7-10-2013

Inferring Policy Diffusion Networks in the American States

Bruce Desmarais
University of Massachusetts - Amherst

Jeffrey J. Harden
University of Colorado Boulder

Frederick J. Boehmke
University of Iowa

DOI: https://doi.org/10.17077/ra4p-kxo9

Comments

Subsequently published as "Persistent Policy Pathways: Inferring Diffusion Networks in the American States" in American Political Science Review / Volume 109 / Issue 02 / May 2015, pp 392-406. DOI: https://doi.org/10.1017/10.1017/S0003055415000040

Inferring Policy Diffusion Networks in the American States*

Bruce Desmarais† Jeffrey J. Harden‡ Frederick J. Boehmke§

July 10, 2013

Abstract

For decades scholars of state politics have studied the ways in which innovations in public policy diffuse across the states. Several studies indicate that policy diffusion is an explicitly dyadic process whereby states learn and adopt policies from their neighbors in geographic, social, economic, and political space. This dyadic diffusion process implies the existence of a policy diffusion network among the states. Using a dataset consisting of 189 policies, we introduce and apply algorithms designed to directly infer a diffusion network from a sample of policy adoption sequences. In addition to presenting the network inference algorithm, we offer three substantive contributions with regard to research on policy diffusion in the American states. First, we summarize and analyze the structure of the inferred diffusion network and assess the ways in which it has changed over the last several decades. Second, we demonstrate how the inferred diffusion network can be integrated into conventional statistical models of state policy adoption. Third, we estimate models to explain the pattern of diffusion ties and test a variety of theoretical expectations about who states choose to emulate.

*Prepared for presentation at the 30th Annual Meeting of the Society for Political Methodology, July 18–20, 2013, Charlottesville, VA.

†Assistant Professor, Department of Political Science, University of Massachusetts–Amherst, 420 Thompson Hall, 200 Hicks Way, Amherst, MA 01003, desmarais@polsci.umass.edu.

‡Assistant Professor, Department of Political Science, University of Colorado Boulder, 136 Ketchum, UCB 333, Boulder, CO 80309, jeffrey.harden@colorado.edu.

§Professor, Department of Political Science, University of Iowa, 341 Schaeffer Hall, Iowa City, IA 52242, frederick-boehmke@uiowa.edu.

1 Introduction

How should a state tax businesses? How should healthcare professionals be licensed? Should undocumented immigrants be charged in or out-of-state tuition at state universities? When state governments consider important policy questions such as these, they often look to related efforts undertaken by other states. Due to myriad competitive, cooperative, and imitative forces, policy innovations regularly spread throughout the American states. This notion has informed decades of research on policy adoption. Recent work moves beyond the foundational monadic models of policy adoption and characterizes policy diffusion as a mixed process of independent adoption and dyadic emulation (Volden 2006; Boehmke 2009).Though theory on policy diffusion is well developed, empirical operationalizations of influence pathways remain rudimentary at best. Empirically, diffusion ties are typically assumed to exist almost exclusively between geographically contiguous states. In this paper we explicitly conceptualize a diffusion network connecting the states. The ties in this network serve as pathways along which policies spread. We then propose and apply a method for inferring the ties in this network.

Equating geographic contiguity with a diffusion connection is a logical starting point in operationalizing a state-to-state policy diffusion network. In order to keep residents who could easily relocate without substantial disruption to the rest of their lives, neighboring states regularly compete when establishing public policy. Neighbors also co-operate to assure regional consistency in policy regimes. And, of course, neighbors have unrivaled access to each others' policymaking environments. For all of these reasons and more, it makes sense that scholars would use geographic contiguity as a proxy for the presence of an influence tie between states. As confirmation of this measurement decision, several studies have shown that the likelihood of a state adopting a novel policy increases with the number of its neighbors that have previously adopted (see, e.g., Berry and Berry 1990; Mooney 2001; Shipan and Volden 2006). However, although scholars typically operationalize influence ties via geographic contiguity, anecdotal evidence indicates that policies do *not* only spread between neighboring states.

Indeed, many states lead and follow other states throughout the country. Consider the recent carbon emissions policies crafted by states based on those in place in California. California is considered both a prolific policy innovator in general (Boehmke and Skinner 2012) and a leader in energy and environmental policy specifically (Ghanadan and Koomey 2005). As evidence of this, in 2007 New Jersey and Maryland both modeled automotive greenhouse gas emissions limits on restrictions recently implemented in California.[1] In 2009, Rhode Island implemented a cap-and-trade program applied to greenhouse gas emissions of power plants that was based on a California program.[2] These are just a few examples, but the coast-to-coast reach of diffusion highlights the shortcomings of using contiguity to operationalize diffusion ties.

Diffusion based on contiguity is just one pathway for states to learn about or feel pressure to adopt new policies. Often these pressures emphasize economic forces as contiguity facilitates movement by people to gamble (Boehmke and Witmer 2004) or buy lottery tickets (Berry and Baybeck 2005) in adjacent states or to move for more generous welfare benefits (e.g., Volden 2002). But other forces likely influence diffusion, such as social learning or comparison to peer networks or states facing similar policy problems. Walker (1969) originally identified national and regional policy leaders, but the focus moved mostly towards contiguity, particularly given the ease of measuring in the context of event history analysis. Scholars have begun to again look more broadly for policy diffusion, for example by examining whether states emulate the policies of states that have proven successful in addressing the underlying policy problem. Most prominently is Volden's (2006) introduction of dyadic event history analysis that allows each state to emulate the policy of any other state and to include variables that characteristics of both states that might be emulated and those that do the emulating.

Simultaneous to this emerging approach for expanding our understanding of policy diffusion, scholars have also returned to exploiting information on the timing of policy adoptions across dozens, and occasionally hundreds, of policies. This broadens our ability to learn about policy innovativeness and diffusion by moving from policy-specific results to learning about consistent

[1]Nussbaum, Alex. "Scientists: State Must Act on Warming Now." *The Record*. February 21, 2007. Wagner, John. "Bills to Cut Pollution, Aid Bay, are Signed into Law." *The Washington Post*. April, 25 2007.

[2]"Cutting our Carbon." *Providence Journal*. January, 26 2009.

trends across policies (e.g., Nicholson-Crotty 2009; Boushey 2010) and states (Boehmke and Skinner 2012). An advantageous consequence of the accumulation of data on state policy adoptions is that proposed methods of measuring state-to-state diffusion networks can be evaluated against abundant adoption data. Boehmke and Skinner (2012) introduce a dataset that records the times at which states adopted over 180 policies. If states really do learn from their neighbors, then in all or most of these policy areas, states should be more likely to adopt a given policy once their neighbors have already adopted that policy.

However, these rich data also open the door to evaluating proposed diffusion ties more broadly. If we think of a diffusion tie as linking a state to those states which it typically emulates, then a state should be more likely to adopt a given policy once the states it typically emulates have adopted that policy. Following this line of reasoning one step forward, we can utilize this empirical postulate regarding the policy diffusion process to infer the network of state-to-state diffusion ties. Specifically, we can infer the network such that states are likely to emulate those states to which they are connected via diffusion pathways in the network. This is the core objective of the current research; we use data on policy diffusions to directly infer the latent diffusion network connecting the states. In what follows we describe a recently developed machine learning algorithm that can be used to infer a latent diffusion network from data consisting of binary diffusion "cascades," present our application of diffusion network inference to state policy adoptions, illustrate the use of the inferred diffusion network in conventional monadic policy adoption studies, and present an analysis of the factors that predict the formation of diffusion ties between states.

2 Diffusion Network Inference with State Policy Adoption Data

Gomez-Rodriguez, Leskovec, and Krause (2010) consider the problem of inferring latent diffusion pathways connecting units (e.g., states) based only on data recording the times at which those units adopted or were infected with some attribute (e.g., policy), over several attributes. Two examples are data on when a collection of people fell ill over several ailments and data on when news websites reported a given story over several stories. These *cascades*, as they are termed, may record the operation of a hidden diffusion network connecting the units under study. State policy adoption data that includes several policies constitutes data of the type addressed by Gomez-Rodriguez, Leskovec, and Krause (2010). We use their latent network inference algorithm, which they term `NetInf`, to infer policy diffusion networks connecting the states. The `NetInf` algorithm is derived and described in detail in the Appendix.

In what follows we apply `NetInf` to a moving window of policy adoptions on the 189 policies included in the database introduced by Boehmke and Skinner (2012) to infer an evolving state-to-state policy diffusion network for the years 1960–2009.[3] Before presenting our application further, we define some useful terminology. We infer a different network in each year (t). The diffusion ties that we infer, which we refer to as *edges*,are directed, identifying for each pair of states (i, j), whether policies diffuse from i to j, from j to i, both, or neither. For a directed edge $i \rightarrow j$, which indicates that policies diffuse from i to j, we refer to i as the *source* and j as the *recipient*. Thus, if the edge $i \rightarrow j$ exists in the network at time t, then we say i is one of j's sources at time t.

2.1 Network Inference over Time

Following Boehmke and Skinner (2012), we adopt the position that influence pathways connecting states vary over time. As such, we seek to estimate diffusion networks that vary over time. There are many ways we could divide the data in order to use `NetInf` to infer a different network for each year. We made our decision with an eye toward how the networks and measures computed on them would likely be used by state politics scholars in the future. We expect, and later in this manuscript suggest, that scholars will use the diffusion networks in the same way they use geographic neighbors in statistical models of the adoption of new policies. That is, we expect and suggest that scholars will use the number of state s's sources that have adopted the policy prior to t to predict whether s will adopt that policy at time t.

To eliminate obvious problems with endogeneity that would arise in policy diffusion studies, we specify our time-varying network inference to assure that only policy adoptions prior to time t are used to inform the structure of the diffusion network at time t. As such, an edge from i to

[3]See the online appendix to Boehmke and Skinner (2012) for more information about the data: http://myweb.uiowa.edu/fboehmke/Papers/boehmke-skinner2012preprint.pdf. The data themselves are available on the Dataverse Network at http://dvn.iq.harvard.edu/dvn/dv/boehmke.

j at t can be interpreted as indicating that j has frequently emulated i in the period immediately preceding t. This way, we can be certain that a state's policy adoption at time t is not used to first infer an edge, then used to predict that same policy adoption at time t in a future study. Below we address the question of how many years preceding t should be used to infer the network for time t.[4]

2.2 Parameter Definition

We must set three parameters to define in order to infer a diffusion network at time t. First, we need to define the set of policy adoptions (denoted A) that will be used to infer the network for time t. Second, we need to define the number of edges (E) we want to infer in each time period. Third, we need to tune a rate parameter λ of the exponential distribution used by `NetInf` to calibrate how long it takes for policies to diffuse from one state to another. Higher rates place a higher penalty on the addition of edges to the network along which it takes a long time for policies to diffuse.

Absent reasonable theoretical priors regarding the appropriate values of these parameters, we take a data-driven approach to finding optimal values. We use the conventional discrete-time event history modeling methodology to evaluate the performance of the network in predicting future adoptions measured at different parameterizations. For each unique combination of parameters $\{A, E, \lambda\}$, we fit a pooled (across all policies in the dataset) logistic discrete-time event history model predicting policy adoption. The model contains three classes of regressors. For state s still in the data at time t for policy p, the regressors are

1. *States Adopting*: The number of other states that have adopted by time $t-1$,

2. *Sources Adopting*: In a network inferred on all adoptions between $t-k$ and $t-1$, where k is

[4]There may be concern that we infer one diffusion network at each time point, which models the diffusion of all policy adoptions within the respective time window. After all, some types of policies (e.g., based on whether they are more or less subject to economic competition or social learning effects (Boehmke and Witmer 2004)) may diffuse in systematically different patterns than other types of policies. In the appendix we present diagnostics to evaluate whether there exist multiple classes of policies that systematically effect the ties inferred in the diffusion networks. We find very strong evidence against the idea that there are multiple classes of diffusion patterns in our dataset of policies.

varied to define A, the number of states that influence s in the network that have adopted p (i.e., the number of states adopting p previously that have influenced s recently).

3. *Policy Area:* A dummy variable that models the unique rate of adoption for each policy.

In this design, all of the adoptions used to infer the network used to predict adoptions at time t occurred prior to t. We use a simple grid search to find best-fitting values of $\{A, E, \lambda\}$. We search over $\lambda \in \{0.125, .25, .5, 1\}$, which corresponds to mean diffusion times of 8, 4, 2, and 1 years, respectively, $k \in \{5, 10, \ldots, 50\}$, and $E \in \{100, 200, \ldots, 1000\}$. We use the Bayesian Information Criterion (BIC) to evaluate the fit of each combination of parameters and search for the combination of parameters that best fits the data (i.e., results in the lowest BIC).

Figure 1 depicts the BIC values for all of the parameter combinations that we consider. The rate of the exponential distribution – λ – does not have much effect on the fit of the models. The two influential parameters are k and E. The network that results in the best predictive fit, across all values of λ is one with 300 edges and defined over 35 years of policy adoptions. We analyze networks inferred with these optimal parameter values going forward.

[Insert Figure 1 here]

Figure 2 gives the number of unique policies and the total number of adoption instances used to infer the diffusion network in each year. The networks inferred toward the end of the time series are generally based on more data than those earlier on in the series.

[Insert Figure 2 here]

3 Descriptive Analysis of the Policy Diffusion Network

We have inferred networks over time that characterize the diffusion relations between the states in each year from 1960 to 2009. These networks are based on the co-adoption sequence patterns in the preceding 35 years.[5] In this section we conduct basic analyses of the networks to validate

[5]We also inferred networks based on 10-year periods for use in the application to policy adoption models (see below).

their structures according to dominant theoretical paradigms that characterize the policy diffusion literature. We first show that strictly monadic tie formation does not account for the network structures we observe. We then demonstrate that the network is quite distinct from a set of relations recording geographic contiguity. Lastly, we summarize the outgoing and incoming diffusion ties of each state over five-year periods.

Implicit in our analysis is that policy diffusion is best characterized by a system of dyadic relations. The significance of the dyad in our approach means that diffusion patterns cannot be explained with reference to the innovators and emulators alone. Rather, the patterns of policy diffusion are driven by unique innovator-emulator pairs. This departs from the monadic approach to the study of policy diffusion, which has dominated attempts to measure and characterize the diffusion system based on policy adoption patterns. Specifically, the state innovativeness scores introduced by Walker (1969) and extended by Boehmke and Skinner (2012) characterize the diffusion process by measuring different innovativeness (Walker 1969) and adoption rate (Boehmke and Skinner 2012) scores for each state and even over time (Boehmke and Skinner 2012).

3.1 A Monadic Process?

The first analysis we conduct is designed to assess whether the structure of connectivity in the networks departs significantly from a network generated through a strictly monadic process. By monadic process we mean a process by which a certain number of policies diffuse from each state and a certain number of policies diffuse to each state, but there is no systematic bias in the particular state-to-state dyadic paths that policies traverse. In other words, under a strictly monadic process, any two dyadic network structures that result in the same number of ties sent and received by each state are equally likely, and the specific state-to-state pairings are strictly random.

We use the concept of transitivity in networks to assess whether the inferred state-to-state diffusion networks depart markedly in structure from a strictly monadic process. If a network is transitive, then there is a heightened probability of a tie between A and B if A and B are both tied to a third partner, C. Newman (2003, 183) singles out network transitivity as a “clear deviation” from random network structure. Network transitivity (or clustering, C) is defined as

$$C = \frac{3 \times \text{the number of triangles in the network}}{\text{number of connected triples in the network}}, \tag{1}$$

where a triangle is a triple of states $\{A, B, C\}$ in which there is a tie connecting each pair of states and a connected triple exists if at least one of the three states is connected to the other two. To assess whether the diffusion networks we identify are highly transitive, we simulate, for each year, 5,000 networks from a *null* distribution over networks such that each network in which each state sends and receives the same number of ties it sends and receives in the empirically inferred network is equiprobable. We then compare the transitivity in the empirically inferred network to the transitivity of the simulated networks.[6] The results, depicted in Figure 3, show that there is a statistically high degree of transitivity in the diffusion networks relative to the strictly monadic network. The level of transitivity in the inferred networks is in each year well above the median transitivity level in the monadically simulated networks and above the 95$^{\text{th}}$ percentile of monadically simulated transitivity in the vast majority of years. These results indicate that there is substantial complex dyadic structure in the diffusion networks that cannot be explained by its monadic properties. There is more to the policy diffusion process than followers following leaders - the specific leader-follower combinations play a significant role in the structure of the networks.

[Insert Figure 3 here]

3.2 Geographic Contiguity

Another descriptive feature of the diffusion networks that we consider is whether they are accurately approximated by a network of geographic contiguity relations among states. We consider this aspect of the network because geographic contiguity has been used in several studies as a proxy for the potential for policy diffusion between states. Figure 4 plots the percentage of contiguity relations between states that are identified as diffusion ties and the percentage of inferred diffusion

[6] This is a special case of a conditional uniform graph test (Anderson, Butts, and Carley 1999). We used the `igraph` package (Csardi and Nepusz 2006) in R to perform the simulations and compute transitivity.

ties that are between contiguous states. Both of these percentages hover between ten and twenty percent between 1960 and 2009. This indicates that the overwhelming majority of policy diffusion relations exist between states that are *not* geographically contiguous. Thus, though a good first start, geographic contiguity relations do not comprise a comprehensive proxy for the policy diffusion network.

[Insert Figure 4 here]

3.3 Influential States

We have demonstrated that our network inference procedure is identifying diffusion relationships that depart from previous characterizations of the state policy diffusion process: (1) the structure of the policy diffusion networks is much more complex than a strictly monadic process and (2) geographic contiguity does not provide a good proxy for the policy diffusion network. In order to examine whether our results overlap with previous empirical rankings of the influence and innovativeness of states, we now present rankings of states based on the number of states they influence (Table 1) and the number of sates influencing them (Table 2) over five-year periods. In their time-aggregated measures of policy innovativeness, Walker (1969) and Boehmke and Skinner (2012) find {CA, NJ, OR, NY, CT} and {CA, NJ, IL, NY, OR} to be the top five states, respectively. Many of these states are at the top of our list in each five year period. A comparison of Tables 1 and 2 indicates, however, that the states most frequently representing the origin of diffusion ties are not those most frequently representing the destination of diffusion ties. This means that our network inference is not simply distinguishing between active and inactive states. Rather, we are distinguishing between innovators and imitators.

[Insert Table 1 here]

[Insert Table 2 here]

3.4 Media-based Validation of State Policy Emulation

We have inferred policy diffusion networks that connect states to those states that they *appear* to regularly emulate based on policy adoption patterns. However, we have not yet connected the diffusion ties we have inferred with any real-world instances of state-to-state policy emulation. Given the high profile status of several areas in state law, selected major policy decisions at the state level are afforded in-depth press coverage (Tan and Weaver 2009). As we show below, newspaper articles regularly indicate when a substantial portion of a state law has been modeled after another state's policy. We identify accounts of policy emulation in journalistic coverage of state policymaking by searching Lexis Nexis Academic for newspaper articles containing the phrase, "modeled after a/an ***', where '***' is the name of a state, for all fifty states. Lexis Nexis covers newspaper articles going back to 1981. From the search results we derive a count of the number of stories that report the emulation of each states' policies. These documented instances of policy emulation can serve as a qualitative validation test for the inferred networks. If the news media accurately reports some (possibly biased) sample of actual policy emulation instances, then we should observe a positive association between the number of diffusion ties sent by a state and the number of media reports of that state being emulated by others.

[Insert Figure 5 here]

Figure 3 depicts the bivariate relationship between the number of emulation stories identified and the average number of ties sent by each state in the inferred diffusion network, averaged over 1981–2009. On the linear scale, there is a moderate-strong correlation of $r = 0.70$. However, NY and CA are both large positive outliers, with approximately twice as many emulation stories as any other state, so it is prudent to consider the correlation on the log-scale, which is a moderate 0.497. This positive correlation is statistically significant at the 0.01 level based on both the Pearson's correlation coefficient and Spearman's rank-based correlation. The positive relationship between emulation reports in the media and average ties sent in the inferred diffusion networks indicates that the diffusion relationships we identify align with in-depth journalistic accounts of state-to-state policy diffusion.

4 Applying the Inferred Network to Models of Policy Diffusion

Most policy diffusion studies examine the influence of state-level features on the adoption of new policies as well as the influence states have on one another—primarily via contiguity relations. Having estimated networks of policy diffusion across the fifty states, our data provide a novel opportunity to account for cross-state dependencies in policy adoption studies. In this section we apply the inferred policy diffusion networks to published empirical analyses of diffusion for four separate policies: lotteries (Berry and Berry 1990), Indian gaming (Boehmke 2005), capital punishment (Boehmke 2005), and restaurant smoking bans (Shipan and Volden 2006).[7] We show in these examples that incorporating information from our inferred diffusion network can improve these models both theoretically and empirically.

We focus on these particular models for several reasons. First, they represent a wide variety of policies, which provides the opportunity to examine whether the diffusion network has a broad or narrow, policy-specific impact on adoption. Second, the studies presenting the original models are well-known in the policy diffusion literature, having each garnered at least 50 citations according to Google Scholar.[8] Finally, the four separate models are comparable because they all use similar event history analysis (EHA) empirical specifications. The dependent variable in each is coded "1" if a state adopted the policy in a given year and "0" otherwise, with states that have already adopted dropping out of the data beginning in the year after adoption.[9]

The theoretical frameworks explaining adoption of each of the four policy areas have their own unique characteristics. To conserve space, we refer readers to the original studies for detailed discussions of each one. We focus here on comparing the effect of the diffusion network on adoption to that of a factor that consistently appears in these models: the influence of neighboring states. Nearly all studies of policy diffusion include in their models either the number of or percentage

[7]Specifically, we replicate the following models: Berry and Berry (1990, 409), Table 1, model 1; Boehmke (2005, 85 and 89), Tables 4.2 and 4.4; Shipan and Volden (2006, 839), Table 3, model 9.

[8]In fact, Berry and Berry (1990) is included on the "high impact" list of most influential articles appearing in the *American Political Science Review* (Sigelman 2006).

[9]The Berry and Berry (1990) and Boehmke (2005) models are estimated with probit and the Shipan and Volden (2006) model is estimated with logistic regression.

of neighboring states that have previously adopted the policy. The expectation for this variable is that, due to economic competition and/or policy learning, as more neighbors adopt, the probability of a state adopting increases (see, for example, Berry and Berry 1990, 403–404; Boehmke 2005, chapter 4; Shipan and Volden 2006, 828).

While the role of economic competition is likely limited to neighboring states, it is not necessarily the case that states can only learn from states with whom they share a border. Indeed, Berry and Berry (1990) point out that there are many plausible means of state-to-state influence, including shared borders, a shared region, or even shared culture. This discussion suggests that it would be useful to have a measure of which states a state tends to "follow" in policy adoption. With information on "predesignated leader states" in regions, the authors "would hypothesize that a state's probability of adopting a lottery increases after one or more states with a reputation as a leader within its region adopt it" (Berry and Berry 1990, 403). However, the authors go on to acknowledge that they have no means of measuring this concept because there is no "reliable data about which states are perceived. . . to be regional leaders in a policy area" (Berry and Berry 1990, 403).

4.1 Including Network Information

Our inferred policy diffusion network provides those data that previous scholars of policy diffusion have not had available. In fact, beyond simply measuring regional leaders, the network gives information on any state that tends to be a leader, or "source," of policy innovation for another state. In our replications we incorporate information from the estimated diffusion network by creating a variable on the same scale as *Neighbors Adopting*: the number of a state's sources in a given year that previously adopted the policy. We use the inferred network to produce a list of states that influence the state in a specified time period immediately preceding a given year.[10] This list represents all of that state's sources at that time. Next, to create the variable *Sources Adopt-*

[10] We constructed a version using 35-year periods and one with 10-year periods. Results between the two are very similar for all four models. For each model we used the version that produced the lowest AIC value. For lotteries, capital punishment, and restaurant smoking bans this was the 35-year version and for Indian gaming it was the 10-year version.

ing we count the number of states from that list that have previously adopted the policy.[11] After creating this variable, we then add it to the EHA model for each policy area.[12]

4.1.1 Model Fit

We first examine the extent to which the inclusion of *Sources Adopting* instead of *Neighbors Adopting* improves model fit.[13] Table 3 reports three model fit statistics computed with one variable or the other included for our four replications.[14] The first two columns report the AIC values of each specification. In two of the four models (lotteries and capital punishment), the specification with *Sources Adopting* produces a value that is lower by at least 2, indicating better fit (Burnham and Anderson 2002). In the other two (Indian gaming and restaurant smoking bans), the values between the two specifications are within 1-2 units of each other, which suggests neither model fits better than the other. The next two columns report BIC values. By this criteria the model with *Sources Adopting* fits the data better in three of the four replications (lotteries, Indian gaming, and capital punishment), and in the fourth the two values are too close to declare either one a better fit.

[Insert Table 3 here]

The last two columns of Table 3 report the percentage of observations correctly classified by each specification. We compute this measure via leave-one-out cross-validation, which involves iteratively dropping one observation, estimating the model, computing an expected probability

[11] This could also be computed as a percentage, as with studies that compute the percentage of *Neighbors Adopting* (e.g., Shipan and Volden 2006). The two approaches represent very different views on the diffusion process. The percentage measure specifies a diffusion process where the non-adopting neighbors (sources) have just as much influence as the adopting neighbors (sources) and the state ends up being pulled between the two. The count-based measure assumes that non-adopting neighbors (sources) do not influence a state's decision to adopt. We see the decision of which assumption to make and which measure to use as context-dependent, and thus best left to the analyst working on a specific policy area. We use a count measure in all of our replications because it is the most commonly used in this literature.

[12] We include all policies in the construction of the network used to produce *Sources Adopting*, including the policy of interest in the EHA model. We avoid endogeneity problems because we only use adoptions that occurred before a given year to measure the network for that year. However, as a robustness check we also estimated the models after having removed the policy area of interest and found results that are virtually identical to what we present below.

[13] We also performed this analysis by keeping *Neighbors Adopting* in the model and assessing whether adding *Sources Adopting* improves fit. Those results, presented in the appendix, are substantively similar. The question of whether *Sources Adopting* should replace or complement *Neighbors Adopting* is context-dependent. If a researcher's goal is to measure the number of states that influence a state, we recommend using *Sources Adopting* instead of *Neighbors Adopting*. If the goal is more narrow, such as operationalizing economic pressure surrounding a state, we recommend including both *Neighbors Adopting* and *Sources Adopting*.

[14] We report the coefficient estimates and standard errors of these models in the appendix.

from that model for the left-out observation, then generating a predicted value of the dependent variable based on a single draw from the Bernoulli distribution with that expected probability. We then compute the percentage of the observations for which the prediction matches the actual dependent variable value. Thus, unlike information-based measures of fit such as AIC and BIC, this measure assesses each specification's capacity to make out-of-sample predictions.[15] Here we see that the specification with *Sources Adopting* improves fit for all four of the models: lotteries (+15 percentage points), Indian gaming (+11 percentage points), capital punishment (+3 percentage points), and restaurant smoking bans (+12 percentage points).

Overall, Table 3 provides good evidence that *Sources Adopting* can improve the fit of policy diffusion EHA models. For two policies (lotteries and capital punishment) all three model fit statistics point toward the model with *Sources Adopting* as the better fit. For the other two policies, either one (smoking bans) or two (Indan gaming) of the three fit statistics favor the model with *Sources Adopting* while the others are indeterminate. None of the fit statistics select the model with *Neighbors Adopting* as the better fit across the four policies. Given this evidence that *Sources Adopting* is a useful addition to diffusion models, our next step is to examine its substantive impact on policy adoption.

4.1.2 Marginal Effects

We examine the substantive implications in Figure 6 by graphing the average marginal effects of *Neighbors Adopting* (top row) and *Sources Adopting* (bottom row) in each policy area on the probability scale. All estimates are computed from the model that includes either *Neighbors Adopting* or *Sources Adopting*.[16] We employ the "observed value" method of Hanmer and Kalkan (2013) in these computations. Rather than setting the other variables in the models to particular values (e.g., their means or modes), we allow them to vary naturally over the observed values for every case in the data, then compute the average expected probability for each observed value of *Neighbors Adopting* and *Sources Adopting*, respectively.

[15]Cross-validation methods are common in statistics and other related fields and have recently become more prominent in political science (e.g., Jensen and Cohen 2000; Ward, Greenhill, and Bakke 2010; Harden and Desmarais 2011; Desmarais and Harden 2012; Carsey and Harden 2013).

[16]Results with both included in the same model are substantively similar (see the appendix).

[Insert Figure 6 here]

The first point to note from Figure 6 is the effect of the count of *Neighbors Adopting* (lotteries, Indian gaming, and capital punishment) and percentage of *Neighbors Adopting* (restaurant smoking bans) is positive. Consistent with the expectation that states react to economic competition and/or policy learning, more neighboring states with the policy corresponds with an increases in the probability of adoption. The magnitude and level of uncertainty varies somewhat across the policies, but the effect is consistently in the positive direction.

Moving to the bottom row of Figure 6, note that when substituted for *Neighbors Adopting*, the effect of *Sources Adopting* is also positive in all four policy areas; as the number of sources adopting the policy increases, so too does probability of a state adopting the policy. From the minimum (0) to the maximum (lotteries: 4, Indian gaming: 10, capital punishment: 10, restaurant smoking bans: 9) of *Sources Adopting*, the probability of adoption increases by the following percentage points, on average: 8 (lotteries), 24 (Indian gaming), 49 (capital punishment), and 15 (restaurant smoking bans). As with the effect of *Neighbors Adopting*, the confidence intervals indicate varying degrees of uncertainty around these estimates. Nonetheless, these graphs show that *Sources Adopting* exerts a substantively significant impact on the probability of adoption across four different policies.

Moreover, these positive effects remain even after controlling for *Neighbors Adopting*. In the appendix we show graphs identical to Figure 6, but computed from models that include both *Neighbors Adopting* and *Sources Adopting*. The magnitudes of the effects weaken slightly in a few cases, but they are substantively very similar. In short, these replication results show that information from our policy diffusion networks can make a valuable contribution to policy adoption studies. We show examples of four policy areas in which states utilize a persistent set of diffusion sources to guide their policymaking decisions.

5 Understanding the Inferred Network

In the previous section we demonstrated that accounting for previous adoption activity by source states in the policy diffusion networks improves a number of existing event history analyses of state policy diffusion. This indicates that latent diffusion networks play an important role in the cross-state policymaking process. With this evidence in hand, we now seek to determine whether the inferred policy diffusion network fits with various theoretical expectations about which states will be viewed as leaders, which states will be viewed as followers and the extent to which the choice of which states to follow differs across potential followers.

Existing research suggests a number of state characteristics that should increase the chance that certain states become sources for other states. As we noted earlier, in fact, identifying states that were policy leaders was one of the goals of Walker's (1969) trailblazing study, yet the literature largely moved away from this question until the introduction of the dyadic EHA model by Volden (2006). The motivation for interstate diffusion of public policy innovations begins with the idea that states do not have complete information about the costs and benefits of policies that they may consider adoption. Walker and others suggest that this leads to a process of satisficing (e.g., May 1992; Boehmke and Witmer 2004; Mooney 2001; Volden 2006), to use Simon's (1976) terminology, through which states use a series of heuristics to decide whether to adopt a policy without becoming completely informed and without determining whether it is the best policy, but rather whether it is good enough. An easy shortcut for making such a determination involves observing what other states have decided, particularly a group of peer states that may face similar policy circumstances. If none of the peer states has adopted the policy innovation, then a state will tend to go along and not adopt it itself; as the number of peers that have the policy in question increases, however, the state becomes more likely to decide that it, too, should adopt.

Walker (1969) focused largely on regional clusters of states with a small number of them serving as leaders within the cluster. More developed theories have emerged over the years, though. Many focus on the role of contiguity explicitly, whether as a source of information transmission about public opinion (Boehmke 2005; Pacheco 2012) or as a facilitator of economic flows as citizens cross state borders for desired goods or services (Berry and Baybeck 2005; Baybeck, Berry, and Siegel 2011).

Beyond contiguity, recent work has begun to focus on broader patterns of diffusion. For example, a number of studies (e.g., Volden, Ting, and Carpenter 2008; Grossback, Nicholson-Crotty, and Peterson 2004; Volden 2006) highlight the importance of ideological similarity in determining whether a state will copy the policy adopted by another state. Ideological similarity may therefore serve as a measure of a state's peer network and help determine a state's sources in our estimated diffusion network. Finally, Volden's (2006) introduction of the dyadic EHA model brings a focus squarely on identifying which states serve as leaders and for which states. The dyadic EHA model examines whether a state moves it policy towards the policy in each other state, thereby identifying the characteristics under which one state will emulate the policy of other states. In addition to the previous factors, then, Volden theorizes about the general role that political, demographic, and budgetary similarity play in determining how states choose which other states to follow.[17]

Bringing these arguments and findings together then, we can construct a set of measures to help us study the estimated diffusion network. This network represents the general pattern of policy leadership and followership over one hundred different policy diffusion episodes and therefore gives us a general sense of the set of peers for each state. Our task involves testing whether that set of peers depends on characteristics of states in the way that previous research suggests. Since our dependent variable here measures whether one state regularly emulates the policies of another state, we have dyadic data. Thus we can include characteristics of each state as it looks to other states as possible sources, characteristics of those potential source states, as well as relative characteristics of each pair of states.

First, we want to consider the characteristics of potential source states. Under the satisficing

[17] At this point it is prudent to emphasize how our analysis departs from Volden's (2006) approach, because of important overlaps. The dependent variable that Volden (2006) studies is whether a state A moves policy in the direction of state B's policy at time t, for all combinations of A, B, and t. This approach identifies policy specific emulation of B by A. Of course, if several states have the same policy as B, Volden's approach cannot determine which state A is emulating. In contrast, `NetInf` searches for a network of edges that represent regular and reliable diffusion pathways over many policies, meaning that our approach is capable of identifying the state(s) that A persistently emulates. However, our approach is not capable of identifying policy-specific diffusion ties between states – only ties that manifest consistently over many policies.

or informational theories of policy diffusion we can look to Walker's (1969) discussion of slack resources as a means to become more informed. He focuses on population and income and argues that larger, wealthier states more often have the resources and motivation to learn about policies on their own. If this leads to greater information then we would expect other states to look to them as potential sources. Further, such states may also have a greater capacity to observe what other states have done and would therefore be more likely to consider other states as sources. Thus the value of these variables in a state as well as their value in a potential source state should both increase the probability of that state being a source. Beyond these effects, however, we also want to capture Walker's idea of peer states, which fits more with the idea of satisficing than information. States may not just look to the wealthiest states to identify sources, they may also focus on similar states that have similar characteristics and whose choices may reflect more upon their specific circumstances. To capture this we include a variable measuring the absolute difference between these variables for each observation; we expect its effect to be negative.

We also include a variable for diversity within each state. More diverse states often face a greater array of public policy problems and may tend to be more innovative in response to those challenges. Thus more diverse states may provide policy information to other states as their diversity increases over time. We therefore expect diversity to increase the chance that a state is a source for other states and for diverse states to need to look to other states more often to help them identify policy solutions. Again, however, we also expect that states will look more often to states that are more similar, suggesting a negative effect for the difference in diversity.

Second, we consider two political variables: ideology and legislative professionalism. Ideology has a clear link to innovativeness on any particular policy, but here we ask whether it has an independent effect on sources across multiple policies, some of a more liberal nature and some of a more conservative nature. Our primary interest in ideology rests in its effect on identifying peers: do states consider ideological distance when determining their sources? Given that we include ideological difference, we also include its two distinct effects for the pair of states in each dyad. Legislative professionalism also has a long history in the study of policy diffusion.

More professional legislatures have great staff resources and time to learn about new policies and investigate the value of potential new policies. Again drawing from the slack resources argument, we would expect more professional states to more often be sources and to have more sources, with states also favoring more similar states.

Finally, we also address the critical role of geography. Contiguity remains the workhorse variable for interstate diffusion, so we start with it. But we also include a measure of distance to test whether states have a local bias in determining their peers; we measure this by the distance between state capitals since that is where most policymaking occurs.

In order to evaluate these predictions, we estimate a multi-level, over-time, logit model of our estimated diffusion network. As noted earlier, each observation corresponds to whether one state considers a second state as a source. We therefore have dyadic data, which facilitates the inclusion of characteristics of each state separately as well as their relative characteristics. In order to account for dependence between observations we include two (non-nested) random effects: one for each state when it is the one choosing its peer network and another when it is a potential source for other states. We also include, but do not report, a set of fixed effects for each year. We report the results of this estimation in Table 4.[18]

[Insert Table 4 here]

Overall, these results provide strong confirmation of our expectations regarding the structure of the diffusion network. The results for slack resources stand out as especially strong, with wealthier and more populous states more likely to serve as sources and more likely to identify other states as sources. Further, we find strong evidence of a similarity effect, with the larger absolute differences between states decreasing the probability of each state choosing the other as a source. Interestingly, though, the results for legislative professionalism do not conform to this pattern. The effects for the

[18]We recognize that network data may exhibit more complex dependencies than directed vertex random effects (Ward, Siverson, and Cao 2007; Berardo and Scholz 2010; Cranmer and Desmarais 2011; Desmarais and Cranmer 2012). As such, we used quadratic assignment procedure (Krackardt 1987)—a permutation testing method designed for network data—to validate the hypothesis tests presented in table 4. The QAP tests were consistent with the conclusions drawn from our hierarchical logit model.

two states are far from statistical significance and the difference term has a positive and significant effect, indicating that states rely more on states with different values of professionalism.

Our measure of citizen ideology also produces results consistent with expectations. In particular, the ideological distance has a negative and significant effect, indicating that states tend to find sources more among ideologically similar states. We also find that more liberal states have fewer sources and that liberal states tend to be sources less often, though the ideology of potential sources does not have a significant effect. We find some evidence of ideological consistency with the government as well. Unified democratic states have similar states as sources more often than states with divided government, though states with unified democratic government tend to be sources less often. No effects emerge among unified republican states, whether as potential sources or states looking for sources.

Minority diversity, which taps into the scope of potential public policy problems as well as similarity, also produces a number of findings. More diverse states get identified as sources more often and identify more sources as well. Still, states also look to states with similar levels of diversity when identifying their sources.

Finally, we turn to the important geographic variables. In the first model that does not include additional covariates, we find strong effects of contiguity and distance, but once we include characteristics of the states we find no effect of contiguity above and beyond that captured by distance. These results still suggest that nearness matters — states look to their geographic neighborhood when identifying their peers, but they do not limit themselves solely to states with shared borders.

[Insert Figure 7 here]

In order to substantively interpret these coefficient estimates, we present a series of figures that translate them into predicted probabilities that another state is chosen as a source. We first examine the variables that have an absolute difference interpretation in Figure 7. To calculate these probabilities we put every continuous variable at its mean value and every dichotomous variables at its modal value in 1985, which lies about halfway between the beginning and end of our period of analysis. We set the estimated random effects at their mean of zero, largely for convenience.

We then present partial effects for each of the five variables: one changing just the value in the state seeking sources, one changing the value in potential sources, and one changing the absolute difference between the two states. This offers a way to decompose the overall effect of each variable into its theoretically interesting pieces, but it does not represent an actual manipulation that could occur in practice since changing any of the three ways we include each variable necessarily involves changing at least one of the others. For example, holding ideology fixed in one state but changing it in the potential source leads the difference to change as well. The overall effect can be represented roughly by the sum of these two partial effects but is less interesting theoretically.

Consider first the top left figure for the effects of ideology. The baseline condition involves citizen ideology at its mean value, represented by the vertical line. If we change its value in a state seeking sources the probability of choosing another state as a source decreases when the state becomes more liberal and increases when it becomes more conservative. A similar result obtains when we manipulate the ideology of the potential source state: more liberal states get chosen less often and more conservative states more often. Of course, both of these manipulations would also increase the ideological distance, which has a negative effect on source selection. The combined effect of making the potential source more liberal would then lead to an even greater decrease than either on its own whereas the effect of making it more conservative would lead to a decrease, but one less severe than the effect of distance on its own. In terms of magnitude, the effects appear large relative to the baseline probability that the hypothetical potential source is chosen as a peer—about 15%—with the partial effects ranging from zero to about 30% relative to the baseline.

The other plots show similar patterns obtaining for per capita income, population, and minority diversity. Note that the scales of the graphs differ to enhance readability. The magnitudes do differ quite a bit, though, with population showing very large effects for a handful of large states and diversity producing a relatively small effect. Interestingly for all of the first four variables, the own state effect appears largest, the similarity effect in the middle in three cases and the potential source state effect the smallest. The results for legislative professionalism do not fit our expectations as suggested by the coefficients and the pattern is completely different. We find these results puzzling.

[Insert Figure 8 here]

We interpret the other variables in Figure 8. We present these results differently given that unified government control is binary and distance is relational only. Unified government control does have an effect, but it appears to be quite small, generally less than one or two percentage points. The biggest effects occur for same unified governments, with Democratic states most likely to choose other Democratic states as sources, but Republican states less likely to choose other Republican states. The bottom graph shows the effect of geographic distance and contiguity. Increasing distance by a thousand miles leads to an approximately a two and a half percentage point drop in the probability of choosing a state as a source whereas contiguity leads to a minuscule change once we account for distance—the small capped bar at the minimum of 40 miles represents the estimated additional effect of contiguity.

6 Conclusions

Policy diffusion is the process by which initial policy innovations become universal or modal state-level practices across the U.S. At its core, policy diffusion is a dyadic relational phenomena connecting states with those they emulate. Where do states look to when they decide which public policies to adopt? Previous quantitative research has focused almost exclusively on policy diffusing between contiguous states, but theoretical arguments such as social learning from peers suggest that diffusion will likely occur between noncontiguous states. To date, the research that has examined broader diffusion patterns has focused on a single policy area at time (Volden 2006) and has focused on specific relationships between states based on observed features such as ideological similarity. We take a different approach by estimating the latent diffusion network using a sample of over one hundred different policies that diffused over the twentieth century. We find that these networks exhibit considerable complex structure that goes well-beyond a monadic collection of leader and follower states. Moreover, standing in contrast to the literature on policy diffusion, we find that the overwhelming majority of diffusion ties connect states that are not geographic neighbors.

High on the list of influential states that we identify are those that others have found to be prolific policy innovators including New York, California, and Florida. Other states, such as South Dakota, Oklahoma influence relatively few states. We also see evidence of states such as Pennsylvania serving as leaders early on—rising to tenth in the early 1970s—and then falling dramatically to 48th by the twenty-first century. In future work we hope to refine our estimates of source states in a few ways. For example, we can consider whether such networks differ across broad policy areas. With over 100 policies, we could consider differences for social and economic policy diffusions. Gathering additional data would allow us to perform cross-validation of our estimated diffusion network in an independent sample of new policies to see how they compare.

We also show that the policy diffusion networks we infer stand to advance event history analyses of policy adoption within specific policy areas. Through a series of four replications of previously published diffusion models, we find that when a state's policy diffusion network sources adopt a policy, the likelihood of adoption in the future increases. The effect of source adoption is statistically significant and generally comparable in magnitude to adoption by contiguous states. Focusing on contiguity therefore misses a substantial part of policy diffusion and we believe that future studies should account for previous adoptions by source states in much the same way they account for adoption by contiguous neighbors today.

Our models that explain this inferred networks, by modeling the existence of directed ties between states, provide support for a number of theoretical perspectives on diffusion. Perhaps most interestingly, they highlight the role of internal capacity and pairwise similarity, which tend to dominate. States with greater resources tend to have more peers, but all states favor other states that share similar features. We also find evidence of leadership, with larger and wealthier states more often chosen as sources. One of the strengths of our approach lies in the fact that these relationships emerge based on the pattern of leadership across over a hundred policies. In future work we hope to be able to simultaneously estimate the network and its correlates by combining these two currently separate steps.

Appendix

A Latent Network Inference

The derivation of the `NetInf` algorithm begins with the definition of a probabilistic model describing how attributes would cascade through a diffusion network. To clarify application to state policy diffusion, we refer to the units and attributes in the model as states and policies, respectively. Denote a single policy cascade—the years in which states adopted a given policy—as c. The model is derived in three steps. First, we construct the probability that state u spreads a policy to state v: $P_c(u,v)$. Second, given these dyadic spread probabilities, we build the probability that a policy spreads through the states in a given *tree* pattern $P(c|T)$, where T specifies which states influence which other states. Third, we define $P(c|G)$, which is the probability of cascade c given the diffusion network (i.e., graph) connecting the states G. With these three quantities defined, we can define a proper likelihood of the policy cascades given a proposed diffusion network by evaluating the probability of each cascade on that diffusion network.

The `NetInf` algorithm assumes that diffusion occurs in continuous time and that diffusion time has an exponential distribution. If state u adopts a policy at time t_u and state v adopts a policy at time t_v ($t_v \geq t_u$) and u spreads the policy to v, then the probability of the diffusion time ($t_v - t_u$) is given by

$$P_c(u,v) = \lambda \exp\left(\frac{-(t_v - t_u)}{\lambda}\right), \tag{2}$$

where λ is the rate parameter of the exponential distribution. Given this, the probability of observing a cascade that propagates in a given pattern over the states, represented by the tree T that encodes (i, j) pairs listing which states were influenced by which other states, is

$$P(c|T) = \prod_{(i,j)\in T} P_c(i,j). \tag{3}$$

The diffusion network G places a constraint on the possible tree structures T along which the policy can spread. That is, a policy cannot spread from i to j if there is not a diffusion pathway from i to j in G. Thus, to build the probability of a cascade c given the diffusion network G, we average the probability of the cascade c over all possible tree structures in G, denoted $\mathscr{T}(G)$.

$$P(c|G) = \frac{1}{|\mathscr{T}(G)|} \sum_{T \in \mathscr{T}(G)} P(c|T), \tag{4}$$

where $|\mathscr{T}(G)|$ is the number of tree structures that can be constructed from G. Given a set of policy cascades (C), the likelihood of the cascade data given a proposed diffusion network G is:

$$P(C|G) = \prod_{c \in C} P(c|G). \tag{5}$$

A.1 Inferring the Network

With the probabilistic model of diffusion along a diffusion network defined, the task of inferring a diffusion network is to find a network structure G under which we would have been highly likely to observe the set of policy cascades C. Ideally, we would identify the network structure that maximized the likelihood of observing C. Likelihood maximization in this case, however, turns out to be a computationally intractable task. Among the 50 states, there are 2×2^{1225} possible network structures. Moreover, Gomez-Rodriguez, Leskovec, and Krause (2010) show that every network structure would need to be evaluated to assure that the optimal network had been identified.

As a more computationally tractable alternative, Gomez-Rodriguez, Leskovec, and Krause (2010) derive an approach to approximation of the optimal G. They also demonstrate analytically and through simulations that this method is capable of inferring a very-close-to-optimal network structure inference within feasible compute times. Their departures from exhaustive optimization are two. First, instead of computing the likelihood of a cascade given a network structure by enumerating all possible propagation trees represented by that network structure, they simply focus on the most likely propagation tree for each cascade within a given network structure—a shortcut which they refer to as *lazy evaluation*. Second, they adopt a greedy (i.e., local) optimization approach that iteratively adds diffusion ties to the network structure G such that the k^{th} diffusion tie added to the network improves the likelihood function more than any other tie that could be added to the network, given the $k-1$ ties already in the network.

B Checking for Heterogeneity in Diffusion Classes

As the heterogeneity in the results from models of policy diffusion in the state politics literature suggests, there is considerable variation in the processes that drive the diffusion of different policies. It is therefore important to check whether we are inappropriately pooling policies to infer a single diffusion network. Though we know that policies vary in terms of the patterns and predictors of diffusion, we must evaluate whether this variation is policy-specific and idiosyncratic with respect to the underlying diffusion network, or whether there are systematic and consistent cross-policy differences. In other words, we need to check whether there are different classes of policies in terms of the underlying diffusion network.

We use a probabilistic mixture modeling approach (Imai and Tingley 2012) to examine whether there are multiple classes of policies in terms of their effects on the inferred diffusion network. We iteratively remove each policy from the dataset and infer a new network with 300 edges that spans the entire time period in our data. For each policy, we have a network inferred without that policy included. If two (or more) policies affect the diffusion network in the same way, the inferred network should change in systematically similar ways when those two policies are removed from the dataset. Using a policies$\times$potential edges –189$\times$2,450 – observation dataset, we fit a Bernoulli mixture model with the likelihood

$$l(\mathbf{y}, \alpha, \pi) = \prod_{i=1}^{50} \prod_{j \neq i} \prod_{p=1}^{189} \sum_{a=1}^{k} \alpha_{ap} \pi_{ija}^{y_{ijp}} (1 - \pi_{ija})^{(1-y_{ijp})},$$

where y_{ijp} is an indicator of whether there is a diffusion tie from i to j when policy p is removed from the dataset, k is the number of classes (i.e., mixture components) included in the model, α_{ap} is the probability that policy p is a member of class a, and π_{ija} is the probability that there is an edge from i to j in networks inferred excluding policies in class a.

We estimate models with $k \in \{1, 2, \ldots, 15\}$. The R package `flexmix` (Leisch 2004) is used to fit the models. Estimation also requires an initial assignment of the component membership probability for each policy. We use k-means clustering to identify initial cluster memberships, then assign the component membership cluster probability for each policy according to $\alpha_{ap}^{0} = \frac{\lambda^{\mathbf{1}(c_p=a)}}{\sum_{i=1}^{k} \lambda^{\mathbf{1}(c_p=i)}}$, where c_p is the initial cluster assignment of policy p and λ is a weight that controls the entropy in the initial component assignment probabilities, with higher values of λ corresponding to lower entropy. We evaluate models with 10 values of λ, varied equally between 1 and 5.

[Insert Figure 9 here]

The model fit results of the mixture modeling are presented in Figure 9. Following Fraley and Raftery (1998), we evaluate the fit of each model using the BIC. Across all values of λ, the best fitting model is clearly the one with only one component. This indicates that, insofar as removing individual policies changes the results of the network inference, the network is changed in ways that are idiosyncratic with respect to the other policies. In other words, policies do not appear to affect the network in patterns that can be efficiently grouped into a discrete number of classes, aside from the overall patterns that cut across all policies. These results support our use of a single diffusion network to model the diffusion patterns across all of the policies in our dataset.

C Complete Replication Results

C.1 Replication Model Results

Table 5 presents coefficient estimates and standard errors (in parentheses) for *Neighbors Adopting* and *Sources Adopting* and AIC and BIC values in three specifications of the replication models. The first column presents the original results, the second column presents results with *Sources Adopting* substituted for *Neighbors Adopting*, and the third column presents results with both variables included.

[Insert Table 5 here]

C.2 Model Fit with *Neighbors Adopting* and *Sources Adopting*

Table 6 compares the fit of our replication models with and without *Sources Adopting* added to a specification that includes *Neighbors Adopting*. The first column gives the χ^2 statistic from a likelihood ratio (LR) test comparing the two specifications. For three of the four models the statistic is statistically significant (lotteries and capital punishment: $p < 0.05$; Indian gaming: $p < 0.10$), indicating the model with *Sources Adopting* fits the data better. For the restaurant smoking ban model the test statistic is not significant ($p = 0.21$).

The next two columns give AIC values. In two of the four models (lotteries and capital punishment), the specification with *Sources Adopting* produces a value that is lower by at least 2, indicating better fit (Burnham and Anderson 2002). In the other two (Indian gaming and restaurant smoking bans), the values between the two specifications are within 1-2 units of each other, which suggests neither model fits better than the other.

Finally, the cross-validated percent correctly classified indicates that the model with *Sources Adopting* improves fit for three of the four models (lotteries, capital punishment, and smoking band) and is equivalent to the model without *Sources Adopting* for one model (Indian gaming). As before, these results indicate that *Sources Adopting* is a useful addition to policy adoption models.

[Insert Table 6 here]

C.3 Marginal Effects with *Neighbors Adopting* and *Sources Adopting*

Figure 10 presents the average marginal effects of *Neighbors Adopting* and *Sources Adopting* on the expected probability of adoption, *controlling for the other* (i.e., from the models in column 3 of Table 5). Note that results are substantively similar to those in Figure 6 in the main text, which presents results from models with one variable or the other.

[Insert Figure 10 here]

D Fixed Effects Logit Results

[Insert Table 7 here]

References

Anderson, Brigham S, Carter Butts, and Kathleen Carley. 1999. "The interaction of size and density with graph-level indices." *Social Networks* 21(3): 239 – 267.

Baybeck, Brady, William D. Berry, and David A. Siegel. 2011. "A Strategic Theory of Policy Diffusion Via Intergovernmental Competition." *The Journal of Politics* 73(01): 232–247.

Berardo, Ramiro, and John T. Scholz. 2010. "Self-Organizing Policy Networks: Risk, Partner Selection, and Cooperation in Estuaries." *American Journal of Political Science* 54(3): 632–649.

Berry, Frances Stokes, and William D Berry. 1990. "State Lottery Adoptions as Policy Innovations: An Event History Analysis." *American Political Science Review* 84(2): 395–415.

Berry, William D., and Brady Baybeck. 2005. "Using Geographic Information Systems to Study Interstate Competition." *American Political Science Review* 99(04): 505–519.

Boehmke, Frederick J. 2005. *The Indirect Effect of Direct Legislation: How Institutions Shape Interest Group Systems*. Columbus, OH: The Ohio State University Press.

Boehmke, Frederick J. 2009. "Policy Emulation or Policy Convergence? Potential Ambiguities in the Dyadic Event History Approach to State Policy Emulation." *The Journal of Politics* 71: 1125–1140.

Boehmke, Frederick J., and Paul Skinner. 2012. "State Policy Innovativeness Revisited." *State Politics & Policy Quarterly* 12(3): 303–329.

Boehmke, Frederick J, and Richard Witmer. 2004. "Disentangling diffusion: The Effects of Social Learning and Economic Competition on State Policy Innovation and Expansion." *Political Research Quarterly* 57(1): 39.

Boushey, Graeme. 2010. *Policy Diffusion Dynamics in America*. New York City: Cambridge University Press.

Burnham, Kenneth P., and David R. Anderson. 2002. *Model Selection and Multi-Model Inference*. 2nd ed. New York: Springer.

Carsey, Thomas M., and Jeffrey J. Harden. 2013. *Monte Carlo Simulation and Resampling Methods for Social Science*. Thousand Oaks, CA: Sage.

Cranmer, Skyler J., and Bruce A. Desmarais. 2011. "Inferential Network Analysis with Exponential Random Graph Models." *Political Analysis* 19(1): 66–86.

Csardi, Gabor, and Tamas Nepusz. 2006. "The igraph software package for complex network research." *InterJournal* Complex Systems: 1695.

Desmarais, Bruce A., and Jeffrey J. Harden. 2012. "Comparing Partial Likelihood and Robust Estimation Methods for the Cox Regression Model." *Political Analysis* 20(1): 113–135.

Desmarais, Bruce A., and Skyler J. Cranmer. 2012. "Micro-Level Interpretation of Exponential Random Graph Models with Application to Estuary Networks." *Policy Studies Journal* 40(3): 402–434.

Fraley, C., and A. E. Raftery. 1998. "How Many Clusters? Which Clustering Method? Answers Via Model-Based Cluster Analysis." *The Computer Journal* 41(8): 578–588.

Ghanadan, Rebecca, and Jonathan G. Koomey. 2005. "Using energy scenarios to explore alternative energy pathways in California." *Energy Policy* 33(9): 1117 – 1142.

Gomez-Rodriguez, M., J. Leskovec, and A. Krause. 2010. Inferring Networks of Diffusion and

Influence. In *The 16th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*.

Grossback, Lawrence J., Sean Nicholson-Crotty, and David A. M. Peterson. 2004. “Ideology and Learning in Policy Diffusion.” *American Politics Research* 32(5): 521 –545.

Hanmer, Michael J., and Kerem Ozan Kalkan. 2013. “Behind the Curve: Clarifying the Best Approach to Calculating Predicted Probabilities and Marginal Effects from Limited Dependent Variable Models.” *American Journal of Political Science* 57(1): 263–277.

Harden, Jeffrey J., and Bruce A. Desmarais. 2011. “Linear Models with Outliers: Choosing Between Conditional-Mean and Conditional-Median Methods.” *State Politics & Policy Quarterly* 11(4): 371–389.

Imai, Kosuke, and Dustin Tingley. 2012. “A Statistical Method for Empirical Testing of Competing Theories.” *American Journal of Political Science* 56(1): 218–236.

Jensen, David D., and Peter R. Cohen. 2000. “Multiple Comparisons in Induction Algorithms.” *Machine Learning* 38(3): 309–338.

Krackardt, David. 1987. “{QAP} partialling as a test of spuriousness.” *Social Networks* 9(2): 171 – 186.

Leisch, Friedrich. 2004. “FlexMix: A General Framework for Finite Mixture Models and Latent Class Regression in R.” *Journal of Statistical Software* 11(10): 1–18.

May, Peter J. 1992. “Policy Learning and Failure.” *Journal of Public Policy* 12(October): 331–354.

Mooney, Christopher Z. 2001. “Modeling Regional Effects on State Policy Diffusion.” *Political Research Quarterly* 54(March): 103–124.

Newman, Mark. 2003. “The Structure and Function of Complex Networks.” *SIAM Review* 45(2): 167–256.

Nicholson-Crotty, Sean. 2009. “The Politics of Diffusion: Public Policy in the American States.” *The Journal of Politics* 71(01): 192–205.

Pacheco, Julianna. 2012. “The Social Contagion Model: Exploring The Role of Public Opinion on the Diffusion of Anti-Smoking Legislation across the American States.” *The Journal of Politics* 74(01): 187–202.

Shipan, Charles R., and Craig Volden. 2006. “Bottom-Up Federalism: The Diffusion of Antismoking Policies from U.S. Cities to States.” *American Journal of Political Science* 50(4): 825–843.

Sigelman, Lee. 2006. “Top Twenty Commentaries.” *American Political Science Review* 100(4): 667–687.

Simon, Herbert Alexander. 1976. *Administrative Behavior: A Study of Decision-Making Processes in Administrative Organization*. 3d ed ed. New York: Free Press.

Tan, Yue, and David H. Weaver. 2009. “Local Media, Public Opinion, and State Legislative Policies: Agenda Setting at the State Level.” *The International Journal of Press/Politics* 14(4): 454–476.

Volden, Craig. 2002. “The Politics of Competitive Federalism: A Race to the Bottom in Welfare Benefits?” *American Journal of Political Science* 46(April): 352–363.

Volden, Craig. 2006. “States as Policy Laboratories: Emulating Success in the Children’s Health Insurance Program.” *American Journal of Political Science* 50(2): 294–312.

Volden, Craig, Michael M. Ting, and Daniel P. Carpenter. 2008. “A Formal Model of Learning and

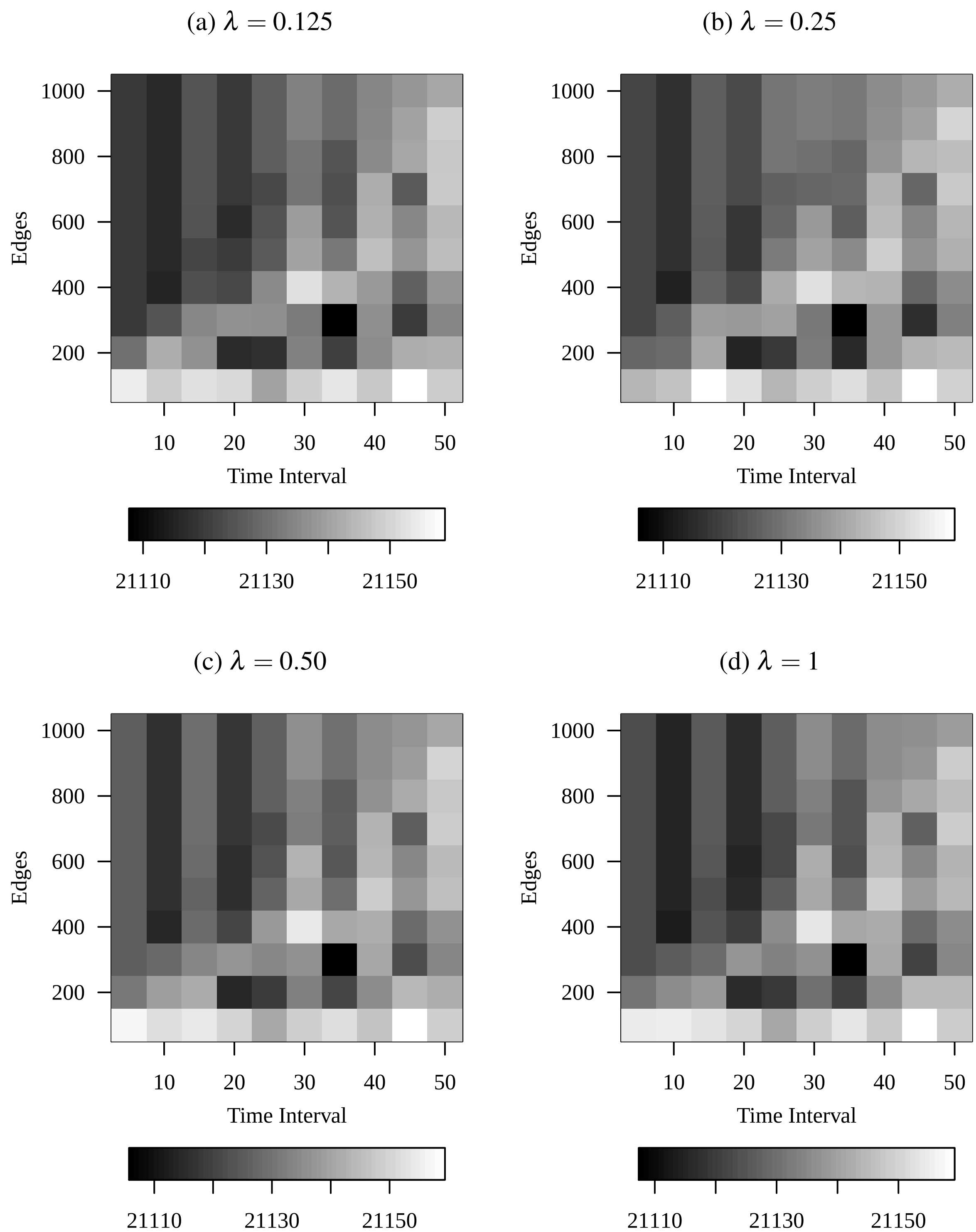
Policy Diffusion." *American Political Science Review* 102(03): 319–332.

Walker, Jack L. 1969. "The Diffusion of Innovations Among the American States." *American Political Science Review* 63(3): 880–899.

Ward, Michael D., Brian D. Greenhill, and Kristin M. Bakke. 2010. "The Perils of Policy by *p*-value: Predicting Civil Conflicts." *Journal of Peace Research* 47(4): 363–375.

Ward, Michael D., Randolph M. Siverson, and Xun Cao. 2007. "Disputes, democracies, and dependencies: a re-examination of the Kantian Peace." *American Journal of Political Science* 51(3): 583–601.

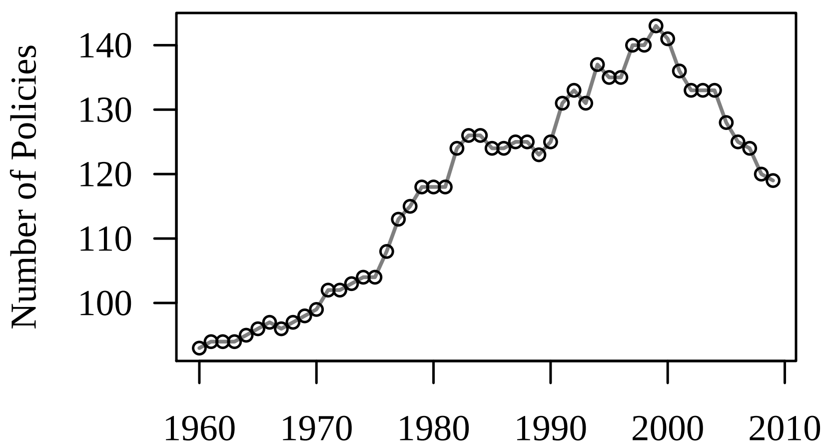
Figure 1: BIC of the Pooled Discrete Time Event History Models



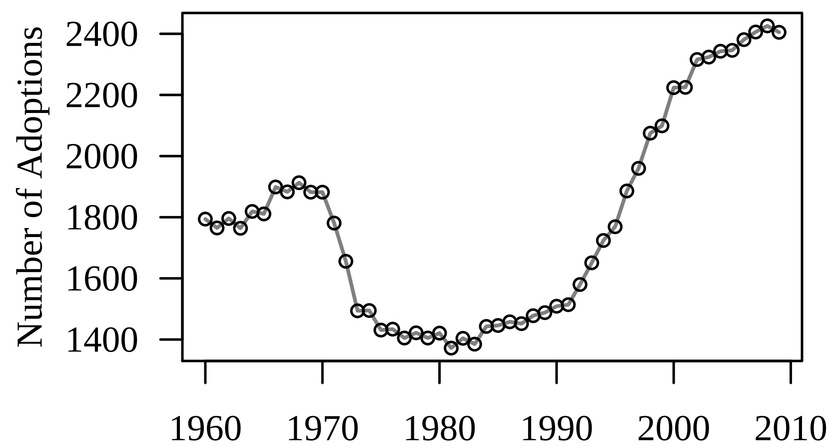
Note: There are 189 policies and 65,885 observations in each model.

Figure 2: Number of Policies and Adoption Instances by Year

(a) Policies Included

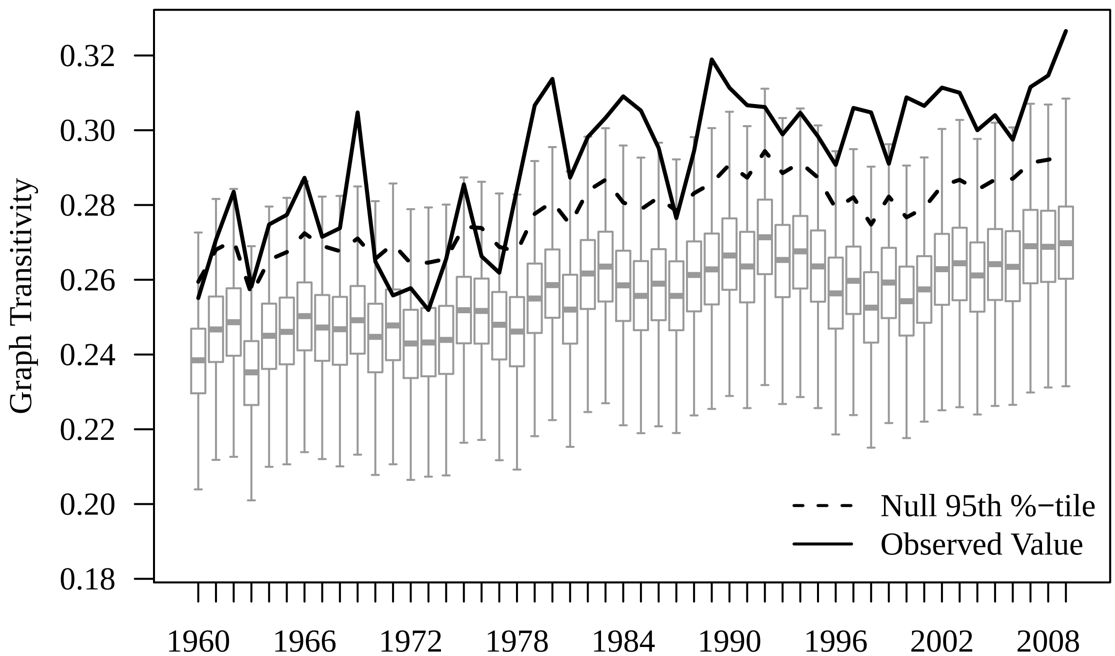


(b) Adoption Instances



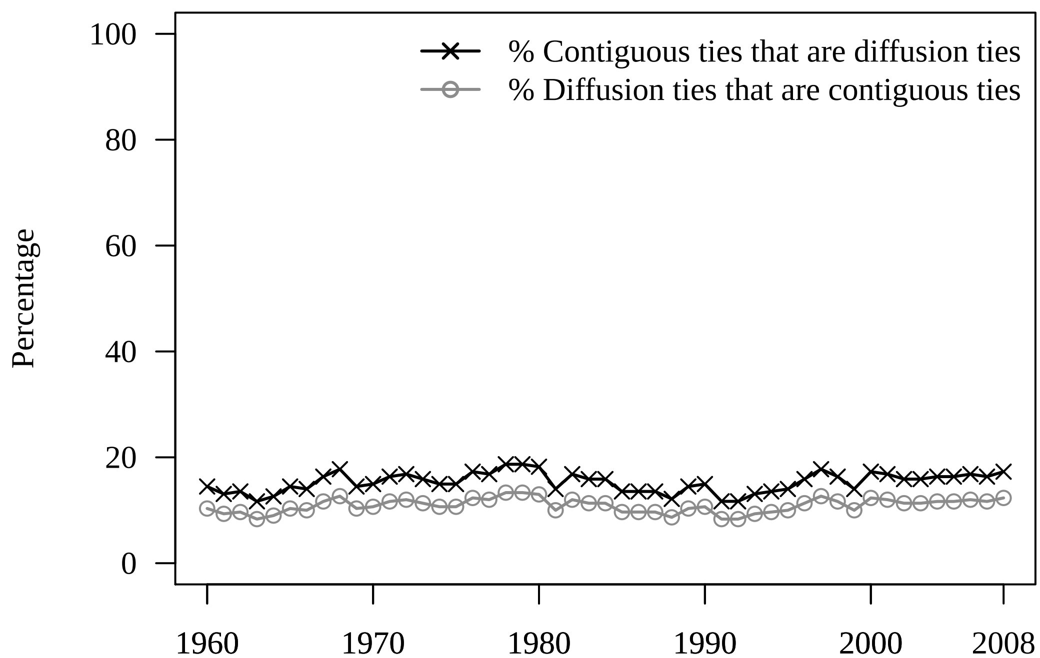
Note: The graphs present the number of policies (panel a) and number of adoption instances (panel b) included in the data used to infer the network in each year.

Figure 3: Test for Significantly High Transitivity Relative to a Network Formed Through Strictly Monadic Processes



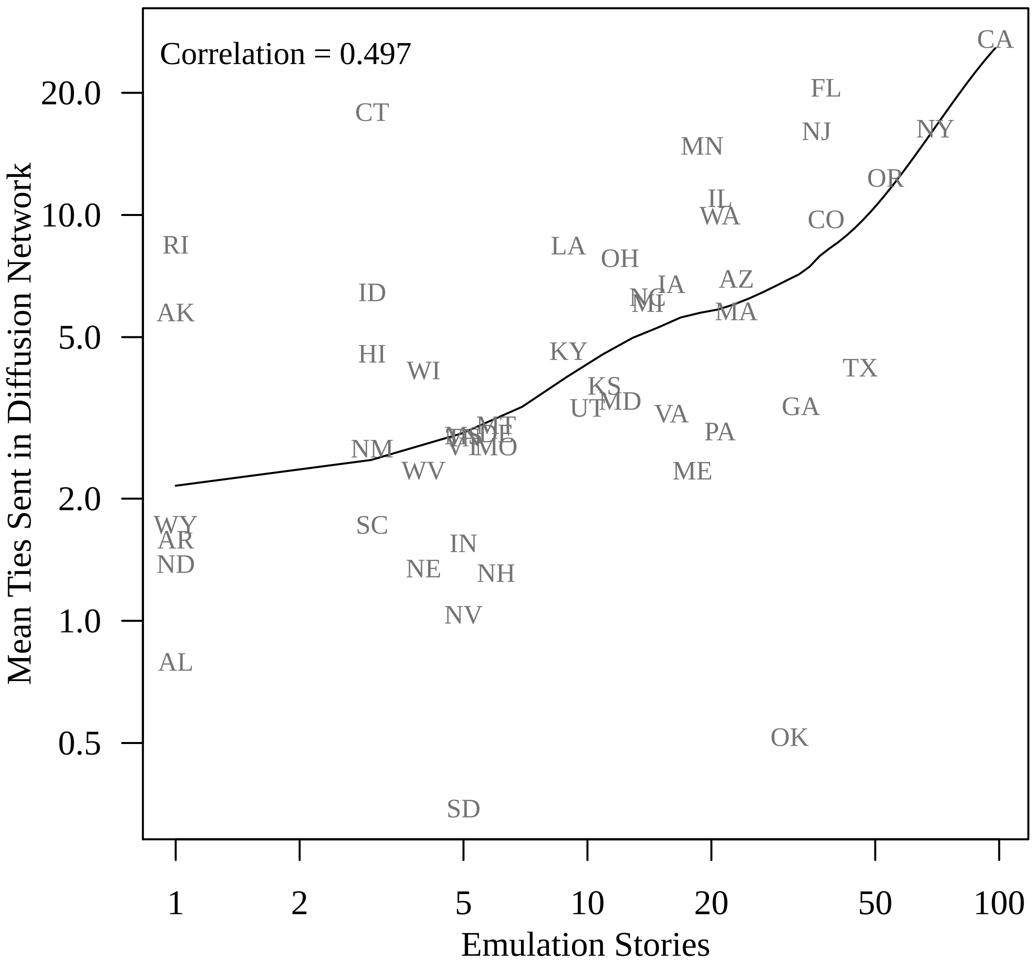
Note: Boxplots depict the distribution of simulated transitivity values under the null hypothesis of strictly monadic tie formation. A dashed line is drawn at the 95$^{\text{th}}$ percentile of simulated null transitivity values. The solid black line is drawn at the observed transitivity in the diffusion network.

Figure 4: Comparison of Diffusion Relations with Geographic Contiguity



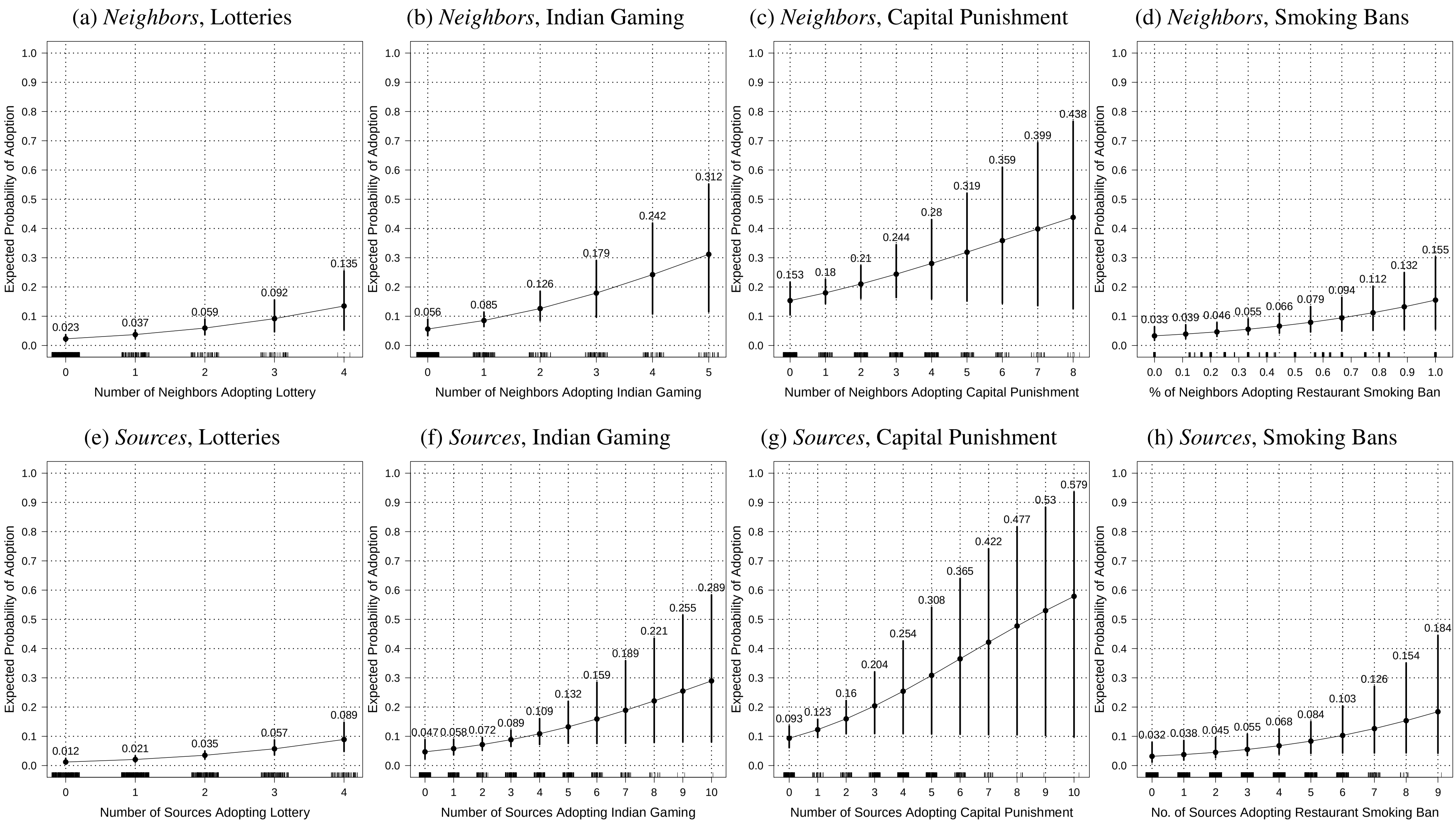
Note: The graph presents the percentage of contiguity relations between states that are identified as diffusion ties and the percentage of inferred diffusion ties that are between contiguous states.

Figure 5: Association Between Inferred Diffusion Ties and Media Reports of Emulation



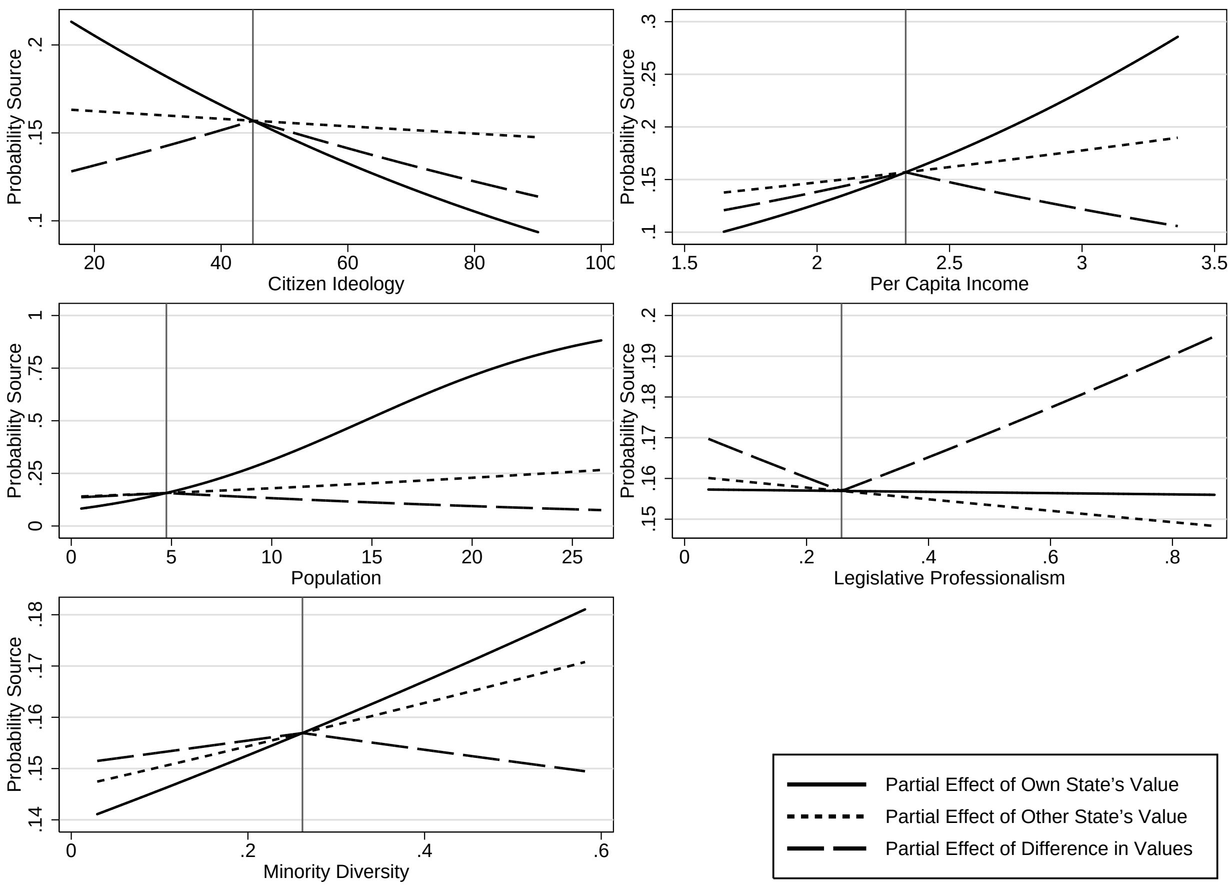
Note: Both axes are on the natural logarithm scale. Since NY and CA are large positive outliers on the linear scale, the correlation is also computed on the natural log scale. The correlation on the linear scale is 0.70. The line depicts a loess regression fit.

Figure 6: Average Marginal Effects of *Neighbors Adopting* and *Sources Adopting* on the Adoption of Four Policies



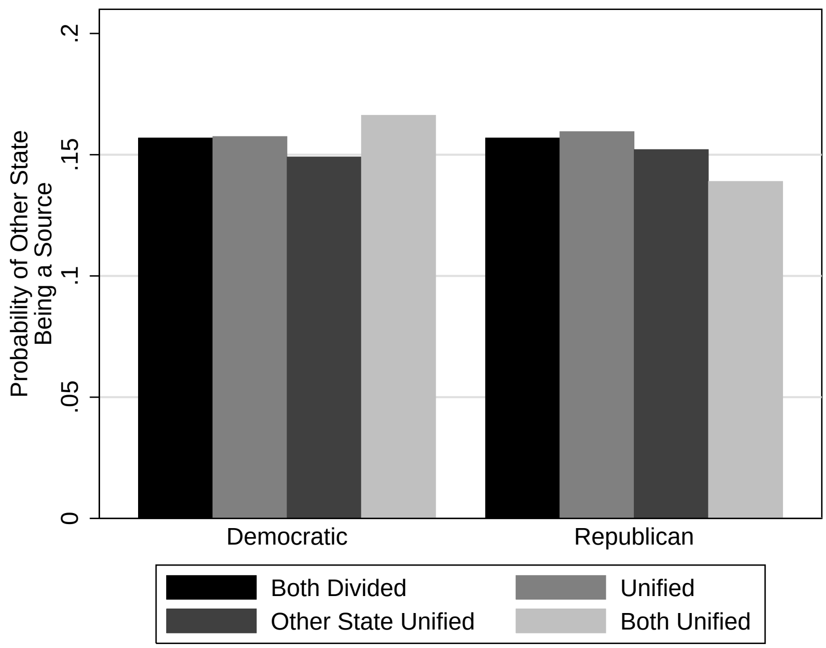
Note: The graphs present the average marginal effects of *Neighbors Adopting* (top row) and *Sources Adopting* (bottom row) on the adoption of four policies: lotteries (Berry and Berry 1990), Indian gaming (Boehmke 2005), capital punishment (Boehmke 2005), and restaurant smoking bans (Shipan and Volden 2006). Points represent expected probability point estimates and vertical lines represent 95% confidence intervals.

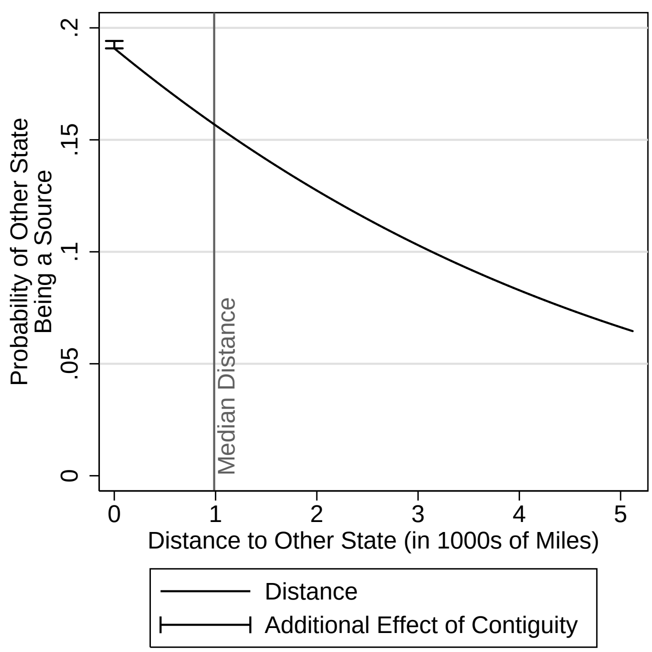
Figure 7: Estimated Substantive Effects of Absolute Difference Variables



Note: Figures generated using Model 2. All other variables set to their mean (continuous variables) or mode (binary variables) in 1985. Calculations set the random effects to zero.

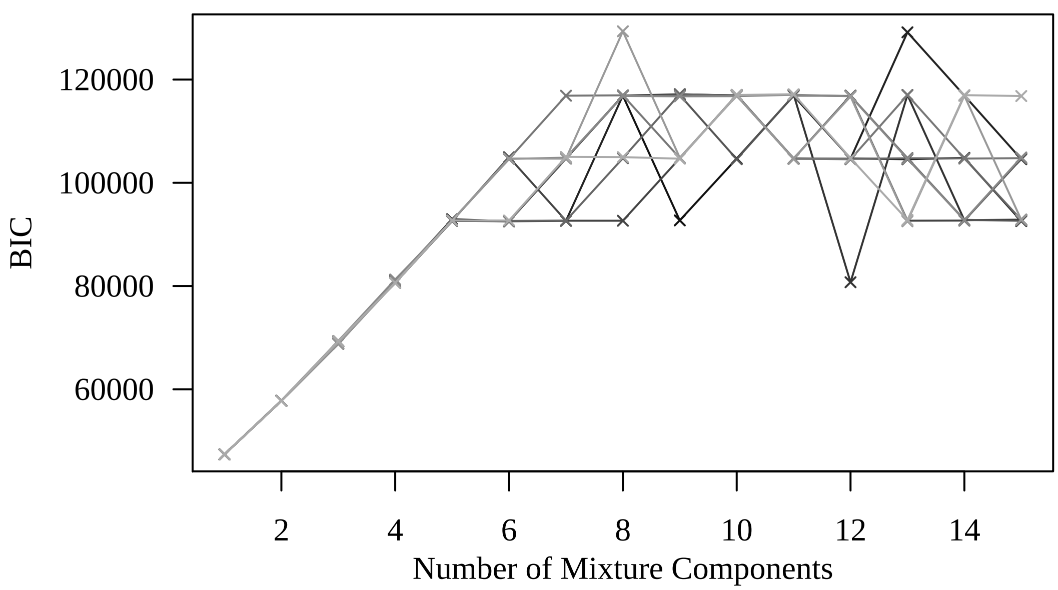
Figure 8: Estimated Substantive Effects of Selected Variables





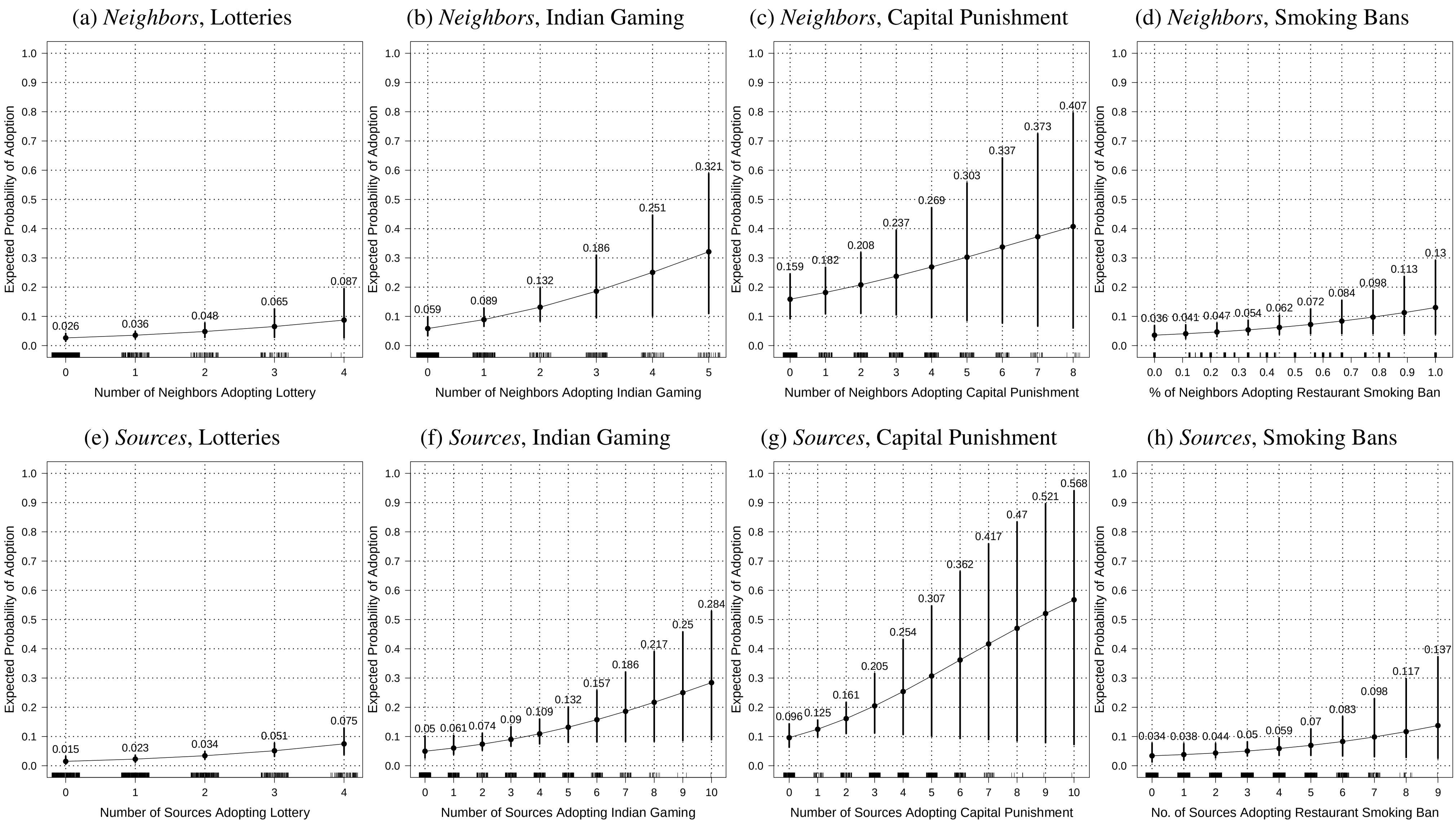
Note: Figures generated using Model 2. All other variables set to their mean (continuous variables) or mode (binary variables). Calculations set the random effects to zero.

Figure 9: Mixture Model Fit



Note: Different line shades correspond to different values of λ, the parameter that controls the entropy in the initial cluster assignment probabilities.

Figure 10: Average Marginal Effects of *Neighbors Adopting* and *Sources Adopting*, Controlling for the Other, on the Adoption of Four Policies



Note: The graphs present the average marginal effects of *Neighbors Adopting* (top row) and *Sources Adopting* (bottom row) on the adoption of four policies: lotteries (Berry and Berry 1990), Indian gaming (Boehmke 2005), capital punishment (Boehmke 2005), and restaurant smoking bans (Shipan and Volden 2006). Points represent expected probability point estimates and vertical lines represent 95% confidence intervals.

Table 1: States Ranked Based on the Total Number of Diffusion Ties Sent to Other States within Five-Year Periods

Rank	60–64	65–69	70–74	75–79	80–84	85–89	90–94	95–99	00–04	05–09
1	NY	NY	NY	NY	NY	FL	FL	CA	CA	CA
2	KY	KY	FL	FL	FL	NY	NY	CT	CT	CT
3	CA	SC	CO	NJ	NJ	CA	CA	NJ	FL	NJ
4	MN	AL	RI	MN	MN	MN	CT	FL	WA	FL
5	AL	CO	CT	OR	RI	OR	OR	NY	NJ	WA
6	SC	NM	MN	IL	OR	NJ	MN	MN	IL	IL
7	RI	MN	MI	CO	CO	RI	NJ	OR	MN	MN
8	MI	OH	NJ	AK	CA	CT	CO	WA	AZ	AZ
9	VT	NJ	NE	NH	AK	AK	OH	LA	IA	LA
10	NJ	WA	PA	RI	IL	IL	RI	CO	NC	IA
11	IL	MI	LA	AR	LA	CO	IL	IA	OR	OH
12	WA	RI	AL	CT	MI	ID	AK	AZ	CO	NC
13	MD	MD	OR	MI	CT	MI	LA	NC	HI	CO
14	OH	PA	MD	DE	ID	OH	MI	OH	LA	WI
15	MS	VT	AR	MS	PA	KS	ID	ID	OH	UT
16	AR	AR	IL	WY	MA	LA	MA	MI	ID	OR
17	LA	CA	VA	LA	IA	KY	NC	IL	NY	MA
18	MA	ME	NH	CA	KY	TN	IA	HI	VA	HI
19	NM	VA	SC	PA	NM	PA	KY	MA	MA	NY
20	GA	FL	MS	IA	KS	TX	MS	ME	MD	VA
21	AZ	CT	KY	ID	WI	MA	MT	RI	RI	ID
22	CT	NE	NM	GA	GA	MS	DE	ND	WI	TX
23	PA	LA	AK	KS	WY	AZ	KS	DE	UT	MI
24	ME	OR	DE	MO	NH	VT	PA	KY	GA	MD
25	WI	KS	MO	NM	AR	MD	WA	WI	TX	RI
26	CO	MA	KS	SC	DE	SC	MO	MO	WV	WV
27	FL	IL	CA	VT	MS	NM	TN	AK	IN	GA
28	NV	TN	TN	VA	TX	GA	TX	UT	MI	IN
29	UT	AK	WY	MA	UT	DE	GA	MD	KS	MO
30	KS	UT	IN	OH	VT	MT	ME	NE	MO	KY
31	NC	NH	VT	KY	OH	NC	VA	TX	VT	MT
32	IN	GA	ID	AL	SC	WA	AZ	VA	AK	ND
33	VA	MS	UT	UT	HI	WY	MD	MT	KY	KS
34	ID	AZ	WV	HI	MT	IA	WY	TN	MT	ME
35	TN	DE	GA	WV	AZ	MO	WV	VT	NM	SC
36	WV	IN	NC	TX	WA	HI	AR	IN	AL	VT
37	OR	NC	TX	WI	NC	VA	HI	NM	ME	DE
38	WY	MT	SD	SD	SD	WV	WI	GA	NE	NE
39	MO	WV	HI	NE	WV	UT	NE	WV	NV	NH
40	ND	NV	WA	WA	NV	ME	SC	MS	TN	NV
41	NE	WI	ME	ME	NE	AR	NH	NV	DE	TN
42	NH	WY	NV	ND	MD	OK	NM	AR	AR	MS
43	TX	HI	OK	TN	ME	AL	VT	KS	ND	AK
44	AK	SD	ND	AZ	OK	NV	IN	AL	NH	NM
45	IA	TX	IA	NC	TN	NH	OK	WY	MS	AL
46	SD	MO	MA	NV	AL	WI	ND	PA	SC	AR
47	HI	ND	AZ	MT	IN	NE	NV	SC	OK	OK
48	OK	IA	OH	OK	MO	SD	UT	NH	PA	PA
49	DE	ID	WI	MD	VA	IN	AL	OK	SD	SD
50	MT	OK	MT	IN	ND	ND	SD	SD	WY	WY

Table 2: States Ranked Based on the Total Number of Diffusion Ties Received From Other States within Five-Year Periods

Rank	60–64	65–69	70–74	75–79	80–84	85–89	90–94	95–99	00–04	05–09
1	MN	NE	MD	MD	DE	MO	FL	CA	NC	ME
2	AK	CA	HI	OH	MI	IL	LA	NC	CA	AZ
3	AR	AK	CO	NV	IA	GA	CA	AZ	IL	MD
4	CO	CO	IN	ME	MD	RI	DE	IN	VA	NV
5	OR	IL	NC	OR	OR	TX	WA	LA	AZ	CO
6	CA	NC	CA	AK	OH	OK	MN	CT	NV	DE
7	IL	DE	ID	DE	TX	TN	AZ	DE	OK	WA
8	ME	MD	VT	VA	RI	DE	GA	MI	IN	IL
9	OH	ID	IL	AR	NM	MA	IA	KS	GA	IN
10	ID	OH	AR	CO	NV	AL	ME	CO	ME	CT
11	MT	IN	KS	MT	HI	AZ	CO	TX	NH	NC
12	TN	MI	OK	ND	KY	OH	MI	FL	DE	NY
13	MA	AR	DE	RI	VA	WV	CT	NM	WI	LA
14	NE	IA	TN	TN	CT	CA	TN	WI	CT	TX
15	CT	WV	AK	ID	WI	NM	IL	RI	FL	NE
16	DE	MO	NM	IL	CO	NY	MO	ID	KS	NJ
17	FL	MA	IA	MI	ME	SC	NC	IL	RI	OK
18	VT	OK	MA	NJ	MN	CO	SC	NE	TX	WI
19	MD	VT	MI	NY	TN	FL	WY	OH	CO	NH
20	OK	TN	WV	TX	IL	KS	MA	VA	NE	GA
21	IA	AZ	NY	WI	ND	NC	NM	WA	MI	KY
22	PA	NH	ND	FL	NY	WA	NY	GA	MA	NM
23	RI	HI	NE	HI	PA	WY	OR	SC	TN	OR
24	MI	KS	NJ	IN	WA	IA	SD	MA	MD	AR
25	KS	OR	NV	KY	AK	OR	AR	SD	MO	HI
26	KY	CT	GA	NC	NC	AK	MT	OK	OR	MO
27	IN	PA	SD	WA	AR	MN	OH	UT	AK	PA
28	NC	MN	MT	MN	MT	VA	NJ	MS	HI	RI
29	TX	ND	RI	NM	OK	NJ	WV	VT	PA	WY
30	WV	SD	CT	IA	VT	IN	AL	MD	UT	CA
31	NY	WY	MN	WY	ID	ME	IN	MN	ID	TN
32	MS	GA	UT	LA	MA	WI	VA	NH	MS	OH
33	WI	MT	WA	UT	AL	MT	KS	MT	NM	MS
34	AZ	ME	NH	CT	CA	PA	UT	NJ	WA	MT
35	GA	NY	OH	PA	KS	CT	WI	TN	OH	VA
36	MO	RI	PA	VT	NJ	LA	RI	AK	AR	FL
37	NH	VA	SC	NE	LA	MS	VT	IA	LA	IA
38	NM	WA	TX	OK	AZ	HI	NE	MO	NJ	SD
39	SC	LA	AZ	CA	GA	VT	TX	NV	SC	KS
40	WA	WI	MO	KS	SC	KY	HI	WV	KY	MN
41	NJ	TX	VA	MA	UT	MI	OK	AR	MN	ND
42	LA	NM	WY	AZ	MO	MD	MD	HI	WV	VT
43	VA	KY	FL	GA	NE	NE	PA	OR	MT	AL
44	WY	NJ	OR	SC	FL	NH	KY	WY	VT	MA
45	SD	SC	WI	NH	WV	AR	MS	ME	AL	UT
46	UT	FL	AL	WV	WY	ND	ND	PA	NY	AK
47	AL	MS	KY	MS	IN	ID	AK	AL	SD	SC
48	NV	AL	ME	MO	MS	NV	NH	ND	WY	MI
49	HI	NV	LA	AL	SD	SD	ID	NY	ND	WV
50	ND	UT	MS	SD	NH	UT	NV	KY	IA	ID

Table 3: Model Fit Statistics with *Neighbors Adopting* or *Sources Adopting*

	AIC		BIC		Cross-validated % Correctly Classified	
Policy	*Neighbors*	*Sources*	*Neighbors*	*Sources*	*Neighbors*	*Sources*
Lotteries	195.12	191.02	233.15	229.05	8%	23%
Indian Gaming	144.25	144.45	241.68	237.98	76%	87%
Capital Punishment	204.53	200.58	283.31	279.35	3%	6%
Restaurant Smoking Bans	248.57	249.36	328.36	329.14	16%	28%

Cell entries report three model fit statistics computed with *Neighbors Adopting* or *Sources Adopting* included in the specification for the adoption models of lotteries (Berry and Berry 1990), Indian gaming (Boehmke 2005), capital punishment (Boehmke 2005), and restaurant smoking bans (Shipan and Volden 2006). The first two columns report AIC. The next two columns report BIC values. The last two columns report the percentage of observations correctly classified by each specification, computed via leave-one-out cross-validation.

Table 4: Multi-Level Logit Models of State Policy Diffusion Ties

Follower State Characteristics:				
Citizen Ideology			-0.013^{**}	(0.002)
Legislative Professionalism			-0.012	(0.230)
Minority Diversity			0.538^{*}	(0.233)
Per Capita Income			0.745^{**}	(0.086)
Population			0.170^{**}	(0.011)
Unified Democratic Government			0.004	(0.035)
Unified Republican Government			0.020	(0.041)
Potential Source Characteristics:				
Citizen Ideology			-0.002	(0.001)
Legislative Professionalism			-0.109	(0.216)
Minority Diversity			0.316^{+}	(0.187)
Per Capita Income			0.224^{**}	(0.076)
Population			0.031^{**}	(0.009)
Unified Democratic Government			-0.060^{+}	(0.034)
Unified Republican Government			-0.037	(0.039)
Relative Follower/Source Characteristics:				
Contiguous	0.190^{**}	(0.034)	0.021	(0.040)
Distance	-0.263^{**}	(0.017)	-0.240^{**}	(0.020)
Citizen Ideology (Absolute Difference)			-0.008^{**}	(0.001)
Legislative Professionalism (Absolute Difference)			0.429^{**}	(0.134)
Minority Diversity (Absolute Difference)			-0.180^{+}	(0.104)
Per Capita Income (Absolute Difference)			-0.442^{**}	(0.047)
Population (Absolute Difference)			-0.038^{**}	(0.004)
Unified Democratic (Product)			0.125^{**}	(0.046)
Unified Republican (Product)			-0.125	(0.083)
Constant	0.217^{**}	(0.025)	0.230^{**}	(0.033)
σ_{u_1} (Follower Random Effect)	0.809^{*}	(0.082)	0.828^{+}	(0.089)
σ_{u_2} (Potential Source Random Effect)	0.217^{**}	(0.025)	0.230^{**}	(0.033)
N	122,500		94,080	

Observations are dyadic. The dependent variable indicates whether potential source state is a source for a follower state. Source state values are taken from the predictive network with 300 edges over 35 years of policy adoptions. $^{+}$ $p < 0.10$; * $p < 0.05$; ** $p < 0.01$.

Table 5: Estimates for *Neighbors Adopting* and *Sources Adopting* in the Replication Models

	Neighbors (Original Model)	*Sources*	*Neighbors* and *Sources*
Berry and Berry (1990): Lotteries (Probit, N = 857)			
Neighbors Adopting	0.27*		0.17
	(0.09)		(0.10)
Sources Adopting		0.29*	0.23*
		(0.09)	(0.09)
AIC	195.12	191.02	190.24
BIC	233.15	229.05	233.02
Boehmke (2005): Indian Gaming (Probit, N = 364)			
Neighbors Adopting	0.42*		0.42*
	(0.17)		(0.19)
Sources Adopting		0.20+	0.21*
		(0.10)	(0.10)
AIC	144.25	144.45	143.54
BIC	241.68	237.98	244.86
Boehmke (2005): Capital Punishment (Probit, N = 227)			
Neighbors Adopting	0.16		0.21+
	(0.10)		(0.11)
Sources Adopting		0.13	0.17
		(0.13)	(0.13)
AIC	204.53	200.58	201.70
BIC	283.31	279.35	283.89
Shipan and Volden (2006): Restaurant Smoking Bans (Logit, N = 807)			
% Neighbors Adopting	1.92*		1.54
	(0.86)		(0.95)
Sources Adopting		0.24+	0.18
		(0.14)	(0.15)
AIC	248.57	249.36	249.02
BIC	328.36	329.14	333.50

Cell entries report coefficient estimates and standard errors (in parentheses) for *Neighbors Adopting* and *Sources Adopting* and AIC and BIC values in three specifications of the replication models. * $p < 0.05$; + $p < 0.10$.

Table 6: Model Fit Statistics with and without *Sources Adopting*

	LR Test	AIC		Cross-validated % Correctly Classified	
Policy		No *Sources Adopting*	*Sources Adopting*	No *Sources Adopting*	*Sources Adopting*
Lotteries	6.88^{*}	195.12	190.24	8%	22%
Indian Gaming	2.72^{+}	144.25	143.54	76%	76%
Capital Punishment	4.84^{*}	204.53	201.70	3%	5%
Restaurant Smoking Bans	1.55	248.57	249.02	16%	18%

Cell entries report three model fit statistics computed with and without *Sources Adopting* included in the specification for the adoption models of lotteries (Berry and Berry 1990), Indian gaming (Boehmke 2005), capital punishment (Boehmke 2005), and restaurant smoking bans (Shipan and Volden 2006). The first column reports likelihood ratio χ^2 test statistics comparing the two models (with 1 degree of freedom). The next two columns report AIC values. The last two columns report the percentage of observations correctly classified by each specification, computed via leave-one-out cross-validation. * $p < 0.05$; $^{+}$ $p < 0.10$.

Table 7: Fixed Effects Logit Models of State Policy Diffusion Ties

Follower State Characteristics:				
Citizen Ideology			-0.014^{**}	(0.002)
Legislative Professionalism			-0.049	(0.234)
Minority Diversity			0.448^{+}	(0.241)
Per Capita Income			0.738^{**}	(0.088)
Population			0.191^{**}	(0.012)
Unified Democratic Government			0.011	(0.035)
Unified Republican Government			0.015	(0.041)
Potential Source Characteristics:				
Citizen Ideology			-0.005^{**}	(0.002)
Legislative Professionalism			-0.323	(0.243)
Minority Diversity			-0.032	(0.243)
Per Capita Income			0.084	(0.088)
Population			0.082^{**}	(0.014)
Unified Democratic Government			-0.047	(0.035)
Unified Republican Government			-0.047	(0.040)
Relative Follower/Source Characteristics:				
Contiguous	0.175^{**}	(0.034)	-0.000	(0.040)
Distance	-0.284^{**}	(0.018)	-0.266^{**}	(0.020)
Citizen Ideology (Absolute Difference)			-0.008^{**}	(0.001)
Legislative Professionalism (Absolute Difference)			0.383^{**}	(0.136)
Minority Diversity (Absolute Difference)			-0.222^{*}	(0.104)
Per Capita Income (Absolute Difference)			-0.422^{**}	(0.047)
Population (Absolute Difference)			-0.041^{**}	(0.004)
Unified Democratic (Product)			0.126^{**}	(0.046)
Unified Republican (Product)			-0.121	(0.083)
constant	-2.390^{**}	(0.120)	-2.682^{**}	(0.216)
N	122,500		94,080	

Observations are dyadic. The dependent variable indicates whether potential source state is in fact a source for a follower state. Source state values are taken from the predictive network with 300 edges over 35 years of policy adoptions. Fixed effects for each potential source and each potential follower are included as a set of indicator variables, but not reported. $^{+}$ $p < 0.10$; * $p < 0.05$; ** $p < 0.01$.