

Persistent Policy Pathways: Inferring Diffusion Networks in the American States*

Bruce A. Desmarais[†] Jeffrey J. Harden[‡] Frederick J. Boehmke[§]

May 30, 2014

Under Revise & Resubmit at *The American Political Science Review*

Abstract

Policy diffusion has been a focus of scholars studying both national and subnational governments for the last half century. In the American context, diffusion is commonly conceptualized as an explicitly dyadic process whereby states adopt policies from other states. This dyadic diffusion process implies the existence of a *policy diffusion network* connecting the states. Using a dataset consisting of 187 policies, we introduce and apply algorithms capable of inferring this network based on persistent patterns of diffusion. In addition to presenting and applying the algorithms for the first time in political science, we offer three substantive contributions to state policy diffusion research. First, we summarize and analyze the structure of the inferred diffusion network. Second, we demonstrate how the network can improve conventional statistical models of state policy adoption. Third, we model the inferred diffusion pathways in the network to test a variety of theoretical expectations about the policy connections among states.

*Complete replication materials for the analyses presented here will be made available online upon publication.

[†] 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

One critical element influencing the policy choices that governments make is the set of choices made by other peer governments. A considerable amount of scholarship demonstrates the processes by which policies diffuse across national and subnational boundaries (for a review, see Shipan and Volden 2012). For instance, trade liberalization policy (Meseguer 2006), hospital finance policy (Gilardi, Füglistner and Luyet 2009), and even armed conflict (Most and Starr 1980) have been shown to disseminate from country to country. Moreover, the institution of federalism provides an ideal environment for such processes by encouraging member governments to learn from one another. The American states represent an important example of such an environment (e.g., Walker 1969; Gray 1973; Berry and Berry 1990; Shipan and Volden 2006; Boushey 2010).

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, while more recent work has moved beyond the foundational monadic models of policy adoption to characterize policy diffusion as a mixed process of independent adoption and dyadic emulation (Volden 2006; Boehmke 2009).¹

Although theory on policy diffusion is well developed, empirical operationalizations of diffusion pathways remain rudimentary. Empirically, diffusion ties are nearly always assumed to exist exclusively between geographically contiguous states. Equating geographic contiguity with a diffusion connection is a reasonable 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. Contiguity emphasizes economic forces. For example, the policies in neighboring states might facilitate

¹Note that the term “dyadic” in our general discussion refers to pairs of states, as it relates to the network analytic understanding of the term. However, “dyadic EHA models,” which commonly appear in this literature, refer to event history models in which the dependent variable measures whether one state moves policy toward or away from every other previously adopting state.

movement by people to buy lottery tickets (Berry and Baybeck 2005) or to move for more generous welfare benefits (e.g., Volden 2002). Neighbors also cooperate to assure regional consistency in policy regimes. Experience with policy in neighboring states by citizens can lead to public opinion spillover effects (Pacheco 2012). 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).

Despite the focus in the literature on geographic contiguity, diffusion ties regularly form between states dispersed throughout the country. For example, California is considered both a prolific policy innovator in general (Volden 2006) and a leader in energy and environmental policy specifically (Ghanadan and Koomey 2005). The states of New Jersey and Maryland have both recently implemented policies explicitly modeled after energy and emissions policies in California (Nussbaum 2007; Wagner 2007). This represents a coast-to-coast instance of diffusion that would not be captured via contiguity. Several non-geographic forces facilitate diffusion, such as social learning or comparison to peer networks or states facing similar policy problems.

Recently, scholars have begun to look anew for broader forms of policy diffusion by examining, for example, whether states emulate the policies of states that have proven successful in addressing the underlying policy problem. Most prominent in this line of work is Volden's (2006) introduction of dyadic event history analysis. This modeling approach adheres to the assumption in Gray's (1973) model of state policy diffusion that policymakers across the states are "completely intermixed" (1176), meaning that all current adopters of a policy have the potential to influence all states that have not yet adopted. In dyadic event history analysis, the characteristics of both adopters and non-adopters are used to model implicit diffusion via sequential policy adoption.

Simultaneously, scholars have returned to exploiting information on the timing of policy adoptions across samples of policies. This broadens our ability to learn about policy innovativeness and

diffusion by moving from policy-specific results to learning about consistent trends across large databases of policies (e.g., Nicholson-Crotty 2009; Boushey 2010) and states (Boehmke and Skinner 2012*b*). The availability of such extensive data opens the door to evaluating proposed diffusion ties more broadly. If we think of a diffusion tie as linking two states across which policy innovations commonly diffuse in sequence, then a state should be more likely to adopt a given policy once the states to which it has tied have adopted that policy. We are not the first to have postulated that diffusion patterns could be represented as a network. In discussing possible extensions to her model in which every state influences every other state, Gray (1973, 1176) noted that:

More elaborate models could be constructed. . . in which there is incomplete mixing of the population, e.g., regional or professional communication networks may produce distinctive diffusion patterns.

Following this line of reasoning, we argue that patterns in policy diffusion can be used to infer the network such that states become increasingly likely to adopt a policy as more of the states to which they are connected (via diffusion pathways) adopt that policy. This conjecture underpins the core objective of the current research; we use data on policy diffusions to directly infer the latent diffusion network connecting the states. This latent network provides the first (to our knowledge) measure of state-to-state policy diffusion influence. We show that the introduction of this latent network represents a critical advancement in the study of policy diffusion. Indeed, seminal works in this literature have alluded to, but never measured, the network we infer (e.g., Walker 1969; Gray 1973; Berry and Berry 1990). Moreover, while we focus on state policy diffusion, the technology we use for network inference has applications in a variety of settings in political science.

In what follows we demonstrate the significance of our policy diffusion network in detail. We first describe our method for constructing the network: a recently developed machine learning algorithm that can be used to infer a latent diffusion network from data consisting of binary diffusion “cascades.” Next we present our application of diffusion network inference to state policy adoptions. Then we illustrate the use of the inferred diffusion network in conventional monadic

policy adoption studies. Finally, we 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 or countries) based only on data recording the times at which those units adopted or were infected with some attribute (e.g., a policy), over several attributes. Two examples are data on when a collection of people fell ill over several ailments or 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. Information on policy adoption for several states or countries and several policies also constitutes data of this type. Here we use Gomez-Rodriguez, Leskovec and Krause’s (2010) latent network inference algorithm, called `NetInf`, to infer policy diffusion networks connecting the American states.

The `NetInf` algorithm is derived and described in detail in the online appendix. Here we give a broad overview of its major steps. The inferential task addressed by `NetInf` is the identification of a latent, directed network (i.e., each tie has a sender and a receiver) that can be used to explain a dataset with several cascades, where each cascade is a recording of when units (e.g., states) exhibited some dichotomous attribute (e.g., a policy adoption). Each cascade is stylistically represented as a tree, in which there is a branch for each diffusion instance whereby the attribute (e.g., policy) spreads from the origin (i.e., sender) of the branch to the destination (i.e., receiver). The network being inferred constrains the trees that can be used to construct the cascade such that only edges in the network can be used to construct the trees. The network is tied to the set of cascades in that the algorithm will attempt to find edges that can be used in trees to explain many cascades. The structure of this algorithm actually fits quite closely with Walker’s (1969) description of the ideal way to represent state-to-state policy diffusion:

At the top of the tree would be a set of pioneering states which would be linked together

in a national system of emulation and competition. The rest of the states would be sorted out along branches of the tree according to the pioneer, or set of pioneers, from which they take their principal cues (892–893).

Here we apply `NetInf` to a moving window of policy adoptions on the 187 policies included in the database introduced by Boehmke and Skinner (2012b) to infer an evolving state-to-state policy diffusion network for the years 1960–2009. Before presenting our application further, we define some useful terminology. We infer a different network in each year (t). The diffusion ties (i.e., edges) that we infer 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*. 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

Our approach permits the structure of diffusion pathways 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 base our approach on how the networks and measures computed on them would likely be used in future research. We expect, and later 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, statistical models 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 avoid endogeneity in the use of the network at t to predict adoption at t , 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 . An edge from i to j at t can be interpreted as indicating that the policy has frequently spread from j to 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, via the inferred network, to predict that same policy adoption at time t . Below we address the question of how many years preceding t should be used to infer the network for time t .²

²There may be concern that we infer one diffusion network at each time point, which models

2.2 NetInf Parameter Tuning

We set three parameters in the network inference procedure. First, we need to define the number of preceding years of adoptions (denoted k) 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. A policy can only diffuse from i to j if there is an edge from i to j in the inferred network. The exponential distribution gives the distribution of diffusion times between states, provided that there is an edge connecting them. 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. This prevents any given adoption by one state that happens to fall later in time than adoption by another state from contributing to the formation of a tie between the two states.

We take a data-driven approach to finding optimal values of these parameters. 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 $\{k, E, \lambda\}$, we fit a pooled (across all policies in the data) 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 $t - 1$, the number of s 's sources in the network that have adopted p .
3. *Policy Area*: A dummy variable that models the unique rate of adoption for each policy.

the diffusion of *all* policy adoptions within the respective time window. Indeed, some types of policies may diffuse in systematically different patterns than do other types of policies. In the online appendix we present diagnostics to evaluate whether there exist multiple classes of policies that systematically affect the ties inferred in the diffusion networks. We find very strong evidence that there are *not* multiple classes of diffusion patterns in our dataset of policies.

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 $\{k, E, \lambda\}$. We search over $\lambda \in \{0.125, .25, .5, 1\}$, which corresponds to mean diffusion times of 8, 4, 2, and 1 years, respectively, $k \in \{5, 10, \dots, 50\}$, and $E \in \{100, 200, \dots, 1000\}$. 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). 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.³ The fit is not particularly sensitive to the rate parameter, but the network using a rate of 0.5 results in the best fit. This means that policies diffuse, on average, in two years. An average of approximately 1900 adoption instances over an average of approximately 120 policies is used to infer the network for each year.⁴

3 Descriptive Analysis of the Policy Diffusion Network

In this section we conduct descriptive and exploratory analyses of the networks we have inferred to evaluate their structures. First, we demonstrate that the network is quite distinct from a set of relations recording geographic contiguity. Second, we summarize the outgoing and incoming diffusion ties of each state over five-year periods. Third, we provide an external empirical validation of the network by comparing it to newspaper reports of state-to-state emulation during the same time period.

3.1 Geographic Contiguity

The first descriptive feature of the diffusion networks that we consider is whether they are accurately approximated by a network of geographic contiguity relations among states. Figure 1 plots the percentage of contiguity relations between states that are identified as diffusion ties (black

³We also use a network based on 400 edges and 10-year periods for use in one application to a policy adoption model (see below).

⁴The online appendix presents the complete model fit results from the grid search over `NetInf` parameters as well as the number of adoption instances and policies used for each network-year.

line) and the percentage of inferred diffusion ties that are between contiguous states (gray line). 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. Therefore, although geographic contiguity represents a good first start, ties between neighboring states do *not* comprise a comprehensive proxy for the policy diffusion network.

[Insert Figure 1 here]

3.2 State-Level Activity in Diffusion Pathways

Ranking states based on their innovativeness is a research problem that dates back at least to Walker (1969). We now present the top 15 states based on the number of states to which they send diffusion ties (Table 1) over five-year periods. In their time-aggregated measures of policy innovativeness, Walker (1969) and Boehmke and Skinner (2012*b*) 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. In terms of leading policy innovators, the state that emerges in our analysis as an outlier with respect to previous rankings is Florida. Walker (1969) and Boehmke and Skinner (2012*b*) rank Florida as 13th and 12th, respectively, whereas we find Florida to be in the top five for nearly every five year period, and at the top of the list for a decade.

[Insert Table 1 here]

To venture an explanation as to why Florida emerges as an innovator in our analysis, but not in previous studies, we present Table 2, which breaks down both how often each of the three top innovators (New York, California, and Florida) were first adopters, and also how often the other two did not adopt. We see from this table that, even though Florida is the least frequent first adopter among the three, the policies for which it is the first adopter are, at a very high rate, never adopted by New York or California. Thus, although Florida does not stand out as a notably frequent first adopter, it is often placed at the root of cascade trees because other frequent adopters are not

innovators in policy areas led by Florida. This inference regarding Florida highlights a primary strength of Net Inf : *a state will not be deemed innovative based solely on the speed with which it adopts policies*. Rather, a state is deemed innovative if its adoption serves to explain adoptions by other states that cannot be explained with reference to other early adopters.

[Insert Table 2 here]

3.3 Media-based Validation of the Policy Diffusion Networks

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 often indicate when a substantial portion of a state law has been modeled after another state’s policy. We identified accounts of policy emulation in journalistic coverage of state policymaking by searching LexisNexis Academic for newspaper articles containing the phrase, “modeled after a/an ***”, where “***” was the name of a state, for all fifty states. LexisNexis covers newspaper articles going back to 1981. From the search results we derived a count of the number of stories that report the emulation of each states’ policies. These documented instances of policy emulation can serve as the basis for 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.

Figure 2 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, we find a strong correlation of $r = 0.70$. However, two outliers—New York and California—have approximately twice as many emulation stories as any other state, so we also consider the correlation on the log-scale, which produces a slightly more moderate correlation of 0.497. Both the Pearson’s correlation coefficient and Spearman’s rank-based correlation reach statistical significant at the 0.01 level. 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.

[Insert Figure 2 here]

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).⁵

Two primary contributions come from our application of the inferred diffusion networks to models of policy adoption. First, we illustrate how the diffusion networks can be integrated into conventional adoption models and demonstrate that their use improves the performance of those models. The second contribution stems from the fact that `NetInf` does not condition on covariates, making it possible that the ties inferred by `NetInf` arise from some underlying covariates that induce regular patterns of policy diffusion. By observing whether the use of the inferred networks improves models of policy adoption, we test whether the inferred ties are simply an artifact of the covariates already known to influence policy adoption, or if our diffusion networks are substantively important for diffusion models.⁶

⁵Specifically, 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.

⁶We validated this characteristic of `NetInf` with a simulation experiment in which we generated policy adoption data based solely on state covariates. `NetInf` produced network estimates

In addition to these policy-specific event history analysis (EHA) models we replicate Boehmke and Skinner’s (2012a) “pooled event history analysis” (PEHA) model by combining data on 151 different policies diffusing over the period 1960–1999 (see also Boehmke 2009). This approach stacks the data from different policies and estimates a unified model with a common set of independent variables (including state, year, and policy fixed effects). Pooling the data does result in fewer independent variables than for any single policy, but it provides insight into what factors affect diffusion most broadly across the issue spectrum of American politics. We show below that information from our inferred diffusion networks is one of those factors.

4.1 Model Details

We focus on these five models for several reasons. First, the four policy-specific models represent a wide variety of policies, and by definition the pooled model represents an even wider range. This provides the opportunity to examine whether the diffusion networks we infer have a broad or narrow, policy-specific impact on adoption. Second, the original studies presenting the policy-specific models are well-known in the policy diffusion literature, having each garnered at least 50 citations according to Google Scholar.⁷ Finally, the models all use similar EHA empirical specifications, enhancing comparability. 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.⁸

The theoretical frameworks behind our replication models each have their own unique characteristics. To conserve space, we refer readers to the original studies for detailed discussions of each that were consistent with patterns in covariate values, which indicates that consistent effects of covariates can give the appearance of diffusion ties between states.

⁷In 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).

⁸The Berry and Berry (1990) and Boehmke (2005) models are estimated with probit and the Shipan and Volden (2006) and Boehmke and Skinner (2012a) models are estimated with logistic regression.

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 geographic contiguity. Nearly all studies of policy diffusion include in their models either the number of or percentage 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. As with the quotes from Walker (1969) and Gray (1973) given above, this discussion—found in a seminal study from the state policy diffusion literature—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 are no “reliable data about which states are perceived... to be regional leaders in a policy area” (Berry and Berry 1990, 403).

4.1.1 Including Network Information

Our inferred policy diffusion networks provide those data that previous scholars of policy diffusion have not had available. In fact, beyond simply measuring regional leaders, the networks give 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 networks 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 networks to produce a list of states

that influence the state in a specified time period immediately preceding a given year.⁹ This list represents all of that state's sources at that time. Next, to create the variable *Sources Adopting* we count the number of states from that list that have previously adopted the policy.¹⁰ After creating this variable, we then add it to each of the five replication models.¹¹

4.1.2 Estimates and Model Fit

We first examine the extent to which the inclusion of *Sources Adopting*—instead of or in addition to *Neighbors Adopting*—improves model fit.¹² Table 3 reports coefficient estimates and standard errors for the two variables as well as model fit statistics for three specifications: (1) the original model with *Neighbors Adopting* (plus the authors' other covariates), (2) a model with

⁹As mentioned above, we constructed a version using 35-year periods and one with 10-year periods. Results between the two are substantively similar. For each model we used the version that produced the lowest AIC value. For all policies besides Indian gaming, we used the 35-year version.

¹⁰This 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 use a count measure in all of our replications because it is the most commonly used in this literature.

¹¹We include all policies in the construction of the networks used to produce *Sources Adopting*, including the policy of interest in the EHA model. Recall from above that we avoid endogeneity problems because we only use adoptions that occurred before a given year to measure the network for that year. 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.

¹²The question of whether *Sources Adopting* should replace or complement *Neighbors Adopting* is context-dependent. We focus on model fit here, but theoretical expectations should also be an important guide.

Sources Adopting substituted for *Neighbors Adopting* (plus the other covariates), and (3) a model with both *Neighbors Adopting* and *Sources Adopting* (plus the other covariates). In all cases the coefficients are positive (as expected), though statistical significance varies somewhat across specifications and replications. We assess the substantive impact of these effects in section 4.1.3.

[Insert Table 3 here]

To compare model fit we compute AIC, BIC, and cross-validated percent correctly classified. We compute this last measure via leave-one-out cross-validation, which involves iteratively dropping one observation, estimating the model, computing an expected probability 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.¹³ In Table 3, the values in bold indicate the best-fitting model according to each statistic.

The AIC and BIC values support the inclusion of *Sources Adopting* in all but the restaurant smoking ban model, where the original model and the model with *Sources Adopting* produce AIC and BIC values within 2 units of each other (indicating equal fit, see Burnham and Anderson 2002). The cross-validated percent correctly classified measure also generally supports the inclusion of *Sources Adopting*. In four of the five replication models the percent correctly classified in one or both models with *Sources Adopting* increases from the original model with *Neighbors Adopting* (the restaurant smoking ban model is again the lone exception). These improvements are somewhat small in magnitude—ranging from +1 to +3 percentage points across the different models. Nonetheless, they consistently point to the models that include *Sources Adopting* in the specification as the best fit.

¹³Cross-validation methods are common in other fields and have recently become more prominent in political science (e.g., Ward, Greenhill and Bakke 2010).

Overall, Table 3 provides good evidence that *Sources Adopting* can improve the fit of policy diffusion EHA models, either in place of or in addition to *Neighbors Adopting*. Importantly, across the five models, *none* of the fit statistics decisively selects the original model with *Neighbors Adopting* as the better fit. 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.

4.1.3 Marginal Effects

We examine the substantive implications of including *Sources Adopting* in Figure 3 by graphing the average marginal effects of *Neighbors Adopting* (top row) and *Sources Adopting* (bottom row) in each model on the probability scale.¹⁴ All estimates are computed from the specifications that include either *Neighbors Adopting* or *Sources Adopting*.¹⁵

[Insert Figure 3 here]

The first point to note from Figure 3 is the effect of the count of *Neighbors Adopting* (lotteries, Indian gaming, capital punishment, and pooled model) 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 increase in the probability of adoption. The magnitude and level of uncertainty varies somewhat across the models, but the effect is consistently in the positive direction.

Moving to the bottom row of Figure 3, note that when substituted for *Neighbors Adopting*, the effect of *Sources Adopting* is also positive in all five models; as the number of sources adopting the policy increases, so too does probability of a state adopting the policy. From the minimum (0) to

¹⁴We 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.

¹⁵Results with both included in the same model are substantively similar (see the online appendix).

the maximum (lotteries: 7, Indian gaming: 10, capital punishment: 10, restaurant smoking bans: 9, pooled model: 15) of *Sources Adopting*, the probability of adoption increases by the following percentage points, on average: 24 (lotteries), 24 (Indian gaming), 48 (capital punishment), 15 (restaurant smoking bans), and 13 (pooled model). As with the effect of *Neighbors Adopting*, the confidence intervals indicate varying degrees of uncertainty around these estimates.¹⁶ Nonetheless, these graphs show that *Sources Adopting* exerts a substantively significant positive impact on the probability of adoption across many different policies.

Moreover, these positive effects remain even after controlling for *Neighbors Adopting* (see the online appendix). 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 from four specific policy areas and a 151-policy pooled model in which states utilize a persistent set of diffusion sources to guide their policymaking decisions.

5 Understanding the Inferred Network

Having demonstrated that accounting for previous adoption activity by source states in the policy diffusion network improves a number of existing event history analyses of state policy diffusion we now seek to evaluate the structure of this network through the lens of extant theoretical expectations about the identities of leaders and followers. To do so we specify logit models to explain source-recipient ties over the period 1960–2009.

Incomplete information underpins Walker’s (1969) theory of policy diffusion and much of the subsequent research (e.g., May 1992; Mooney 2001; Volden 2006). States do not have the time or resources to fully evaluate all possible policies solutions to their pressing policy problems. Walker and others therefore suggest that states may act according to Simon’s (1976) concept of satisficing, in which they attempt to identify policies that will improve their lot even if they may not constitute

¹⁶This is at least partially due to the fact that policy adoption models tend to have many independent variables (the median is 19 in the four policy-specific replications). Each new variable adds more overall error to the model, because each coefficient is estimated with uncertainty.

the optimal policy. To accomplish this, states rely on a set of heuristics to identify policies for possible adoption. Most importantly, states will look to the actions of other states as a source of information. These may be neighboring states, states with similar characteristics and therefore similar policy needs, or states with more extensive resources that act as leaders by investigating new policies that have not yet been widely adopted.

We follow Walker's (1969) slack resources approach to understanding states' ability to investigate new policies or to learn about existing policies adopted by other states. 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. To this we add the role of legislative professionalism, which diffusion scholars have more recently used as a measure of legislative capacity (see, e.g., Shipan and Volden 2006).

Since previous EHA studies overwhelmingly focus on monadic policy diffusion, scholars typically estimate the effect of slack resources on policy adoption to test whether greater resources lead states to adopt new policies faster. Because we seek to explain their effect on the diffusion network, however, we have the opportunity to separate their distinct effects on leaders and followers. If diffusion occurs according to an informational process, then the slack resources approach suggests that states that score high on such resources will tend to be leaders since they can investigate policies on their own more thoroughly. This same logic also suggests that states with greater resources can also process more information and consider policy solutions in more states simultaneously. We therefore expect states with more resources to be more likely to a source, but also to identify other states as sources.

Beyond resource effects, however, we also want to capture Walker's idea of peer states. When identifying sources, states may look beyond the wealthiest states to states that have similar characteristics and whose choices may reflect more upon their specific circumstances. The identity of peer states likely goes beyond the concepts connected to slack resources, however, so we also consider the role of factors for which similarity may matter in and of itself. In particular, we consider the similarity between states in terms of ideology and racial diversity. Ideology plays a crucial role

in the types of policies states seek to adopt. With incomplete information, then, states may look to the policies adopted by ideologically similar states rather than to those of dissimilar states since the former has a greater chance of providing a solution consistent with the preferences of its citizens. A number of studies have demonstrated the important role that ideology plays in determining whether a state will copy the policy adopted by another state (e.g., Grossback, Nicholson-Crotty and Peterson 2004; Volden 2006; Volden, Ting and Carpenter 2008). We also consider the role of racial and ethnic diversity. States with more heterogeneous populations face distinct policy challenges so we expect that states will use diversity in defining their peer network.

The most studied concept of peer states remains the geographic-based one. While 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. 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 cross border economic activity as citizens search for desired goods or services (Berry and Baybeck 2005; Baybeck, Berry and Siegel 2011). While contiguity remains the workhorse variable for interstate diffusion, we also want to leverage the fact that our network considers the relationship between all pairs of states to examine the role of geographic proximity above and beyond contiguity. To do this we include a measure of distance between states' capitals to test whether states have a regional tendency when determining their peers.

In order to test for the effects of slack resources and similarity on the leader-follower relationship, we include variables corresponding to each and enter them into our model in three ways. We start with variables on total state population and income from the Bureau of Economic Affairs, legislative professionalism from King (2000),¹⁷ Berry, Ringquist, Fording and Hanson's (1998) state citizen ideology measure (the revised 1960–2008 series) as well as partisan control of state government from Klarner (2003),¹⁸ and racial diversity using Hero and Tolbert's (1996) formula applied

¹⁷We use this instead of Squire's (2007) measure because it goes back to the 1960s.

¹⁸These data come from <http://www.indstate.edu/polisci/klarnerpolitics.htm>.

to Census data. For each variable, we include its value in the potential source state to model which states tend to be emulated, its value in the potential recipient state to capture the tendency of states to identify sources, and as a relative measure using either the absolute difference between the values in potential source and recipient states (for continuous variables) or the product of their values (for the partisan control variables). We expect that the first three measures of slack resources have positive effects; that the relative measures of ideology and diversity exert negative effects (since larger values correspond to greater difference between the two states); and that shared borders and geographic proximity have positive effects. Of course, similarity likely extends beyond ideology and diversity, so we also expect that the absolute difference between these variables has a negative effect. We have no specific expectation about the role of ideology on its own in the source or recipient state.

In order to evaluate these predictions, we estimate a multi-level, over-time, logit model of the diffusion network.¹⁹ In accordance with the structure of this network, 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 ef-

¹⁹At 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) uses 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.

fects for each year.²⁰ Finally, recall that as we noted in the previous section the `NetInf` algorithm does not condition on underlying covariates. As such, anything that would predispose two states to prefer the same policies might induce the appearance of diffusion ties among them. Measures of partisanship and political ideology would be chief among these common exposures when it comes to policymaking, which suggests some initial caution in interpreting these results.

[Insert Table 4 here]

We report the results of this estimation in Table 4. Overall, these results indicate the importance of slack resources, political similarity and geographic proximity. 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 sources and recipients are not statistically significant and the difference term has a positive effect, which is only significant according to the parametric p -values, 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

²⁰We recognize that network data may exhibit more complex dependencies than directed vertex random effects (Ward, Siverson and Cao 2007; Cranmer and Desmarais 2011). As such, we used quadratic assignment procedure (Krackardt 1987)—a permutation testing method designed for network data—to replicate the hypothesis tests presented in Table 4. The QAP was run for 500 iterations. We use the variant of QAP in which the rows and columns of the adjacency-matrix-valued dependent variable are permuted.

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, but the effect is only statistically significant according to the parametric p -values. No effects emerge among unified Republican states.

In order to substantively interpret these coefficient estimates, we present a series of graphs that translate them into expected probabilities that another state is chosen as a source. We first examine the variables that have an absolute difference interpretation in Figure 4. 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.

[Insert Figure 4 here]

Consider first the top left graph 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 choosing sources the probability of identifying another state as a source decreases when the state becomes more liberal and increases when it becomes more conservative. A similar result appears 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. In contrast, the effect of making it more conservative would lead to a decrease, though this effect would be less severe than the effect of distance on its own. In terms of magnitude, the effects are large relative to the baseline probability that the hypothetical potential source is chosen as a peer (about 15%). Indeed, the partial effects range from zero to about 30% relative to the baseline.

The other graphs show similar patterns for per capita income, population, and minority diversity, though the magnitudes do differ quite a bit, with population showing very large effects for a handful of large states and diversity producing a relatively small effect (note that the scales of the graphs differ to enhance readability). 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 is 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 and hope to explore them in future work.

We interpret the other variables in Figure 5. 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 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.

[Insert Figure 5 here]

6 Conclusions

A considerable amount of research over the last fifty years examines the causes and consequences of policies diffusing across national and subnational boundaries. However, until now scholars have not had an ideal means of measuring the precise patterns through which policies are expected to diffuse. Geographic contiguity is one important factor, but does not capture the complete network of policy diffusion. In this paper, we employ a new methodology to infer such a network from diffusion data.

Using a recently compiled database of adoptions for over 150 state-level policies and a network inference algorithm, we infer latent policy diffusion networks connecting the American states. We offer three broad contributions related to the inferred networks. First, in contrast to common assumptions in the literature on policy diffusion, we find that the overwhelming majority of diffusion ties connect states that are not geographic neighbors.

Second, we 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 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 inclusion of a source adoption variable improves model fit and its effect is statistically significant and generally comparable in magnitude to adoption by contiguous states.

Third, we present modeling results that explain the ties in the inferred networks. This analysis provides support for a number of theoretical perspectives on diffusion. Perhaps most interestingly, the results 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 demographic and political features. We also find evidence of leadership, with larger and wealthier states more often chosen as sources.

The current research opens the door to several future directions. First, and chief among them, is extending the `NetInf` algorithm to simultaneously infer covariate-based commonalities in policy adoptions as well as the underlying diffusion network. This would present the opportunity to clearly differentiate between diffusion ties and patterns that are attributable to covariates. A second worthwhile extension of `NetInf` would be to incorporate whether policy innovations succeed or fail by some metric, which would allow us to evaluate the degree to which diffusion depends upon the result of the innovation. Third, our tracing of diffusion accounts in LexisNexis is rather limited, but demonstrates the feasibility of defining diffusion networks through the broad-based analysis of textual sources. Finally, though our work utilizes what is, to our knowledge, the most comprehensive database of state policy adoptions currently available, there are many more policies that could

be traced through the states. An expanded policy database would permit fine-grained inference of policy-specific diffusion networks connecting the states and to identify possibly different structures by policy type.

A final contribution of our analysis is the introduction of the `NetInf` technology to political science. While we illustrate its applicability in the context of state policy adoption, it has potential for use in other areas as well. For instance, diffusion studies are not limited to the states; several works examine how policy travels across national boundaries as well (e.g., Most and Starr 1980; Meseguer 2006; Gilardi, Füglistner and Luyet 2009). Furthermore, the `NetInf` algorithm could be useful for other areas of research with data that exhibit the “cascade” structure described here. This might include how media sources pick up stories from each other (e.g., Hamilton 2011), how support for political candidates or the choice to participate in politics travels through citizens’ social networks (e.g., Sinclair 2012; Makse and Sokhey 2013), or legislative cue-taking (e.g., Matthews and Stimson 1975). Political science routinely confronts the fact that the individual decisions of a collection of actors affect, and are affected by, other actors. We show here that inferring a network reflecting those decisions can play a crucial role in understanding political processes.

References

- Baybeck, Brady, William D. Berry and David A. Siegel. 2011. "A Strategic Theory of Policy Diffusion Via Intergovernmental Competition." *Journal of Politics* 73(1):232–247.
- 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(4):505–519.
- Berry, William D., Evan J. Ringquist, Richard C. Fording and Russell L. Hanson. 1998. "Measuring Citizen and Government Ideology in the American States, 1960–93." *American Journal of Political Science* 42(1):327–348.
- 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." *Journal of Politics* 71(3):1125–1140.
- Boehmke, Frederick J. and Paul Skinner. 2012a. "The Determinants of State Policy Innovativeness." Presented at the 12th Annual Conference on State Politics and Policy, February 16–18, 2012, Houston, TX.
- Boehmke, Frederick J. and Paul Skinner. 2012b. "State Policy Innovativeness Revisited." *State Politics & Policy Quarterly* 12(3):303–329.
- Boushey, Graeme. 2010. *Policy Diffusion Dynamics in America*. New York: Cambridge University Press.
- Burnham, Kenneth P. and David R. Anderson. 2002. *Model Selection and Multi-Model Inference*. 2nd ed. New York: Springer.
- Cranmer, Skyler J. and Bruce A. Desmarais. 2011. "Inferential Network Analysis with Exponential Random Graph Models." *Political Analysis* 19(1):66–86.

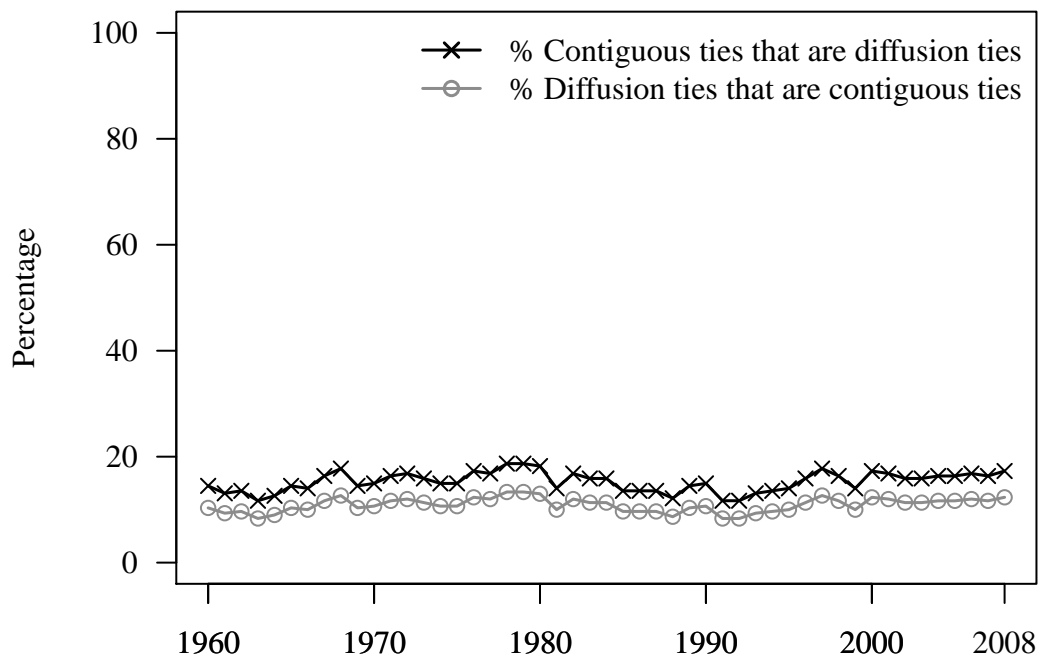
- Ghanadan, Rebecca and Jonathan G. Koomey. 2005. "Using Energy Scenarios to Explore Alternative Energy Pathways in California." *Energy Policy* 33(9):1117–1142.
- Gilardi, Fabrizio, Katharina Füglister and Stephane Luyet. 2009. "Learning from Others: The Diffusion of Hospital Financing Reforms in OECD Countries." *Comparative Political Studies* 42(4):549–573.
- 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)*.
- Gray, Virginia. 1973. "Innovation in the States: A Diffusion Study." *American Political Science Review* 67(4):1174–1185.
- 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.
- Hamilton, James T. 2011. *All the News That's Fit to Sell: How the Market Transforms Information into News*. Princeton, NJ: Princeton University Press.
- 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.
- Hero, Rodney E. and Caroline J. Tolbert. 1996. "A Racial/Ethnic Diversity Interpretation of Politics and Policy in the States of the U.S." *American Journal of Political Science* 40(3):851–871.
- King, James D. 2000. "Changes in Professionalism in U. S. State Legislatures." *Legislative Studies Quarterly* 25(2):327–343.
- Klarner, Carl. 2003. "The Measurement of the Partisan Balance of State Government." *State Politics & Policy Quarterly* 3(3):309–319.
- Krackardt, David. 1987. "{QAP} Partialling as a Test of Spuriousness." *Social Networks* 9(2):171–186.
- Makse, Todd and Anand E. Sokhey. 2013. "Not in My Front Yard! The Displaying of Yard Signs

- as a Form of Political Participation.” Forthcoming, *Political Behavior*.
- Matthews, Donald R. and James A. Stimson. 1975. *Yeas and Nays: Normal Decision-Making in the U.S. House of Representatives*. New York: Wiley.
- May, Peter J. 1992. “Policy Learning and Failure.” *Journal of Public Policy* 12(4):331–354.
- Meseguer, Covadonga. 2006. “Rational Learning and Bounded Learning in the Diffusion of Policy Innovations.” *Rationality and Society* 18(1):35–66.
- Mooney, Christopher Z. 2001. “Modeling Regional Effects on State Policy Diffusion.” *Political Research Quarterly* 54(1):103–124.
- Most, Benjamin A. and Harvey Starr. 1980. “Diffusion, Reinforcement, Geopolitics, and the Spread of War.” *American Political Science Review* 74(4):932–946.
- Nicholson-Crotty, Sean. 2009. “The Politics of Diffusion: Public Policy in the American States.” *Journal of Politics* 71(1):192–205.
- Nussbaum, Alex. 2007. “Scientists: State Must Act on Warming Now.” *The Washington Post*. February 21.
- Pacheco, Julianna. 2012. “The Social Contagion Model: Exploring The Role of Public Opinion on the Diffusion of Anti-Smoking Legislation across the American States.” *Journal of Politics* 74(1):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.
- Shipan, Charles R. and Craig Volden. 2012. “Policy Diffusion: Seven Lessons for Scholars and Practitioners.” *Public Administration Review* 72(6):788–796.
- 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*. New York: Free Press.
- Sinclair, Betsy. 2012. *The Social Citizen: Peer Networks and Political Behavior*. Chicago: Uni-

versity of Chicago Press.

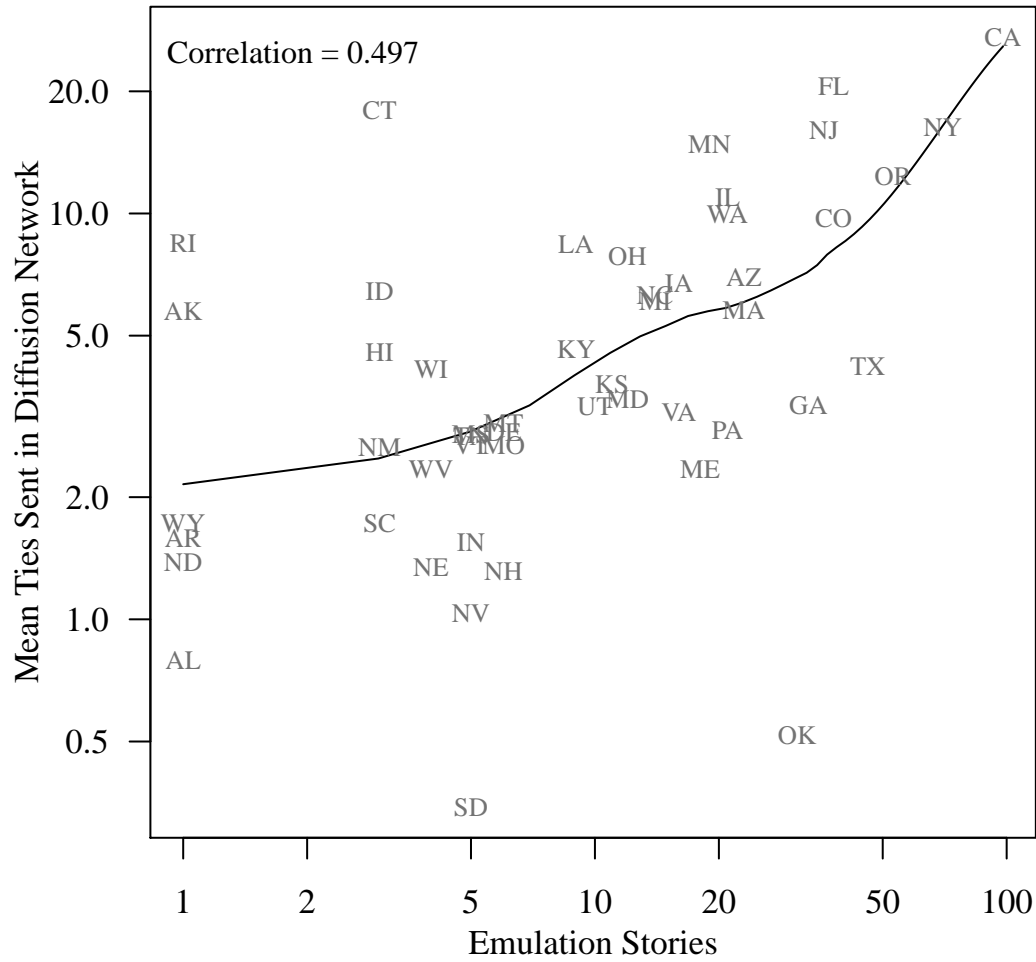
- Squire, Peverill. 2007. "Measuring State Legislative Professionalism: The Squire Index Revisited." *State Politics & Policy Quarterly* 7(2):211–227.
- Tan, Yue and David H. Weaver. 2009. "Local Media, Public Opinion, and State Legislative Policies: Agenda Setting at the State Level." *The International 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(2):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(3):319–332.
- Wagner, John. 2007. "Bills to Cut Pollution, Aid Bay, are Signed into Law." *The Washington Post*. April 25.
- 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: Comparison of Diffusion Relations with Geographic Contiguity



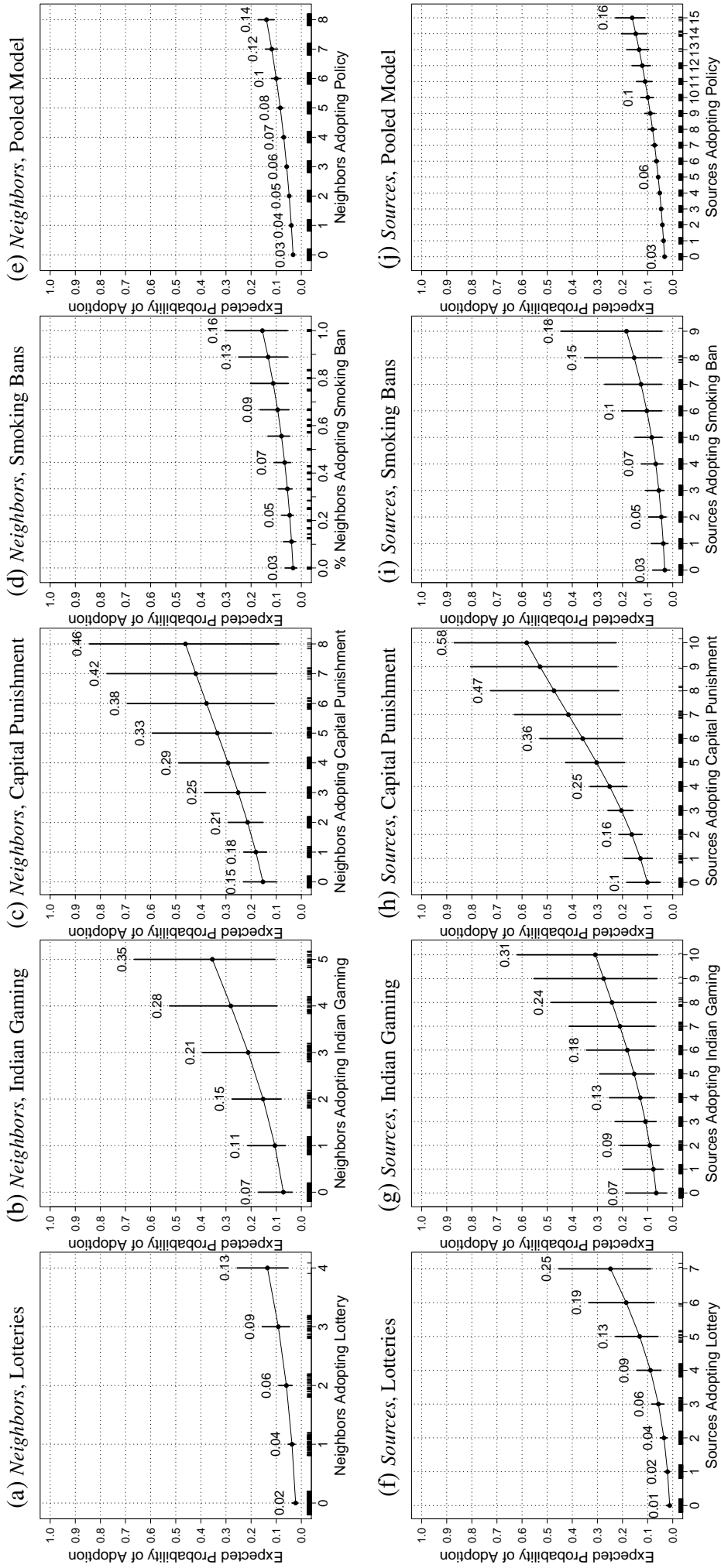
Note: The graph presents the percentage of contiguity relations between states that are identified as diffusion ties and the percentage of inferred diffusion ties that are between contiguous states.

Figure 2: Association Between Inferred Diffusion Ties and Media Reports of Emulation



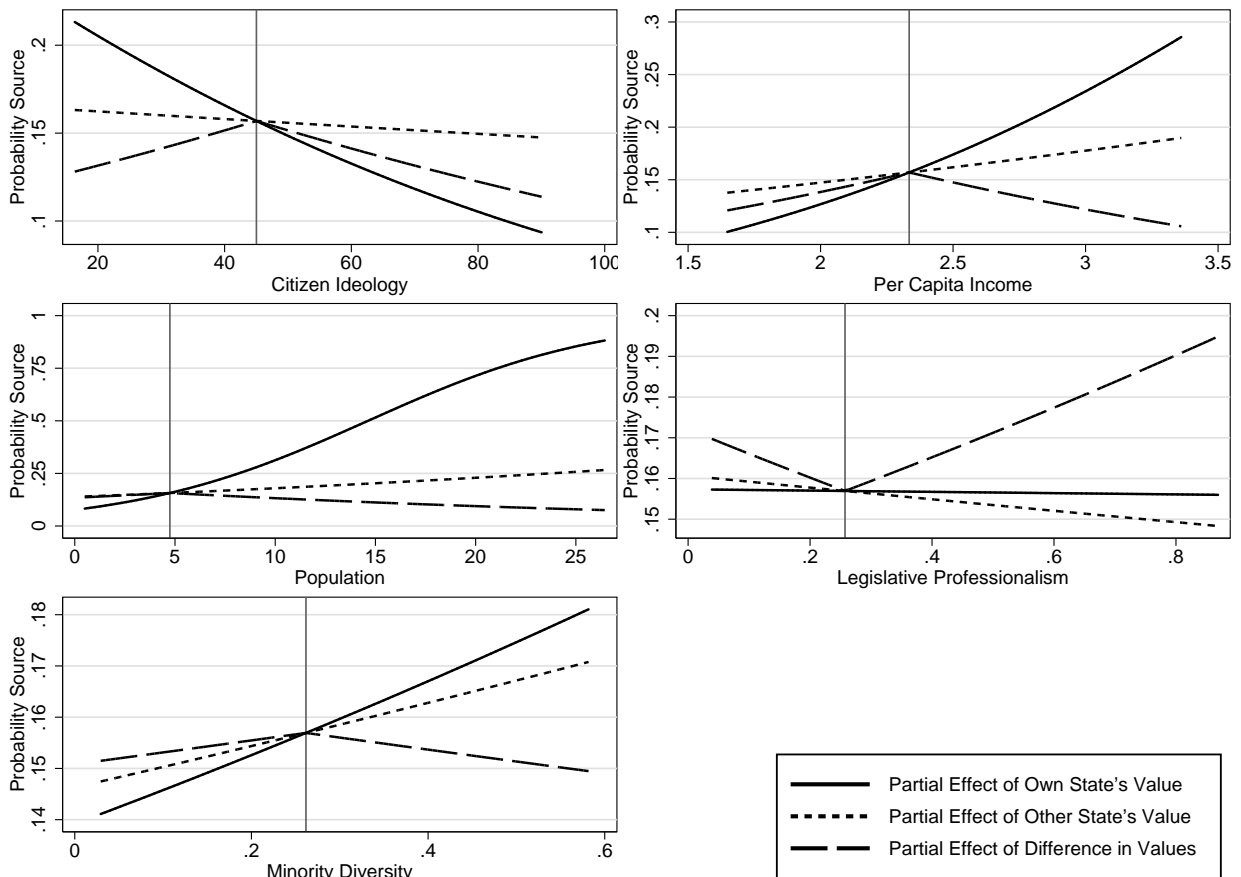
Note: Both axes are on the natural logarithm scale. Since New York and California are large positive outliers on the linear scale, the correlation is also computed on the natural log scale. The correlation on the linear scale is 0.70. The line depicts a loess regression fit.

Figure 3: Average Marginal Effects of *Neighbors Adopting* and *Sources Adopting*



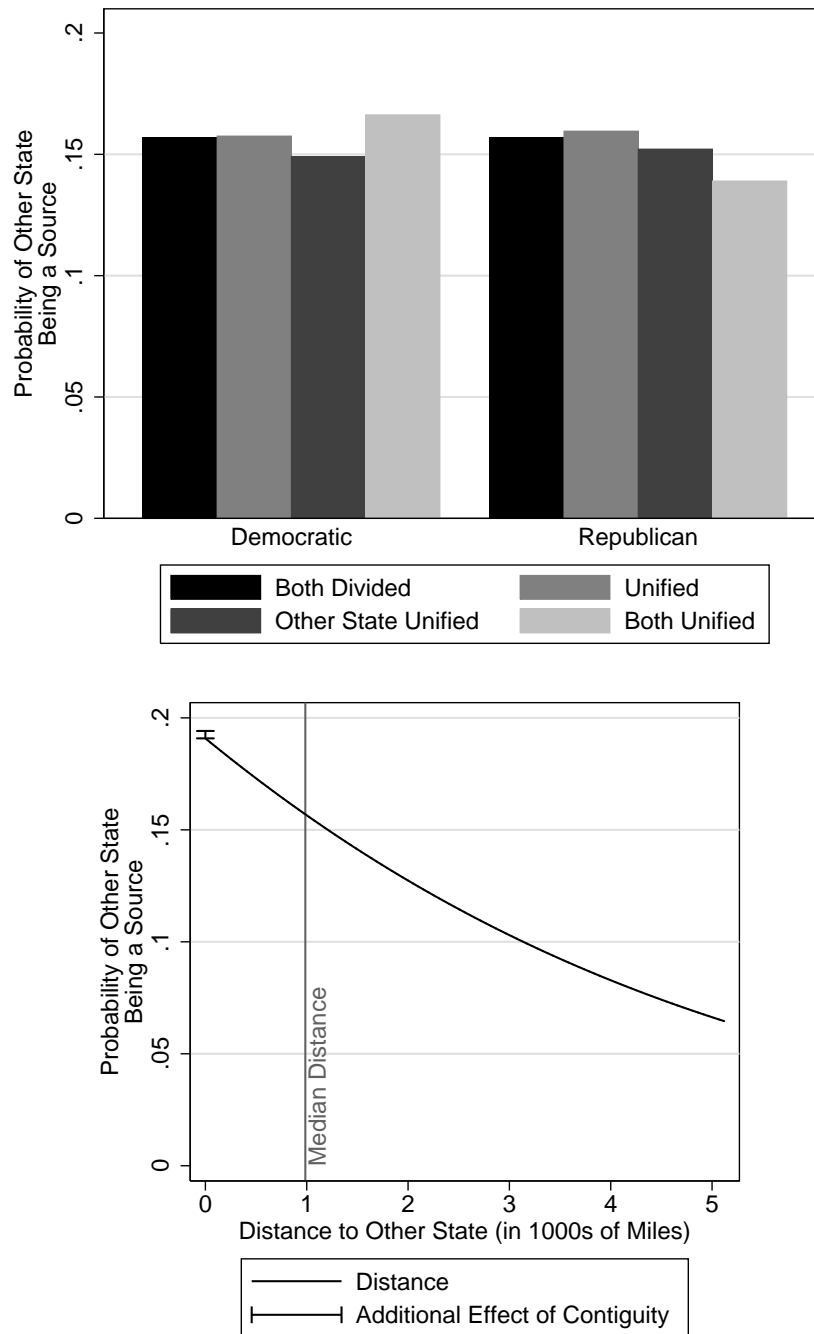
Note: The graphs present the average marginal effects of *Neighbors Adopting* (top row) and *Sources Adopting* (bottom row) in the five replication models: lotteries (Berry and Berry 1990), Indian gaming (Boehmke 2005), capital punishment (Boehmke 2005), restaurant smoking bans (Shipan and Volden 2006), and the pooled model (Boehmke and Skinner 2012a). Points represent expected probability point estimates and vertical lines represent 95% confidence intervals.

Figure 4: Estimated Substantive Effects of Absolute Difference Variables



Note: The graphs present the effects of each variable on the probability scale using Model 2's estimates. All other variables are set to their mean (continuous variables) or mode (binary variables) in 1985 and the random effects are set to zero.

Figure 5: Estimated Substantive Effects of Selected Variables



Note: The graphs present the effects of each variable on the probability scale using Model 2's estimates. All other variables are set to their mean (continuous variables) or mode (binary variables) and the random effects are set to zero.

Table 1: Top 15 States 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

Table 2: Top Innovators from the Inferred Diffusion Networks

	FL	NY	CA	First Adopter
FL	–	7	5	13
NY	7	–	3	18
CA	6	14	–	24

The entry in row i , column j of the state \times state elements of this table gives the number of policies for which state i was the first adopter and state j never adopted. The last column gives the total number of policies for which state i was the first adopter.

Table 3: Estimates and Model Fit Statistics for *Neighbors Adopting* and *Sources Adopting* in the Replication Models

	Only <i>Neighbors</i> (Original Model)	Only <i>Sources</i>	<i>Neighbors</i> and <i>Sources</i>
<hr/> Berry and Berry (1990): Lotteries (Probit, N = 857)			
<i>Neighbors Adopting</i>	0.27* (0.09)		0.17 (0.10)
<i>Sources Adopting</i>		0.29* (0.09)	0.23* (0.09)
AIC	195.12	191.02	190.24
BIC	233.15	229.05	233.02
CV % Correctly Classified	94%	96%	95%
<hr/> Boehmke (2005): Indian Gaming (Probit, N = 364)			
<i>Neighbors Adopting</i>	0.42* (0.20)		0.42* (0.21)
<i>Sources Adopting</i>		0.20+ (0.12)	0.21+ (0.13)
AIC	144.25	144.45	143.54
BIC	241.68	237.98	244.86
CV % Correctly Classified	89%	91%	90%
<hr/> Boehmke (2005): Capital Punishment (Probit, N = 227)			
<i>Neighbors Adopting</i>	0.16 (0.14)		0.14 (0.14)
<i>Sources Adopting</i>		0.23* (0.10)	0.22* (0.10)
AIC	204.53	200.58	201.70
BIC	283.31	279.35	283.89
CV % Correctly Classified	75%	78%	78%
<hr/> Shipan and Volden (2006): Restaurant Smoking Bans (Logit, N = 807)			
% <i>Neighbors Adopting</i>	1.92* (0.86)		1.54 (0.95)
<i>Sources Adopting</i>		0.24+ (0.14)	0.18 (0.15)
AIC	248.57	249.36	249.02
BIC	328.36	329.14	333.50
CV % Correctly Classified	94%	93%	94%
<hr/> 151-Policy Pooled Model (Logit, N = 62,290)			
<i>Neighbors Adopting</i>	0.22* (0.02)		0.19* (0.02)
<i>Sources Adopting</i>		0.13* (0.02)	0.07* (0.02)
AIC	17030.64	17087.50	17017.81
BIC	19263.41	19320.27	19259.62
CV % Correctly Classified	93%	93%	94%

Cell entries report coefficient estimates and standard errors (in parentheses) for *Neighbors Adopting* and *Sources Adopting* and AIC, BIC, and cross-validated percent correctly classified values in three specifications of the replication models. All other variables from the original models are included, but those estimates are not shown to conserve space. Numbers in bold identify the best-fitting model for each fit statistic. * $p < 0.05$; + $p < 0.10$.

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

<i>Follower State Characteristics:</i>				
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)
<i>Potential Source Characteristics:</i>				
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)
<i>Relative Follower/Source Characteristics:</i>				
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. We use the network 300 edges over 35 years of policy adoptions. + indicates statistical significance at the 0.05 level (two-tailed) according to just the parametric p -values from the multilevel logit. * indicates statistical significance at the 0.05 level according to the QAP p -values and the parametric p -values. QAP p -values derived from 500 network permutations.