CESifo WORKING PAPERS

7867
2019

September 2019

Dynamics of Policy Adoption with State Dependence

David R. Agrawal, Gregory A. Trandel

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 7867*
Category 1: Public Finance

Dynamics of Policy Adoption with State Dependence

Abstract

We study the dynamics of policy diffusion when a first-moving jurisdiction that legalizes an activity reduces the probability of legalization in nearby later-acting jurisdictions. If a jurisdiction's firms can sell to neighboring residents, but if the good is competitively sold at every location, then policies converge: all jurisdictions legalize or all jurisdictions ban. If firms have some market power, and if the location of firms depends on the order of legalization, an early-adopting government may legalize, but an otherwise identical, but later-acting, neighboring government might not. This possible asymmetry is due to state dependence resulting from the initial distribution of firms following the first-mover's legalization. Empirically, counties that legalize the sale of fireworks first have more firework vendors just inside their border than counties that legalize later. Furthermore, counties have a longer duration to legalize fireworks if nearby counties have already adopted. State dependence resulting from a first-mover advantage contributes to the policy divergence of regulatory policies.

JEL-Codes: H700, K200, L500, R300, R500.

Keywords: dynamics, fiscal competition, state dependence, externalities, borders.

David R. Agrawal
University of Kentucky
Martin School of Public Policy &
Department of Economics
433 Patterson Office Tower
USA – Lexington, KY 40502-0027
dragrawal@uky.edu

Gregory A. Trandel
University of Georgia
Department of Economics
Brooks Hall
USA – Athens, GA 30602

September 2019
Following the death of Greg Trandel, the paper was completed by David Agrawal with the support of Trandel's family. I wish to thank the Trandel family for their support and I hope that the paper remains true to Greg's vision. We thank Adib Bagh, Makoto Hasegawa, Masayoshi Hayashi, William Hoyt, Ryo Ishida, Yulia (Paramonova) Kuchumova, Todd Nesbit, Maximilian von Ehrlich, David Wildasin and conference participants at the International Institute of Public Finance and the Southern Economic Association for helpful comments on an earlier version of this paper. The editor, Stephen Ross, and two anonymous referees greatly improved the paper. Cole Rakow provided excellent research assistance. Agrawal takes full responsibility for any errors contained in it.

1 Introduction

Current state and local policies are important determinants of the location of economic activity. However, states that differ in their policies today may have differed in their policies in the past. The persistence of these differences may be a result of reasons no longer relevant today. Thus, state dependence resulting from initial policy decisions made long ago may still play an important role today. This raises several questions. What is the role of dynamics in the policy process and how does the *historical* order of adoption affect the equilibrium location of firms and policies *today*? Does the initial order of legalization (and thus historical policy differentials) between jurisdictions influence the location of firms and thus have persistent effects on the policy choices of governments today via state dependence?[1]

As an example, Holmes (1998) considers how pro-business and anti-business policies influence firm location decisions. Holmes (1998) shows, theoretically and empirically, that when adjacent states adopt different policies, a sharp discontinuity in manufacturing activity – with extra activity on the pro-business side – occurs at the border. However, the theoretical model in Holmes (1998) is static: it considers "bunching" of firms on one side of the border as a result of existing business policies. Although current day policies are used in the model, in his conclusion, Holmes (1998) acknowledges the importance of past policies on current day firm locations: "even if differences [in the location of businesses] at the border are due to state policies, it may be policies from long ago that have nothing to do with a state's current policies toward business."

In this paper, we consider several modification to Holmes (1998) in order to emphasize the role of past policies. We consider the role of state dependence resulting from initial firm location decisions following the adoption of policies in early-moving states as an explanation of why policy discontinuities persist at some state borders and not others. To do this, we introduce dynamics into a model of fiscal competition with state dependence and derive conditions under which identical jurisdictions that differ *only in the order of legalization* may not adopt the same policies. Critically, we show that the initial order of adoption, which may no longer be relevant for policy-making today, may discourage the entry of firms near the border within the later-moving jurisdiction's boundary. If these effects are strong enough, a later-acting jurisdiction may not adopt the same policy as its neighbor because the first-moving jurisdiction can obtain producer surplus from non-residents and can tax export while the later-moving jurisdiction cannot. Adding dynamics and state dependence requires two

[1]By state dependence, we mean that past choices influence the probability of future outcomes, making some outcomes more likely and others less likely (Heckman 1981). Path dependence can be defined as a sequence of state dependent outcomes converging on a particular equilibrium (David 1985). See Whatley (1990) for definitions.

major modifications to Holmes (1998). First, we endogenize the policies set by jurisdictions by allowing for fiscal competition. Second, in Holmes (1998), the costs of one firm does not depend upon the location of another firm. Our model is closer to Hotelling (1929); in such a setting, due to firms competing for shoppers, not all firms may want to locate in a state with favorable policies because they may forgo a larger market where they have fewer nearby spatial competitors. Empirically, unlike Holmes, who studies firms near state borders that differ in their policy, we look at firms near county borders that have the *same* policy *today*, but had *different* policies in the *past*. As a result, we test a dynamic model, rather than Holmes' static model, which results from the incorporation of state dependence in a dynamic context. In our empirical setting, current day policies are the same and therefore difference out, allowing us to identify the effect of one jurisdiction legalizing first.

To study policy dynamics with state dependence, we focus on a particular policy – whether a jurisdiction legalizes an activity or not – and then test the theory using data on regulatory policies. Why do nearby jurisdictions that are otherwise *similar* set *different* regulatory policies concerning the legalization or prohibition of a particular activity? As an example, firework stores often advertise their products on billboards in states where fireworks are illegal. These advertisements are usually for stores that set up shop strategically just over the border to sell fireworks; the goal is to attract customers to buy types of fireworks that are legally sold where the firm is located, but are not available in the neighboring state.[2] There are many instances where U.S. states sharing a border impose very different regulations on the sale of fireworks. In many cases, the differing state policies — including those between states that don't appear to have widely dissimilar political environments — have been in existence for many years. Regulatory differences across states are not unique to fireworks, however, some regulatory policies converge across states. For example, the spatial correlation in regulatory policies such as fireworks and casinos is non-existent or negative, but is positive for lottery sales.[3]

To answer the question of why nearby jurisdictions that are otherwise similar may set different regulatory policies, we model regulatory policy choices in a spatial setting in the spirit of Braid (1987) and Kanbur and Keen (1993). Various authors have already used theoretical spatial models to consider why the policy choices (concerning, say, continuous variables such as tax rates) of neighboring jurisdictions might differ. Abstracting away from possible differences in political considerations, these authors have explained varying policies

[2]Figure A.1 shows an example of firework vendors just south of the North/South Carolina border.

[3]For a survey of lottery and casino gambling, see Kearney (2005*a*). Knight and Schiff (2012) show that lottery sales per capita are more responsive to prices in small states with densely populated borders while Kearney (2005*b*) focuses on consumer behavior. Tosun and Skidmore (2004) show that border competition is a determinant of lottery sales and Brown and Rork (2005) show evidence of spatial competition in prizes.

by looking at the incentives that result, for instance, from jurisdictions having differing sizes or population densities. If the jurisdictions in these models have identical parameter values, the analysis predicts the policy choices of the jurisdictions will match (be the same) and the equilibrium will be symmetric. Although precise comparisons to the prior literature are reserved for section 2, we summarize the four defining features of our model:

1. Governments endogenously set policies by engaging in fiscal competition;
2. Firms interact strategically when setting prices;
3. Policies are set in a dynamic setting with otherwise identical jurisdictions; and
4. Past policy adoptions influence the probability of future outcomes.

In a model with a negative consumption externality, an early adopting jurisdiction, and firms with some permanence of location and market power, we show that *otherwise-identical* neighboring jurisdictions can make differing policy choices that persist — some jurisdictions legalize, while others don't. Policies do not converge. On the other hand, policies will match and be consistent when the good is sold competitively at many locations. To that end, the model *may* explain why states have converged for some regulatory policies and not others.

Our model has two major differences relative to the tax competition literature. First, the model doesn't consider governments choosing a continuous variable like a tax rate. Rather, it analyzes governments that make a discrete choice about whether or not to legalize a certain activity (fireworks) that has a negative externality in the jurisdiction where the consumer resides (noise). This binary decision may result in large and salient policy discontinuities at borders, which create dramatic incentives for firms to locate just inside the border in order to capture customers from the neighboring jurisdiction where sale of the commodity is illegal. More specifically, this paper considers a case in which the externality is large enough that the net welfare effect of use by a jurisdiction's own residents is negative. In spite of this harmful impact, though, a first-moving welfare-maximizing government might legalize if it has the potential to collect tax revenue (and, if its firms can earn producer surplus) from border-crossing residents of another jurisdiction, while not experiencing the (full) externality created by their consumption. That externality is instead felt by the home jurisdiction of the border-crossers.

Second, the paper considers jurisdictions with identical ex ante characteristics, but allows one jurisdiction to — for whatever reason, possibly low costs of policy experimentation — make its policy choice first. As noted previously, this first-mover status affects ultimate outcomes *only if* firms possessing market power arrange themselves (with some permanence) in a manner that depends on early policy choices. In particular, with infinitely many firms in perfect competition, regulatory competition results in all jurisdictions legalizing, but total (world) welfare is lower than if no jurisdictions had legalized; this is because the binary policy-setting game between decentralized jurisdictions has the characteristics of a prisoners' dilemma. In contrast, policies need not converge if firms have market power. With market power, if a first-moving jurisdictional legalizes, a later-moving jurisdiction may not legalize due to the market power of the seller located just within the first-mover's borders. Even if the late moving government eventually legalizes, no seller of similar size may want to locate just inside that jurisdiction's borders due to competition from the rival firm already established in the other state. This results in a late moving jurisdiction continuing to ban the commodity because it obtains no producer surplus and tax revenue from non-residents because cross-border shopping flows into the first-moving jurisdiction due to the favorable location of its firms created by early legalization.

We then show that the theoretical model is consistent with two empirical questions for the albeit selective example of firework policy decentralization to the counties in New York. Will less firework establishments locate in late adopting counties? Is a county less likely to adopt fireworks if a nearby county has already adopted?

To answer these questions, we consider the state of New York, which allows counties to determine if they want to legalize or ban fireworks. We collect data on the address of all vendors with firework licenses and calculate the shortest driving time to the nearest county border road crossing. Knowing the exact date of legalization, we classify each store as being in an "early" adopting county if the county it is in legalized prior to the nearest neighboring county. Using the McCrary (2008) test, which identifies discontinuities in densities, we show that more firms are near the border in early adopting counties than in late adopting counties, despite the fact that population near the border in early adopting counties is lower. The extra firms on the side of the early moving jurisdiction, combined with the persistence of this distribution even after a neighboring county legalizes, demonstrates the joint importance of borders and the timing of initial policy adoption. The empirical analysis suggests that first mover jurisdictions have an advantage at attracting firms and that this advantage persists even if the policy advantage becomes obsolete. Then, we turn to a hazard model to explain legalization based on the number of neighboring jurisdictions that have adopted the policy already. Overall, we conclude that having an additional neighbor legalize previously correlates with a delay in the time to adoption. A county that has one more neighbor legalizing previously proceeds to adopt at a rate approximately 50% slower than a county with one less neighbor adopting. This combined with the extra mass of firms inside the firs-mover's borders suggests that early adoption of a policy creates state dependence that delays similar jurisdictions from legalizing.

In addition to the fireworks example presented above, this model may be relevant for activities such as casino gambling, as long as (i) some of the negative externalities – perhaps pathologies associated with addictive behavior (Guryan and Kearney 2010) affecting family members – occur in the gambler's jurisdiction of residence and not in the casino, and (ii) a jurisdiction that becomes a gambling center before its neighbors is likely to remain a prominent tourist attraction. Our model may also apply to marijuana legalization, gun shops, and racetracks. Other examples abound so long as the externality is partially born in the place of residence and path dependence persists in the location that is the early adopter.

2 Contributions

We make contributions to three distinct literatures: policy and firm sorting, state and path dependence, and fiscal competition. We discuss each of these in turn.

A large literature attempts to study firm sorting across jurisdictions based on taxes or regulations. Some of this literature considers "local" movements around borders, while other parts of the literature consider "global" movements of firms across states.[4] As discussed above, Holmes (1998) considers firms that sort across a border because states have pro/anti-business policies broadly defined Although he uses a classification based on right to work laws as his designation, these regulatory policies may also arise from other policies such as environmental regulations. The environmental policy literature has reached varying conclusions regarding the effect of regulatory policies on new firm births and firm location decisions (Jeppesen and List 2002; List et al. 2003; Jaffe et al. 1995; Levinson 1996). Additionally, tax policy – corporate, individual income tax, and sales tax policy – have been shown theoretically to influence the location of economic activity at state borders (Agrawal and Hoyt 2018; Rohlin, Rosenthal and Ross 2014; Kanbur and Keen 1993).[5] In the Kanbur and Keen (1993) class of sales tax competition models, firms are simply responding to consumer demand, and although these models focus on consumer behavior, firms must implicitly be selling the goods where the consumer buys. Of course, sorting across borders need not only apply to state borders: local policies (Agrawal 2015) and borders within municipalities due to tract-specific policies may also influence firm locations (Busso et al. 2013).[6] Unlike many of these models, in our model, one explanation for observed sorting is the order of legalization.

[4] We focus on the literature that considers firm sorting at boundaries. Papers considering the location of firms across states and countries include Fajgelbaum et al. (2019) and Giroud and Rauh (2019).

[5] Empirically, see: the effect of sales/excise tax differentials (Fox 1986; Thompson and Rohlin 2012; Rohlin and Thomson 2018; Hurtado 2018), business tax differences (Rathelot and Sillard 2008; Chirinko and Wilson 2008), place-based policies (Hanson and Rohlin 2011), minimum wage differences (Dube, Lester and Reich 2010), personal income tax differentials (Young et al. 2016) or spending differences (Peltzman 2016).

[6] Bidding for firms, that is, firm-specific policies targeting large firms also influence the distribution of economic activity (Black and Hoyt 1989; Slattery 2019).

With respect to regulation and timing, Hansen, Miller and Weber (2018) focus on marijuana and provide evidence that jurisdictions may engage in a "race to legalize;" in their model, a later-moving government can recapture some of its consumers, but whether the later-moving jurisdiction remains at a disadvantage in terms of tax revenue and potentially firm profits remains an open question. More generally, in the environmental policy literature on regulatory races, dynamics are important but these models often predict policy convergence.[7] In our model, policies may also diverge.

Although government institutions may evolve slowly (North 1991), past choices may also affect the probability of policy adoption or reform. Recently, a large literature in urban economics has focused on the role of path dependence. Rosenthal and Ross (2015) survey this literature and delineate three factors of metropolitan persistence: self-reinforcing effects of agglomeration, natural advantages, and the effects of culture. The first of these is the most important in for this paper. Rosenthal and Ross (2015) give the example of a featureless plain where economic activity first appears at one site as opposed to an alternative site. Then, agglomeration may foster productivity gains that make the first chosen site have an advantage over the later chosen site; these agglomerations may be self-reinforcing. As such, the first site will be forever larger than the second even though the reason explaining why one site developed first is no longer relevant. An important example is Bleakley and Lin (2012) and Bleakley and Lin (2015) who document that portage paths were important for the development of cities, and despite their lack of relevance today, populations remain clustered at these points. Such path dependence may also arise from historical policy decisions that may alter the distribution of firms, which may in turn explain economic density well into the future. For example, von Ehrlich and Seidel (2018) shows that temporary place based policies in Germany result in much higher firm densities well after those policies are terminated.[8] The state dependence in Bleakley and Lin (2012) and von Ehrlich and Seidel (2018) are a result of location fundamentals or one jurisdiction receiving federal transfers in the past. In our setting, policy non-convergence will result from one jurisdiction acting first.

The spatial tax competition literature — examples in addition to those referenced previously include Trandel (1992), Braid (1993), Trandel (1994), Haufler (1996), Ohsawa (1999),

[7] Carruthers and Lamoreaux (2016) write about regulatory races: "The theory's [of regulatory arbitrage] core prediction is that regulatory policies in open economies will tend to converge over time. The basic cause of this convergence is the mobility of capital, but the actions of political decision-makers matter as well. A regulatory race thus requires (1) an initial difference in regulatory rules across jurisdictions, (2) the credible threat that firms' locational decisions will be affected by that difference, and (3) a move by jurisdictions with more burdensome rules to change their regulations in response to the threat."

[8] See also Kline and Moretti (2014), Freedman (2017), Lichter, Löffler and Siegloch (2018), and Brooks and Lutz (2019) for other examples of policies with persistent effects.

Nielsen (2001), Behrens et al. (2009), Agrawal (2015), and Agrawal (2016) [9] — often focuses on differences in size, preferences or densities.[10] This literature is primarily static and does not feature state dependence. In an exception, Wang (1999) considers a model in which a Stackelberg-leader jurisdiction is the first of two jurisdictions — with different population densities — to choose its tax policy. The focus in Wang (1999) is on comparing the Nash and Stackelberg outcomes; state dependence cannot arise because infinitely many firms are assumed to be able to freely enter at any point on the line segment. Non-spatial tax competition models have focused on leadership and dynamics more than the spatial literature. Several recent papers have endogenized leadership (Kempf and Rota-Graziosi 2010; Eichner 2014; Hindriks and Nishimura 2015) and conclude that depending on the nature of the asymmetry, leadership by the large or small jurisdiction can be the risk dominant equilibrium.[11] For example, in Kempf and Rota-Graziosi (2010), both jurisdictions may have a second-mover advantage if they are ex ante identical because they benefit from capital flows from the leader; thus, the model does not feature state dependence in the tax base. Unlike models of endogenous leadership, all jurisdictions are ex ante identical and moving first is the only source of asymmetry in our model, which allows us to focus uniquely on its effects.

Our paper also has some parallels to capital tax competition with stock effects or agglomeration, which also relates to the mobility of capital literature. For example, Wildasin (2003) demonstrates in a dynamic context, that "old" capital may be very difficult to relocate, but new capital may be highly tax sensitive. Thus, a tax increase will result in a large mechanical increase in tax revenue for jurisdictions with a large stock of "old" capital. Thus, policy differentials may persist if governments with large stocks of "old" capital maintain high tax rates while capital poor countries cut their taxes to attract "new" capital (see Janeba and Peters 1999; Marceau, Mongrain and Wilson 2010). "Old" capital might be viewed as the initial distribution of firms in our model.

3 Stylized Fact: Some Policies Do Not Match

Do regulatory policies exhibit positive, negative, or no spatial correlation? Or does it depend on the type of regulatory policy?

Before turning to the theory, we show that several binary policies do not "match." When

[9] For surveys of fiscal competition, see Keen and Konrad (2013), Wildasin (2006) and Wilson (1999).

[10] There is also an existing empirical literature (see Brueckner 2003) that shows, controlling for other influences, that differences in policy choices concerning taxes and spending tend to lessen over time. Examples include Case, Hines and Rosen (1993), Brueckner and Saavedra (2001), Brueckner (1998) and Brueckner (2000). In the case of discrete policy choices, some papers have shown that states are more likely to adopt a lottery if neighboring states have already done so (Alm, McKee and Skidmore 1993; Caudill et al. 1995).

[11] Altshuler and Goodspeed (2014) empirically find evidence that some jurisdictions are leaders.

policies do not match, similar nearby states set the opposite policy (one legalizes and one bans). Matching policies occur when both jurisdictions legalize or both jurisdictions ban.[12] Figure 1 shows the regulatory policies of counties in two states that allow these policies to be set at the local level – where dramatic differences in firework preferences may be unlikely. The figure indicates no strong – or perhaps a negative – relationship between the policies of one county and the policies of neighboring counties. The lack of an obvious spatial correlation in firework regulatory policies also persists at the state level. Figure 2 shows fireworks laws at the state level; we define a state as banning fireworks if there is a complete ban on aerial fireworks sales to residents and nonresidents. Seven states ban aerial fireworks to their residents but allow for sales to nonresidents, suggesting the presence of a negative externality at the destination rather than the origin. The appendix (figures A.2, A.3, and A.4) show maps for lottery, casino, and marijuana policy.[13]

While informative, the maps do not provide a formal test of the degree to which these regulatory policies do not "match." Thus, we construct Moran's I for variable x as

$$I = \frac{N}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij}(x_i - \overline{x})(x_j - \overline{x})}{\sum_i (x_i - \overline{x})^2} \quad (1)$$

where N is the number of jurisdictions, w_{ij} is the spatial weight between jurisdiction i and j and $\overline{x}$ is the mean of x. We implement the test using contiguity between jurisdictions as the basis for the spatial weights. In the absence of spatial correlation, then we expect Moran's I to take on a value of $-1/(N-1)$. Implementing this for each of our binary regulatory variables, Table 1 shows that no spatial correlation is evident for firework policies, commercial casino policy and marijuana policy. Positive correlation exists for racetrack gambling and lotteries. In New York and Washington state, firework policy is decentralized to counties and localities. At the county level, negative spatial correlation in firework policy is evident, which suggests that counties that legalize are neighbored by counties that ban. This negative and null relationship is not usually found in standard fiscal variables where policy choices are continuous (for example, Brueckner 1998; Figlio, Koplin and Reid 1999; Brueckner 2000; Eugster and Parchet 2019). The apparent lack of matching for some regulatory policies raises the question of why these policies might differ across otherwise similar jurisdictions.

Moran's I is designed to motivate the theoretical model to follow and not to explain policy differences. Although there could be many reasons explaining these policy choices (fiscal competition, yardstick competition, fiscal spillovers, learning), we focus on dynamic

[12] In our terminology, matching does not relate to a political-economy model where policies are aligned with the preferences of the electorate. Rather, matching refers to the similarity of policies across space.

[13] The policy is whether the state has a lottery. The choice of games (such as Powerball represents a competition effect). As of 2018, all states that did not adopt Powerball, also do not have state lotteries.

policymaking and policy experimentation as a reason for non-convergence. We proceed by constructing a model of fiscal competition – broadly defined as in Wildasin (2006).

4 Model with Competitive Firms

We construct a theoretical model of regulatory competition for a good that imposes a negative externality in the place of consumption. As noted above, one example is fireworks. Consumers are arranged with uniform density equal to unity around a circle (Salop 1979; Trandel 1992; Agrawal 2015) with circumference of $2n$.[14] The circle is divided into two equally sized jurisdictions of length n. Each jurisdiction decides whether or not to allow the sale of a certain good. If legalized, the sales of that good within a jurisdiction are subject to the same tax t that is imposed on all other transactions. Because the tax isn't unique to the studied good and because we focus on goods that are only a very small fraction of the total tax base, its level is assumed exogenous — and identical in all jurisdictions — throughout this paper.[15] For this reason, jurisdictions will not compete by setting their sales tax rates and will only engage in regulatory competition. Of course, jurisdictions may have access to multiple policy instruments (Braid 2013; Bucovetsky and Wilson 1991) that may result in convexification of the discrete policy choice (e.g., licensing fees, policing intensity or enacting various degrees of zoning regulations). If such other instruments target many different products, similar to the tax, we can argue they are exogenous. If instruments are specifically designed to target the commodity in question, then our model is a useful representation of the first stage where jurisdictions decide to legalize or not.

As in the standard spatial model, every consumer is interested in buying only a single unit of the good, and will do so if the effective price (monetary price plus travel expense) to him or her is less than some exogenous maximum value v, which is the same for all consumers.

The use of the good creates a uniform negative externality e; this externality affects the jurisdiction in which the consumer resides, not the one in which he or she buys the good. In the fireworks illustration noted above, the externality can capture either the annoyance to, or the possible endangerment of, neighbors, or the extent to which a jurisdiction's taxpayers are responsible for some of the medical/policing expenses that result from use of the product. In other examples, this externality may include increased crime or traffic fatalities.

[14]The use of a circle rather than a line will become critical in the next section. If using a line segment, the finite nature of the line segment makes it so firms at the ends face competition only from one neighbor. Further, the model does not feature agglomeration forces (Agrawal and Wildasin 2017), so as to isolate the effect of first mover advantages.

[15]For example, the state of New York only obtains tax revenue on fireworks from the general sales tax. This is common in most, but not all states; some states levy an excise tax specifically on fireworks.

In this section of the paper, it is assumed that firms are perfectly competitive, firms have no fixed costs of supplying the product, and can freely enter anywhere in a legalizing jurisdiction. Firms may locate at any point on the circle, so that there are infinitely many firms; it is also assumed that any quantity of customers can be accommodated at any point along the circle. In combination, these assumptions imply that the good will always be sold at its constant and uniform marginal cost c; profits and producer surplus are entirely competed away. Consumers can buy the good from any firm, but the negative externality accrues to the place of consumption (residence).[16] When buying the good from a firm located at the point where one lives, no transport costs are incurred. Thus, it is assumed that consumers' value for the good exceeds its tax-inclusive price so that the consumer will buy the good if $v > c + t$. If a jurisdiction doesn't legalize selling the good, people interested in obtaining it must travel to the nearest location (i.e., just across the border) in a neighboring jurisdiction that has legalized. Traveling to (and back from) the purchasing location requires any buyer to pay a travel cost equal to k per unit distance.

Governments select their regulatory policy in order to maximize welfare. Welfare is the sum of producer surplus, consumer surplus, and tax revenue net of any costs of policy experimentation; we assume the government values all of these equally in the welfare function. Although all jurisdictions are in ex ante identical situations, political and bureaucratic characteristics that differ among the jurisdictions may allow one to act before the other. Certain political leaders might also be willing to be more "experimental" or risky in their policy choices.[17] This difference in policy experimentation is incorporated in the model by assuming that the two jurisdictions $j = \{F, S\}$ have a (welfare) cost ψ_j of policy experimentation (e.g., worry about uncertainty of the policy's effect, perhaps over the size of the externality). We assume one jurisdiction has $\psi_j = 0$ regardless of if it moves first or not. The other jurisdiction has ψ_j such that the welfare cost is infinite if it move first, but once its neighbor has legalized, this cost is zero because the government may have learned about the policy. Then, it is easy to show that the jurisdiction with no cost will always move first if the social welfare of legalizing is positive. Without loss of generality suppose that the jurisdiction with no first moving costs of legalizing when its neighbors have not legalized is $j = F$ (Fast /

[16] In our model we assume that *selling* the good can be legal or illegal, but *using* the good is (effectively) legal. For fireworks, both the sale and use of fireworks are often both simultaneously illegal or legal. The sale of fireworks can easily be enforced but the use of fireworks is more difficulty to enforce. Thus, our model's assumption regarding legal use is based on the fact that although use is technically illegal, the difficulty of enforcing this makes it *de fact*o legal, which results in the externality being realized at home.

[17] The fiscal federalism system in the United States often calls states as "laboratories" for the federal government. Some state may – for whatever reason – propose a new policy innovation not yet passed in any other states. Reinganum (1981) models the decision of a firm to adopt a technology and shows that it is "never a Nash equilibrium for identical firms to bring the new technology on line at the same date."

First) and has borders that are at 0 and some positive integer n (so that its population is n). Jurisdiction $j = S$ (Slow / Second) has high costs of moving first. Figure 3 shows the geography of these two jurisdictions.[18]

Before proceeding, we fully specify the sequence of events.

1. Governments decide who, if any, will move first.
2. One government acts as the first-mover and chooses its regulatory policy.
3. If the first mover legalizes, firms locate at various points in the jurisdiction.
4. The neighboring jurisdiction chooses its policy.
5. If the late mover legalizes, firms locate in the second moving jurisdiction.

In the case of infinitely many perfectly competitive firms, firms will enter at all points on the line segment where the sale of the good is legal.

To summarize, if a resident buys the good from a store located where she lives, she obtains consumer surplus of $v - c - t$, that consumption inflicts an externality e on other residents, but provides tax revenue of t. With respect to welfare, when a jurisdiction decides whether or not to legalize a good, the potential net benefit from consumption by one of its residents is $v - c - e$ (where consumer surplus and tax revenue are assumed to be valued equally). Because this paper considers situations in which a welfare-maximizing government might or might not legalize, it is assumed that $v - c - e < 0$. If borders are closed (so that legalization means selling only to a jurisdiction's own residents – and selling to all of them), a welfare-maximizing government will have never choose to legalize and have a total welfare equal to 0.

When borders can be crossed, however, a jurisdiction may gain tax revenue — and not experience any externality cost — from sales made to residents of a neighboring non-legalizing jurisdiction. Residents of such a location will travel into a legalizing jurisdiction as long as the private benefit from consumption exceeds the tax-inclusive, travel-inclusive cost.

Under these assumptions, the fast jurisdiction will legalize first if its welfare of legalizing is positive. Following legalization, under the perfect competition assumption, firms will locate at all points within the jurisdiction. At the 0 border, some consumers will travel into the jurisdiction to purchase the good; the last consumer willing to do this is the one located at $\bar{b} < 0$, which is defined by $v - c - t - \left|\bar{b}\right| k = 0$. Therefore, $\left|\bar{b}\right| = (v - c - t)/k$. The legalizing jurisdiction will gain tax revenue from border-crossing consumers, but will not accrue the externality (nor the consumer value) associated with the usage of the good. By symmetry, a corresponding formula applies at the n integer border. Thus, when the jurisdiction has

[18] For ease of notation, in jurisdiction S, it will be easiest if we refer to a person living n units from point zero or a firm at n as being at point $-n$.

legalized, therefore, choosing to legalize brings in $2t\left|\bar{b}\right|$ in tax revenue from non-residents. We assume that $n > 2\left|\bar{b}\right|$ such that not all residents cross-border shop.

Depending on the relative values of the various parameters, a jurisdiction that would not legalize in the absence of border-crossers may do so in their presence. Specifically, this will occur if the net benefit from consumption by all of its residents plus the tax revenue from border-crossers is greater than the welfare from not legalizing (i.e., zero). The first-mover legalizes if:

$$W_F^L \equiv n(v-c-e) + 2t\left|\bar{b}\right| > 0 \equiv W_F^N, \tag{2}$$

where we let subscripts $\{F, S\}$ on welfare W denote the jurisdiction and superscripts $\{L, N\}$ denote welfare under legalization and no legalization, respectively.

Consider now the case of the slow jurisdiction that hasn't legalized, but has seen its neighbor so do so. Such a jurisdiction has some of its residents traveling across each border to buy the good. In the absence of legalizing, the jurisdiction's welfare is the sum of three elements. First, ignoring their cost of traveling, the $2\left|\bar{b}\right|$ people who buy across a border pay a monetary price of $c+t$; each therefore experience $v-c-t$ personal net benefit. The cost of travel for a border-crosser located at $\tilde{b}$ is $\tilde{b}k$, where $\tilde{b}$ ranges from 0 to $(v-c-t)/k$. The total travel cost of both sets of border-crossers is thus $2\left|\bar{b}\right|(1/2)(v-c-t) = (v-c-t)^2/k$. Finally, border-crossers impose a total externality cost of $2\left|\bar{b}\right|e$ on their home jurisdiction, but contribute no tax revenue at home. If the slow jurisdiction does non-legalize, it thus experiences welfare equal to

$$\begin{aligned} W_S^N &\equiv 2\left|\bar{b}\right|(v-c-t) - \tfrac{(v-c-t)^2}{k} - 2\left|\bar{b}\right|e \\ &= 2\left|\bar{b}\right|(v-c-t-e) - \tfrac{(v-c-t)^2}{k} \end{aligned} \tag{3}$$

If the slow jurisdiction instead legalize, no border-crossing would exist because firms would locate at all points, ψ_S is zero given the first mover has legalized, and welfare would be consumer surplus plus tax revenue, $n(v-c-e)$. Given that the jurisdiction's neighbor legalized, legalizing will raise this jurisdiction's welfare only if

$$\begin{aligned} W_S^L = n(v-c-e) &> 2\left|\bar{b}\right|(v-c-t-e) - \tfrac{(v-c-t)^2}{k} = W_S^N \\ &\Leftrightarrow \\ (n-2\left|\bar{b}\right|)(v-c-e) + 2t\left|\bar{b}\right| + \tfrac{(v-c-t)^2}{k} &> 0 \end{aligned} \tag{4}$$

Remembering that $(v-c-e) < 0$, inequality (4) will hold whenever inequality (2) does. Equations (2) and (4) thus establish the result of this section.

Proposition 1. *When the good is sold at all locations because firms are perfectly competitive:*

(a) and when first-moving jurisdictions gain by legalizing, then later-moving jurisdictions will gain by legalizing. Regulatory policies match – all jurisdictions legalize.

(b) and when first-moving jurisdictions do not gain by legalizing, then no jurisdiction will legalize. Regulatory policies match – all jurisdictions ban.

Proof. Given that $v - c - e < 0$, all terms in equation (4) are positive except for $n(v - c - e)$. A first moving jurisdiction will gain from legalizing if equation (2) is positive, that is, if $2t\left|\bar{b}\right| > -n(v - c - e) > 0$. Thus, if equation (2) is positive, equation (4) must also be positive. If equation (2) is negative then it immediately follows no jurisdiction will legalize. □

In either case, equilibrium policy choices are match — all jurisdictions legalize or none do so. Importantly, in the all-legalize outcome, each jurisdiction experiences a lower welfare than it would have had none legalized. In other words, when a first-mover finds it beneficial to legalize and the good can be bought at all points, decisions about whether or not to legalize have the characteristics of a prisoners' dilemma. Regulatory competition results in all jurisdictions legalizing, but total (world) welfare is lower than if no jurisdictions had legalized.

Intuitively, all jurisdictions are identical with respect to their characteristics. The symmetry of jurisdictions and presence of firms everywhere means that the gains to legalization (or banning) are identical for all jurisdictions: they realize the same sized externality, obtain the same profits (none) from firms and raise the same tax revenues. Once the first-mover legalizes, then both jurisdictions have the same cost of policy experimentation. The prisoners' dilemma effect on social welfare arises because in equilibrium, if all jurisdictions legalize, then no cross-border shopping will occur in equilibrium because of symmetry. This entirely erodes any advantage from tax exporting, but a later moving jurisdiction will still legalize to avoid a case where they only incur the negative externality from border-crossers. Similarly, the symmetry of jurisdictions means that if one jurisdiction finds it against its interest to legalize then all other jurisdictions will also ban; this would still be the case even if the later-moving jurisdiction had a near-zero cost of policy experimentation. The only source of asymmetry (ψ_j) simply determines which jurisdiction may move first.

The key result is that under the standard perfect competition assumption of Kanbur and Keen (1993), with infinitely many firms located potentially everywhere, all jurisdictions adopt the same binary regulatory policy. Regulatory policies should be observed to have perfect (and positive) spatial correlation given the prisoners' dilemma nature of the game. However, as noted above in our discussion of stylized facts, some regulatory policies lack this spatial pattern (fireworks), which motivates the modifications below.

5 Model with Market Power

In this section, and following Braid (1987), we assume that firms are only able to locate at a finite number of fixed locations, perhaps due to location-specific fixed costs. Once a firm exists at such a location, it remains there, which is likely true in the short-run. [19] We can state this setup more formally as follows: (1) there are a fixed number of possible firm locations corresponding to each integer i and (2) each location has an indivisible plot of land owned by resident landowners, which is rented to a firm at R_i. In this case, when a government legalizes, a firm will be willing to bid for land up to a rent exactly equal to the profits that accrue to the firm.[20] Given that firms will compete à la Bertrand and the parcel is indivisible, in equilibrium, each integer can support only one firm.[21] Then, the timing of the game remains the same as in the prior section, but we clarify the implications of the presence of land rents on various stages of the game.

In stage 3 there is now free entry of firms, but only at the fixed integer points of commercial property. Firms bid for land and are willing to pay up to a rent that earns the firm zero profits. Thus, when the government legalizes, land rents rise at each integer until profits equal zero at each location.[22] Given the indivisibility of land, one firm will locate at each location i within the jurisdiction and no new firms have an incentive to enter. Under the assumptions above, in the third stage of the game, $n + 1$ firms enter at each integer in the first-moving jurisdiction. In particular, firms (denoted z_i) will locate inside the first legalizing jurisdiction starting a point 0 until point n — Firm 0 (z_0) at point 0, Firm z_1 at point 1, etc., until firm z_n at point n. Thus, two firms – Firm 0 and n – locate (effectively) on the relevant border, just inside the legalizing jurisdiction. The solid hash marks in Figure 4 illustrate this situation following the end of stage 3 of the game, where notice that one firm locates right on the commercial land located at each border. The dotted hash marks are the market borders, defined below, between firms (b_i); there is full coverage within a legalizing jurisdiction. Finally, in stage 5, given commercial land is only available at each

[19] For simplicity, there are a finite number of locations and firms cannot enter at any other points. This simplicity of the model might be replaced with endogenous firm location decisions, discussed below.

[20] Profits are not necessarily zero in models of spatial competition, even in the presence of entry. See Vickrey (1964), Eaton (1976), Eaton and Lipsey (1978), Capozza and Order (1980), and Braid (1987).

[21] The integers corresponding to the jurisdictional boundary always will have the commercial plot of land inside the first moving jurisdiction. Location specific sunk costs prevent landowners and firms from moving if a later moving jurisdiction legalizes. And Bertrand competition prevents there from being two firms located in different states at a given integer i because price would equal to marginal cost, implying losses.

[22] Although profits are competed down to zero, the profits accrued to the firm (total revenue minus total variable cost) remain positive and are equal to land rents. In our model, this total revenue minus total variable cost could be interpreted as profits accruing to the firm, land rents, or producer surplus. For simplicity, we will refer to it a producer surplus. Given this producer surplus acts as a transfer to resident landowners, it will appear in the welfare function unlike in the perfectly competitive case.

integer in the later-moving jurisdiction and firms will freely enter these points. Then, $n-1$ firms freely enter at each integer in the second moving jurisdiction. As point 0 and n already contain one firm and the parcel of commercial land is indivisible, no new firms will enter these points. Thus, the new firms will locate only at points -1 to $-(n-1)$ in the second moving jurisdiction. In this case, no firms in the neighboring jurisdiction will locate precisely at the border; location specific sunk costs prevent landowners and firms from moving and Bertrand competition prevents there from being two firms located in different states at a given point i.

Now we can solve the game. By the assumptions on ψ_j, if any jurisdiction legalizes first, it will be the fast one (F). As in the previous section, it is assumed that legalizing the sale of the good when borders cannot be crossed is welfare reducing. In this section, however, the net benefit of purchasing the good – even in the presence of closed borders – must include the effects of travel cost. Legalization with closed borders implies that consumers in a jurisdiction that has legalized must travel between a distance between 0 and 1/2 to get to the nearest firm; the average travel distance is therefore 1/4 unit, and the average travel cost is $k/4$. The assumption that legalization without border crossing is welfare reducing thus amounts to assuming that $v-c-e-k/4<0$.

Within a legalizing jurisdiction, consumer value v is assumed to always exceed the tax-inclusive, travel-cost inclusive cost. The market border b_1, for example, between the firms located at points 0 and 1 is the point at which a consumer is indifferent between buying at the two locations. In other words, the full price — monetary plus travel expense — of buying from those two firms must be equal:

$$\begin{aligned} p_0+t+kb_1 &= p_1+t+k(1-b_1) \\ b_1 &= \tfrac{1}{2}+\tfrac{p_1-p_0}{2k}. \end{aligned} \tag{5}$$

Other market borders are defined similarly. Then, Firm 1 — which has active firms located on both sides of it — sells to all customers located between b_1 and b_2, so its profit is

$$\begin{aligned} \Pi_1 &= (b_2-b_1)(p_1-c)-R_1 \\ &= \left(\tfrac{3}{2}+\tfrac{p_2-p_1}{2k}-\tfrac{1}{2}-\tfrac{p_1-p_0}{2k}\right)(p_1-c). \end{aligned} \tag{6}$$

Differentiating the above expression with respect to p_1 and setting the resulting expression equal to zero produces the first-order condition for profit maximization. When rearranged, this expression yields Firm 1's profit-maximizing reaction function for price:

$$p_1=\frac{c}{2}+\frac{k}{2}+\frac{p_0}{4}+\frac{p_2}{4}. \tag{7}$$

The reaction functions of other firms not on the jurisdiction border are defined similarly.

Firm 0 faces a different situation if the neighboring jurisdiction has not legalized. This firm faces no competing firm on one side, so its market border on that side is the point at which a consumer gets no surplus from traveling to and buying from Firm 0. This occurs at location b_0, where $v - t - p_0 - |b_0| k = 0$, which implies that this border is defined as

$$b_0 = -\frac{v - t - p_o}{k} < 0. \tag{8}$$

Firm 0 sells to all consumers located between b_0 and b_1,so its profit is

$$\begin{aligned} \Pi_0 &= (b_1 - b_0)(p_0 - c) - R_0 \\ &= \left(\tfrac{1}{2} + \tfrac{p_1 - p_0}{2k} + \tfrac{v - t - p_0}{k}\right)(p_0 - c). \end{aligned} \tag{9}$$

Differentiating, setting equal to zero, and rearranging produces Firm 0's reaction function:

$$p_0 = \frac{v - t}{3} + \frac{c}{2} + \frac{k}{6} + \frac{p_1}{6}. \tag{10}$$

By symmetry, this reaction function similarly holds for Firm n at the other border.

Similar to the result in Braid (1987), the reaction functions given above have a closed-form solution, which can be expressed as a weighted average of two potential profit-maximizing prices. The first potential price is the price that would hold in equilibrium if there was an active firm at every integer spot on the line. If all firms have reaction functions of the form given by (7), the market equilibrium price is $\hat{p} = c + k$.[23] The second potential price is the price that would be charged by a single profit-maximizing firm that is unconstrained by any neighboring firms (in other words, a monopoly firm), which is $\tilde{p} = (v - t + c)/2$.[24] Intuitively, the relevance of these two prices arises from the fact that Firm 0 effectively faces the monopoly situation on one side of its location, and faces the the filled line on the other side. In Braid (1987), firms are located starting at point zero on an infinite line segment, but because jurisdiction size is critical for us, we have a finite number of firms. Thus, in our setting, with a finite number of firms, the firm located at n may influence the pricing decision of firm 0 if n is small. However, as n becomes large, this influence becomes trivially small. If the jurisdictions were on circle with an infinite circumference ($n \to \infty$), the Appendix shows that the Nash equilibrium prices, when firms are located at all positive integer points $i \geq 0$,

[23] An in-jurisdiction consumer who travels the maximum distance to reach a firm incurs a travel cost of $k/2$. The derivation of the equilibrium price $\hat{p}$ requires a consumer value for the good v exceeds the travel-cost, tax-inclusive price $p + (k/2) + t$ for all buyers, which is $c + k + (k/2) + t$ for interior buyers.

[24] The monopolist maximizes $2 |b_0| (p - c) = 2(p - c)(v - t - p)/k$.

are:

$$p_i^* = \alpha_i \tilde{p} + (1 - \alpha_i)\hat{p}, \qquad \text{where} \quad \alpha_i = \frac{4}{4+\sqrt{3}}\left(2-\sqrt{3}\right)^i. \tag{11}$$

To gain intuition, figure A.5 shows an example of how prices changes across integer positions. A clear price gradient emerges where prices at the interior converge to $\hat{p}$. The price gradient will, of course, be sensitive to various parameter values.[25]

However, as we wish to study the role of jurisdiction sizes in a two jurisdiction framework, this solution will not explicitly characterize the pricing solution. As shown in the appendix, for a finite number of firms, an analytical solution requires an explicit formula for the inverse of a tridiagonal matrix of dimensionality of n by n. While such solutions can be obtained for a given n, as shown in Mallik (2001) and Kiliç (2008), the solution for provides no additional economic intuition beyond the approach we follow. First, notice from figure A.5 that prices at the center converge and are barely influenced by firm 0 even at small distances away from this firm. As will become apparent, because the government's problem values firm producer and consumer surplus equally, the only relevant prices necessary to solve the government's problem are p_0^* and p_n^*, which by symmetry must equal each other. Thus, to proceed, we assume that n is large enough[26] such that (11) gives the approximation

$$p_0^* = p_n^* \approx \left(\frac{4}{4+\sqrt{3}}\right)\tilde{p} + \left(\frac{\sqrt{3}}{4+\sqrt{3}}\right)\hat{p}. \tag{12}$$

As can be seen, this approximation simply relies on how close α_0 in (11) is to the true weight on the "monopoly" price given by the analytic solution for the finite n case. We show in figure 5 that this approximation holds even for relatively small values of n. In particular, notice that the analytic solution for six firms ($n = 5$) has an error of 0.0001 relative to the infinite solution. By 20 firms, the approximation is accurate to 10 decimal places. Thus, for n sufficiently large, p_0^* is well approximated by evaluating (11) at $i = 0$. Intuitively, this approximation holds because as n increases, the firm at point n has little influence on firm 0. To a first order, prices are determined by the direct effect of neighboring firms.

Again consider that jurisdictions don't make simultaneous decisions about legalization. If one jurisdiction legalizes first, it will be the fast one. Then, the firm found at the border between the two jurisdictions will locate at the commercial property just inside the legalizing jurisdiction. Crucially, under the sunk cost assumptions above, this firm would remain in the same jurisdiction even if the neighboring jurisdiction later legalized. In particular, if

[25] At some parameter values, the "monopoly" price can be lower than the "oligopoly" price. This outcome can occur because a monopoly faces more elastic demand than does a firm with neighboring sellers.

[26] Such an approach is common. For example, a continuum of types is often used as an approximation to make claims about a large, but finite number of types. The same is often true in Bertrand games.

a government legalizes later, firms will locate at every integer position, but will not enter at the border points (point 0 and n) because a firm is already located at that point and Bertrand competition would compete price down to marginal cost.

Consider the choices facing a welfare-maximizing government. Given the costs of policy experimentation ψ_j, the jurisdiction with low legalization costs will legalize first, if at all. Recall, if no jurisdiction legalize, a jurisdiction will experience welfare equal to 0. In contrast, if it does legalize, all its home consumers will purchase the good, as will some border-crossers from the neighboring jurisdiction. On sales to its own residents, summed over all jurisdiction residents, the first-moving jurisdiction's welfare is expressed as: consumer value minus firm price minus taxes minus travel costs of $n(v - p_i^* - t - \frac{k}{4})$ plus firm surplus (from residents of the jurisdiction) of $n(p_i^* - c)$ plus tax revenue from residents of nt minus externality costs of ne. In total, the expression equals $n(v - c - e - (k/4))$ which is invariant to the prices. On sales to the two sets of border-crossers, a jurisdiction's firms earn surplus and the government collects tax revenue such that the total impact is $2\,|b_0|\,(p_0 - c + t)$. Note that profits that accrue to the firm are simply the land rents, which act as a transfer from the firm to the resident landowner. If consumption by a jurisdiction's own residents has only a slightly negative externality, the gain from border-crossers might be large enough to make legalization welfare-improving. This situation arises if the welfare of the first-moving jurisdiction is positive:

$$W_F^L \equiv n\left(v - c - e - \frac{k}{4}\right) + 2\,|b_0|\,(p_0^* - c + t) > 0 \equiv W_F^N. \tag{13}$$

Now consider the slow jurisdiction, which never legalizes first. Some of its residents already cross the border, ($|b_0|$ in either direction), and in the aggregate they experience — temporarily ignoring travel cost — a benefit of $2\,|b_0|\,(v - p_0^* - t)$. Travel costs range from effectively 0 (for consumers located at the border) to $v - p_0^* - t$ (for consumers just at the margin of buying). The average travel cost for border-crossers is thus $(v - p_0^* - t)/2$, and aggregate travel cost is $|b_0|\,(v - p_0^* - t)$. The cross-border buyers also impose an aggregate externality cost of $2\,|b_0|\,e$ on their home jurisdiction. If the slow jurisdiction doesn't legalize, even though its neighbor has done so, it experiences welfare of

$$W_S^N \equiv |b_0|\,(v - p_0^* - t - 2e) < 0. \tag{14}$$

If the slow jurisdiction in this situation chooses to legalize, its n residents will experience $n(v - \hat{p} - t)$ personal benefit (ignoring travel cost), will create externality cost of ne, and will experience travel cost of $nk/4$. As this jurisdiction moved second, the firms at its borders are located *inside* neighboring jurisdictions, so that only $(n-1)$ firms enter. Those firms collect a total surplus of $(n-1)(\hat{p}-c)$ and total tax revenue of $(n-1)t$. Thus, the jurisdiction that is last to legalize experiences welfare of

$$W_S^L \equiv n\left(v - c - e - \frac{k}{4}\right) - (\hat{p} - c + t) < 0. \tag{15}$$

Given that the fast jurisdiction has legalized, the remaining one will experience higher welfare by not legalizing if

$$W_S^N \equiv |b_0|\,(v - p_0^* - t - 2e) > n\left(v - c - e - \frac{k}{4}\right) - (\hat{p} - c + t) \equiv W_S^L. \tag{16}$$

At any parameter values that satisfy both both expressions (13) and (16), there can be long-term differences in policy choices between the fast and slow jurisdictions. Jurisdictions that have low costs of policy experimentation may find that the opportunity to affect firm location, and to draw in firm surplus and tax revenue (but not externality costs) from some neighboring residents is attractive. In contrast, a jurisdiction that acts late because of high first-mover costs, and that — even if it legalizes — will still see some of its residents purchase elsewhere, may decide not to experience the negative welfare effect of legalizing.

For many combinations of parameter values, equations (13) and (16) – along with the other parameter restrictions introduced throughout the paper – are not simultaneously satisfied. For such parameter values, the prisoners' dilemma nature of the model illustrated in Section 2 also holds in this version of the model and policies match.

However, if equations (13) and (16) are satisfied, an outcome will emerge where policies do not match. Jurisdictions that can legalize early will find the opportunity to affect firm location attractive because firms will locate just inside the legalizing jurisdiction border. This location of the firms inside the border places the later acting jurisdiction at a disadvantage and the later acting jurisdiction will not legalize. In particular, although we have made the simplification that number of firms is given, the order of legalization results in the later-moving jurisdiction having less (i.e., no) firms right on its border due to the incentives from Bertrand competition. This creates a sort of state dependence that results in policies remaining different in the two jurisdictions. The historical order of adoption influences the distribution of firms, which in turn has persistent effects on the late moving jurisdiction's policy choice.

Proposition 2. *When the good is not sold at all locations*

(a) and expressions (13) and (16) hold: although the first-moving jurisdiction legalizes, a later moving jurisdiction will not legalize. Then, regulatory policies will not match and

jurisdictions adopt different policies despite being ex ante identical.

(b) and expression (13) holds, but (16) does not hold: then a first-moving jurisdiction gains by legalizing and a later-moving jurisdiction will also gain by legalizing. Regulatory policies match – all jurisdictions legalize.

(c) and expression (13) does not hold: then no jurisdiction will legalize. Regulatory policies match – all jurisdictions ban.

Proof. If equation (13) holds, a first moving jurisdiction will legalize. If equation (16) holds a later moving jurisdiction will not legalize. There do exist parameter values for which both equations will hold and all other constraints on the problem are satisfied. Parts (b) and (c) follow similarly. □

It is easy to verify that there do exist, however, parameter combinations for which inequalities (13) and (16) – and the other relevant parameter restrictions – hold. For example, a some-but-not-all-legalize outcome arises for the following parameter values: $v = 2.8$, $c = 1$, $k = 1$, $t = .2$, $e = 1.6$, and $n = 24$. These values imply that $p_0^* = 1.860$, $|b_0| = .740$, $W_F^L = 0.368 > 0$ and $W_S^N = -1.820 > -2.400 = W_S^L$. The presence of open borders means that a jurisdiction may get lower welfare than if borders were closed.

The intuition resembles that of oligopoly models with entry decisions. Suppose there are two potential firms and they incur certain fixed costs when entering a market. If one firm is already in the market, the other firm may not enter if operating profits are not sufficiently large to cover the fixed costs. If no firm is in the market, one firm will enter and choose the production level that can prevent the other firm from entering. Put differently, in models of oligopolistic competition there exists the possibility that entry decisions may not match. This intuition can be generalized to the model in this paper where a jurisdiction corresponds to a firm, discrete policy choices correspond to entry decisions, tax revenues correspond to firms' operating profits, and costs of negative consumption externality corresponds to entry costs. Hence, the intuition of the theoretical result is similar to the existing results obtained in the oligopoly (including models of strategic entry deterrence) literature. This similarity is a feature (not a bug) of the model because *a priori* it is not clear that the intuition of an outcome of a welfare maximizing government would mimic the intuition of a profit maximizing firm.

5.1 Determinants of Policy Convergence

We wish to show the effect of particular parameters on whether or not policies match or not. To this end, it is useful to show the comparative statics concerning the effect of various parameters on W_F^L and on $\Delta_S \equiv W_S^N - W_S^L$. Equation (13) and (16) are more likely to hold if W_F^L increases and Δ_S increases. Plugging in all equilibrium expressions, it can be show with respect to jurisdiction size that:

$$\frac{\partial \Delta_S}{\partial n} = v - c - e - \frac{k}{4} = -\frac{\partial W_F^L}{\partial n}. \tag{17}$$

Given the opposite sign of $\partial\Delta_S/\partial n$ and $\partial W_F^L/\partial n$, this implies that it is more likely equation (13) and (16) simultaneously hold for intermediate values of n. Figure 6 shows that for the given sets of parameter values, an intermediate n results in policies not matching. Intuitively, if jurisdictions are "small", a late-moving jurisdiction is already experiencing the negative externality from enough of its population that it gains by legalizing. In contrast, if jurisdictions are "large", the tax revenue from border-crossers is relatively small to entice any government into legalizing.

Conducting a similar exercise, we can see that

$$\frac{\partial W_F^L}{\partial e} = -n < 0 \tag{18}$$

but that

$$\frac{\partial \Delta_S}{\partial e} = -(1+\sqrt{3})\frac{v-c-k-t}{k} - \alpha_0 k + n \lesseqgtr 0 \tag{19}$$

where $v - c - k - t = v - \hat{p} - t > v - \hat{p} - t - k/2 > 0$ because it is assumed everyone buys one unit of the good if legalized. Then, the sign of the above equation is ambiguous but we can see that it will be be positive if n is sufficiently large. The fact that $\partial W_F^L/\partial e$ and $\partial\Delta_S/\partial e$ are opposite signs for large n suggests that the two expressions will only hold for intermediate values of e. However, n must not be too large because at values of large n the slopes of Δ_S and W_F^L become very steep. Figure 7 shows values of e that satisfy both equations as n varies; higher n implies that the region where policies do not match require a smaller e but the region substantially shrinks in size. When e is high, a first moving jurisdiction will not legalize but when e is low, the later moving jurisdiction will legalize. Intuitively, when e is very large a first-moving jurisdiction will not want to legalize but when e is small a later-moving jurisdiction will want to legalize. Thus, only for intermediate e will a first-mover legalize and a later moving jurisdiction not legalize.

Finally, we can show the effect of taxes. Comparative statics with respect to t yield:

$$\frac{\partial W_F^L}{\partial t} = \frac{2(2+\sqrt{3})\left((v-c-k-t) - k - t - \frac{4}{\sqrt{3}}t\right)}{(4+\sqrt{3})^2 k} \lesseqgtr 0 \tag{20}$$

and

$$\frac{\partial \Delta_S}{\partial t} = \frac{-(14+8\sqrt{3})(v-c-e-t-k)+(11+4\sqrt{3})k+(8+4\sqrt{3})e}{(4+\sqrt{3})^2 k} > 0. \qquad (21)$$

As above, note that $v-c-k-t>0$ so the first expression is more likely to be negative the higher is k and t. The second expression is unambiguously positive because $v-c-e-t-k < v-c-e-(k/4) < 0$. The implication, as shown in Figure 8, is that higher taxes will result in policies that are more likely to match because the benefit (added revenue) from legalizing is likely to be small because few residents of the neighboring jurisdiction are willing to cross-border shop into a very high-tax jurisdiction.

Of course, it is possible to state a more formal proposition containing precise cutoff values based on size, taxes or the size of the externalities, but given these cutoff rules for when policies do not match are a (non-linear) function of the other parameters, little additional intuition would arise from knowing the precise values. Nonetheless, they can be derived by solving equations (13) and (16) for the variables of interest, noting that $|b_0|$ and p_0^* are a function of some of these variables.

An outcome in which jurisdictions choose different policies is more likely when (*i*) the welfare effect of legalization per unit population when borders are closed is, while negative, small in absolute value, and (*ii*) jurisdictions are of intermediate size. If jurisdictions are "small", a late-moving jurisdiction is already experiencing the negative externality from enough of its population and therefore, it gains by legalizing. In contrast, if jurisdictions are "large", the tax revenue from border-crossers is relatively small to entice any government into legalizing. This also implies that (*iii*) when t is too high, an outcome where policies do not match is less likely because tax revenue from border-crossers is also likely to be very small; the high tax-inclusive price will deter cross-border activity. Of course, measuring what is intermediate or large depends on other parameters in the model. Nonetheless, the model allows for relative comparisons across different jurisdiction pairs (where taxes or size may matter) and across different industries (where the sizes of the externality may matter).

5.2 Extensions

The model contains several assumptions – a fixed number of commercial lots in the model with fixed costs, the lack of Pigouvian taxes to reduce externalities, and no heterogeneity in local taxes – that we discuss and conjecture how they may affect the results.

First, future work might try to model regulatory policies with endogenous location decisions and entry of firms, thus relaxing the fixed (exogenous) number locations. In our setting, no firm enters just within the border of the late moving jurisdiction because Bertrand competition would compete prices down to marginal cost, which would imply losses. Thus, the distribution of firms can be viewed as endogenous under the given assumptions. If we were to allow for entry in the model, as in the extension of Braid (1987), stores would locate sequentially, after the good is legalized, in a pattern that assures they make positive profits but that no entrant could make profits. Thus, if stores locate sequentially from the boundary, stores would establish themselves sufficiently far away from each other to make positive profits, but to make any entry in between the two firms impossible.[27] In this setting, the first-moving jurisdiction would have its first stores locate at the border; the later-moving jurisdiction would still not have firms entering on the border. Instead, firms in the second moving jurisdiction would enter just far enough away from the to deter any additional profitable entry. Thus, the main mechanism at our model would remain at work, but existence of an equilibrium need be verified.

Second, although we have argued that the tax in the model is a broad-based tax and thus is not subject to tax competition for firework shoppers, an alternative assumption could be a product-specific excise tax. By omitting a product-specific tax, the model then omits a standard policy for dealing with externalities: Pigouvian taxes. Of course, taxes would not be set at their Pigouvian levels, if the tax revenue is valued, as in this model. Nonetheless, ignoring the valuation of revenue, a further complication in this setting is that the externality is not incurred in the place of sale. If the Pigouvian taxes were destination-based, the taxes would be imposed where the harm takes place. However, enforcing destination-based taxes is especially difficult (Agrawal and Mardan 2019) and for goods with externalities, the place of use often does not equal the place of purchase (Merriman 2010). Thus, with origin-based taxes, although a jurisdiction would set taxes related to the marginal damage within the jurisdiction, the first mover would likely also ignore the marginal damage to the nearby jurisdiction and would still place value on tax exporting due to border-crossers.

Finally, we assumed a state tax rate as we assumed all jurisdictions are symmetric. In practice counties, and towns can levy local taxes (Agrawal 2014). Intuitively, as noted above, if the first mover has a higher tax rate, the effect on legalizing is ambiguous.

6 Evidence on the Model

Does legalizing the sale of a product early result in firms locating right near the border – giving the first-mover a persistent advantage resulting from state dependence? And do these firm concentrations persist after the policy advantage is obsolete because a neighboring jurisdiction legalizes later? If so, this would provide evidence consistent with vendors having

[27] Braid (1987) discusses "optimistic" entrants that do not anticipate price reductions from firms in response to its entry and "sophisticated" entrants that rationally anticipate all price changes induced by entry.

some market power to deter subsequent entrants. We proceed in two parts. First, we test for differences in the distribution of firm locations around borders due to differences in the order of legalization. Second, we conduct a hazard model analysis of policy adoption.

6.1 A Recent Legalization Decentralization

The state of New York has long had a ban on fireworks. However, starting in 2015, the state decentralized the authority to pass firework rules. In particular, the state allowed county governments to legalize certain ground level fireworks.[28] Given we observe the date of legalization and the precise location of firms, we can study if the locations of firms are sensitive to the county's time of legalization. The local nature of the data has the advantage of allowing us to look within the state of New York and to focus on similar nearby county comparisons.

We obtain data on the day of legalization from the state. Following the state level reform, Essex county was the first to legalize on March 9, 2015. Then throughout March, April, May and June, a steady trickle of jurisdictions legalized fireworks. Given the staggered order of legalization, firms may not want to wait to set up shop in a different county that may legalize later. We expect firework vendors to have some market power and permanence of location because the state requires vendors to obtain special firework licenses, which in the case of New York requires payment of a fee. Following the first July 4th celebration, a couple of counties have legalized since, but these additional legalizations have been few and far between – suggesting that Figure 1 is near the settled equilibrium. Figure A.6 shows the Kaplan-Meier survivor function, which is a non-parametric estimation of the probability a jurisdiction will adopt. Given the short time period between the reform and July 4 (arguably the most important day for firework usage), only a couple of months difference in legalization may have important effects.

We then scrape the New York database of licensed firework vendors. This database gives us the addresses of over 1000 vendors that are allowed to sell fireworks; we geocode each of these addresses using googlemaps. The median county in which firework sales are legal has approximately 25 firework vendor licenses; as a comparison, the median county has 260 liquor licenses, perhaps suggesting firework vendors may have more market power.

[28] Most states neighboring New York do not allow the sale of aerial fireworks. This implies that for most residents of New York, ground level fireworks are the only fireworks accessible to them.

6.2 Location of Firms Precisely on Borders

We first wish to see if jurisdictions that move first in legalizing have an excessive amount of firms just with their borders relative to a neighboring jurisdiction that moves later. It is well known that kinks or notches will result in bunching on the favored side of a tax bracket (Slemrod 2013, Kleven 2016), however, in the case of geographic bunching, we cannot observe the counterfactual distribution of firms in jurisdictions that do not legalize. Thus, to test our model, we compare the distribution of firms in counties that legalize early with the distribution of firms in neighboring counties that legalized later. Perhaps more importantly, this also allows us to test a dynamic model rather than the static model in Holmes 1998. In order to do this, we first calculate the driving time and distance, from every fireworks vendor f to the *nearest* county border. Given we care about densities in the local region of the border, we do not calculate the distance to all borders.[29] This measure of distance is designed to capture the "true" cost of cross-border shopping for fireworks. Once each firework store is matched to the nearest county border, we classify the firework store as being treated if the date of legalization for the store's own-county occurs before the date of legalization of the nearest neighboring county. Let d be the number of days since January 1, 2015 that it took for the county to legalize. Then, for each vendor f, the treatment variable is given by

$$E_f = \begin{cases} 1 & \text{if } d_f < \tilde{d}_f \\ 0 & \text{if } d_f \geq \tilde{d}_f \end{cases} \tag{22}$$

where d_f are the days it took firm f's county to legalize and $\tilde{d}_f$ is the days it took to legalize for the nearest neighboring county to firm f. Notice that within a particular county j, E_f can vary for different firms in the county depending if the nearest county to that firm moved earlier or later. Letting the driving time to the nearest border be our running variable, r_f, we let the running variable be positive if $E_f = 1$ and negative if $E_f = 0$. The county border acts as the point of discontinuity where we can test if the distribution of firms is higher on the early mover side than on the late mover side.

In our preferred sample, we only use county pairs where both counties legalize by the end of the sample; thus, unlike Holmes (1998), treated and control jurisdiction have the same policies today, but differed in their policies in the past. This provides potential evidence on

[29] We do this by modifying the procedure in Agrawal (2015). As in Lovenheim (2008), we first identify the intersection of all major road crossings and county borders. Then, using the locations of the fireworks vendors as "origins" and the locations of all county border road crossings as "destinations" we find the driving route that minimizes the time from the origin to the destination. We follow a hierarchical system of driving, preferring larger roads over smaller roads. Once the route that minimizes the driving time to the nearest destination is found, we obtain how many minutes it takes to drive and the road distance in miles.

the theory relating to the state dependence of firms given the historical policy differences no longer directly matter by the date of our firm location data. We implement the analysis using the McCrary (2008) test.[30] In particular, letting $g(r)$ be the density of firework vendors at distance r from the border, we define the parameter of interest θ as

$$\theta = \ln lim_{r \downarrow 0}\, g(r) - \ln lim_{r \uparrow 0}\, g(r) \tag{23}$$

which is the log difference in the density of firms near the border on the early adopter side of the border relative to the late adopter side of the border.

Table 2 shows the results using driving time as the running variable and table A.1 shows the results using driving distance. In the text, we focus on driving time and the results using the optimal bandwidth selection procedure. The first column shows the difference in density using all county borders that are not a state or international border. The results in column (1) may be consistent with more firms locating on the early adopter side of the border; however, this result is confounded by the fact that the observations with $E_v = 0$ also includes firms in counties that never adopt. To eliminate this possibility, in all subsequent columns, we only focus on border pairs where both counties allow firework sales by June 30, 2016.[31] In column (1'), we use all border pairs where both counties allow sales at some point in time. Here we find no statistically significant differences in the density of firms; however, this sample includes border pairs where both counties legalized within the same week. When restricting the sample to border pairs where one county legalized at least a month in advance (column 2) or at least two months in advance (column 3), we notice significantly more firework licenses just within the first mover's border. Although two months remains a short period, given fireworks can only be used (by state law) around July 4 and New Year's, a couple of months represents an important first mover advantage in the firework industry, especially given the short period between the state reform and July 4.[32] Given the legalization experiment analyzed, we study the first-mover advantage resulting from short time differences; larger time differences would likely generate even larger effects than what we identify. The magnitudes in are economically large and meaningful. In column (2) the results indicate a 120% increase in the density of firms on the first mover side of the border. This indicates a more than doubling of the number of firms just over the border.

[30] We follow the optimal bandwidth selection procedure, but also show results for larger bandwidths because firms are unlikely to locate right on the border because of border costs. The optimal bandwidth may result in dips near the border simply because a firm can rarely be right on the border based on our driving algorithm. Jacob and McMillen (2015) also find dips near the border.

[31] We include firms in the two counties that legalized in May 2016 even though our firms database is dated June 30, 2016. Dropping these two counties increases the magnitude of the results.

[32] New York law regulates the days fireworks can legally be used or set off, so time to establish in advance of the high demand period is likely important.

Thus, late moving jurisdictions remain disadvantaged in terms of tax revenue and firms producer surplus, because firms near the border in the early moving jurisdiction maintain their advantage even after both counties legalize. This provides indirect evidence that early entering firms have some market power to discourage entry on the later-moving side.

Figure 9 shows the results for columns (2) and (3) visually with driving time as the running variable; figure A.7 shows the results using driving distance as the running variable. There is a discontinuity in the density function at the border.[33] Comparing subfigures, the gap in the densities is increasing as the length of the time advantage of the first mover. These figures reduce geographic space to one dimension, so that there may be many firms located at a given distance because they are located at *different points along* the border. This is consistent with the theoretical model. Thus, any "bunching" we identify is many firms locating at the same distances from the border, but likely not at the same point.

A concern is that jurisdictions that legalized early may have more non-firework firms located near the border. This is a concern because an already established large retail store like Walmart or Target is allowed to obtain a firework license. Although we cannot determine if a store only sells fireworks, we restrict our sample to stores that have the words "Firework", "Fireworks", "TNT", or "Keystone" (a known firework chain) in the legal name of the business. We then run the McCrary test separately for these businesses knowing that they could not have been located in the county prior to legalization. Figure 10 shows larger amounts of firms on the first moving side of the border. On the late moving side of the border, a "pothole" emerges where few firms are located. This suggests that the border firms on the early mover side are able to deter entry of firms on the late moving side. In fact, the estimates using the optimal bandwidth suggest a discontinuity of $\theta = 1.845$ for counties that legalized more than one month earlier than the neighbors. In figure 11 we exclude all firework specialty stores and focus on the retailers that obtained firework licenses but who also sell other products. This figure shows no statistical difference in the density at the border, suggesting that the time of adoption was not driven by pre-existing (non-firework) retail firms near the border. This is also reassuring because it rules out the possibility of firework firms locating on a given side of the border simply because other retailers are there. These two figures taken together confirm that early moving jurisdictions will result in an excess mass of firms that specialize in selling the legalized good within their border.

A second concern might be that the side of the border with more firms also has more people and thus more demand for fireworks. We show that this is false. To do this, we take

[33] The densities decline after 15 minutes from the border because only firms in very large (area) counties are left in the data more than 15 minutes from the border. This is because we only calculate the distance to the nearest border and not the distance to every border. Thus, we can only make inference near the border.

Census block data on population and follow the same procedure to calculate the shortest driving time from every Census Block point (approximately 250,000 blocks) to the nearest intersection of a major road crossing and a county border. Then, assuming that every person in the Census block has that same distance, we conduct the McCrary test for the same set of borders used in the analysis above. The early moving jurisdictions have a lower population density near the border than the late moving jurisdictions. In addition, early moving jurisdictions appear to have a lower population in general.[34] This makes the firm level results even more convincing because all else equal, a firm locating on the border of the early adopter region would have likely preferred to be on the other side (closer to the larger population mass). More dense areas may have stronger zoning restrictions for firms; we cannot rule out this possibility but note that firework stores themselves are not likely to impose negative externalities on surrounding residences.

The results are consistent with a model where first-moving jurisdictions have more firms located just inside their border than second-moving jurisdictions. In particular, the excess mass of firms on the early-moving side of the border provides evidence that firework vendors have some market power which deters subsequent entry on the other side of the border. Our model predicts that state dependence arises because of firms near the border. The empirical analysis does not inform us as to the distribution of firms at the interior of a county, but theoretically, only the limiting argument is necessary to result in policies that may not match. The empirical analysis cannot yet inform us regarding the dynamics of policy adoption. For this, we turn to the next section.

6.3 The Role of the Neighbor's Policy

Is a county less likely to allow fireworks if nearby counties have already adopted? We conduct a hazard analysis that explains whether a county legalizes fireworks as a function of detailed (mostly pre-determined) county level variables and a time varying covariate that captures whether neighbors have adopted the policy already. The set of control variables include (log) area, (log) population, demographics (fraction senior, fraction white, fraction male, education, income, housing controls), whether the county borders another state/country, the number of neighbors, and local sales tax rates. For our explanatory variable of interest, we construct two variants, due to the fact that counties border multiple counties. First, in every time period, we calculate the number of neighbors that have previously legalized fireworks. This variable increases as the number of counties legalize, but its maximum value may be larger for counties with many neighbors. Although the hazard model accounts for the

[34]Firework externalities may have public goods characteristics: the noise and danger affects all exposed population, and the more you have – the denser the area –the more likely that the social cost is large.

number of neighbors directly, we also construct a normalized metric. Second, this normalized variable takes the prior variable and divides by the total number of neighbors such that it is interpreted as the fraction of neighbors with prior legalization. This variable is bounded by zero and one.

Given we observe county adoptions at the daily frequency, we work with a continuous time hazard model. The presence of time varying coefficients adds some complexity to the hazard model and we follow Jenkins (2005). Define the hazard rate as $\varphi(t, X)$, where t is duration and X is a vector of jurisdictional characteristics. Then, define $\varphi_0(\cdot)$ as the baseline hazard which summarizes the pattern of duration dependence common to all jurisdictions. Thus, the accelerated failure time (AFT) model provides the linear relationship between the log of survival time T and X, such that, if all covariates were constant, $ln(T) = \beta' X + \epsilon$.[35] This equation cannot be easily modified in the presence of time-varying covariates, but the hazard can easily be generalized. Define $\phi_t = e^{\beta' X_t}$, so the hazard can be written as

$$\varphi(t, X_t) = \phi_t \varphi_0(t\phi_t). \tag{24}$$

Of course, estimating the hazard model requires certain parametric assumptions. We use the Akaike Information Criterion (AIC) to select the parametric survival distribution.[36] The AIC selects a lognormal parametric form, which we verify by visual inspection of the Kaplan-Meier survivor function and the hazard rate implied by it. The lognormal distribution is estimated using the accelerated failure time (AFT) representation.

Table 3 shows the estimates using the two different independent variables. In order to interpret the results, note that the median county has five neighbors. At the date of adoption, the median county has two neighbors that have adopted or 42% of its neighbors that have adopted. We present various parameters of interest from the hazard model. First, we present the coefficient from the AFT model, which can be interpreted as the semi-elasticity of the expected duration with respect to the covariate. Thus, positive values indicate that the duration dependence is increasing in the covariate. Second, under the AFT representation, the time ratio of jurisdiction i and j with a one unit difference in a particular characteristic k is given by $(T_i/T_j) = e^{\beta_k(X_{ik} - X_{jk})} = e^{\beta_k}$. Finally, we present the average partial derivatives of the conditional mean (and median) duration function, which is given in days to adoption. While the mean duration may be relevant for policy making, as noted in Eberwein, Ham and LaLonde (2002), interpretation of this parameter is more difficult in the presence of

[35] We use the accelerated failure time representation because, as demonstrated below, model selection tools yield the log normal parametric form, which can only be interpreted in AFT form.

[36] Of course, the Cox proportional hazard model is an alternative, but it is not well-suited for our setting. We show robustness to the Cox model in the appendix.

censoring because the tail of the distribution may have large effects on the mean; in this case, the choice of the parametric hazard will determine what happens past the censoring point. Looking at the conditional median duration solves the problems related to sensitivity of the conditional mean function beyond censoring, but the conditional median function may be less useful for policymakers.

Table 3 indicates the coefficient from the AFT hazard model is positive, which implies an increase in the number (and fraction) of neighbors adopting increases the duration at which a county will adopt. To get a better idea of the magnitude, the time ratio for column (2) suggests that a country that has one more neighbor previously adopting proceeds to adopt at a rate approximately 80% (i.e., 1/1.211) slower than a country with one less neighbor adopting. Calculating the average partial derivatives of the conditional duration function, this suggest a 170 and 206 day increase time to adoption for the median and mean duration functions, respectively. Given we utilize approximately two years of data and the average duration is less than a year, this is a considerable increase. In turn, having a second neighbor adopt would double this duration. Although using the number of neighbors already legalizing is transparent, this metric will be larger for places with more neighbors. As an alternative, we calculate the fraction of neighbors that have adopted at any point in time. Focusing on column (4), a one unit increase in this variable indicates the effect of going from 0% of one's neighbors adopting to 100% of one's neighbors adopting. For this reason, the estimates are larger than the first two columns. To make the partial derivatives comparable, recall the median (and also, mean) county has five neighbors implying a 0.2 unit change in the fraction of neighbors variable; this then suggest a 189 and 225 day increase to adoption (per neighbor) for the median and mean duration functions, respectively.

Should these parameters be interpreted as causal effects? The number of neighbors adopting at a particular duration is an external covariate, that is, it is defined even after a county approves firework legislation. External covariates must still satisfy strict exogeneity in the hazard context. This implies that conditional on the set $X(t)$ of all time varying covariates through duration t, the fact that the spell lasts until $t + h$ provides no additional information about $X(t + h)$. Given that these policies may be determined in a competitive game, if fiscal competition is at work, this assumption may not be satisfied. As a result, the estimates should be interpreted as policy correlations rather than causal effects.

As an alternative to the parametric hazard model above, the Cox proportional hazard model is common in the literature because, by eliminating the estimation of the baseline hazard, it is robust to misspecification on this dimension. Table A.2 presents results from the Cox model. The Cox model is not estimated in AFT form and for this reason, we present hazard ratios. Qualitatively, the sign of the effects are similar. In all cases, the hazard ratio is less than one. For example, one additional neighbor adopting, lowers the hazard of adopting fireworks legislation by 25%. In this setting, the Cox model has several limitations. First, as shown in table 4 and 5, the Cox model performs extremely poorly on our model selection criteria. Second, the partial likelihood estimates are derived only using information at each "failure" time. As a result, time varying covariates are only evaluated during estimation at the failure time and it does not matter what happens in between; this is a substantial limitation.

Finally, tables 4 and 5 show robustness to the top three parametric models on the basis of the AIC model selection. As can be seen, the results for the loglogistic and Weibull form are similar to the lognormal results presented in the text. This provides evidence that the results are not being driven by the parametric form.

Overall, we conclude that having an additional neighbor legalize fireworks previously correlates with a delay in the time to adoption. Such increases in the time to adoption may result in medium-term and potentially long-term differences in regulatory policies between otherwise similar jurisdiction. As such, the dynamics of legalization – that is, which jurisdictions legalize first – may induce state dependence resulting from firm market power that discourages second-moving jurisdictions from adopting, even if the initial distribution of firms is no longer relevant for current day optimal firm locations. Although our theoretical model provides one mechanism, we do not deny that other mechanisms may be at work as well. Had we found policy convergence, yardstick competition, for example, may have been a factor, but this alternative mechanism is less likely given our findings.

7 Conclusion

This paper uses a spatial model to consider whether otherwise identical jurisdictions adopt different policies based entirely on the order in which policy decisions are made. If early policy decisions have no impact on firm locations, early legalization is always followed by universal legalization. On the other hand, if the first jurisdiction to legalize an activity gains a long-lasting advantage in firm location — if firms locate where their product is legal and (at least for awhile) stay there even if neighboring jurisdictions legalize — government policies may differ in the long-run. Under this assumption, an early-moving government may legalize because of the ability to attract nonresident consumers, while a later-moving government may not legalize because, even if it did so, some of its residents would continue purchasing outside its borders. Thus, our theoretical model provides a mechanism consistent with the well-known observations that firms often locate on one side of a border if jurisdictions set different policies. The empirical evidence that we identify on firm location using county level firework policy in New York suggests that early moving jurisdictions have an advantage resulting from state dependence. In particular, early moving jurisdictions are likely to have more firms just within their border than later-moving jurisdictions.

This paper began with a discussion of activities that might have negative externalities associated with them. Considering some possible examples, it seems reasonable that a jurisdiction that first legalizes casino gambling could gain a long-lasting advantage by becoming a tourism center. In regards to fireworks, when two jurisdictions have dramatically different laws, large fireworks stores tend to locate just inside the border of the legalizing jurisdiction. It is possible that those stores might well maintain some of their advantage in size or site selection, and thus might continue to draw some cross-border buyers even if legalization became universal. To at least some degree, then, the model in this paper might explain why differences in policy concerning these activities could persist.

In contrast, the technology to sell cigarettes or lottery tickets can be easily placed in many small retail stores. For these products, universal legalization would quickly eliminate any first-mover advantage, and the paper's model predicts policy convergence in the presence of perfect competition.

References

Agrawal, David R. 2014. "LOST in America: Evidence on Local Sales Taxes from National Panel Data." *Regional Science and Urban Economics*, 49: 147–163.

Agrawal, David R. 2015. "The Tax Gradient: Spatial Aspects of Fiscal Competition." *American Economic Journal: Economic Policy*, 7(2): 1–30.

Agrawal, David R. 2016. "Local Fiscal Competition: An Application to Sales Taxation with Multiple Federations." *Journal of Urban Economics*, 91: 122–138.

Agrawal, David R., and David E. Wildasin. 2017. "Sales Taxes, Spatial Agglomeration and the Internet." Working Paper.

Agrawal, David R., and Mohammed Mardan. 2019. "Will Destination-based Taxes Be Fully Exploited When Available? An Application to the U.S. Commodity Tax System." *Journal of Public Economics*, 169: 128–143.

Agrawal, David R., and William H. Hoyt. 2018. "Commuting and Taxes: Theory, Empirics, and Welfare Implications." *Economic Journal*, 128(616): 2969–3007.

Alm, James, Michael McKee, and Mark Skidmore. 1993. "Fiscal Pressure, Tax Competition, and the Introduction of State Lotteries." *National Tax Journal*, 46(4): 463–476.

Altshuler, Rosanne, and Timothy J. Goodspeed. 2014. "Follow the Leader? Evidence on European and US Tax Competition." *Public Finance Review*, 1: 1–20.

Behrens, Kristian, Jonathan Hamilton, Gianmarco Ottaviano, and Jaques-Francois Thisse. 2009. "Commodity Tax Competition and Industry Location Under the Destination and Origin Principle." *Regional Science and Urban Economics*, 39: 422–433.

Black, Dan A., and William H. Hoyt. 1989. "Bidding for Firms." *American Economic Review*, 79(5): 1249–1256.

Bleakley, Hoyt, and Jeffrey Lin. 2012. "Portage and Path Dependence." *Quarterly Journal of Economics*, 127(2): 587–644.

Bleakley, Hoyt, and Jeffrey Lin. 2015. "History and the Sizes of Cities." *American Economic Review*, 105(5): 558–563.

Braid, Ralph M. 1987. "The Spatial Incidence of Local Retail Taxation." *Quarterly Journal of Economics*, 102(4): 881–892.

Braid, Ralph M. 1993. "Spatial Competition Between Jurisdictions Which Tax Perfectly Competitive Retail (or Production) Centers." *Journal of Urban Economics*, 34(1): 75–95.

Braid, Ralph M. 2013. "State and Local Tax Competition in a Spatial Model with Sales Taxes and Residential Property Taxes." *Journal of Urban Economics*, 75: 57–67.

Brooks, Leah, and Byron Lutz. 2019. "Vestiges of Transit: Urban Persistence at a Micro Scale." *Review of Economics and Statistics*, 101(3): 385–399.

Brown, Ryan P., and Jonathan C. Rork. 2005. "Copycat Gaming: A Spatial Analysis of State Lottery Structure." *Regional Science and Urban Economics*, 35(6): 795–807.

Brueckner, Jan K. 1998. "Testing for Strategic Interaction Among Local Governments: The Case of Growth Controls." *Journal of Urban Economics*, 44(3): 438–467.

Brueckner, Jan K. 2000. "Welfare Reform and the Race to the Bottom: Theory and Evidence." *Southern Economic Journal*, 66(3): 505–525.

Brueckner, Jan K. 2003. "Strategic Interaction Among Governments: An Overview of Empirical Studies." *International Regional Science Review*, 26(2): 175–188.

Brueckner, Jan K., and Luz A. Saavedra. 2001. "Do Local Governments Engage in Strategic Property-Tax Competition?" *National Tax Journal*, 54: 203–230.

Bucovetsky, Sam, and John Douglas Wilson. 1991. "Tax Competition with Two Tax Instruments." *Regional Science and Urban Economics*, 21(3): 333–350.

Busso, Matias, Jesse Gregory, , and Patrick Kline. 2013. "Assessing the Incidence and Efficiency of a Prominent Placed Based Policy." *American Economic Review*, 103(2): 897–947.

Capozza, Dennis R., and Robert Van Order. 1980. "Unique Equilibria, Pure Profits, and Efficiency in Location Model." *American Economic Review*, 70: 1046–1053.

Carruthers, Bruce G., and Naomi R. Lamoreaux. 2016. "Regulatory Races: The Effects of Jurisdictional Competition on Regulatory Standards." *Journal of Economic Literature*, 54(1): 52–97.

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

Caudill, Steven B., Jon M. Ford, Franklin G. Mixon Jr., and Ter Choa Peng. 1995. "A Discrete-time Hazard Model of Lottery Adoption." *Applied Economics*, 27(6): 555–561.

Chirinko, Robert S., and Daniel J. Wilson. 2008. "State Investment Tax Incentives: A Zero-sum Game?" *Journal of Public Economics*, 92(12): 2362–2384.

David, Paul A. 1985. "Clio and the Economics of QWERTY." *American Economic Review*, 75(2): 332–337.

Dube, Arindrajit, T. William Lester, and Michael Reich. 2010. "Minimum Wage Effects Across State Borders: Estimates Using Contiguous Counties." *Review of Economics and Statistics*, 92(4): 945–964.

Eaton, B. Curtis. 1976. "Free Entry in One-Dimensional Models: Pure Profits and Multiple Equilibria." *Journal of Regional Science*, 16: 21–33.

Eaton, B. Curtis, and Richard G. Lipsey. 1978. "Freedom of Entry and the Existence of Pure Profit." *Economic Journal*, 88: 455–469.

Eberwein, Curtis, John C. Ham, and Robert J. LaLonde. 2002. "Alternative Methods of Estimating Program Effects in Event History Models." *Labour Economics*, 9: 249–278.

Eichner, Thomas. 2014. "Endogenizing Leadership and Tax Competition: Externalities and Public Good Provision." *Regional Science and Urban Economics*, 46: 18–26.

Eugster, Beatrix, and Raphaël Parchet. 2019. "Culture and Taxes." *Journal of Political Economy*, 127(1): 296–337.

Fajgelbaum, Pablo, Eduardo Morales, Juan Carlos Suárez Serrato, and Owen Zidar. 2019. "State Taxes and Spatial Misallocation." *Review of Economic Studies*, 86(1): 333–376.

Figlio, David N., Van W. Koplin, and William E. Reid. 1999. "Do States Play Welfare Games?" *Journal of Urban Economics*, 46(3): 437–454.

Fox, William F. 1986. "Tax Structure and the Location of Economic Activity Along State Borders." *National Tax Journal*, 39(4): 387–401.

Freedman, Matthew. 2017. "Persistence in Industrial Policy Impacts: Evidence from Depression-era Mississippi." *Journal of Urban Economics*, 102: 34–51.

Giroud, Xavier, and Joshua Rauh. 2019. "State Taxation and the Reallocation of Business Activity: Evidence from Establishment-Level Data." *Journal of Political Economy.*

Guryan, Jonathan, and Melissa Kearney. 2010. "Is Lottery Gambling Addictive?" *American Economic Journal: Economic Policy*, 2(3): 90–110.

Hansen, Benjamin, Keaton Miller, and Caroline Weber. 2018. "Federalism, Partial Prohibition, and Cross Border Sales: Evidence from Recreational Marijuana." Working Paper.

Hanson, Andrew, and Shawn Rohlin. 2011. "Do Location-Based Tax Incentives Attract New Business Establishments?" *Journal of Regional Science*, 51(3): 427–449.

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.

Heckman, James J. 1981. "Studies in Labor Markets." , ed. Sherwin Rosen, Chapter Heterogeneity and State Dependence, 91–140. University of Chicago Press.

Hindriks, Jean, and Yukihiro Nishimura. 2015. "A Note on Equilibrium Leadership in Tax Competition Models." *Journal of Public Economics*, 121.

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.

Hotelling, Harold. 1929. "Stability in Competition." *The Economic Journal*, 39(153): 41–57.

Hurtado, Carlos. 2018. "Behavioral Responses to Spatial Tax Notches in the Retail Gasoline Market." Working Paper.

Jacob, Benoy, and Daniel McMillen. 2015. "Border Effects in Suburban Land Use." *National Tax Journal*, 68(3S): 855–874.

Jaffe, Adam B., Steven R. Peterson, Paul R. Portney, and Robert N. Stavins. 1995. "Environmental Regulation and the Competitiveness of U.S. Manufacturing: What Does the Evidence Tell Us?" *Journal of Economic Literature*, 33(1): 132–163.

Janeba, Eckhard, and Wolfgang Peters. 1999. "Tax Evasion, Tax Competition and the Gains From Nondiscrimination: the Case of Interest Taxation in Europe." *The Economic Journal*, 109(452): 93–101.

Jenkins, Stephen P. 2005. *Survival Analysis.* Book Manuscript.

Jeppesen, Tim, and John A. List. 2002. "Environmental Regulations and New Plant Location Decisions: Evidence from a Meta-Analysis." *Journal of Regional Science*, 42(1): 19–49.

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.

Kearney, Melissa. 2005*a*. "The Economic Winners and Losers of Legalized Gambling." *National Tax Journal*, 58(2): 281–302.

Kearney, Melissa. 2005*b*. "State Lotteries and Consumer Behavior." *Journal of Public Economics*, 89(11-12): 2269–2299.

Keen, Michael, and Kai A. Konrad. 2013. "The Theory of International Tax Competition and Coordination." *Handbook of Public Economics*, 5: 257–328.

Kempf, Hubert, and Grégoire Rota-Graziosi. 2010. "Endogenizing Leadership in Tax Competition." *Journal of Public Economics*, 94(9-10): 768–776.

Kiliç, Emrah. 2008. "Explicit Formula for the Inverse of a Tridiagonal Matrix by Backward Continued Fractions." *Applied Mathematics and Computation*, 197(1): 345–357.

Kleven, Henrik Jacobsen. 2016. "Bunching." *Annual Review of Economics*, 8: 435–464.

Kline, Patrick, and Enrico Moretti. 2014. "Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority." *Quarterly Journal of Economics*, 129(1): 275–331.

Knight, Brian, and Nathan Schiff. 2012. "Spatial Competition and Cross-border Shopping: Evidence from State Lotteries." *American Economic Journal: Economic Policy*, 4(4): 199–229.

Levinson, Arik. 1996. "Environmental Regulations and Manufacturers' Location Choices: Evidence from the Census of Manufactures." *Journal of Public Economics*, 62: 5–29.

Lichter, Andereas, Max Löffler, and Sebastian Siegloch. 2018. "The Long-Term Costs of Government Surveillance: Insights from Stasi Spying in East Germany." Working Paper.

List, John A., Daniel L. Millimet, Per G. Fredriksson, and W. Warren McHone. 2003. "Effects of Environmental Regulations on Manufacturing Plant Births: Evidence from a Propensity Score Matching Estimator." *Review of Economics and Statistics*, 85(4): 944–952.

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.

Mallik, Ranjan K. 2001. "The Inverse of a Tridiagonal Matrix." *Linear Algebra and its Applications*, 325(1-3): 109–139.

Marceau, Nicolas, Steeve Mongrain, and John D. Wilson. 2010. "Why Do Most Countries Set High Tax Rates on Capital?" *Journal of International Economics*, 80(2): 249–259.

McCrary, Justin. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics*, 142(2): 698–714.

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.

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

North, Douglass C. 1991. "Institutions." *Journal of Economic Perspectives*, 5(1): 97–112.

Ohsawa, Yoshiaki. 1999. "Cross-border Shopping and Commodity Tax Competition among Governments." *Regional Science and Urban Economics*, 29(1): 33–51.

Peltzman, Sam. 2016. "State and Local Fiscal Policy and Growth at the Border." *Journal of Urban Economics*, 95(1-15).

Rathelot, Roland, and Patrick Sillard. 2008. "The Importance of Local Corporate Taxes in Business Location Decisions: Evidence from French Micro Data." *The Economic Journal*, 118(527): 499–514.

Reinganum, Jennifer F. 1981. "On the Diffusion of New Technology: A Game Theoretic Approach." *Review of Economic Studies*, 48(3): 395–405.

Rohlin, Shawn M., and Jeffrey P. Thomson. 2018. "Local Sales taxes, Employment, and Tax Competition." *Regional Science and Urban Economics*, 70: 373–383.

Rohlin, Shawn, Stuart S. Rosenthal, and Amanda Ross. 2014. "Tax Avoidance and Business Location in a State Border Model." *Journal of Urban Economics*, 83: 34–49.

Rosenthal, Stuart S., and Stephen L. Ross. 2015. "Change and Persistence in the Economic Status of Neighborhoods and Cities." *Handbook of Regional and Urban Economics*, 5B: 1047–1120.

Salop, Stephen C. 1979. "Monopolistic Competition with Outside Goods." *The Bell Journal of Economics*, 10(1): 141–156.

Slattery, Cailin Ryan. 2019. "Bidding for Firms: Subsidy Competition in the U.S." UVA Working Paper.

Slemrod, Joel. 2013. "Buenas Notches: Lines and Notches in Tax System Design." *eJournal of Tax Research*, 11(3): 259–283.

Thompson, Jeffrey P., and Shawn Rohlin. 2012. "The Effect of Sales Taxes on Employment: New Evidence from Cross-Border Panel Data Analysis." *National Tax Journal*, 65(4): 1023–1042.

Tosun, Mehmet Serkan, and Mark Skidmore. 2004. "Interstate Competition and State Lottery Revenues." *National Tax Journal*, 57(2): 163–178.

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

Trandel, Gregory A. 1992. "Evading the Use Tax on Cross-Border Sales: Pricing and Welfare Effects." *Journal of Public Economics*, 49(3): 313–331.

Vickrey, William S. 1964. *Microstatics.* New York, NY:Brace and World.

von Ehrlich, Maximilian, and Tobias Seidel. 2018. "The Persistent Effects of Place-Based Policy: Evidence from the West-German Zonenrandgebiet." *American Economic Journal: Economic Policy*, 10(4): 344–374.

Wang, You-Qiang. 1999. "Commodity Taxes under Fiscal Competition: Stackelberg Equilibrium and Optimality." *American Economic Review*, 89(4): 974–981.

Whatley, Warren C. 1990. "Getting a Foorin the Door: "Learning," State Dependence, and the Racial Integration of Firms." *Journal of Economic History*, 50(1): 43–66.

Wildasin, David E. 2003. "Fiscal Competition in Space and Time." *Journal of Public Economics*, 87(11): 2571–2588.

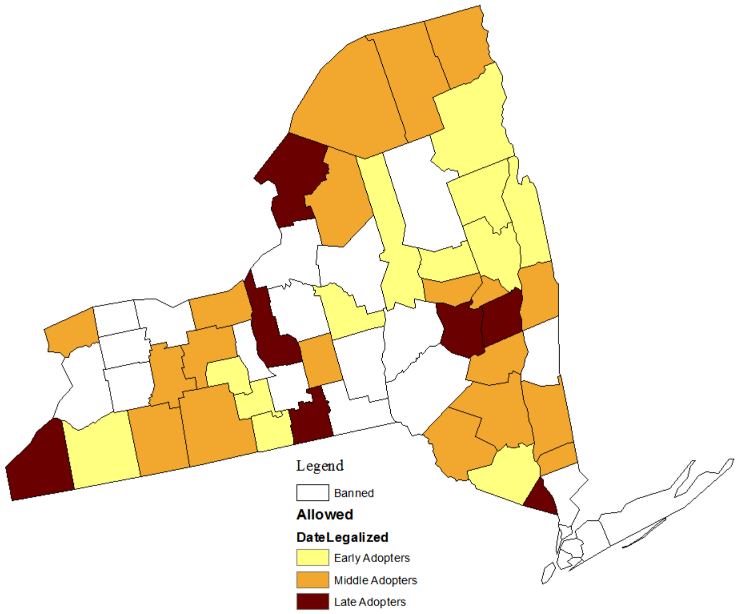
Wildasin, David E. 2006. "Fiscal Competition." In *Handbook of Political Economy.* , ed. Barry R. Weingast and Donald A. Wittman. Oxford, UK:Oxford University Press.

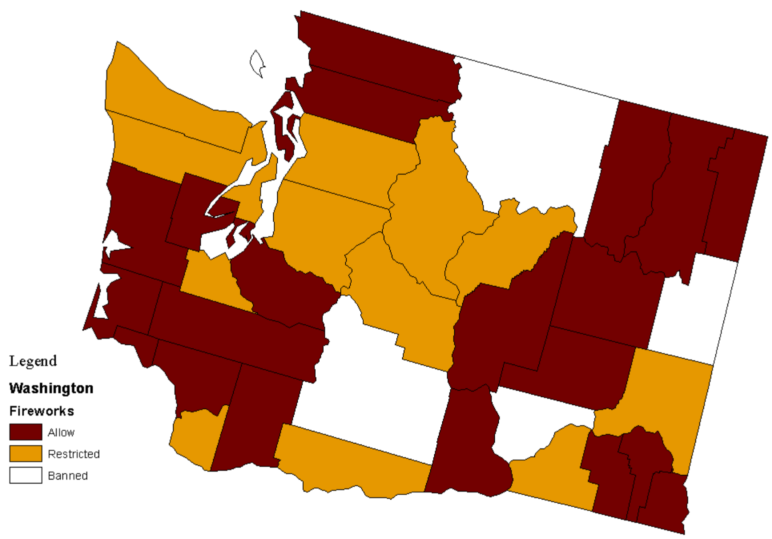
Wilson, John D. 1999. "Theories of Tax Competition." *National Tax Journal*, 52(2): 268–304.

Young, Cristobal, Charles Varner, Ithai Lurie, and Rich Prisinzano. 2016. "Millionaire Migration and Taxation of the Elite: Evidence from Administrative Data." *American Sociological Review*, 81(4): 421–446.

Figure 1: Firework Bans within States

(a) New York

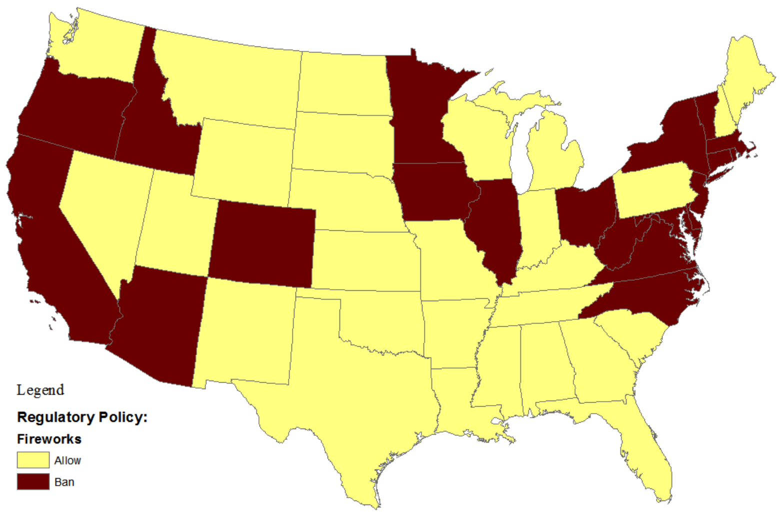




(b) Washington

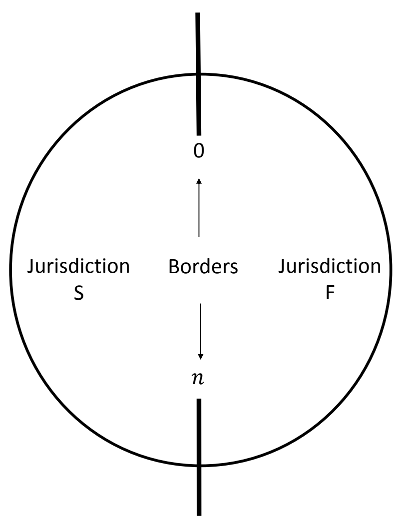
The figures show county level firework policies in the state of New York and Washington. In the New York figure, shaded counties have legalized; the darker colors show late legalizers while the lighter shades show early legalizers. For purposes of this figure, an early adopter legalizes in the first 120 days following the decentralization, a middle adopter legalizes within 120-180 days, and a late adopter legalizes after 180 days. In the Washington figure, shaded counties have legalized, with darker shades having more generous legalization rules and lighter shaded counties imposing some restrictions.

Figure 2: Aerial Firework Bans



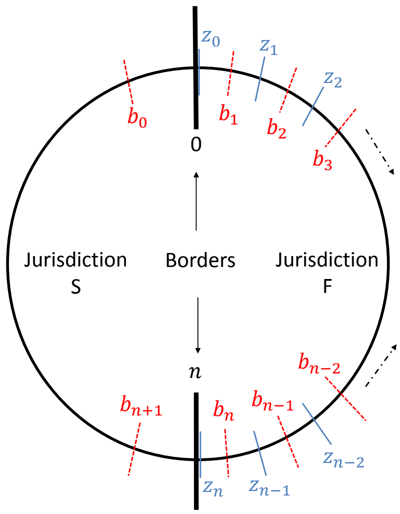
"Allow" means that the states allow for aerial fireworks sales (although possibly sales only to nonresidents). "Ban" means the state bans aerial fireworks sales to both residents and nonresidents.

Figure 3: Geographic Layout of the Model



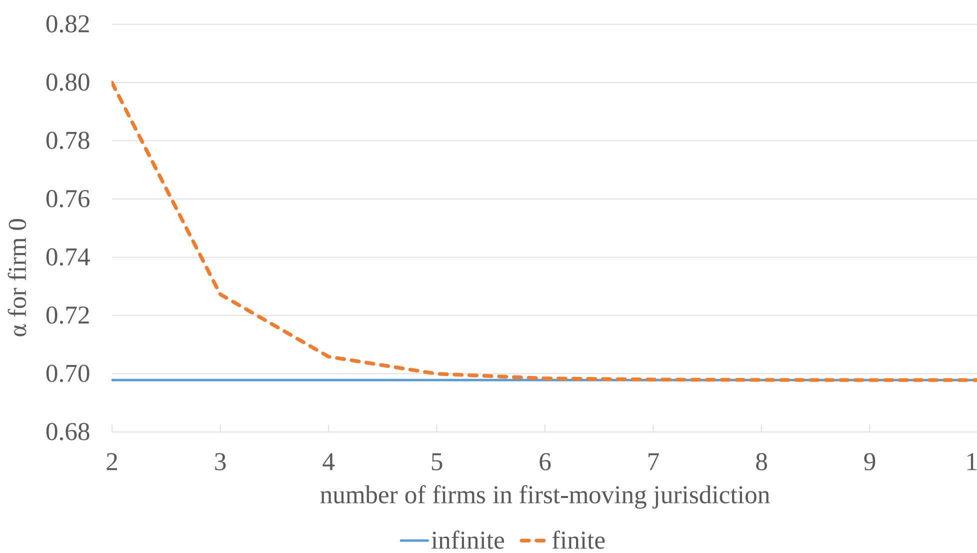
The figure shows the geography of the model. Let n denote the length of a jurisdiction. Then jurisdiction F is located from point 0 to point n moving clockwise on the circle and jurisdiction S is located from point 0 to n moving counterclockwise. For ease of notation, it will be easiest to think of jurisdiction S as occupying points $-n$ to 0 and jurisdiction F as occupying points 0 to n.

Figure 4: Geographic Layout of the Model with Market Power



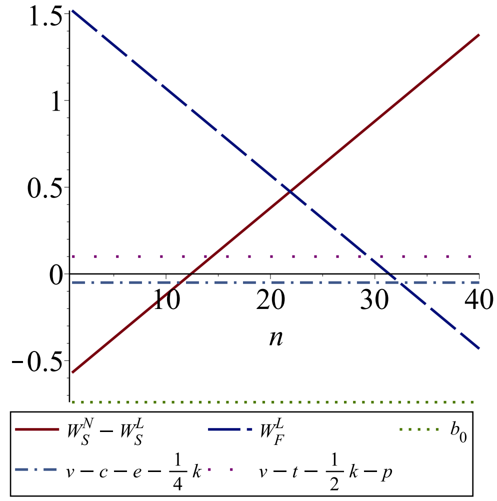
The figure shows the geography of the model with market power under the assumption that only jurisdiction F has legalized the activity. Let z_i denote a firm location at the solid blue lines, and b_i denote the market border of the firm at the dotted red line. For simplicity, we suppress the firms (and their boundaries) between z_2 and z_{n-2}.

Figure 5: Accuracy of Approximating α_0 for Different Numbers of Firms



The figure plots the analytical value for α_0 for a finite number of firms in the first-moving jurisdiction (dashed line) relative to the approximation derived for an infinite number of firms (solid line) in equation(11). The horizontal axis is the number of firms in the jurisdiction, which always equals the jurisdiction length plus one $(n+1)$. We show the case of two firms even though formulation of the model with at least one firm having two competitors involves at least three firms.

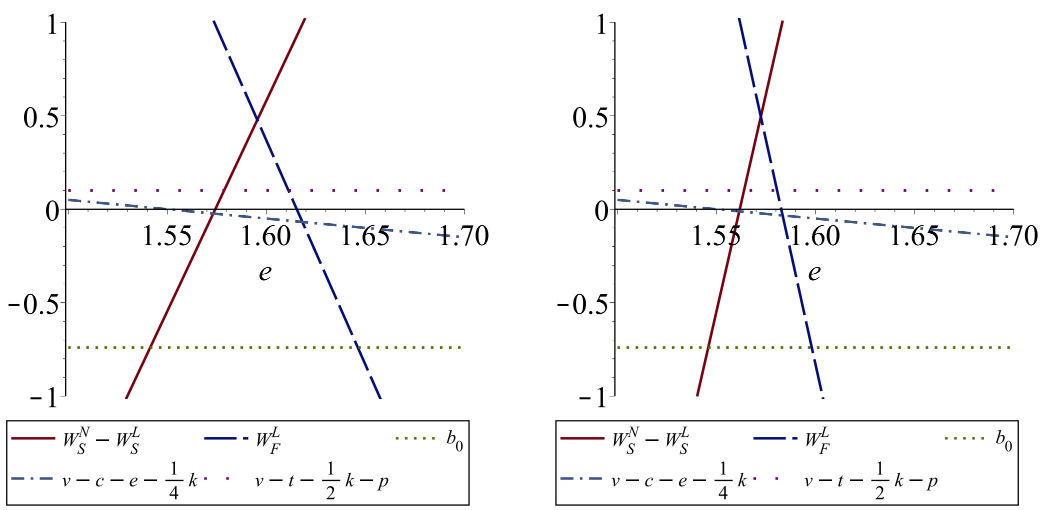
Figure 6: The Role of Size



The vertical axis shows the values of $W_S^N - W_S^L$ and W_F^L, b_0 when $v = 2.8$, $c = 1$, $t = .2$, $k = 1$, $e = 1.6$ and n is allowed to vary. An equilibrium where policies do not match will arise if $W_S^N - W_S^L > 0$ and $W_F^L > 0$. The other assumptions of the model must hold: the parameters must satisfy $b_0 < 0$, $v - c - e - \frac{k}{4} < 0$ and $v - t - \frac{1}{2}k - p > 0$ where p is the maximum value of p_i.

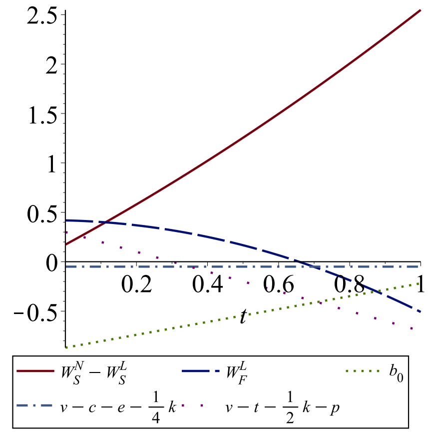
Figure 7: The Role of the Externality

(a) Smaller n (b) Larger n



The vertical axis shows the values of $W_S^N - W_S^L$ and W_F^L when $v = 2.8$, $c = 1$, $t = .2$, $k = 1$, $e = 1.6$ and e is allowed to vary. In the left figure $n = 24$ and the right figure $n = 48$. The scale of both horizontal axes are equivalent to facilitate interpretation. An equilibrium where policies do not match will arise if $W_S^N - W_S^L > 0$ and $W_F^L > 0$. The other assumptions of the model must hold: the parameters must satisfy $b_0 < 0$, $v - c - e - \frac{k}{4} < 0$ and $v - t - \frac{1}{2}k - p > 0$ where p is the maximum value of p_i.

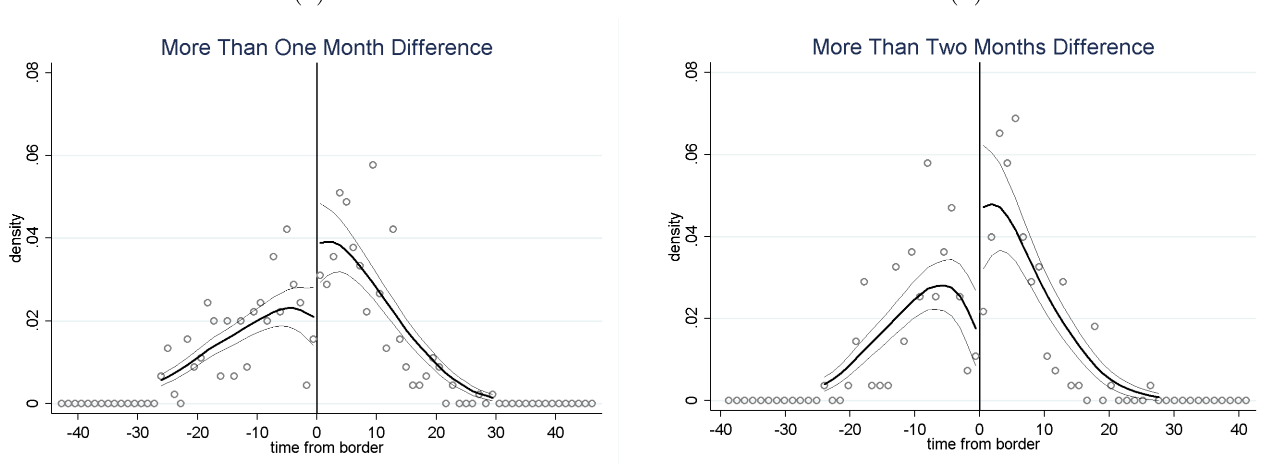
Figure 8: Role of Taxes



The vertical axis shows the values of $W_S^N - W_S^L$ andW_F^L when $v = 2.8$, $c = 1$, $n = 24$, $k = 1$, $e = 1.6$ and t is allowed to vary. An equilibrium where policies do not match will arise if $W_S^N - W_S^L > 0$ and $W_F^L > 0$. The other assumptions of the model must hold: the parameters must satisfy $b_0 < 0$, $v - c - e - \frac{k}{4} < 0$ and $v - t - \frac{1}{2}k - p > 0$ where p is the maximum value of p_i so that all people buy the good when legalized. Note, however, for very large t, the condition $v - t - \frac{1}{2}k - p > 0$ would not be satisfied.

Figure 9: Density of Firms Around Borders

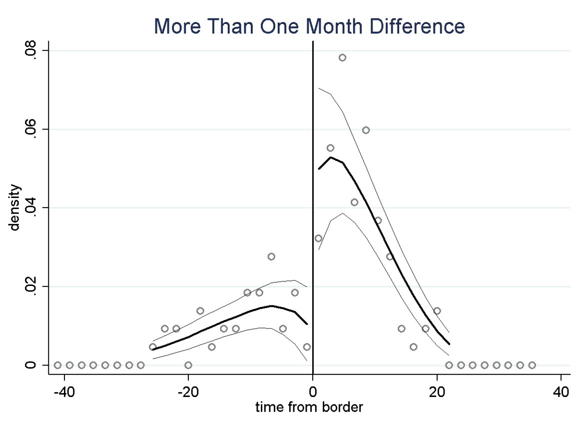
(a) (b)



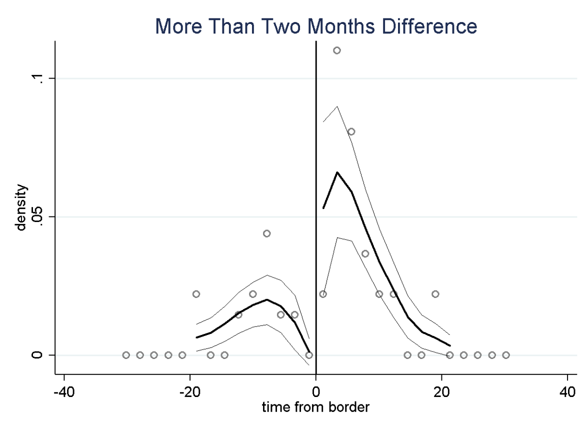
To the left of the zero line is the distribution of firms (authorized to sell fireworks) on the side of the border that was late to legalize. To the right of the zero line is the distribution of firms on the side of the border that was first to legalize. The running variable is driving time in minutes. The left figure compares firms in jurisdictions where the date of legalization differs by more than one month. The right figure compares firms in jurisdictions where the date of legalization differs by more than two months.

Figure 10: Density of Specialty Firework Firms Around Borders

(a)



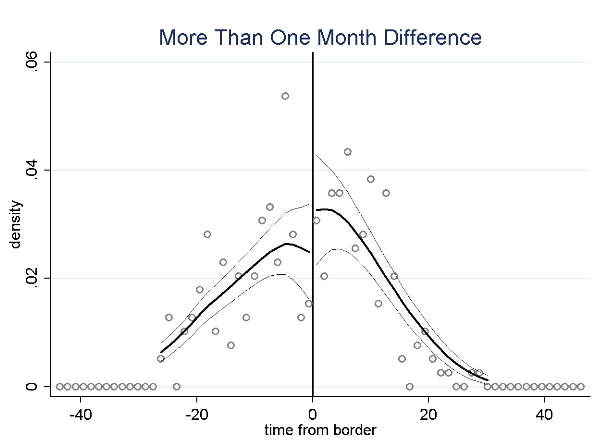
(b)



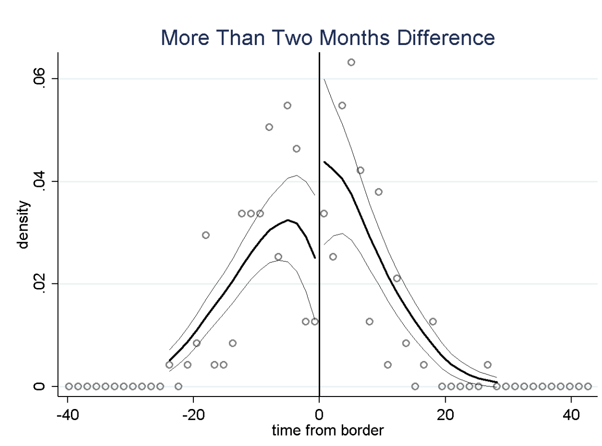
This figure focuses only on specialty stores that sell fireworks, that is, stores that could not have been in that location before the law changed. To do this, we select firms that have the words "Firework", "Fireworks", "TNT" or "Keystone" in their name. To the left of the zero line is the distribution of firms on the side of the border that was late to legalize. To the right of the zero line is the distribution of firms on the side of the border that was first to legalize. The running variable is driving time in minutes. The left figure compares firms in jurisdictions where the date of legalization differs by more than one month. The right figure compares firms in jurisdictions where the date of legalization differs by more than two months.

Figure 11: Density of Non-Specialty Firework Firms Around Borders

(a)



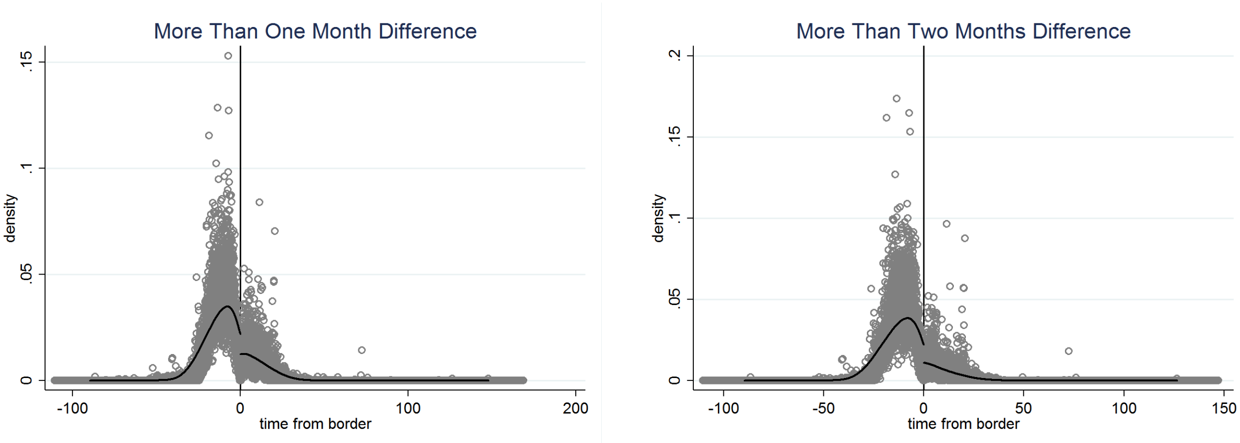
(b)



This figure focuses only on non-specialty stores (standard retailers such as Walmart or Target) that sell fireworks, that is, stores that could have been in that location before the law changed. To do this, we exclude firms that have the words "Firework", "Fireworks", "TNT" or "Keystone" in their name. To the left of the zero line is the distribution of firms on the side of the border that was late to legalize. To the right of the zero line is the distribution of firms on the side of the border that was first to legalize. The running variable is driving time in minutes. The left figure compares firms in jurisdictions where the date of legalization differs by more than one month. The right figure compares firms in jurisdictions where the date of legalization differs by more than two months.

Figure 12: Density of Population Around Borders

(a) (b)



To the left of the zero line is the distribution of population on the side of the border that was late to legalize. To the right of the zero line is the distribution of population on the side of the border that was first to legalize. The running variable is driving time in minutes. The left figure compares firms in jurisdictions where the date of legalization differs by more than one month. The right figure compares firms in jurisdictions where the date of legalization differs by more than two months.

Table 1: Moran's I for Policies

	Counties		States				
	Fireworks (WA)	Fireworks (NY)	Fireworks	Commercial Casino	Marijuana	Racetrack	Lottery
Moran's I	-0.206**	-0.093	.128	.062	-.001	.158*	.142*
	(.104)	(.084)	(.093)	(.092)	(.093)	(.093)	(.083)

This table shows the value of Moran's I test of spatial correlation using counties in the first two columns and the 48 contiguous states in the remaining columns. The significance levels indicate if the statistic is significantly different from the null of no spatial autocorrelation. ***99%, **95%, *90%.

Table 2: McCrary Density Test

		(1) All Borders (Including Bans)	(1') All Borders (No Bans)	(2) First Mover Advantage of 1 month or more	(3) First Mover Advantage of 2 months or more
Optimal Bandwidth	Difference (s.e.)	.911*** (.270)	.297 (.302)	1.203*** (.459)	1.935* (1.063)
1.5 × Bandwidth	Difference (s.e.)	.677*** (.187)	.111 (.208)	.800*** (.296)	1.479*** (.548)
2 × Bandwidth	Difference (s.e.)	.618*** (.150)	.120 (.165)	.636*** (.230)	1.147*** (.366)
	Bandwidth	8.35	8.43	8.30	7.03

This table shows the result of the McCrary test where a store is classified as "treated" if it is in a county that allowed firework sales before the nearest neighboring jurisdiction. Column (1) looks at the density of firms near all county borders within New York including those that do not legalize. Column (1') looks at all county borders within the state, but drops borders where one side does not legalize. Column (2) compares firms in counties that legalized more than a month in advance of the nearest neighbor; both counties must legalize at some point to be in this sample. Column (3) compares firms in counties that legalized more than two months in advance of the nearest neighbor; both counties must legalize at some point to be in this sample. The running variable is driving time to the nearest border (minutes). Each row presents various bandwidths. ***99%, **95%, *90%.

Table 3: Time to Adoption: Parametric Hazard Model

	Number of Neighbors		Percent of Neighbors	
	(1)	(2)	(3)	(4)
coefficient	.214** (0.102)	.192** (0.077)	1.183** (0.570)	1.026** (0.417)
time ratio	1.238** (0.126)	1.211** (0.093)	3.266** (1.262)	2.790** (1.162)
partial derivatives of the median duration	81.033* (48.985)	170.135* (93.487)	452.048 (279.291)	946.453* (544.778)
partial derivatives of the mean duration	119.711* (66.16)	205.701* (113.679)	651.108* (359.519)	1129.402* (651.065)
Time-invariant Covariates?	N	Y	N	Y
Counties	62	62	62	62
"Failures"	38	38	38	38

This table shows the result of the hazard model explaining time to adoption of firework policies in New York as a function of pre-determined covariates and a time varying covariate concerning neighbor adoption. In the first two columns, the independent variable is the number of neighboring counties that have already adopted at each duration; in the second two columns, the independent variable is the fraction of neighboring counties that have adopted. Column (1) and (3) include no other covariates while columns (2) and (4) include time-invariant covariates. We use the lognormal distribution for duration, which is an accelerated failure time representation. Each row presents various interpretations of the model. The first row presents the coefficient from the hazard model, the second row presents the time ratio, the third row presents the average partial derivatives of the conditional median duration function, and the fourth row presents the average partial derivatives of the conditional mean duration function. Standard errors are clustered at the county level. ***99%, **95%, *90%.

Table 4: Time to Adoption: Robustness of Parametric Hazard Model

	Number of Neighbors			
	(1)	(2)	(3)	(4)
	lognormal	loglogistic	Weibull	Cox
coefficient	.192**	0.174**	0.191**	
	(0.077)	(0.080)	(0.082)	
time ratio	1.211**	1.190**	1.210**	
	(0.093)	(0.956)	(0.100)	
partial derivatives of the median duration	170.135*	165.988	242.760	
	(93.487)	(115.758)	(167.475)	
partial derivatives of the mean duration	205.701*	208.762	259.082	
	(113.679)	(145.954)	(178.624)	
hazard ratio				0.761
				(0.147)
Time-invariant Covariates?	Y	Y	Y	Y
Counties	62	62	62	62
"Failures"	38	38	38	38
AIC	148.993	149.866	154.736	279.420

This table shows the result of the hazard model explaining time to adoption of firework policies in New York as a function of pre-determined covariates and a time varying covariate concerning neighbor adoption. Each of the first three column represents various parametric forms of the hazard model, each in accelerated failure time representation; the fourth column is for the Cox model. The independent variable is the number of neighboring counties that have already adopted at each duration. All models include this time-varying variable and fixed covariates. Each row presents various interpretations of the model. The first row presents the coefficient from the hazard model, the second row presents the time ratio, the third row presents the average partial derivatives of the conditional median duration function, and the fourth row presents the average partial derivatives of the conditional mean duration function. The final row, for the Cox model, presents the hazard ratio. The AIC criterion for model selection is presented in the final row. Standard errors are clustered at the county level. ***99%, **95%, *90%.

Table 5: Time to Adoption: Robustness of Parametric Hazard Model

	Percent of Neighbors			
	(1)	(2)	(3)	(4)
	lognormal	loglogistic	Weibull	Cox
coefficient	1.026**	0.987*	1.206***	
	(0.417)	(0.509)	(0.442)	
time ratio	2.790**	2.683*	3.339***	
	(1.162)	(1.368)	(1.475)	
partial derivatives of the median duration	946.453*	940.495	1308.585*	
	(544.778)	(658.537)	(769.088)	
partial derivatives of the mean duration	1129.402*	1161.714	1367.100*	
	(651.065)	(807.261)	(802.107)	
hazard ratio				0.155*
				(0.073)
Time-invariant Covariates?	Y	Y	Y	Y
Counties	62	62	62	62
"Failures"	38	38	38	38
AIC	147.568	148.745	151.682	277.798

This table shows the result of the hazard model explaining time to adoption of firework policies in New York as a function of pre-determined covariates and a time varying covariate concerning neighbor adoption. Each of the first three column represents various parametric forms of the hazard model, each represented in accelerated failure time representation; the fourth column is for the Cox model. The independent variable is the fraction of neighboring counties that have already adopted at each duration. All models include this time-varying variable and fixed covariates. Each row presents various interpretations of the model. The first row presents the coefficient from the hazard model, the second row presents the time ratio, the third row presents the average partial derivatives of the conditional median duration function, and the fourth row presents the average partial derivatives of the conditional mean duration function. The final row, for the Cox model, presents the hazard ratio. The AIC criterion for model selection is presented in the final row. Standard errors are clustered at the county level. ***99%, **95%, *90%.

A Appendix (online only)

Online appendix for "Dynamics of Policy Adoption with State Dependence"

David R. Agrawal and Gregory A. Trandel

A.1 Stylized Facts

Figure A.1 shows an example of firms locating at a state border. Figures A.2, A.3 and A.4 show the state regulatory policies for lotteries, casinos (commercial) and marijuana usage. Figure A.5 shows possible price configurations from the model for one half of a jurisdiction. Figure A.6 shows the time of adoption for New York counties. Figure A.7 and table A.1 are analogous to those in the main text of the paper. The only difference is that these results use driving distance rather than driving time as the running variable.

A.2 Theory: Infinite Solution

Here we consider the case where $n \to \infty$. Equations in the body of the paper show that $p_0 = (v-t)/3 + c/2 + k/6 + p_1/6$ and $p_i = c/2 + k/2 + p_{i-1}/4 + p_{i+1}/4$ (for $i \geq 1$) are the reaction functions for firm prices. This appendix establishes that these reaction functions are satisfied at the p_i equilibrium prices that are given by the following weighted average:

$$p_i^* = \alpha_i \left(\frac{v-t+c}{2}\right) + (1-\alpha_i)(c+k), \qquad \text{where} \quad \alpha_i = \frac{4}{4+\sqrt{3}}\left(2-\sqrt{3}\right)^i. \tag{A.1}$$

Showing that this is indeed the formula for the equilibrium prices entails showing that plugging both p_0^* and p_i^* into the relevant reaction functions produce consistent results.

Considering p_0^* first, equation (A.1) with $i = 0$ yields

$$p_0^* = \frac{2}{4+\sqrt{3}}(v-t+c) + \frac{\sqrt{3}}{4+\sqrt{3}}(c+k). \tag{A.2}$$

Similarly,

$$p_1^* = \frac{2(2-\sqrt{3})}{4+\sqrt{3}}(v-t+c) + \left(\frac{-4+5\sqrt{3}}{4+\sqrt{3}}\right)(c+k). \tag{A.3}$$

Substituting the latter expression into the p_0 reaction function produces

$$p_0^* = \frac{v-t}{3} + \frac{c}{2} + \frac{k}{6} + \left(\frac{2-\sqrt{3}}{3(4+\sqrt{3})}\right)(v-t+c) + \left(\frac{-4+5\sqrt{3}}{6(4+\sqrt{3})}\right)(c+k) \tag{A.4}$$

Setting p_0^* equal to p_0 yields

$$p_0^* = \frac{2(v-t)}{4+\sqrt{3}} + \frac{2c}{4+\sqrt{3}} + \frac{\sqrt{3}(c+k)}{4+\sqrt{3}} = \left(\frac{1}{3} + \frac{2-\sqrt{3}}{3(4+\sqrt{3})}\right)(v-t)+$$
$$\left(\frac{1}{2} + \frac{2-\sqrt{3}}{3(4+\sqrt{3})} + \frac{-4+5\sqrt{3}}{6(4+\sqrt{3})}\right)c + \left(\frac{1}{6} + \frac{-4+5\sqrt{3}}{6(4+\sqrt{3})}\right)k = p_0 \quad \text{(A.5)}$$

or

$$p_0^* = \frac{2(v-t)}{4+\sqrt{3}} + \frac{(2+\sqrt{3})c}{4+\sqrt{3}} + \frac{\sqrt{3}k}{4+\sqrt{3}} = \left(\frac{6}{3(4+\sqrt{3})}\right)(v-t)+$$
$$\left(\frac{3(4+\sqrt{3})+2(2-\sqrt{3})-4+5\sqrt{3}}{6(4+\sqrt{3})}\right)c + \left(\frac{4+\sqrt{3}-4+5\sqrt{3}}{6(4+\sqrt{3})}\right)k = p_0 \quad \text{(A.6)}$$

or

$$p_0^* = \frac{2(v-t)}{4+\sqrt{3}} + \frac{(2+\sqrt{3})c}{4+\sqrt{3}} + \frac{\sqrt{3}k}{4+\sqrt{3}} = \frac{2(v-t)}{(4+\sqrt{3})} + \frac{(12+6\sqrt{3})c}{6(4+\sqrt{3})} + \frac{(6\sqrt{3})k}{6(4+\sqrt{3})} = p_0. \quad \text{(A.7)}$$

The equality between the left-hand side and right-hand side of the above expression establishes that the equilibrium prices p_0^* and p_1^* are consistent with reaction function p_0.

Turning now to equilibrium price p_i^* ($i \geq 1$), equation (A.1) equals

$$p_i^* = \frac{4}{4+\sqrt{3}}(2-\sqrt{3})^i\left(\frac{v-t+c}{2}\right) + \left[1 - \frac{4}{4+\sqrt{3}}(2-\sqrt{3})^i\right](c+k) \quad \text{(A.8)}$$

Plugging p_{i-1}^* and p_{i+1}^* into the p_i reaction function produces

$$p_i = \frac{c}{2} + \frac{k}{2} + \frac{1}{4}\left[\frac{4}{4+\sqrt{3}}(2-\sqrt{3})^{i-1}\frac{v-t+c}{2}\right] + \frac{1}{4}\left[1 - \frac{4}{4+\sqrt{3}}(2-\sqrt{3})^{i-1}\right](c+k)+$$
$$\frac{1}{4}\left[\frac{4}{4+\sqrt{3}}(2-\sqrt{3})^{i+1}\frac{v-t+c}{2}\right] + \frac{1}{4}\left[1 - \frac{4}{4+\sqrt{3}}(2-\sqrt{3})^{i+1}\right](c+k) \quad \text{(A.9)}$$

Rewriting this formula yields

$$p_i = \frac{c}{2} + \frac{k}{2} + \left(\frac{1}{4+\sqrt{3}}\right)\left(\frac{v-t+c}{2}\right)\left[(2-\sqrt{3})^{i-1} + (2-\sqrt{3})^{i+1}\right]+$$
$$\frac{1}{4}\left[1 - \frac{4}{4-\sqrt{3}}(2-\sqrt{3})^{i-1} + 1 - \frac{4}{4-\sqrt{3}}(2-\sqrt{3})^{i+1}\right](c+k) \quad \text{(A.10)}$$

or

$$p_i = \left(\frac{1}{4+\sqrt{3}}\right)\left(\frac{v-t+c}{2}\right)(2-\sqrt{3})^{i-1}\left[1+(2-\sqrt{3})^2\right]+$$
$$\left[\frac{1}{2}+\frac{1}{4}+\frac{1}{4}-\left(\frac{(2-\sqrt{3})^{i-1}}{4+\sqrt{3}}\right)\left(1+(2-\sqrt{3})^2\right)\right](c+k) \quad (A.11)$$

or

$$p_i = \left(\frac{1}{4+\sqrt{3}}\right)\left(\frac{v-t+c}{2}\right)(2-\sqrt{3})^{i-1}(8-4\sqrt{3})+\left[1-\left(\frac{(2-\sqrt{3})^{i-1}}{4+\sqrt{3}}\right)(8-4\sqrt{3})\right](c+k) \quad (A.12)$$

or

$$p_i = \left(\frac{4(2-\sqrt{3})^i}{4+\sqrt{3}}\right)\left(\frac{v-t+c}{2}\right)+\left[1-\frac{4(2-\sqrt{3})^i}{4+\sqrt{3}}\right](c+k). \quad (A.13)$$

This last expression matches equation (A.8) which establishes that, for all $i \geq 1$, the formula for equilibrium price p_i^* satisfies reaction function p_i.

A.3 Theory: Finite Solution

For the finite n case, the solution is given by the reaction functions:

$$p_0 = \frac{v-t}{3}+\frac{c}{2}+\frac{k}{6}+\frac{p_1}{6}, \quad (A.14)$$
$$p_i = \frac{c}{2}+\frac{k}{2}+\frac{p_{i-1}}{4}+\frac{p_{i+1}}{4},\ i=1,...,n-1 \quad (A.15)$$
$$p_n = \frac{v-t}{3}+\frac{c}{2}+\frac{k}{6}+\frac{p_{n-1}}{6}. \quad (A.16)$$

These reaction functions can be written in matrix notation as

$$\begin{bmatrix} p_0 \\ p_1 \\ \vdots \\ p_{n-1} \\ p_n \end{bmatrix} = \begin{bmatrix} \frac{v-t}{3}+\frac{c}{2}+\frac{k}{6} \\ \frac{c}{2}+\frac{k}{2} \\ \vdots \\ \frac{c}{2}+\frac{k}{2} \\ \frac{v-t}{3}+\frac{c}{2}+\frac{k}{6} \end{bmatrix} + \begin{bmatrix} 0 & 1/6 & & & & 0 \\ 1/4 & \ddots & 1/4 & & & \\ & 1/4 & & \ddots & & \\ & & \ddots & & 1/4 & \\ & & & 1/4 & \ddots & 1/4 \\ 0 & & & & 1/6 & 0 \end{bmatrix} \begin{bmatrix} p_0 \\ p_1 \\ \vdots \\ p_{n-1} \\ p_n \end{bmatrix} \quad (A.17)$$

and the equilibrium prices are characterized by:

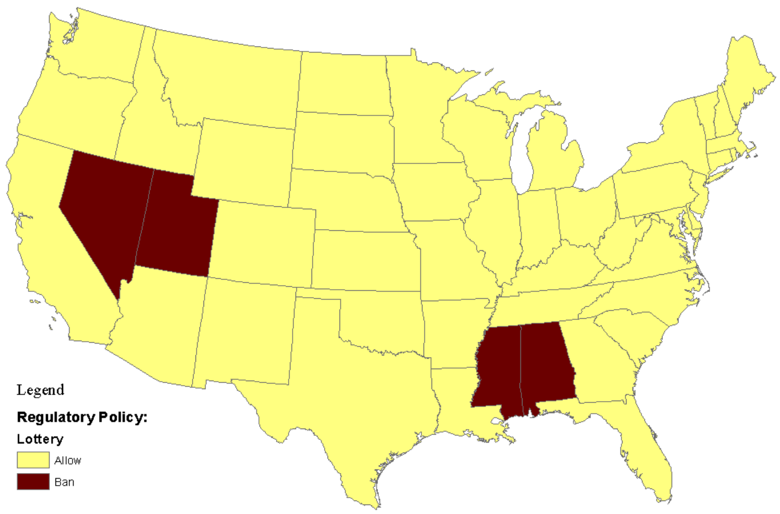
$$\begin{bmatrix} p_0 \\ p_1 \\ \vdots \\ p_{n-1} \\ p_n \end{bmatrix} = \begin{bmatrix} 1 & -1/6 & & & & 0 \\ -1/4 & \ddots & -1/4 & & & \\ & -1/4 & & \ddots & & \\ & & \ddots & & -1/4 & \\ & & & -1/4 & \ddots & -1/4 \\ 0 & & & & -1/6 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \frac{v-t}{3} + \frac{c}{2} + \frac{k}{6} \\ \frac{c}{2} + \frac{k}{2} \\ \vdots \\ \frac{c}{2} + \frac{k}{2} \\ \frac{v-t}{3} + \frac{c}{2} + \frac{k}{6} \end{bmatrix}. \tag{A.18}$$

For a given value of n this can be easily solved using standard software. The solution for a general n, as noted in the text, requires complex mathematics that provides little additional information beyond the approximation we follow. Nonetheless, the finite solution is characterized by solving (A.18).

Figure A.1: Example of Vendors Near Borders

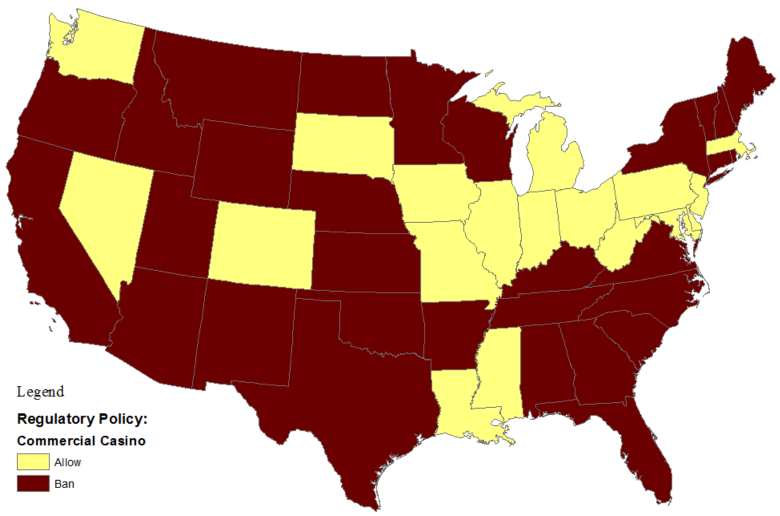
The figure shows a map of the shopping area, South of the Border, which is located just off of I-95 on the border of the North Carolina and South Carolina. Fireworks are legal in South Carolina but illegal in North Carolina. This area contains two large firework stores – Rocket City and Fort Pedro – both of which have very large structures and signage to attract cross-border shoppers.

Figure A.2: Lottery Bans



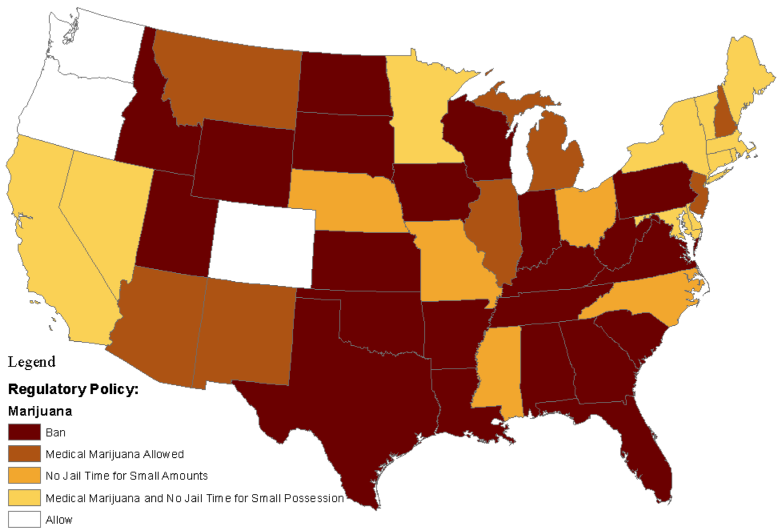
"Allow" means that the states allow for lottery ticket sales in 2018. "Ban" means that the states ban for lottery ticket sales on state lotteries or Powerball.

Figure A.3: Commercial Casino Bans



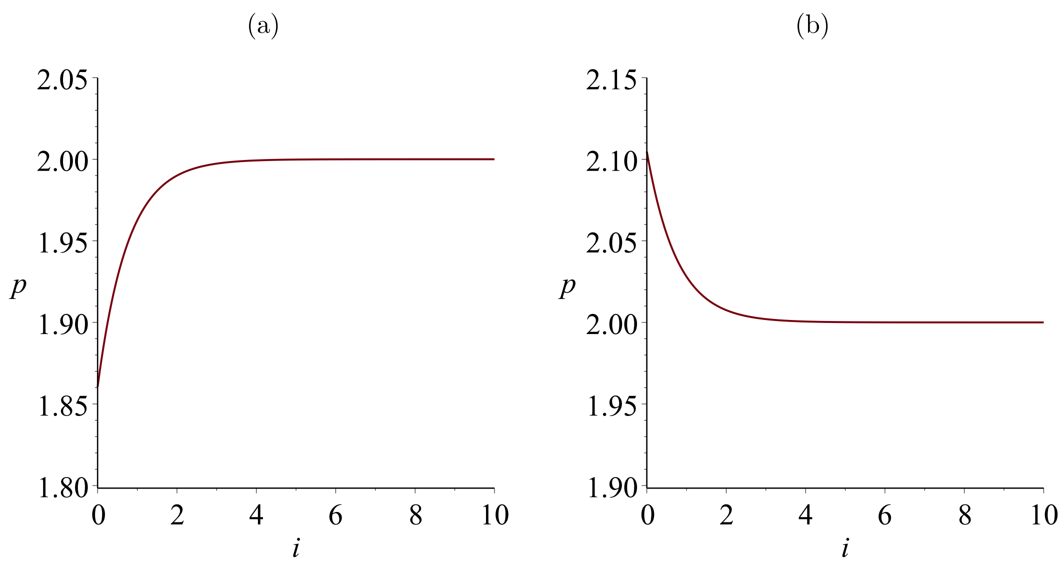
The figure shows states that ban and allow commercial casino gambling. Tribal casino gambling is not included.

Figure A.4: Marijuana Bans



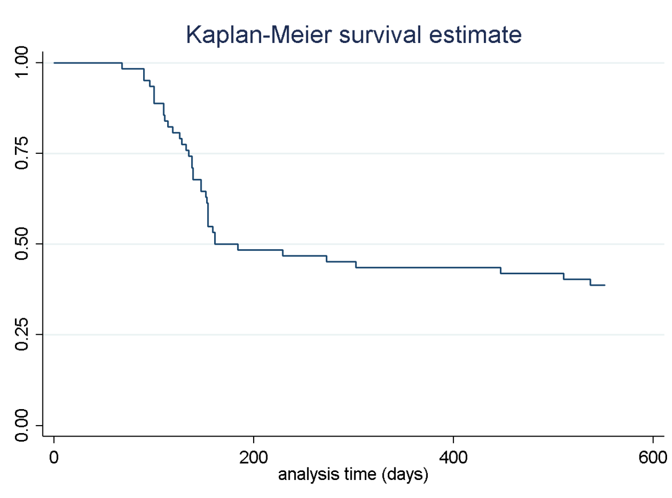
The figure shows states that ban and allow marijuana (for various purposes).

Figure A.5: Equilibrium Prices



The left figure plots equilibrium prices for each firm located at integer i (distance from the border) when parameter values are $v = 2.8$, $c = 1$, $t = .2$, and $k = 1$. The right figure keeps all parameters fixed, but raises $v = 3.5$. Firms are located only at integers, but we plot a continuous path so the pattern is easily visible. If jurisdictions have a finite length, prices are symmetric around $i = n/2$.

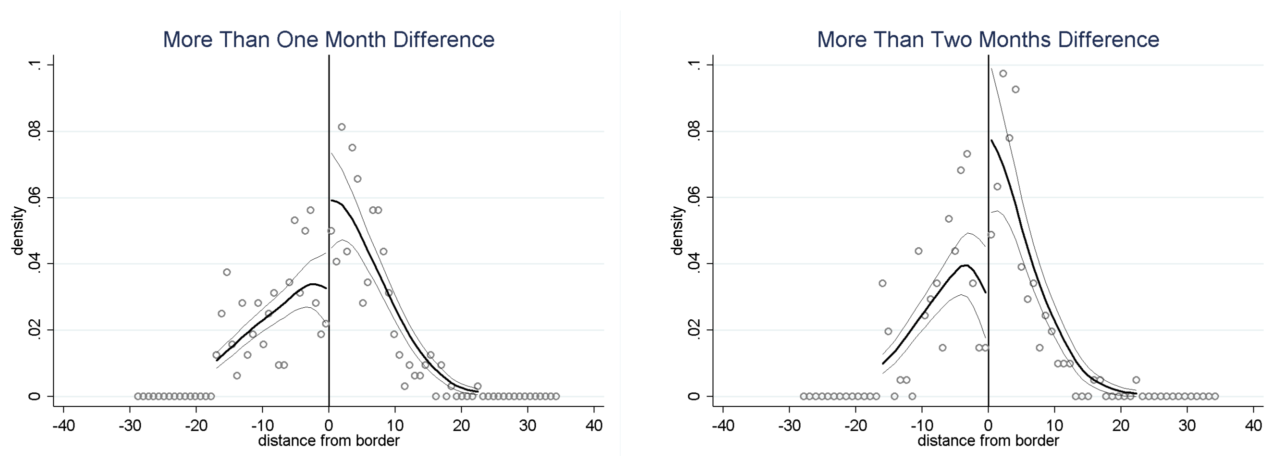
Figure A.6: County-level Legalization Timing



The figure shows the survival function for the counties in New York. A "failure" is defined as a county legalizing firework sales and thus survivors are counties that do not legalize.

Figure A.7: Density of Firms Around Borders: Driving Distance

(a) (b)



To the left of the zero line is the distribution of firework firms on the side of the border that was late to legalize. To the right of the zero line is the distribution of firms on the side of the border that was first to legalize. The running variable is driving distance in miles. The left figure compares firms in jurisdictions where the date of legalization differs by more than one month. The right figure compares firms in jurisdictions where the date of legalization differs by more than two months.

Table A.1: McCrary Density Test: Running Variable is Driving Distance

		(1) All Borders (Including Bans)	(1') All Borders (No Bans)	(2) First Mover Advantage of 1 month or more	(3) First Mover Advantage of 2 months or more
Optimal Bandwidth	Difference (s.e.)	.922*** (.238)	.334 (.253)	1.083*** (.412)	2.611** (1.024)
1.5 × Bandwidth	Difference (s.e.)	.690** (.167)	.132 (.181)	.720*** (.277)	1.319*** (.407)
2 × Bandwidth	Difference (s.e.)	.665*** (.137)	.156 (.147)	.603*** (.223)	.998*** (.298)
	Bandwidth	6.32	6.58	5.62	5.59

This table shows the result of the McCrary test where a store is classified as "treated" if it is in a county that allowed firework sales before the nearest neighboring jurisdiction. Column (1) looks at the density of firms near all county borders within New York including those that do not legalize. Column (1') looks at all county borders within the state, but drops borders where one side does not legalize. Column (2) compares firms in counties that legalized more than a month in advance of the nearest neighbor; both counties must legalize at some point to be in this sample. Column (3) compares firms in counties that legalized more than two months in advance of the nearest neighbor; both counties must legalize at some point to be in this sample. The running variable is driving distance to the nearest border (miles). Each row presents various bandwidths. ***99%, **95%, *90%.

Table A.2: Cox Hazard Model

	Number of Neighbors		Percent of Neighbors	
	(1)	(2)	(3)	(4)
coefficient	-0.125 (0.158)	-0.273 (0.193)	-0.943 (0.888)	-1.867* (1.040)
hazard ratio	0.882 (0.139)	0.761 (0.147)	0.389 (0.346)	0.155* (0.073)
Time-invariant Covariates?	N	Y	N	Y
Counties	62	62	62	62
"Failures"	38	38	38	38

This table shows the result of the Cox hazard model explaining time to adoption of firework policies in New York as a function of pre-determined covariates and a time varying covariate concerning neighbor adoption. In the first two columns, the independent variable is the number of neighboring counties that have already adopted at each duration; in the second two columns, the independent variable is the fraction of neighboring counties that have adopted. Column (1) and (3) include no other covariates while columns (2) and (4) include time-invariant covariates. We use the Cox model with a time-varying covariate. Each row presents various interpretations of the model. The first row presents the hazard model coefficient and the second row presents the hazard ratio. Standard errors are clustered at the county level. Significance of hazard ratios is for differences from one. ***99%, **95%, *90%.