At home, do you or any member of this household access the Internet?".

Figure 4: County Borders Placebo Tests



The first panel focuses on all towns; the second panel is large towns. The vertical line in the graph shows the difference in the effect of the Internet in a low-tax county from the effect of the Internet in a high-tax county using the sample of borders where county tax rates are different (Table 7). The blue curve shows the distribution of the difference using only the set of county borders where county tax rates are the same at the border. The distribution is from 600 random assignments of each side of the border to be high-tax or low-tax. The distribution indicates that, on average, the effect of the Internet is the same on both sides of the border when no tax differentials exist at the border.

Figure 5: Correlation: Lightning and Internet Use



The graph shows the relationship between the response to the CPS question "At home, do you or any member of this household access the Internet?" and lightning flash density at the state level.

Table 1: Regression of Usage on Penetration

Standard Errors in ( ) and R^2 in [ ].

Each cell represents the coefficient on the listed variables from a **separate** regression.

	(1) Usage	(2) Usage	(3) eBay	(4) eBay
Any Service	.19	.21	.04***	.02*
	(.56)	(.24)	(.009)	(.01)
	[.002]	[.733]	[.144]	[.605]
Speed $\geqslant$ 768k	.03	.22	.04***	.02*
	(.25)	(.24)	(.01)	(.01)
	[.001]	[.733]	[.143]	[.605]
Speed $\geqslant$ 3000k	.03	.19	.02***	.01*
	(.25)	(.14)	(.003)	(.007)
	[.001]	[.738]	[.165]	[.616]
Providers$\geqslant$ 1	-.37	.10	.05***	.02
	(.44)	(.28)	(.006)	(.02)
	[.006]	[.731]	[.111]	[.593]
Providers $\geqslant$ 2	.23	.10	.03***	.02*
	(.32)	(.14)	(.936)	(.01)
	[.01]	[.732]	[.162]	[.606]
Providers $\geqslant$ 3	.41**	.13	.02***	.01***
	(.17)	(.11)	(.007)	(.004)
	[.12]	[.739]	[.248]	[.640]
Providers $\geqslant$ 4	.28***	.14*	.01***	.01***
	(.08)	(.08)	(.003)	(.002)
	[.20]	[.752]	[.312]	[.663]
Providers $\geqslant$ 5	.16***	.08*	.01***	.005***
	(.06)	(.05)	(.001)	(.001)
	[.17]	[.752]	[.327]	[.670]
Providers $\geqslant$ 6	.28***	.04	.003***	.003***
	(.08)	(.02)	(.001)	(.001)
	[.14]	[.744]	[.186]	[.661]
N	51	51	50	50
Unit of Analysis	State	State	State	State
Demographic Controls	N	Y	N	Y
R^2 with no Internet variables	-	.730	-	.582

Each cell represents a different regression. Each row of columns (1) and (3) reports the coefficient on the variable listed, the standard error and the R^2 from a state-based **univariate** regression of the form $I_i^* = \theta + \delta I_i + \nu_i$ with robust standard errors. Columns (2) and (4) adds state level controls for the log of population, the log of income, density, the percent of males, the percent of seniors, the fraction white, education, the age of houses, the percent non-citizen, and housing units. The dependent variable in column (1) and (2) is the fraction of homes in a state with Internet access at home. The dependent variable in (3) and (4) is the per capita number of eBay purchases in the state measured by Einav et al. (2014). ***99%, **95%, *90%

Table 2: The Effect of the Internet on Tax Levels: Baseline

	(1)	(2)	(2')	(3)	(4)	(5)	(6)	(7)
	OLS				Fractional Response		Tobit	
Multiple Providers (I)	-.176***	-.132***	-.018	-.108***	-.209**	-.159***	-.215**	-.206**
Marginal Effect	(.024)	(.030)	(.021)	(.028)	(.022)	(.026)	(.027)	(.054)
N	14,459	14,459	14,459	14,459	14,459	14,459	14,459	14,459
R^2	.821	.645	.641	.809	.810	.648	-	-
Dependent Variable	Town + County	Town	Town	Town	Town + County	Town	Town + County	Town
Controls	Y	Y	Y	Y	Y	Y	Y	Y
State Fixed Effects	Y	Y	Y	N	Y	Y	Y	Y
County Fixed Effects	N	N	N	Y	N	N	N	N
Population Interaction	Y	Y	N	Y	Y	Y	Y	Y
Cluster SE?	Y	Y	Y	Y	Y	Y	Y	Y

The data used in this table ares at the town level. The dependent variable in columns (1), (4) and (6) is the city tax rate plus county tax rate; in all other columns it is only the city tax rate. The measure of Internet penetration is the percent of the population with access to four or more providers; this measure is interacted with the log of population and marginal effects presented. Column (1)-(2) presents the results of equation 7. Column (2') includes population as a control but excludes its interaction with the Internet measure. Column (3) adds county fixed effects. Column (4)-(5) present the mean derivatives (marginal effect) from a fractional response model of equation 8. Columns (6)-(7) use a Tobit model. All magnitudes are scaled to be interpreted as the percentage point change in the tax rate given a change in Internet penetration from no penetration to complete penetration; Dividing by 100 will yield the effect of a one percentage point increase in Internet penetration on the tax rate. Standard errors are clustered at the county level. The Delta Method is used to calculate standard errors. ***99%, **95%, *90%

Table 3: Sensitivity to Controls and County Fixed Effects

	(0)	(1)	(2)	(3)	(4)
	OLS	Fractional Response			
Multiple Providers (I)	-.108***	-.159***	-.155***	-.162***	-.181***
Marginal Effect	(.028)	(.026)	(.026)	(.026)	(.021)
N	14,459	14,459	14,459	14,459	14,459
R^2	.809	.648	.648	.648	.727
Dependent Variable	Town	Town	Town	Town	Town
Controls	All	All Listed in Table Appx.2	(1) Excluding Income, Density, %Agriculture	Add Industry, Employment, and Public Spending Measures	Add All in (3) Plus County Tax Rate
State Fixed Effects	County FE	Y	Y	Y	Y
Cluster SE at County?	Y	Y	Y	Y	Y

The data used in this table ares at the town level. The dependent variable is the city tax rate. The measure of Internet penetration is the percent of the population with access to four or more providers; this measure is interacted with the log of population and marginal effects presented. Column (0) presents the results of equation 7 with county fixed effects. Column (1) presents the baseline results of 8. The subsequent columns add or subtract controls to show the robustness of the results. All magnitudes are scaled to be interpreted as the percentage point change in the tax rate given a change in Internet penetration from no penetration to complete penetration; Dividing by 100 will yield the effect of a one percentage point increase in Internet penetration on the tax rate. Standard errors are clustered at the county level. The Delta Method is used to calculate standard errors. ***99%, **95%, *90%

Table 4: Heterogeneity in the Response to the Internet: Split Sample by Population

	(1)	(2)	(3)	(4)	(5)	(6)
	Total Local Rate Rate		Town Tax Rate			
Multiple Providers (I)	-.091***	-.034	-.075**	-.012	-.314**	-.033
	(.033)	(.022)	(.038)	(.021)	(.134)	(.021)
N	7231	7228	7231	7228	2198	12,261
Jurisdiction Size	Large: Above Median	Small: Below Median	Large: Above Median	Small: Below Median	Large: Above Mean	Small: Below Mean
Controls	Y	Y	Y	Y	Y	Y
State Fixed Effects	Y	Y	Y	Y	Y	Y
Cluster SE at County?	Y	Y	Y	Y	Y	Y

All columns are the result of a fractional response model; population enters as a control but is **not** interacted with any variables. The data used in this table are at the town level. The dependent variable in the first two columns is the city tax rate plus county tax rate; in the final four columns it is only the city tax rate. The measure of Internet penetration is the percent of the population with access to four or more providers. Column (1) and (3) presents the results of equation 8 for large jurisdictions while column (2) and (4) focus on small jurisdictions. A town is defined as small if it is below the median population size. In columns (5) and (6), the split is based on whether the population is above the mean. All magnitudes are scaled to be interpreted as the percentage point change in the tax rate given a change in Internet penetration from no penetration to complete penetration; Dividing by 100 will yield the effect of a one percentage point increase in Internet penetration on the tax rate. Standard errors are clustered at the county level. The Delta Method is used to calculate standard errors in the fractional response model. ***99%, **95%, *90%

Table 5: Response in High- and Low-Tax States

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Marginal Effect at $d \to 0$:	-	-.242*	-.458***	-.095	-.369**	-.209	-.122	-.400***
Low-Tax State		(.141)	(.151)	(.180)	(.149)	(.212)	(.198)	(.152)
Marginal Effect at $d \to 0$:	-	-.077	-.175	-.060	-.075	-.159	.050	-.100
High-Tax State		(.108)	(.163)	(.106)	(.157)	(.220)	(.209)	(.159)
Marginal Effect: Low-Tax	-.072*	-.067*	-.204***	-.011	-.150***	-.125*	-.164**	-.178***
State	(.042)	(.040)	(.063)	(.041)	(.052)	(.067)	(.056)	(.060)
Marginal Effect: High-Tax	-.007	-.004	-.006	.004	-.014	-.039	-.004	.009
State	(.030)	(.030)	(.042)	(.024)	(.043)	(.050)	(.051)	(.042)
N	9792	9792	4707	5085	9792	9792	9792	9792
Tax Rate	Town	Town	Town	Town	Town	Total	Total	Town
Jurisdiction Size	All	All	Large	Small	Large	Large	Large	Large
Restriction / Addition	No Distance Function	None	None	None	None	None	None	Interact w/ Differential
State Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
Border Pair Fixed Effects	N	N	N	N	Y	N	Y	N

All columns are the result of a fractional response model. The dependent variable in all columns except (6) and (7) is the city tax rate; in column (6) and (7) the dependent variable is the city plus county tax rate. The measure of Internet penetration is the percent of the population with access to four or more providers. The sample is restricted to towns within one hundred minutes of a border to reduce the need for an extremely flexible function of distance; towns near same tax borders are also excluded. Population enters as a control but is never interacted. Column (1) presents the results of equation 9 where I restrict all the coefficients on the distance terms to be equal to zero. In column (2), I present results for equation 9 in the text. Columns (3) and (4) estimate equation 9 for the sample split at the median population size (note: I use the median for the full sample to compare to previous tables). All remaining columns use only the sample of large towns. Column (5) adds border pair fixed effects. Column (6) studies the total local rate and column (7) adds border pair fixed effects to the total local rate specification. Column (8) includes additional interactions of the relevant terms with the size of the tax differential at the state border in addition to the terms in the text. All estimates are the marginal effect conditional on the side of the border. All magnitudes are scaled to be interpreted as the percentage point change in the tax rate given a change in Internet penetration from no penetration to complete penetration; dividing by 100 will yield the effect of a one percentage point increase in Internet penetration on the tax rate. Standard errors are clustered at the county level. The Delta Method is used to calculate standard errors in the fractional response model. ***99%, **95%, *90%

Table 6: The Effect By Distance and Tax Differential

Panel A: Distance

	Large Towns		Small Towns	
d	Low Side	High Side	Low Side	High Side
At the	-.458***	-.175	-.095	-.060
border (0)	(.151)	(.163)	(.180)	(.106)
10	-.215**	-.074	-.012	-.007
minutes	(.091)	(.085)	(.086)	(.057)
20	-.105	.004	.026	.018
minutes	(.093)	(.073)	(.069)	(.044)
50	-.250***	.063	.016	.000
minutes	(.087)	(.059)	(.053)	(.029)
90	-.203*	-.072	-.083	.022
minutes	(.104)	(.085)	(.080)	(.038)

Each column presents the mean derivative from column (3) and (4) of the previous table at various distances from the border. Each row evaluates the mean derivative of the distance function at values of 0, 10, 20, 50 and 90 minutes from the border. The standard errors are calculated using the delta method and are clustered at the county level. ***99%, **95%, *90%

Panel B: Tax Differentials

	Large Towns		
b	Low Side	High Side	Difference
Nearly Same	.025	.004	.021
(0)	(.136)	(.080)	
1 percentage	-.093	.008	-.101
point	(.072)	(.053)	
3 percentage	-.246***	.013	-.259***
points	(.072)	(.052)	
5 percentage	-.320***	.017	-.337**
points	(.101)	(.093)	

Each column presents the mean derivative of column (8) from Table 5. The model includes additional interactions of the Internet terms with the size of the state tax differentials at the state border in addition to the terms in the text. Each row evaluates the mean derivative at values of the tax differential at 0, 1, 3, and 5 percentage point differentials. The standard errors are calculated using the delta method and are clustered at the county level. Significance levels on the difference are calculated using nonlinear tests. ***99%, **95%, *90%

Table 7: The Effect of the Internet on Tax Rates Using County Borders

Marginal Effect	(1)	(2)
Low Tax County	-.168**	-.416****
	(.069)	(.100)
High Tax County	.188***	.246***
	(.054)	(.086)
N	5792	3155
Dependent Variable	Town	Town
Sample	All Towns	Large Towns
Borders Used	County Borders Not State Borders	County Borders Not State Borders
County Tax Differentials at Border	Not Zero	Not Zero

All columns are the result of a fractional response model using county borders that are not also state borders. The dependent variable in all columns is the city tax rate. The measure of Internet penetration is the percent of the population with access to four or more providers. The sample is restricted to towns within one hundred minutes of a county border; towns near same tax borders are also excluded. Column (1) presents the results of equation 9 for all towns. In column (2) I present results for equation 9 in the text, using the sub-sample of large towns. All estimates are the marginal effect conditional on the side of the border. All magnitudes are scaled to be interpreted as the percentage point change in the tax rate given a change in Internet penetration from no penetration to complete penetration; dividing by 100 will yield the effect of a one percentage point increase in Internet penetration on the tax rate. Standard errors are clustered at the county level. The Delta Method is used to calculate standard errors in the fractional response model. ***99%, **95%, *90%

Table 8: Border Discontinuity: Response by Nexus in the State

	(1)	(2)	(2')	(3)	(4)	(4')
Marginal Effect at $d \to 0$:	-.946***	-.795***	-.226	-.816***	-.827***	-.188
Low-Tax & Low Nexus	(.236)	(.186)	(.223)	(.190)	(.194)	(.244)
Marginal Effect $d \to 0$:	-.064	.681	.634*	.859	.832	-.043
Low-Tax & High Nexus	(.240)	(.703)	(.361)	(.963)	(.684)	(.148)
Marginal Effect $d \to 0$:	-.257	-.219	-.103	-.143	-.213	-.195
High-Tax & Low Nexus	(.187)	(.138)	(.119)	(.211)	(.171)	(.197)
Marginal Effect $d \to 0$:	.036	.370	.618	.307*	.003	.110
High-Tax & High Nexus	(.249)	(.335)	(.471)	(.187)	(.218)	(.128)
Marginal Effect: Low-Tax	-.369***	-.325***	-.059	-.341***	-.349***	-.037
State & Low Nexus State	(.107)	(.082)	(.051)	(.086)	(.087)	(.056)
Marginal Effect: Low-Tax	-.008	.323	.164***	.215*	.219	.060
State & High Nexus State	(.072)	(.217)	(.048)	(.129)	(.138)	(.030)
Marginal Effect: High-Tax	-.131**	-.090**	-.016	-.075	-.095**	-.035
State & Low Nexus State	(.058)	(.041)	(.027)	(.049)	(.045)	(.030)
Marginal Effect: High-Tax	.139**	.275***	.068	.143*	.149**	.036
State & High Nexus State	(.060)	(.102)	(.042)	(.086)	(.065)	(.037)
N	4707	4707	5085	4707	4707	5085
Tax Rate	Town	Town	Town	Town	Town	Town
Jurisdiction Size	Large	Large	Small	Large	Large	Small
Nexus Definition	Above Mean Total Number	Top Quartile Total Number	Top Quartile Total Number	Above Mean Online Sales	Top Quartile Online Sales	Top Quartile Online Sales
State Fixed Effects	Y	Y	Y	Y	Y	Y

All columns are the result of a fractional response model. The dependent variable is the town tax rate; columns without a prime are the subset of large towns, while columns with a prime are the subset of small towns. The measure of Internet penetration is the percent of the population with access to four or more providers. The sample is restricted to towns within one hundred minutes of a border to reduce the need for an extremely flexible function of distance; towns near same tax borders are also excluded. All specification include additional interactions of the distance function, the high-tax dummy terms, and the Internet penetration term with a dummy variable for the state's nexus status in addition to the terms in the text. Columns (1)-(2) use states that have an above average (or are in the top quartile for) number of firms with nexus as established by the Bruce, Fox and Luna (2014) study. Columns (3)-(4) use the fraction of online sales by firms with nexus to allocate states as being above average or in the top quartile. All estimates are the marginal effect conditional on the side of the border and conditional on whether the state has a high or low amount of nexus firms. The first set of four rows focuses on towns that are located in a region of the border; the second set of four rows focuses on all towns at any distance. All magnitudes are scaled to be interpreted as the percentage point change in the tax rate given a change in Internet penetration from no penetration to complete penetration; dividing by 100 will yield the effect of a one percentage point increase in Internet penetration on the tax rate. Standard errors are clustered at the county level. The Delta Method is used to calculate standard errors in the fractional response model. ***99%, **95%, *90%

Table 9: The Effect of the Internet on Tax Rates Using Lubotsky and Wittenberg (2006)

Marginal Effect LW-Index	(1)	(2)	(3)	(4)	(5)	(6)
Full Sample	-.207***	-.189***	-.179***	-.102***	-	-
	(.022)	(.030)	(.043)	(.023)		
Low Tax State					-.813***	-.320*
$d \to 0$					(.268)	(.175)
High Tax State					-.398**	-.185**
$d \to 0$					(.172)	(.096)
Low Tax State					-.375***	-.134***
					(.089)	(.041)
High Tax State					-.081*	-.046**
					(.049)	(.021)
N	14,459	14,459	7231	7228	4707	5085
R^2	.820	.643	.709	.591	.703	.600
Dependent Variable	Town + County	Town	Town	Town	Town	Town
Sample	All	All	Large	Small	Large	Small

The data used in this table are at the town level. The dependent variable is given for each column and the measure of Internet penetration is constructed using the Lubotsky and Wittenberg (2006) method. The variables used in the analysis include percent of town with access to 0, 1, 2, 3, 4, 5, or more than 6 providers, percent with access to any technology, with access to wireless technology, and with access to download speeds greater than 1500k and greater than 3000k. Column (1) and (2) presents the comparable results to Table 2 column 1 and 2, respectively. Columns (3) and (4) are comparable to Table 4 column 3 and 4. Columns (1) and (2) Interact the Internet measure with population, while (3) and (4) are a split sample analysis without any interactions. Columns (5) and (6) are comparable to Table 5 columns 3 and 4, respectively. For the borders analysis, the first two rows are for towns on the border, while the second two rows are the mean derivatives across all towns on each side of the border. All columns include demographic controls and state fixed effects and are estimated using a linear framework. If applying this procedure to the non-linear fractional response method, for example, I get comparable results. Standard errors are clustered at the county level. ***99%, **95%, *90%

Table 10: The Effect of the Internet on Tax Rates Using Lightning Strikes as an IV (Andersen et al. 2012)

	(0) No IV	(1)	(2)	(3)	(4)	(5)	(6)
First Stage: Internet Use on Lightning	-	-.035***	-.028***	-.037***	-.037***	-.037***	-1.065
		(.006)	(.008)	(.011)	(.011)	(.011)	(.911)
R^2	-	.393	.785	.821	.821	.821	.697
F-Statistic	-	32.34	10.55	11.747	11.747	11.747	1.195
IV Marginal Effect	-.108	-.182	-.415	-.344**	-.191***	-.132	-.012*
	(.088)	(.117)	(.267)	(.175)	(.070)	(.139)	(.006)
R^2	.109	.109	.449	.490	.556	.432	.489
Controls	Y	N	Y	Y	Y	Y	Y
Region Dummies	Y	N	N	Y	Y	Y	Y
N	48	48	48	48	48	48	48
Tax Rate	Total (St.+Local)	Total (St.+Local)	Total (St.+Local)	Total (St.+Local)	Local	State	Total (St.+Local)
Instrument For	Internet at Home	Internet at Home	Internet at Home	Internet at Home	Internet at Home	Internet at Home	Internet Outside Home

This table presents the IV results. The first row represents the reduced form first stage coefficient on the instrument. The row labeled "IV marginal effect" is the second stage marginal effect from a fractional response model. Column (0) presents results for the fractional response model assuming Internet usage at home is exogenous. Column (1) to (3) uses the total tax rate (state plus population weighted municipal and county tax rates) as the dependent variable where each of the columns successively adds controls and region dummies. Column (4) uses only the population aggregated local tax rate in the state and column (5) studies the state rate. In all specifications, the measure of Internet usage is the fraction of people who answer yes to "At home, do you or any member of this household access the Internet?" on the CPS. Column (6) acts as a test by instrumenting for "Do you or any member of this household access the Internet at any location outside the home?". I instrument for the Internet usage from the CPS using the natural log of flash density of lightning strikes in an area. Controls include population, income, density, gender, race, fraction senior, education, public assistance, age of house, non-citizen, rooms, race, and age; regional dummies are as defined by the Census. ***99%, **95%, *90%

A Appendix (For Online Publication)

A.1 Importance of e-Commerce to Local Politicians

An online news search of words such as “sales tax”, “e-commerce”, “revenue lost” combined with political actors such as “mayor” or “council” yields frequent news hits where politicians blame revenue changes on structural changes resulting from the Internet. This provides evidence that policymakers are aware of the problems that e-commerce will generate and shows that they are worried about the future. One example of a recent newspaper article is displayed in Figure Appx.1.

Take Mayor Bartlett at his word, but assume his view of “millions of dollars” means a conservative 3 million dollars. Local sales tax revenue for operating funds were 145 million in FY2013 for the city of Tulsa. This implies a lower bound of an approximately 2% decline (in his view) in revenue from Internet sales, which supposedly “double” every year in Tulsa.

A.2 Derivation of the Nash Equilibrium

A.2.1 No Online Shopping

The revenue functions for the town in the high-tax state and the low-tax state are

$$R_H = t_H(1 - \frac{b + t_H - t_L}{d}),\ R_L = t_L(1 + \frac{b + t_H - t_L}{d}). \tag{A.1}$$

Differentiating with respect to each town’s tax rate yields the best response functions

$$t_H = \frac{1}{2}(d - b) + \frac{t_L}{2},\ t_L = \frac{1}{2}(d + b) + \frac{t_H}{2} \tag{A.2}$$

and after solving the second equation for t_H, it can easily be seen that a Nash equilibrium will exist if d is sufficiently large; I maintain this assumption in the following cases to guarantee positive tax rates.[25] Algebra yields the equilibrium

$$t_H^N = d - \frac{b}{3},\ t_L^N = d + \frac{b}{3}. \tag{A.3}$$

A.2.2 The Internet as a Tax Haven

The revenue functions are now

$$R_H = t_H(1 - \frac{b + t_H - t_L}{d}) - \theta t_H(t_H + T_H),\ R_L = t_L(1 + \frac{b + t_H - t_L}{d}) - \theta t_L(t_L + T_L). \tag{A.4}$$

---

[25] This condition arises because towns cannot be composed of all cross-border shoppers.

Figure Appx.1: Political Salience of e-Commerce

Mayor Bartlett says internet sales are main cause for revenue shortfall

Says Tulsa loses millions in sales taxes to internet purchases

Tulsa Mayor Dewey Bartlett at podium generic

By Steve Berg

Tulsa Mayor Dewey Bartlett says internet sales are the main reason the City of Tulsa is running short on revenue.

"Every year, from what I'm told, the internet sales in Oklahoma, Tulsa in particular, DOUBLE," the Mayor said.

The Mayor says he'll be presenting a lean budget to the City Council in a few days.

He says it's difficult to put an exact number on the amount of sales taxes that are lost to internet sales, but he said it's estimated to be in the millions of dollars.

"That sale represents a replacement of something that would normally be bought in Tulsa, which would normally result in sales taxes," the Mayor said.

Related

Share your opinion on our Facebook page!

Listen to KRMG on iHeartRadio

Follow KRMG on Twitter

Source: KRMG-Tulsa, April 21, 2014

The best response functions are obtained as

$$t_H = \frac{d - b - d\theta T_H}{2 + 2d\theta} + \frac{t_L}{2 + 2d\theta},\ t_L = \frac{d + b - d\theta T_L}{2 + 2d\theta} + \frac{t_H}{2 + 2d\theta} \tag{A.5}$$

Algebra yields the equilibrium

$$t_H^I = \frac{d(3 - 4T_H\theta + T_L\theta) + 2d^2\theta(1 - T_H\theta) - b}{(1 + 2d\theta)(3 + 2d\theta)},\ t_L^I = \frac{d(3 + T_H\theta - 4T_L\theta) + 2d^2\theta(1 - T_L\theta) + b}{(1 + 2d\theta)(3 + 2d\theta)}. \tag{A.6}$$

The tax rates are guaranteed to be positive given the specification of the parameters and it is easy to check that even with the revenue leakage, the revenue functions evaluated at the Nash equilibrium are also positive.

A.2.3 The Internet as an Anti-Haven

The revenue functions are given by

$$R_H = t_H(1 - (1 - \theta)\frac{b + t_H - t_L}{d}),\ R_L = t_L(1 + (1 - \theta)\frac{b + t_H - t_L}{d}). \tag{A.7}$$

The best response functions are obtained as

$$t_H = \frac{d}{2(1 - \theta)} + \frac{t_L - b}{2},\ t_L = \frac{d}{2(1 - \theta)} + \frac{t_H + b}{2} \tag{A.8}$$

Algebra yields the equilibrium

$$t_H^A = \frac{d}{1 - \theta} - \frac{b}{3},\ t_L^A = \frac{d}{1 - \theta} + \frac{b}{3} \tag{A.9}$$

and where as θ approaches one, the border closes and the tax rates are set to extract all the surplus from the consumers.

A.3 Proofs of Propositions

A.3.1 Proof of Proposition 1

Differentiating the Nash equilibrium tax rates given by A.6 with respect to the Internet parameter yields:

$$\frac{\partial t_L^I}{\partial \theta} = -\frac{d(5T_H + 4T_L + 8d^3\theta^2 + 4d^2\theta(6 + T_H\theta) + 2d(9 + 4T_H\theta + 2T_L\theta))}{(3 + 8d\theta + 4d^2\theta^2)^2} < 0 \tag{A.10}$$

and

$$\frac{\partial t_H^I}{\partial \theta} = -\frac{d(4T_H + 5T_L + 8d^3\theta^2 + 4d^2\theta(6 + T_L\theta) + 2d(9 + 2T_H\theta + 4T_L\theta))}{(3 + 8d\theta + 4d^2\theta^2)^2} < 0 \quad \text{(A.11)}$$

which demonstrates that tax rates fall as Internet access increases. The question is then whether taxes fall by more in the low or high-tax state. Tedious algebra then yields a variant of the equation in the text:

$$\left|\frac{\partial t_L^I}{\partial \theta}\right| - \left|\frac{\partial t_H^I}{\partial \theta}\right| = \frac{db}{(3 + d\theta)^2} \geq 0. \quad \text{(A.12)}$$

This demonstrates that taxes fall by more in the low-tax state.

A.3.2 Proof of Proposition 2

Differentiating the Nash equilibrium given by A.9 yields

$$\frac{\partial t_L^A}{\partial \theta} = \frac{\partial t_H^A}{\partial \theta} = \frac{d}{(1 - \theta)^2} > 0 \quad \text{(A.13)}$$

and where it is obvious that $\frac{\partial t_L^A}{\partial \theta} - \frac{\partial t_H^A}{\partial \theta} = 0$.

A.4 The Role of Jurisdiction Size and Proof of Proposition 3

Consider a two jurisdiction model where both localities are located in the *same* state. The jurisdictions differ in their size as in Nielsen (2001). The Nielsen (2001) model yields similar results as when jurisdictions differ based on population density as in Kanbur and Keen (1993). The two localities in the model are on the interval [-1, 1] and the larger of the two towns occupies the space from -1 to ℓ where $1 > \ell > 0$, while the smaller country extends from ℓ to 1. The large jurisdiction sets a tax rate t_l while the small jurisdiction sets a tax rate of t_s. All of the assumptions of the previous model are maintained except I now assume that there are no state tax rate differences across the jurisdictions. In a world without e-commerce, the Nash equilibrium tax rates are given in Nielsen (2001):

$$t_l = d(1 + \frac{\ell}{3}),\ t_s = d(1 - \frac{\ell}{3}) \quad \text{(A.14)}$$

and where it is easy to see that $t_l - t_s = \frac{2\ell d}{3} > 0$. Adding a tax free Internet in the same manner as in the text with a quadratic leakage function yields tax revenue:

$$R_l = t_l(1 + \ell + \frac{t_s - t_l}{d}) - \theta t_l^2,\ R_s = t_l(1 - \ell - \frac{t_s - t_l}{d}) - \theta t_s^2 \quad \text{(A.15)}$$

and the Nash equilibrium is solved as

$$t_l^I = \frac{d(3+\ell+2d\theta(1+\ell))}{(1+2d\theta)(3+2d\theta)},\ t_s^I = \frac{d(3-\ell+2d\theta(1-\ell))}{(1+2d\theta)(3+2d\theta)}. \tag{A.16}$$

and where $t_l^I - t_s^I = \frac{2\ell d}{3+2d\theta} > 0$. More interesting are the comparative statics. Differentiating the Nash equilibrium with respect to the Internet penetration parameter yields:

$$\frac{\partial t_l^I}{\partial \theta} = -\frac{2d^2(\ell(1+2d\theta)^2+(3+2d\theta)^2)}{(3+8d\theta+4d^2\theta^2)^2} < 0 \tag{A.17}$$

and

$$\frac{\partial t_s^I}{\partial \theta} = -\frac{2d^2((3+2d\theta)^2-\ell(1+2d\theta)^2)}{(3+8d\theta+4d^2\theta^2)^2} < 0 \tag{A.18}$$

where the sign of the last inequality follows from the fact that $1 > \ell > 0$ and $(3+2d\theta)^2 > (1+2d\theta)^2$. It is clear that

$$\left|\frac{\partial t_l^I}{\partial \theta}\right| - \left|\frac{\partial t_s^I}{\partial \theta}\right| = \frac{4\ell d^2}{(3+d\theta)^2} > 0. \tag{A.19}$$

Thus, the above model shows that tax rates will fall as more consumers access the Internet when online purchases are tax-free, but tax rates will fall by a larger amount in the large jurisdiction.

If the Internet acts as an anti-haven, the revenue functions are

$$R_l = t_l(1+\ell+(1-\theta)\frac{t_s-t_l}{d}),\ R_s = t_l(1-\ell-(1-\theta)\frac{t_s-t_l}{d}) \tag{A.20}$$

and the Nash equilibrium can be solved for as

$$t_l^A = \frac{d(1+\frac{\ell}{3})}{(1-\theta)},\ t_s^A = \frac{d(1-\frac{\ell}{3})}{(1-\theta)}. \tag{A.21}$$

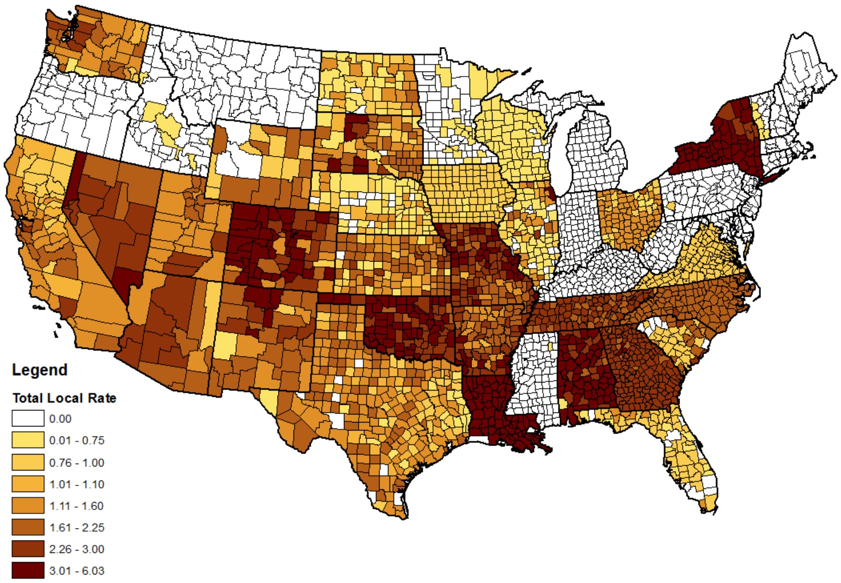
It is easy to see that the Internet will place upward pressure on tax rates

$$\frac{\partial t_l^A}{\partial \theta} = \frac{d(1+\frac{\ell}{3})}{(1-\theta)^2} > 0 \tag{A.22}$$

and

$$\frac{\partial t_s^A}{\partial \theta} = \frac{d(1-\frac{\ell}{3})}{(1-\theta)^2} > 0 \tag{A.23}$$

Figure Appx.2: County Plus Average Municipal Tax Rates



The map shows the county tax rate plus the population weighted municipal tax rate in the county.

but where unlike the case in the text, the size of the effect depends on the jurisdiction's size:

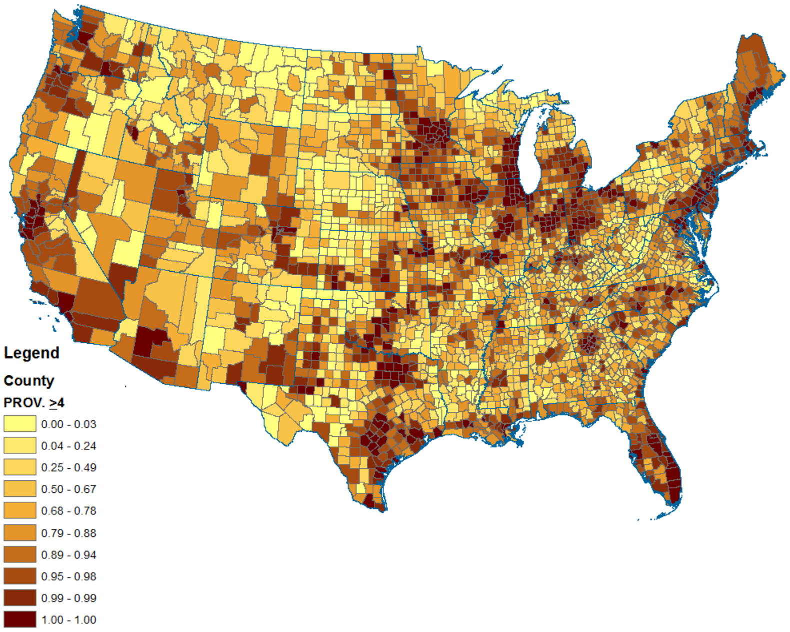
$$\frac{\partial t_l^A}{\partial \theta} - \frac{\partial t_s^A}{\partial \theta} = \frac{2\ell d}{(1-\theta)^2} > 0. \tag{A.24}$$

Intuitively, the anti-haven effect has a larger effect in the big jurisdiction because it allows it to have an even larger size advantage relative to the pre-internet era. The effect was symmetric in the border jurisdictions because both jurisdictions were of the same size and the heterogeneity resulting from the state tax rate affected both jurisdictions similarly.

A.5 Data

Figure Appx.2 and Appx.3 show the spatial distribution of tax rates and Internet penetration. Table Appx.1 shows descriptive statistics on tax changes from 2003 to 2011 after linking local sales tax data across years. Although I have previous data on local sales tax rates, Internet penetration data is only available in 2011.

Figure Appx.3: Percent of Population With Access to 4 or More Providers



The map shows the fraction of the population in each county with access to four or more providers.

Table Appx.1: Town Tax Changes: From 2003 to 2011

	High-Tax State	Low-Tax State	All States
High Penetration	.081	.123	.104
	(.269)	(.357)	(.322)
	[-1.5, 2.75]	[-4, 3.5]	[-4, 3.5]
Low Penetration	.069	.172	.125
	(.277)	(.554)	(.451)
	[-2.75, 2.5]	[-5, 6]	[-5, 6]
All Penetrations	.075	.145	.114
	(.273)	(.457)	(.387)
	[-2.75, 2.75]	[-5, 6]	[-5, 6]

This table shows the average town sales tax rate change from 2003 to 2011 with standard deviations in ( ) and the range in [ ]. The sample is split by towns in relatively high-tax states and low-tax states as well as towns with high vs. low Internet penetration. A town is a high penetration town if the percent of the population with access to four or more providers (in 2011) is equal to or above the median in the full sample. The averages includes towns that are allowed to have LOST, but that do not change their tax rate. Although I have tax data for 2003, I do not have Internet penetration for 2003.

Table Appx.2: Summary Statistics

Averages with Standard Deviations in ( )

	Variable	Census Place	County Level
Tax	Municipal Tax Rate	.77	.81
		(1.15)	(1.08)
	County Tax Rate	1.07	1.14
		(1.15)	(1.14)
Internet	% of Pop: Any Internet Technology	.99	.98
		(.06)	(.05)
	% of Pop: Any Wired Technology	.91	.84
		(.24)	(.15)
	% of Pop: Providers ≥ 4	.71	.67
		(.43)	(.35)
	% of Pop: Download Speed Greater than 3000k	.98	.96
		(.13)	(.09)
Controls	Number of Neighbors	1.81	6.97
		(1.86)	(1.28)
	Area	5.60	102
		(17.65)	(139)
	Perimeter	14.41	148
		(24.06)	(83.66)
	Population	9615	93,196
		(85,836)	(339,842)
	Senior (%)	15.90	15.61
		(7.97)	(4.17)
	Less Than College (%)	81.65	81.47
		(14.37)	(8.06)
	Work in State (%)	96.01	96.04
		(9.03)	(6.99)
	Male (%)	49.06	49.93
		(5.34)	(2.25)
	Ratio of Private to Public School Students	.13	.11
		(.54)	(.07)
	Public Assistance (%)	2.42	2.16
		(3.44)	(1.34)
	Non-Citizen (%)	3.01	3.04
		(5.66)	(3.81)
	White (%)	84.41	83.89
		(20.32)	(15.60)
	Mean Income	57,090	55,728
		(31913)	(12,125)
	Median Age	39.24	39.67
		(7.91)	(4.84)
	Obama Vote Share	44.00	40.01
		(13.68)	(13.87)
	Number of Rooms in Home	5.59	5.54
		(.74)	(.42)
	Average Age of Home	39.24	38.35
		(7.91)	(11.58)
	Density	1208	174
		(1472)	(1601)
	% Agriculture	4.57	7.36
		(7.45)	(7.42)
	Sample Size	14,459	2086

Averages are for the towns and counties within states that allow for local taxes only. All regression also include controls for proximity to water, an international border, and the latitude and longitude of the town.

Table Appx.3: Testing For Internet Discontinuities at Borders

	(1)	(2)	(3)	(4)
	Polynomial Regression			Local Linear
Effect of Switching (Low to High)	-.038	.058	-.029	-.010
	(.039)	(.038)	(.045)	(.030)
R^2	.357	.377	.404	-
Border Fixed Effects	N	Y	Y	N
State Fixed Effects	Y	N	Y	N

This table presents the results of equation A.25. The columns differ in the fixed effects included. The final column uses local linear regression techniques with an optimal bandwidth selection procedure. ***99%, **95%, *90%

With respect to the data, because I combine the tax data with Census data, I restrict the sample to municipalities that are identified Census Places.[26] To do this, I name merge geo-coded data provided by the 2010 American Community Survey (ACS) to the tax data. The sales tax data used in this paper and the process of merging the tax data to Census places are described in detail in Agrawal (forthcoming); the only difference is that I use a cross-section from 2011 rather than 2010.[27] Because I have merged the the jurisdictions in the tax data file to Census places, I also have access to any data in the 2010 ACS. Table Appx.2 describes the summary statistics of the control variables used in the regressions. Using the geo-spatial data constructed in Agrawal (forthcoming), I also have calculated the driving time to the nearest state border. Distance to the border is the time (in minutes) that minimizes the driving time from the population weighted centroid of a town to the closest state border and a major road intersection; by using minutes rather than miles, I am able to capture the true cost of driving to the border. I identify a town as being in a high-tax state if its state that has a higher state sales tax rate than the nearest neighboring state; a town is in a low-tax state if its state sales tax rate is lower than the nearest neighboring state.

A.6 Testing For Differences at Borders

Showing the response of municipalities is asymmetric on the high-tax and low-tax side of borders is most convincing if it can be shown that the Internet penetration rates do not vary based on the side of the border conditional on other explanatory variables. To test this, I implement a fractional response model, for the sub-sample of municipalities not bordering a

---

[26] A Census Place is generally an incorporated place with an active government and definite geographic boundaries such as a city, town, or village. In some western states, a Census Place may be an unincorporated place that has no definite boundaries or government. Census Places contain some locations that may not have legal authority or jurisdiction to set sales taxes.

[27] Although I do have panel data on local sales tax rates until 2011, the Internet penetration data dates only to 2011.

same-tax border:

$$E(I_i|\cdot) = G(\alpha + \beta H_i + \sum_{k=1}^{K} \lambda_k (d_i)^k + \sum_{k=1}^{K} \delta_k H_i (d_i)^k + \zeta + \sum_m X_{im}\gamma_m) \tag{A.25}$$

where all variables are as defined in the text, distance is the driving time to the border and ζ are state fixed effects or border pair fixed effects, so that I am using variation in Internet penetration within a state or within the region of a particular border crossing. As in the text, I also control for latitude and longitude to account for position along the border. I estimate this using a fractional response model given that Internet penetration rates are between 0 and 1 with mass points at both extremes. The results in table Appx.3 indicate that there is no difference in Internet penetration depending on the side of the border in a local region of the border.

I would also like to show that towns on the high state sales tax side of a border are similar to towns on the low-tax side. The best test is to see if any observable characteristics are different at the border. To test for this, I implement the regression discontinuity design for each of the control variables used in the regression. I do this for both the full sample and the split samples based on population. As is evident in table Appx.4, no variables are significant in the full sample and only three variables show significant differences in the split sample. This implies that on average, towns on the high-tax side are similar to towns on the low-tax side. Intuitively, this makes sense given that the sales tax is easily evaded through cross-border shopping, which implies that theoretically, there is little reason to believe that people will sort across borders in order to reduce their sales tax liability. Furthermore, state sales tax rates will only be weakly related to public good provisions given that sales taxes are only one of many taxes that state governments have access to and spending in a local region is not proportional to the amount of sales in that region.

A.7 Other Measures of Internet Use

The results in Table 1 comparing CPS data with measures of broadband penetration suggested that four different measures of Internet penetration have an R^2 greater than .10. In Table Appx.5 I show that the border discontinuity results are robust to using most of these measures as the proxy variable. More detailed specifications reported in the text are also robust. This suggests that the signs are consistent across the set of possible proxy variables. Of course, using Lubotsky and Wittenberg (2006) is the preferred method of verifying robustness because it encompasses multiple variables simultaneously and can say something about the magnitude of the coefficients in the presence of measurement error concerns.

Table Appx.4: Testing For Observable Differences at Borders

Variable	Full Sample	Split Samples	
	All Towns	Large Towns	Small Towns
Number of Neighbors	.102	-.180	.794
	(.359)	(.972)	(.605)
Area	-.380	-.462	-.404
	(.324)	(.473)	(.308)
Perimeter	-.064	-.063	.296
	(.219)	(.362)	(.282)
Population	-.624	-.499	-.712*
	(.410)	(.549)	(.378)
Senior (%)	.348	-7.16**	4.69
	(2.02)	(3.64)	(3.01)
Less Than College (%)	-3.97	-3.86	-3.30
	(3.61)	(4.93)	(3.87)
Work in State (%)	-4.63	-6.79	-3.44
	(4.14)	(6.60)	(4.95)
Male (%)	-1.33	1.25	-2.41
	(1.43)	(1.32)	(1.84)
Ratio of Private to Public School Students	.078	-.020	.163
	(.130)	(.261)	(.157)
Public Assistance (%)	.656	.977	.288
	(.938)	(1.25)	(1.07)
Non-Citizen (%)	-.149	1.49	-2.00
	(1.64)	(3.65)	(1.32)
White (%)	3.89	4.91	3.74
	(3.05)	(4.27)	(3.90)
Mean Income	.242	-.111	.464
	(.236)	(.325)	(.340)
Median Age	-.968	-6.31	1.08
	(19.93)	(3.85)	(2.51)
Obama Vote Share	3.62	2.84	2.09
	(3.09)	(4.56)	(3.18)
Number of Rooms in Home	.483	-1.02	1.02**
	(.360)	(.855)	(.496)
Average Age of Home	.828	-1.04	3.03
	(3.29)	(4.47)	(3.74)
Density	-.288	-.198	-.390
	(.346)	(.670)	(.356)
% Agriculture	-1.80	1.62	-1.86
	(1.89)	(2.57)	(2.22)

Each cell represents the result from a **different** RD design where the outcome variable is given in the row and the running variable is driving time to the nearest border

Table Appx.5: Response in High and Low States Using Other Proxy Variables

	(0)	(1)	(2)	(3)
Marginal Effect at $d \to 0$:	-.242*	-.357*	-.329**	-.352**
Low-Tax State	(.141)	(.191)	(.153)	(.176)
Marginal Effect at $d \to 0$:	-.077	-.028	-.082	-.028
High-Tax State	(.108)	(.134)	(.098)	(.099)
Marginal Effect: Low-Tax	-.067*	-.050	-.103**	-.030
State	(.040)	(.053)	(.047)	(.055)
Marginal Effect: High-Tax	-.004	.015	-.024	.030
State	(.030)	(.034)	(.030)	(.029)
N	9792	9792	9792	9792
Tax Rate	Town	Town	Town	Total
Jurisdiction Size	All	All	All	All
Proxy Variable	Prov ≥ 4	Prov ≥ 3	Prov ≥ 5	Prov ≥ 6
Controls	Y	Y	Y	Y
State Fixed Effects	Y	Y	Y	Y

All columns are the result of a fractional response model. The dependent variable in all columns is the town tax rate; the specification is identical to column (2) in Table 5 and this is reproduced in the first column for comparison. The only difference in the subsequent columns is that this table uses different measures of Internet penetration. All estimates are the marginal effect conditional on the side of the border. All magnitudes are scaled to be interpreted as the percentage point change in the tax rate given a change in Internet penetration from no penetration to complete penetration; dividing by 100 will yield the effect of a one percentage point increase in Internet penetration on the tax rate. Standard errors clustered and the county level. The Delta Method is used to calculate standard errors in the fractional response model. ***99%, **95%, *90%

A.8 Internet Regulation By State

The National Conference of State Legislatures (NCSL) tracks Broadband Statutes by state.[28] I use this database to determine which states have state statutes that encourage broadband penetration in under-served or rural areas. I also determine which states have tax credits or other regulatory incentives for investment in broadband capabilities. I create a dummy variable that equals one if the state has policies to stimulate broadband usage along these two margins.

States that encourage broadband service in under-served areas include: Arkansas, California, Colorado, the District of Columbia, Indiana, Kentucky, Maine, Minnesota, Missouri, North Carolina, Oregon, Tennessee, Utah, and Virginia. States that do not necessarily have policies that explicitly encourage broadband deployment to under-served areas, but that provide tax credits or favorable regulatory regimes include: Mississippi, Nevada, Oklahoma, Vermont, and West Virginia.

I then include additional interactions of this dummy variable with the I, H and d terms in equation 9 to see if the effects of I are affected by looking at state level broadband poli-

---

[28] See http://www.ncsl.org/research/telecommunications-and-information-technology/broadband-statutes.aspx . Please note that the NCSL uses the phrase, “provides examples” of broadband statutes to describe its listing of broadband legislation. This phrase is used because the NCSL cannot be one-hundred percent sure that they have identified all pieces of legislation relating to broadband. The intent of this listing was to identify all bills and the NCSL is confident they have identified most – especially the most significant pieces of legislation.

Table Appx.6: Response by Broadband Policies

	(1)	(2)	(2')
Marginal Effect at $d \to 0$: Low-Tax State & No Broadband Policy	-.450***	-.441**	-.137
	(.162)	(.174)	(.214)
Marginal Effect $d \to 0$: Low-Tax State & Favorable Broadband Policy	-.304	-.300	.167
	(.295)	(.314)	(.315)
Marginal Effect $d \to 0$: High-Tax State & No Broadband Policy	-.167	-.159	-.105
	(.277)	(.255)	(.147)
Marginal Effect $d \to 0$: High-Tax State & Favorable Broadband Policy	-.159*	-.176	.023
	(.094)	(.117)	(.108)
Marginal Effect: Low-Tax State & No Broadband Policy	-.212***	-.196***	-.056
	(.081)	(.072)	(.052)
Marginal Effect: Low-Tax State & Favorable Broadband Policy	-.147**	0.097	.151**
	(.072)	(.087)	(.069)
Marginal Effect: High-Tax State & No Broadband Policy	.025	.030	-.003
	(.069)	(.061)	(.032)
Marginal Effect: High-Tax State & Favorable Broadband Policy	-.025	-.035	.022
	(.032)	(.037)	(.026)
N	4707	4707	5085
Tax Rate	Town	Town	Town
Jurisdiction Size	Large	Large	Small
Broadband Policy Definition	Favorable Broadband Environment & Encourage Broadband Availability	Encourage Broadband Usage	Encourage Broadband Usage
Controls	Y	Y	Y
State Fixed Effects	Y	Y	Y

All columns are the result of a fractional response model. The dependent variable is the town tax rate; columns without a prime are the subset of large towns, while columns with a prime are the subset of small towns. The measure of Internet penetration is the percent of the population with access to four or more providers. The sample is restricted to towns within one hundred minutes of a border to reduce the need for an extremely flexible function of distance; towns near same tax borders are also excluded. All specification interact the distance function, the high-tax dummy terms, and the Internet penetration term with a dummy variable for whether the state has policies that encourage broadband penetration in under-served areas and whether the state encourages broadband deployment by reducing regulation in addition to the terms in the text. All estimates are the marginal effect conditional on the side of the border and conditional on whether the state has a broadband policy. The first set of four rows focuses on towns that are located in a region of the border; the second set of four rows focuses on all towns at any distance. All magnitudes are scaled to be interpreted as the percentage point change in the tax rate given a change in Internet penetration from no penetration to complete penetration; dividing by 100 will yield the effect of a one percentage point increase in Internet penetration on the tax rate. Standard errors are clustered at the county level. The Delta Method is used to calculate standard errors in the fractional response model. ***99%, **95%, *90%

Table Appx.7: Robustness to Polynomial Order

	(1)	(2)	(3)	(4)	(5)	(6)
Marginal Effect at $d \to 0$:	-.458***	-.193*	-.635***	-.511***	-.168	-.485***
Low-Tax State	(.151)	(.104)	(.155)	(.183)	(.144)	(.169)
Marginal Effect at $d \to 0$:	-.175	-.015	.008	.033	-.165	-.200
High-Tax State	(.163)	(.090)	(.223)	(.273)	(.120)	(.173)
Marginal Effect: Low-Tax	-.204***	-.219***	-.202***	-.200***	-.102*	-.076
State	(.063)	(.063)	(.063)	(.063)	(.060)	(.065)
Marginal Effect: High-Tax	-.006	.003	-.004	-.004	-.017	-.015
State	(.042)	(.042)	(.042)	(.042)	(.043)	(.044)
N	4707	4707	4707	4707	6844	6844
Tax Rate	Town	Town	Town	Town	Town	Town
Jurisdiction Size	Large	Large	Large	Large	Large	Large
Robustness Exercise	3rd degree for towns ≤100 minutes from border	1st degree for towns ≤100 minutes from border	4th degree for towns ≤100 minutes from border	5th degree for towns ≤100 minutes from border	3rd degree for towns ≤300 minutes from border	5th degree for towns ≤300 minutes from border
Controls	Y	Y	Y	Y	Y	Y
State Fixed Effects	Y	Y	Y	Y	Y	Y

All columns are the result of a fractional response model. The dependent variable in all columns is the city tax rate. The measure of Internet penetration is the percent of the population with access to four or more providers. The sample is restricted to towns within one hundred minutes of a border in the first four columns and to towns within 300 minutes of a border in the last two columns; towns near same tax borders are also excluded. Column (1) presents the results of equation 9 for the specification in the text where the polynomial is a third degree. Column (2) uses a linear distance function, column (3) uses a fourth degree polynomial and column (4) uses a fifth degree polynomial. Columns (5) and (6) estimate 9 using a third and fifth degree polynomial for the sample of towns within 300 minutes of the nearest border. All estimates are the marginal effect conditional on the side of the border; the first two rows focus on the marginal effects as distance to the border approaches zero, while the second set of rows focuses on the mean derivatives conditional on being on a particular side of the border. All magnitudes are scaled to be interpreted as the percentage point change in the tax rate given a change in Internet penetration from no penetration to complete penetration; dividing by 100 will yield the effect of a one percentage point increase in Internet penetration on the tax rate. Standard errors are clustered at the county level. The Delta Method is used to calculate standard errors in the fractional response model. ***99%, **95%, *90%

cies. The results in table Appx.6 indicate very little difference across states with broadband policies versus states without such policies. This result is reassuring in two respects: (1) it suggest that the Internet penetration measure is doing a good job and that little extra information persists in state level policies and (2) it demonstrates that the cutting the sample will not always produce varying results for different blocks of states.

A.9 Other Robustness Checks

In this section, I show that the results of the border discontinuity design are robust to various polynomial orders and to using the full sample of all distances. I also show the results are robust to various weighting methods, looking at linear borders, more conservative clustering of standard errors, and to spatial lag models. I focus on the large town sample given that is where the theory predicts significant results. Results for small towns are also robust and generally remain insignificant as in the text.

Table Appx.7 shows that the effects are robust to using various polynomial orders. The third degree polynomial was preferred as it was selected by the leave-one out cross-validation procedure. When expanding the sample from towns within one hundred miles of a border to towns within three hundred miles of a border, results are quite similar in magnitude. Note that expanding the distance threshold for the sample should also be accompanied by using a higher order polynomial. The results from the fifth degree polynomial with the full sample and the more restrictive sample and the third degree polynomial are almost identical in a local region of the border.

Table Appx.8: Various Robustness Checks

	(1)	(2)	(3)	(4)	(5)	(6)
Marginal Effect at $d \to 0$:	-.458***	-.458***	-.664***	-.825***	-.512**	-.590***
Low-Tax State	(.151)	(.105)	(.244)	(.180)	(.228)	(.208)
Marginal Effect at $d \to 0$:	-.175	-.178	-.181	-.327	-.281	-.112
High-Tax State	(.163)	(.133)	(.227)	(.223)	(.174)	(.228)
Marginal Effect: Low-Tax	-.204***	-.204***	-.273***	-.257***	-.163**	-.192***
State	(.063)	(.061)	(.095)	(.077)	(.082)	(.073)
Marginal Effect: High-Tax	-.006	-.006	-.007	-.076	-.077	-.003
State	(.042)	(.067)	(.049)	(.044)	(.050)	(.045)
N	4707	4707	4707	4707	1741	4707
Tax Rate	Town	Town	Town	Town	Town	Town
Jurisdiction Size	Large	Large	Large	Large	Large	Large
Robustness Exercise	Baseline Results from Text	Cluster at State Level	Weight by Population	Equal Weight to Each State	Linear Borders	Spatial Lag: IV
Controls	Y	Y	Y	Y	Y	Y
State Fixed Effects	Y	Y	Y	Y	Y	Y

The dependent variable in all columns is the city tax rate. The measure of Internet penetration is the percent of the population with access to four or more providers. The sample is restricted to towns within one hundred minutes of a border; towns near same tax borders are also excluded. Column (1) presents the results of equation 9 for the specification in the text where the polynomial is a third degree while column (2) clusters the standard errors at the state level. Column (3) weights each town by its population. Column (4) gives each town in states with many municipalities less weight than towns in states with few municipalities; this weighting gives each state in the sample equal weight. Column (5) restricts the sample to linear borders. Column (6) includes a spatial lag of the independent variable as a control and instruments for it. All estimates are the marginal effect conditional on the side of the border; the first two rows focus on the marginal effects as distance to the border approaches zero, while the second set of rows focuses on the mean derivatives conditional on being on a particular side of the border. All magnitudes are scaled to be interpreted as the percentage point change in the tax rate given a change in Internet penetration from no penetration to complete penetration; dividing by 100 will yield the effect of a one percentage point increase in Internet penetration on the tax rate. Standard errors are clustered at the county level. The Delta Method is used to calculate standard errors in the fractional response model. ***99%, **95%, *90%

Table Appx.8 shows the results of several robustness checks. The first is clustering standard errors at the state level instead of the county level in case Internet providers make entry decisions at the state level rather than at the neighborhood level. The results are also robust to various weighting schemes including weighting by population and giving states with few towns equal weight to states with many towns. The population weighting results become larger as suggested by the role of population in the theory If I restrict the border analysis to linear borders – likely to be on a featureless plane – the results are also unchanged. Finally, if the reader is worried about spatial auto correlation in the tax rate variable, I can include a spatial lag of the tax rate as a control. I define the spatial lag as the average town tax rate within thirty miles and I instrument for it using the average of the neighboring town's area and perimeter. Justification for the instrument is discussed at length in Agrawal (2013); I use a standard IV framework of Kelejian and Prucha (1998) rather than an a fractional response model with an endogenous regressor for simplicity.

A.10 Using Variation within MSAs Crossing State Borders

State borders can be quite long. One way to address this issue is to exploit a different source of variation by focusing on MSAs that cross-state borders. As noted in Agrawal and Hoyt (2014), approximately 50 MSAs straddle state borders; approximately 75 million people live in these MSAs. MSAs that cross-state borders are defined to have a common and integrated labor market as well as defined commuting patterns. It is likely that amenities within the MSA are quite similar. To show that the results are robust, I use only towns that are in cross-border MSAs and implement:

$$E(\tau_i|\cdot) = G(\alpha_0 + \beta_0 I_i + \beta_1 h_i + \beta_2 I_i h_i + \Theta + \sum_m X_{im}\gamma_m), \tag{A.26}$$

where Θ are MSA fixed effects and h_i is a dummy variable that equals one if the town is on the high-tax side of the state border that cuts through the MSA. I define h_i in two manners. First, h_i is equal to one if the state that the town is located in sets a higher state tax rate relative to the nearest neighboring state (as measured by driving distance); this is the same measure as in the text. Second, I define h_i as equal to one if the state that the town is located in sets the highest state tax rate in the MSA; for two-state MSAs this is the same as the previous measure and it will only differ in MSAs that straddle three or four state borders. All other terms in the specification are as defined in the text and a polynomial in distance is excluded given that distances to borders are relatively small within an MSA. It is important to note that in this specification, the fixed effects used apply to a much smaller geographic area that is highly integrated. The advantage of this approach is that the fixed effects control for possibly omitted factors that are correlated with both tax rates and Internet penetration. The disadvantage of the approach is that the sample become much smaller. Standard errors are clustered at the MSA level. The results are in table Appx.9.

The results follow a similar pattern as in the text – the Internet places downward pressure on towns on the low-tax side of MSAs. This is reassuring because it shows that the results are robust to a common labor market where amenities, industries, and work patterns are likely to be similar.

Table Appx.9: The Effect of the Internet on Tax Rates Using Within MSA Variation

Marginal Effect	(1)	(2)
Low Tax Side of MSA	-.677*	-.610*
	(.349)	(.321)
High Tax Side of MSA	-.031	-.018
	(.171)	(.129)
N	1297	1313
Dependent Variable	Town	Town
Sample	Large Towns	Large Towns
High-Tax State	If state is higher than nearest neighbor	If state is highest in the MSA
Standard Errors	Cluster MSA	Cluster MSA

All columns are the result of a fractional response model with towns in MSAs that cross state borders. The dependent variable in all columns is the city tax rate. The measure of Internet penetration is the percent of the population with access to four or more providers. Towns near same tax borders are also excluded. Column (1) presents the results of equation A.26 for large towns towns where h_i is equal to one if the state that the town is located in sets a higher state tax rate relative to the nearest neighboring state. In column (2) I present results where h_i as equal to one if the state that the town is located in sets the highest state tax rate in the MSA. All estimates are the marginal effect conditional on the side of the border. All magnitudes are scaled to be interpreted as the percentage point change in the tax rate given a change in Internet penetration from no penetration to complete penetration; dividing by 100 will yield the effect of a one percentage point increase in Internet penetration on the tax rate. Standard errors are clustered at the MSA level. The Delta Method is used to calculate standard errors in the fractional response model. ***99%, **95%, *90%

A.11 Nexus By State

The data on the number of firms with nexus in a state was provided to me by William Fox. In a series of projects, nexus information was gathered for a sample of companies. The description of how the data was collected is described in Bruce, Fox and Luna (2014) and relevant portions quoted below:

> "We identify our firms from Internet Retailer's Top 500 Guide and then determine in which states each of these firms claim to have sales tax nexus. ... We derive our measures of nexus from hand-collected information for a large sample of online retailers. Specifically, for a series of prior projects, we have visited each company's web site and attempted to place hypothetical orders from each sales-taxing state. If the company added sales tax to the purchase, that company-state pair is coded as a one (and zero otherwise). We gathered nexus information for the 100 largest companies (in terms of online sales in 2005) in 2006, and the largest 50 and a sample of 50 other top-500 companies in 2008. For the 2010 data, we include any companies from the 2006 and 2008 data that were still operating, plus a number of additional companies, for a total of 179 companies. ... We are left with a matrix of nexus information for a reasonably large sample of firm-state pairs for these five years, which we aggregate to the state level in several ways. First, we simply count the number of companies that collect sales taxes on online purchases by residents of a state (Nexus 1). For example, a state having a value of 50 for this variable indicates that 50 of the companies in our data collected sales taxes on purchases from that state in that year.
>
> Recognizing that online companies are heterogeneous, our second nexus measure (Nexus 2) is the share of total online sales among the companies in our data that occur in companies that collect sales taxes on online purchases by residents of a state. This measure allows larger companies (in terms of the dollar value of sales) to have a larger impact on our nexus measure because we are weighting firm nexus by firms sales and dividing by total sales for firms in our sample. To illustrate, a state having a value of 50 for this measure indicates that the companies collecting sales taxes on online purchases from that state represent 50 percent of the total dollar value of online sales among companies in our sample in that year. Hence, two states with the same number of companies collecting tax could have very different values of our second nexus measure. For the purposes of calculating Nexus 2, we draw online sales data from Internet Retailer's Top 500 Guide in each year. Such sales data are available for a large subset of the companies for which we have nexus data."

The team that collected this research data provided me with statistics by state for the 2010 cross-section. Although I do not have access to the raw data underlying Nexus 1 and Nexus 2, I was provided with information about the state's relative ranking in the distribution. In particular, the authors disclosed whether each state was in the first, second, third, or fourth quartile of the distribution of Nexus 1 or Nexus 2. I was told whether the state was above or below the mean of the respective variable. In 2010, the average state had approximately 70 firms from the sample with nexus. The mean of Nexus 2 is approximately 56.

A.12 Retail Firms Headquartered in a State

As an alternative measure, I use a count of the number of firms in traditional retail sectors that are headquartered within a state. This measure does not directly get at nexus as Bruce, Fox and Luna (2014); rather, it is designed to proxy for nexus by state. The underlying assumption is that states with more retail firms headquartered in the state are more likely to have more firms with nexus. This is simply designed as a check on the Bruce, Fox and Luna (2014) measure. To construct this, I classify firms by their NAICS codes as "retail firms" – firms traditionally remitting retail sales taxes on most purchases; I omit NAICS classifications relating to grocery stores because whether food sales are subject to the retail sales tax varies by state. I count the total number of retail firms headquartered in a state as listed in the Compustat database. Thanks to Jeffrey Hoopes for generating the summary statistics for me. The NAICS codes used to classify firms as being retail include NAICS codes: 441110, 441120, 441210, 441221, 441222, 441229, 441310, 441320, 442110, 442210, 442291, 442299, 443111, 443112, 443120, 443130, 444110, 444120, 444130, 444190, 444210, 444220, 445310, 446120, 446130, 446199, 448110, 448120, 448130, 448140, 448150, 448190, 448210, 448310, 448320, 451110, 451120, 451130, 451140, 451211, 451212, 451220, 452111, 452112, 452910, 452990, 453110, 453210, 453220, 453910, 453920, 453930, 453991, 453998, 454111, 454112, 454113, 454210, 454390, 722110, 722211, 722212, 722213, 722310, 722320, and 722330.

I do not include retail firms for which remitting the sales tax on most products is not necessarily common to all states; this mainly includes grocery stores. However, the results in the following table are robust to including NAICS sectors 445110, 445120, 445210, 445220, 445230, 445291, 445292, 445299, 446110, 446191, 447110, 447190, 452910, 454311, 454312, 454319, 454390, 721110, 721120, 721120, 721191, 721199, 721211, 721214, 721310, 722213, and 722410.

The results of this robustness check are presented in table Appx.10. The results are generally consistent in sign although are much weaker than the results in text, especially for the high-tax results. The intuition for the weaker results is that the number of firms headquartered in a particular state is only a weak proxy for the number of firms with nexus

Table Appx.10: Response by Nexus Based on Headquarters

	(1)	(2)
Marginal Effect at $d \to 0$: Low-Tax State & Low Nexus State	-.765***	-.647***
	(.206)	(.174)
Marginal Effect $d \to 0$: Low-Tax State & High Nexus State	.196	.267
	(.409)	(.593)
Marginal Effect $d \to 0$: High-Tax State & Low Nexus State	-.205	-.193
	(.173)	(.152)
Marginal Effect $d \to 0$: High-Tax State & High Nexus State	.197	.240
	(.237)	(.288)
Marginal Effect: Low-Tax State & Low Nexus State	-.296***	-.250***
	(.091)	(.077)
Marginal Effect: Low-Tax State & High Nexus State	.017	.012
	(.093)	(.131)
Marginal Effect: High-Tax State & Low Nexus State	.002	-.001
	(.051)	(.047)
Marginal Effect: High-Tax State & High Nexus State	.028	.027
	(.066)	(.081)
N	4707	4707
Tax Rate	Town	Town
Jurisdiction Size	Large	Large
Nexus Definition	Top Quartile Retail Headquarters	Top Quartile Retail Headquarters (including food)
Controls	Y	Y
State Fixed Effects	Y	Y

All columns are the result of a fractional response model. The dependent variable is the town tax rate; columns without a prime are the subset of large towns, while columns with a prime are the subset of small towns. The measure of Internet penetration is the percent of the population with access to four or more providers. The sample is restricted to towns within one hundred minutes of a border to reduce the need for an extremely flexible function of distance; towns near same tax borders are also excluded. All specification include additional interactions of the distance function, the high-tax dummy terms, and the Internet penetration term with a dummy variable for the state's nexus status in addition to the terms in the text. Nexus status of the state is proxied for by the number of retail firms listed in Compustat that have a headquarters in the state. A state is defined as a high nexus state if the number of retail firms with headquarters is in the top quartile of the distribution across states. All estimates are the marginal effect conditional on the side of the border and conditional on whether the state has a high or low amount of nexus firms. The first set of four rows focuses on towns that are located in a region of the border; the second set of four rows focuses on all towns at any distance. All magnitudes are scaled to be interpreted as the percentage point change in the tax rate given a change in Internet penetration from no penetration to complete penetration; dividing by 100 will yield the effect of a one percentage point increase in Internet penetration on the tax rate. Standard errors are clustered at the county level. The Delta Method is used to calculate standard errors in the fractional response model. ***99%, **95%, *90%

in a state. Thus, the results using the Bruce, Fox and Luna (2014) index are preferred.

A.13 Lubotsky and Wittenberg (2006)

I have access to N proxy variables denoted p_n. Lubotsky and Wittenberg (2006) show the researcher can run a regression including all of the proxy variables for I^*:

$$\tau = \tilde{\alpha} + \sum_{n=1}^{N} b_n p_n + \zeta + \sum_{m} X_m \tilde{\gamma_m} + \tilde{\epsilon}, \tag{A.27}$$

and then can aggregate up to a single coefficient of interest using:

$$b^{\rho} = \sum_{n=1}^{N} b_n \frac{cov(\tau,\, p_n)}{cov(\tau,\, p_1)}. \tag{A.28}$$

The expression is normalized by $cov(\tau,\, p_1)$ and this means that the procedure is an interpretation procedure where the coefficient is scaled such that a one unit increase in I^* will result in a one unit increase in p_1. In order to be able to compare this procedure to my other results, I select do the normalization such that it the results are comparable to the fraction of the population with access to four or more providers. Lubotsky and Wittenberg (2006) show that attenuation bias will be maximally reduced when estimating b^{ρ}.

The aggregation procedure above is equivalent to creating an index where

$$I_i^{\rho} = \frac{1}{b^{\rho}} \sum_{n=1}^{N} b_n p_n. \tag{A.29}$$

Having an index variable will allow me to estimate an equation where I^{ρ} is also interacted with other covariates as in the border design:

$$\tau_i = \hat{\alpha_0} + \hat{\beta}_0 I_i^{\rho} + \hat{\beta}_1 H_i + \hat{\beta}_2 I_i^{\rho} H_i + \sum_{k=1}^{K} \hat{\delta}_k (d_i)^k + \sum_{k=1}^{K} \hat{\varphi}_k H_i (d_i)^k + \sum_{k=1}^{K} \hat{\rho}_k I_i^{\rho} (d_i)^k + \sum_{k=1}^{K} \hat{\lambda}_k I_i^{\rho} H_i (d_i)^k + \zeta + \sum_{m} X_{im} \hat{\gamma_m} + \hat{\epsilon}. \tag{A.30}$$

As noted in Lubotsky and Wittenberg (2006), the procedure is not a license to include every variable the researcher may think is a proxy variable. Proxy variables can affect other control variables (Bollinger 2003) and "adding proxies that absorb the effects of covariates rather than proxying for the latent variable will be particularly damaging." For this reason, I exclude most of the type of technology variables (dsl, optical fiber, copper, etc.); I also exclude the fraction of people with access to 7 or more providers, 8 or more providers, or higher; I exclude the fraction of people with access to very high speeds of Internet access.

The summary statistics of these variables indicate that many of the scenarios are relatively uncommon in the data. Thus, I use only variables indicating the percent of consumers with access to wireless, any technology, 0 providers, one or more, ... six or more providers, download speeds greater than 1500k, and download speeds greater than 3000k. The results are robust to smaller subsets of the possible proxy variables.

A.14 Appendix References

Agrawal, David R. 2013. "Inter-federation Competition: Sales Taxation with Multiple Federations." WorkingPaper, University of Georgia.

Agrawal, David R., and William H. Hoyt. 2014. "State Tax Differentials, Cross-Border Commuting, and Commuting Times in Multi-State Metropolitan Areas." CESifo Working Paper.

Bruce, Donald, William F. Fox, and LeAnn Luna. 2014. "The Effect of State Tax Policies on Where E-Tailers Collect Sales Tax." University of Tennessee Working Paper.

Bollinger, Christopher R. 2003. "Measurement Error in Human Capital and the Black-White Wage Gap." *The Review of Economics and Statistics*, 85(3): 578-585.

Kanbur, Ravi, and Michael Keen. 1993. "Jeux Sans Frontieres: Tax Competition and Tax Coordination When Countries Differ in Size." *American Economic Review*, 83(4): 877-892.

Kelejian, Harry H., and Ingmar R. Prucha. 1998. "A Generalized 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.

Lubotsky, Darren, and Martin Wittenberg. 2006. "Interpretation of Regressions with Multiple Proxies." *The Review of Economics and Statistics*, 88(3): 549-569.

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