The diffusion of policy perceptions: evidence from a structural topic model

Gilardi, Fabrizio; Shipan, Charles R; Wüest, Bruno

Abstract: Policy diffusion occurs when policies in one unit (e.g., states, cantons, cities) are influenced by the prior adoption of policies in other units. Although numerous studies have convincingly shown that policy adoption is a function of previous adoptions in other units, they have, with very few exceptions, generally ignored a crucial step in the diffusion process—namely, how political units perceive the policies that they are considering adopting. This policy perception plays a crucial role in linking the actions of previous units with the potential actions in other units. In this paper we focus on the link between prior adoptions and policy perceptions, both by identifying the mix of perceptions and by examining the link between prior adoptions and policy perceptions. We study these perceptions in the area of restrictions on smoking in U.S. states. Our analysis draws upon an original dataset of almost half a million articles published in thirty American newspapers between 1996 and 2014 and uses structural topic models to estimate how smoking bans have been perceived and how perceptions changes as a function of policy adoption in nearby states. We find that many of the most prominent topics are indeed a function of prior policy adoptions in other states.

Posted at the Zurich Open Repository and Archive, University of Zurich
ZORA URL: https://doi.org/10.5167/uzh-143867
Draft Version

Originally published at:
The Diffusion of Policy Perceptions: Evidence from a Structural Topic Model

Fabrizio Gilardi†  Charles R. Shipan‡  Bruno Wueest§

August 24, 2015

Abstract

Policy diffusion occurs when policies in one unit (e.g., states, cantons, cities) are influenced by the prior adoption of policies in other units. Although numerous studies have convincingly shown that policy adoption is a function of previous adoptions in other units, they have, with very few exceptions, generally ignored a crucial step in the diffusion process—namely, how political units perceive the policies that they are considering adopting. This policy perception plays a crucial role in linking the actions of previous units with the potential actions in other units. In this paper we focus on the link between prior adoptions and policy perceptions, both by identifying the mix of perceptions and by examining the link between prior adoptions and policy perceptions. We study these perceptions in the area of restrictions on smoking in U.S. states. Our analysis draws upon an original dataset of almost half a million articles published in thirty American newspapers between 1996 and 2014 and uses structural topic models to estimate how smoking bans have been perceived and how perceptions changes as a function of policy adoption in nearby states. We find that many of the most prominent topics are indeed a function of prior policy adoptions in other states.

†Department of Political Science, University of Zurich (http://www.fabriziogilardi.org/).
‡Department of Political Science, University of Michigan (cshipan@umich.edu).
§Department of Political Science, University of Zurich (wueest@ipz.uzh.ch).
Introduction

When political units—states, cities, or even countries—adopt policies, they do not do so in a vacuum, basing their decisions on only internal factors and pressures. Instead, they can observe the actions that other units previously have taken with respect to these policies. Thus, a state that is deciding whether to adopt, say, new gun control laws, or new rules concerning eligibility for various state-funded medical services, can look around to see which other states have adopted such policies, as well as what types of policies these other states have adopted. They can then base their own decisions on what they observe in these other units.

This process, known as policy diffusion, has been the focus of a large and rapidly growing number of studies (Dobbin, Simmons and Garrett, 2007; Gilardi, 2012; Graham, Shipan and Volden, 2013; Maggetti and Gilardi, 2015). These studies have convincingly established, across a wide range of policy areas, that policies do indeed diffuse, with policies in one unit influenced by policies in other units. That is, these studies have demonstrated that when a unit is considering what to do about a policy, the likelihood that it will adopt the policy is influenced by the existence, in other units, of similar policies.

Although the link between new policy adoptions and earlier policy adoptions has been well established, the focus of the vast majority of studies of policy diffusion has been exclusively on the final adoption decision—that is, did the unit adopt the policy, or did it fail to do so? Although this focus is understandable and has produced numerous important insights, it also ignores a key earlier stage in the policymaking process. In particular, the adoption decision arrives only after the unit has considered various aspects of the policy. During this stage, the unit forms policy perceptions. These perceptions can shape the final outcomes, including whether to adopt a policy and what form the policy should take. But these perceptions, as part of the diffusion process, can themselves be shaped by the prior policy adoptions that have taken place elsewhere. Thus, a more complete consideration of the interdependence of policymaking needs to account for the link between earlier adoptions and the way that a unit perceives how policy problems and solutions are defined and understood.

To examine how the perception of policy problems and solutions changes as a function of the adoption of policies elsewhere, we focus on anti-smoking laws—policies restricting or banning smoking in public places—in the United States. Our choice of policy area is motivated by several considerations. First, several American studies (e.g., Shipan and Volden, 2006, 2008, 2014; Pacheco, 2012), as well as abundant anecdotal evidence, indicate that smoking bans have exhibited a diffusion process. This al-
allows us to concentrate on the nature of the process instead of its mere existence. Second, smoking bans have been adopted in a convenient time frame—roughly a ten year period—which is long enough to detect variations and to supply sufficient information but short enough to be practically manageable. Third, the policy has well-defined characteristics and is comparable across units. Fourth, there was significant uncertainty about the potential consequences of the policies along a number of dimensions—economic consequences, popular support, interest group support, ease of implementation, and so on. And finally, this uncertainty over consequences means that the debate over adoption can be perceived in multiple ways.

In our empirical analysis we rely on an original dataset of almost half a million articles published in thirty American newspapers between 1996 and 2014. More specifically, we use structural topic models to identify how these articles have discussed anti-smoking laws and to estimate how these laws have been perceived in the states. We then show how these perceptions change as a function of policy adoption in nearby states.

2 Stages of the Diffusion Process

Policy diffusion occurs if the policy choices of one unit (e.g., countries, states, cities, etc.) are influenced by the policy choices of other units. Although this simple definition captures key elements of the diffusion process, it also omits others. Consider a situation in which State A is deciding whether to adopt a new law. The standard approach, found in most analyses of policy diffusion, is to consider whether State B already has adopted this policy; and then to see whether State B’s adoption affects the likelihood that State A adopts the policy. In effect, then, these studies implicitly model diffusion as a two stage process; what happens in between these two stages is rarely seen as important.

We argue instead that the process of diffusion occurs in three stages, not two. First, State B adopts a policy. Second, State A then forms perceptions of this policy. And third, State A then decides whether to adopt the policy. The middle stage, in which State A forms its policy perceptions, is more than just a transitional stage; it is worthy of attention in its own right. It is at this stage, when states are

---

1Although we refer to “State B,” the earlier adoption can be by a single state, as in analyses that examine dyadic relationships between individual states, or by a set of states, as in studies that look at the number of previous adoptions among a specified set of states.

2It is also possible to consider an earlier stage—namely, the way in which the policy is framed prior to State B’s adoption. Although this is certainly a topic worthy of attention, we leave it for future research.
considering what to do and forming perceptions of a policy, that they might consider some of the factors that scholars refer to as the mechanisms of diffusion (Simmons, Dobbin and Garrett, 2006; Braun and Gilardi, 2006; Dobbin, Simmons and Garrett, 2007; Gilardi, 2012). What can they learn about the political or policy consequences of adoptions in earlier states? Would they be likely to suffer negative economic consequences, or would the reap positive economic benefits, if they adopt such a law? Are there norms in place to which they want to adhere, or would they be acting against prevailing norms by adopting a new policy?

More generally, there are clear links between this second stage, in which a state forms policy perceptions, and the first and third stages. Our focus in this paper is on developing a way to characterize the policy perceptions that exist at the second stage, and to investigate whether there is a connection between the adoptions in the first stage and the perceptions in the second stage. But it is worth noting that this connection is important in part because of the link between the second and third stages. This link between the latter two stages is both straightforward and of obvious importance. Put simply, does the way in which an issue is perceived within a polity have an effect on the likelihood that the polity will adopt a policy? Especially given that policies usually can be framed in multiple ways, does the specific frame that dominates discussion influence the eventual policy choices? Viewed in this light, policy perceptions are important as a cause of policy outcomes.

Our main interest in this paper is instead on the relationship between the first and second stages, which means that we examine policy perceptions as an effect. Given that State A’s consideration of an issue is subsequent to State B’s action, we investigate whether State B’s action influences the policy perceptions in State A. In the area of anti-smoking laws, for example, one state might perceive the policy as being primarily about the health consequences of adoptions restrictions on smoking, while another might concentrate on public support. Does the type of perception change over time? And is this perception in a state influenced by the actions taken earlier in other states? In effect, then, our focus is on whether the diffusion process involves policy perceptions, whereby these perceptions—which might eventually influence further outcomes—are themselves a product of diffusion from the actions of other actors. Thus, instead of focusing on the direct diffusion from one set of policy outcomes to another, our interest in this paper is in establishing whether previous policy outcomes diffuse to policy perceptions—a key aspect of the diffusion process that few studies have recognized, let alone examined.

A notable exception is Pacheco’s (2012) study, which not only examines how prior adoptions influence public opinion,
To assess which policy perceptions exist and are most prevalent, and whether the prevalence of these perceptions is a function of prior adoptions (and thus part of the overall diffusion process), we rely on structural topic models (STMs) [Roberts et al., 2014; Roberts, Stewart and Airoldi, 2015], which we describe in more detail in the following section. This approach allows us to examine, in great detail, which topics dominate the discussion surrounding a policy. We identify and measure the topics by analyzing media coverage of this policy issue.

One question that arises is whether the media coverage we examine reflects how policies are perceived, or whether it influences this perception. On this question we are agnostic. Regardless of whether this coverage reflects or influences perceptions, it can be used as an accurate source for identifying the ways in which smoking bans are perceived in a given unit. Thus, we can use the information derived from structural topic models both to identify the most common perceptions and to identify their distributions, both cross-sectionally and over time. In the analysis in this paper, we will look specifically at whether the prevalence of specific topics is a function of adoptions in other states—that is, whether policy perceptions change as a function of the adoption of smoking bans in other units. But our data could be used to examine several other aspects of policy perceptions, such as whether the mix of these perceptions (e.g., the ratio of different perceptions or another composite measure) varies over time, whether the topics used focus less on economic consequences over time, and whether states exhibit the same topics that are found in similar states.

3 Methodology

3.1 Data sources and preprocessing

Our analysis of policy perceptions as a part of the diffusion process concentrates, as noted earlier, on the adoption of antismoking policies in the U.S. states. The states traditionally have had considerable autonomy in public health areas, and smoking restrictions are no exception. Although smoking-related issues are often discussed at the national level [McCann, Shipan and Volden, 2015], few laws have been passed at this level in the US; rather, the vast majority of policymaking has taken place within the states. Thus, the topic of anti-smoking laws provides an excellent forum for examining the process of diffusion.

The time frame we examine begins in 1996, which is two years before the first statewide smoking
ban was adopted in California. To analyze public discussions and to gain a handle on policy perceptions within a state, we rely on articles published in the newspapers listed in Table 1. Currently we have processed articles from thirty newspapers, but the full construction of the newspaper corpus is still under way and the final corpus eventually will include the largest newspaper in terms of circulation for every state. We use print media rather than television or radio programs partly for technical reasons but especially because they generally report more extensively on political matters than do on-air media (Druckman, 2005, 469).

We retrieved newspaper texts using a simple but broad keyword search from different databases such as LexisNexis. Then we split the texts into paragraphs of a similar length, which produced a corpus containing 3,026,793 paragraphs. A manual evaluation of a random sample of 14,519 paragraphs revealed a very low share of paragraphs actually covering smoking bans—about 1.5% on average. This is due to the looseness of our keyword search, aimed at minimizing the number of articles of smoking bans escaping our search.

We manually coded 14,519 paragraphs as relevant or irrelevant. Relevant paragraphs are those containing information on smoking restrictions—that is, bans or limits on smoking in public places or specific workplaces. This definition includes statements about any kind of restriction of smoking (smoking bans) in public places and/or businesses introduced through legislative action, executive action, or other democratic actions (e.g., direct democratic processes). By contrast, we coded as irrelevant paragraphs discussing, for example, smoking bans introduced by private actors (e.g., companies, businesses), or bans of specific tobacco products (e.g., mentholated cigarettes).

Using the information gained by manual coding, we then classified all 3,026,793 paragraphs in our corpus as relevant or irrelevant using an ensemble classifier, that is, a tool that aggregates the classifications produced by different algorithms. We proceeded in two steps. First, we applied an ensemble...
<table>
<thead>
<tr>
<th>Newspaper</th>
<th>State</th>
<th>N articles</th>
<th>N paragraphs</th>
<th>N filtered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albuquerque Journal</td>
<td>NM</td>
<td>4,953</td>
<td>25,464</td>
<td>1,197</td>
</tr>
<tr>
<td>Argus Leader</td>
<td>SD</td>
<td>3,801</td>
<td>25,339</td>
<td>1,237</td>
</tr>
<tr>
<td>Atlanta Journal-Constitution</td>
<td>GA</td>
<td>23,281</td>
<td>114,843</td>
<td>2,554</td>
</tr>
<tr>
<td>Baltimore Sun</td>
<td>MD</td>
<td>14,096</td>
<td>78,124</td>
<td>1,914</td>
</tr>
<tr>
<td>Charleston Gazette</td>
<td>WV</td>
<td>18,228</td>
<td>116,099</td>
<td>2,236</td>
</tr>
<tr>
<td>Chigaco Tribune</td>
<td>IL</td>
<td>31,855</td>
<td>157,102</td>
<td>4,568</td>
</tr>
<tr>
<td>Clarion Ledger</td>
<td>MS</td>
<td>3,206</td>
<td>17,005</td>
<td>709</td>
</tr>
<tr>
<td>Courier-Journal</td>
<td>KY</td>
<td>10,593</td>
<td>71,887</td>
<td>2,953</td>
</tr>
<tr>
<td>Daily News</td>
<td>NY</td>
<td>14,202</td>
<td>60,828</td>
<td>990</td>
</tr>
<tr>
<td>Daily Oklahoman</td>
<td>OK</td>
<td>12,250</td>
<td>44,793</td>
<td>1,640</td>
</tr>
<tr>
<td>Denver Post</td>
<td>CO</td>
<td>13,088</td>
<td>79,843</td>
<td>1,650</td>
</tr>
<tr>
<td>Des Moines Register</td>
<td>IA</td>
<td>5,750</td>
<td>41,160</td>
<td>1,154</td>
</tr>
<tr>
<td>Detroit Free Press</td>
<td>MI</td>
<td>6,021</td>
<td>54,014</td>
<td>647</td>
</tr>
<tr>
<td>Hartford Courant</td>
<td>CT</td>
<td>14,821</td>
<td>83,980</td>
<td>980</td>
</tr>
<tr>
<td>Las Vegas Review Journal</td>
<td>NV</td>
<td>9,430</td>
<td>56,605</td>
<td>1,199</td>
</tr>
<tr>
<td>Los Angeles Times</td>
<td>CA</td>
<td>29,597</td>
<td>196,061</td>
<td>2,391</td>
</tr>
<tr>
<td>Omaha World-Herald</td>
<td>NE</td>
<td>12,295</td>
<td>72,506</td>
<td>2,474</td>
</tr>
<tr>
<td>Philadelphia Inquirer</td>
<td>PA</td>
<td>18,975</td>
<td>105,861</td>
<td>1,581</td>
</tr>
<tr>
<td>Portland Press Herald</td>
<td>OR</td>
<td>5,374</td>
<td>27,796</td>
<td>819</td>
</tr>
<tr>
<td>Providence Journal</td>
<td>RI</td>
<td>15,264</td>
<td>89,549</td>
<td>1,631</td>
</tr>
<tr>
<td>New Jersey Record</td>
<td>NJ</td>
<td>19,453</td>
<td>95,395</td>
<td>1,923</td>
</tr>
<tr>
<td>St. Louis Post-Dispatch</td>
<td>MO</td>
<td>27,516</td>
<td>137,830</td>
<td>3,241</td>
</tr>
<tr>
<td>Star Tribune Minneapolis</td>
<td>MN</td>
<td>13,693</td>
<td>120,220</td>
<td>2,027</td>
</tr>
<tr>
<td>Tennessean</td>
<td>TN</td>
<td>5,475</td>
<td>36,611</td>
<td>728</td>
</tr>
<tr>
<td>Tribune-Eagle</td>
<td>WY</td>
<td>2,024</td>
<td>13,526</td>
<td>958</td>
</tr>
<tr>
<td>Tribune/Deseret News</td>
<td>UT</td>
<td>15,884</td>
<td>58,817</td>
<td>1,256</td>
</tr>
<tr>
<td>USA Today</td>
<td>NY</td>
<td>11,246</td>
<td>59,637</td>
<td>881</td>
</tr>
<tr>
<td>Wall Street Journal</td>
<td>NY</td>
<td>22,971</td>
<td>139,448</td>
<td>1,517</td>
</tr>
<tr>
<td>Washington Post</td>
<td>DC</td>
<td>58,495</td>
<td>501,552</td>
<td>3,458</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>497,248</strong></td>
<td><strong>3,026,793</strong></td>
<td><strong>52,398</strong></td>
</tr>
</tbody>
</table>

Table 1: *Selected sources for the content analysis.*
composed of a support vector machine (SVM) optimized for recall (average of 0.81 on the heldout set) and of a kernel ridge (KR) regression with good performance in terms of precision (average of 0.79 on the heldout set). 

Ensemble decisions were taken by voting in which we gave the SVM more weight to optimize the recall of positives (average of 0.83) while removing most of the irrelevant paragraphs. Second, we applied the same ensemble classifier again, now giving more weight to the KR regression to optimize precision since the percentage of relevant and irrelevant paragraphs was now much more balanced.

This filter produced a corpus of 52,398 paragraphs containing 42,383 unique terms and 3,199,613 instances of these terms. Evaluated on two randomly sampled heldout sets and a manual check of randomly sampled paragraphs, the filter is able to remove 93-95% of the irrelevant paragraphs, while retaining most of the relevant ones. Moreover, most classification runs we tested agreed with an overall F-Score of 0.95 or higher—a further sign for the consistency and thus reliability of the classification. Therefore, we are confident that our estimations reveal the general trend in the newspapers’ coverage of smoking bans.

Finally, prior to the estimation we pre-processed all documents with standard procedures such as text segmentation into paragraphs and sentences, tokenizing, removal of punctuation, as well as lemmatizing and converting all words to lowercase.

3.2 Estimation

We identify policy perceptions inductively with a structural topic model (STM) (Roberts et al., 2014; Roberts, Stewart and Airoldi, 2015), which produces estimates document-topic and word-topic probabilities (Roberts, Stewart and Airoldi, 2015; Roberts, Stewart and Tingley, 2014). It builds on well-established generative topic models, such as the Latent Dirichlet Allocation (LDA) (Blei, Ng and Jordan, 2003). The LDA is a mixed-membership model, meaning that it assumes that each document consists of a mixture of topics (Grimmer and Stewart, 2013, 283–285). Concretely, the LDA is a hierarchical model in which a document’s proportion of topics has a common prior drawn from a Dirichlet distribution:

---

8 Recall is the fraction of relevant documents that are retrieved.
9 Precision is the fraction of retrieved documents that are relevant.
10 The parameter search in scikit-learn yielded the following setting: No normalization of the document-term matrix, TFIDF transformation, no removal of stopwords, lowercase transformation, no lemmatization, no POS tags, bigrams, and no maximum of features considered.
\( \pi_i \sim \text{Dirichlet}(\alpha). \)

Then, the topic of the \( j \)-th word in the \( i \)-th document is drawn from a multinomial distribution:

\( \tau_{ij} \sim \text{Multinomial}(\pi_i). \)

Finally, the \( j \)-th word in the \( i \)-th document is drawn from a multinomial distribution, conditional on its probability of being drawn from topic \( k \), \( \theta_k \):

\( w_{ij} \sim \text{Multinomial}(\theta_k). \)

The STM’s major innovation is that the prior distribution of topics can be influenced by covariates \( \text{LogisticNormal}(X\beta, \Sigma) \):

\( \pi_i \sim \text{LogisticNormal}(X\beta, \Sigma). \)

Furthermore, covariates can also be specified for the word distribution over topics, that is, not only the probability of topics within documents, but also that of words within topics. For instance, this would allow to see how the language used in a given topic changes as a function of covariates. We will consider this useful option in future work.

Our analysis includes four covariates: (1) month dummies, (2) newspaper IDs, (3) the presence of a smoking ban in a state, and (4) the proportion of a state’s neighbors with smoking bans. The last covariate is the most interesting substantively and theoretically and we use it to estimate diffusion effects. We purchased detailed data on smoking bans, including the dates of the relevant bills, from MayaTech’s Center for Health Policy and Legislative Analysis, which has already proven to be a highly reliable data source \( \text{Shipan and Volden (2006)} \).

We estimated four models differing along two dimensions: first, the number of topics that we assume (10 or 20); second, the type of smoking bans that we take into account (smoking bans in restaurants or in seven areas—restaurants, bars, government worksites, private worksites, hotels, indoor arenas, malls). In Section 4, we focus on the model assuming 10 topics and considering only smoking bans in restaurants, which produced the most convincing results. The results of all models are shown in Appendix A2. In future work we will expand the models in a number of ways, including
the number of topics, the coding of the explanatory variables, and the detection of locations in texts with a Named Entity Recognizer.\textsuperscript{11}

4 Results

4.1 Media attention to smoking bans

Figure 1 shows the frequency of newspaper coverage of smoking bans over time. Points indicate the number of monthly published paragraphs since 1996, while the line and the grey area indicate the loess smoothed trend with its 95% confidence interval.\textsuperscript{12} The data show a clear trend. The coverage of smoking bans soared to around 400 paragraphs per month in the late 1990s, reached another peak in 2003 and most notably in 2007, after which it gradually decreased to around 150 paragraphs per month towards the end of the period. Moreover, in most newspapers coverage was intense in the period prior to the introduction of federal or statewide smoking bans, peaked when legislation was passed, and then decreased.\textsuperscript{13} Further, the peak in the late 1990s correlates with California’s extension of the smoking bans to bars, making it the first US state to enact a complete ban in all enclosed workplaces. Overall, reports on smoking bans in the US spiked again at the end of 2003, which is likely related to the introduction of a statewide smoking ban for all enclosed workplaces in New York, closely followed by the very similar Smoke-Free Air Act in New York City. From 2004 until 2007, finally, several states in our sample introduced state-wide smoking bans (e.g., Connecticut, Rhode Island, District of Columbia, New Jersey, Colorado, Utah, and Minnesota). It seems plausible that US newspapers paid particular attention to the two highest-profile anti-smoking policies of the last two decades.

Overall, it is encouraging for the external validation of our supervised classification that the development of US newspapers’ coverage on smoking bans seems to mirror the proposition, debate, and introduction of major legislative acts. Because our newspaper sample will eventually cover all US regions, we are confident that all major legislative activity related to smoking bans will be covered.

\textsuperscript{11}The Denver Post, for example, likely does not only report on smoking bans in Colorado. The problem is even acuter for national newspapers such as the New York Times, Wall Street Journal, or Washington Post.

\textsuperscript{12}The development of the number of relevant articles over time (not shown) is very similar to the trends in Figure 1.

\textsuperscript{13}Figures for each newspaper are presented in the Appendix.\textsuperscript{A1} The Minnesota Star Tribune represents an exception to the general trend, as its coverage correlates with the non-smoking ordinance in Minneapolis, adopted and implemented in 2005.
4.2 The perception of smoking bans: topic models

As mentioned above, we discuss here the results of a model assuming the existence of ten topics. The results of the four models we estimated are shown in Appendix A2. Of the ten topics that the STM then estimated, five stood out, in the sense that the twenty-five words (stems) most closely associated with each topic corresponded with previously recognized frames for this policy area (Champion and Chapman, 2005). The first of these topics covers Health—that is, the health aspects of smoking restrictions—with common words that are clearly related to the health effects of smoking (e.g., protect, heart, risk, disease, lung, cancer, etc.). The second focuses on Business, Customers, and Workers (e.g., patron, exemption, customer, business, health). The third focuses on Sales Restrictions, which often are related to limiting the access that children have to cigarettes but which also can extend to more general sorts of restrictions (e.g., store, product, license, minor). The fourth topic, Legislative Action, is less about aspects of the policy itself and more about the process by which policies get enacted (e.g., sponsor, amendment, approve, legislation, tax, vote). And the fifth and final topic centers around Groups and Voters regarding these policies (e.g., voter, public, group, support). Table 2 provides a list of the top twenty-five words for each of these topics.

In Figure 2, we provide initial evidence of the relationship between these topics, or perceptions, and
Figure 2: Relationship between topics and presence of smoking bans in restaurants in neighboring states.
Table 2: Top twenty-five words for each topic.

<table>
<thead>
<tr>
<th>Topic</th>
<th>Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health</td>
<td>smoke, smoking, secondhand, health, cancer, study, lung, smoker, american, cause, public, people, disease, child, risk, heart, year, report, protect, tobacco, say, air, smoke-free, nonsmoker, percent</td>
</tr>
<tr>
<td>Business, Customers, and Workers</td>
<td>ban, bar, smoking, restaurant, say, business, owner, state, law, casino, new, health, effect, city, customer, establishment, will, allow, percent, exemption, take, court, patron, smokefree, exempt</td>
</tr>
<tr>
<td>Sales Restrictions</td>
<td>law, use, fine, police, change, require, sale, state, tobacco, fire, rule, sell, license, , new, also, gun, marijuana, allow, court, product, , minor, official, cigarette, store, safety</td>
</tr>
<tr>
<td>Legislative Action</td>
<td>bill, house, senate, state, committee, pass, vote, sen, tax, governor, legislation, rep, legislature, lawmaker, approve, year, session, measure, republican, amendment, gov, budget, legislative, sponsor, assembly</td>
</tr>
<tr>
<td>Groups and Voters</td>
<td>say, ban, support, issue, county, year, will, smoking, measure, group, one, legislation, last, oppose, public, proposal, member, make, government, effort, voter, vote, antismoking, debate, maryland</td>
</tr>
</tbody>
</table>

earlier policy enactments. In each of the figures the x-axis shows the spatial lag, which is effectively the percentage of neighboring states that have enacted limits on smoking in restaurants. At the left end of this scale, no neighbors have adopted restrictions or bans; at the right end, all of them have. The y-axis provides a measure of the prevalence of that topic. Thus, the first chart in the figure shows the prevalence of the Health topic as a function of the percentage of neighboring states that have adoptions restrictions on smoking in restaurants. We see that health consequences are the topic in about nine percent of paragraphs; but we also see that although the relationship have a very slight positive slope, there is very little substantive difference across the range of the spatial lag. Thus, although policy perceptions may be slightly more likely to focus on the health aspects of antismoking laws as more of a state’s neighbors adopt such policies, the overall substantive effect is small. Attention to the health aspects of anti-smoking laws, it seems, are a fairly constant presence.

The other figures evince clearer trends. Two of these figures—the second figure, which portrays the Business, Customers, and Workers frame, and the fifth figure, which covers Groups and Voters, both show a strong and positive slope. Taking the first of these, the figure indicates that business-oriented concerns are dramatically higher in states if the surrounding states have been proactive, in terms of

14In the future we can investigate whether the frames are influenced by states other than neighboring states—for example, states that share similar ideologies or that have demonstrated success in the policy area. [Gilardi 2010, Füglistier 2012, Shipan and Volden 2014].
passing laws to restrict smoking. In some respects, this is surprising, since a fairly general finding has been that these restrictions have not have the sorts of dramatic negative consequences on businesses that some opponents of these laws originally had feared and predicted; and this in turn might have resulted in a negative relationship. At the same time, our current analysis indicates only whether something was the topic of discussion, and not whether it was discussed in a positive or negative light. Thus, it is entirely possible that the discussions centered around the fact that the enactment of laws in other states did not have the negative effects that opponents had worried about. The positive slope for the final figure, on the other hand, is easier to interpret: the more surrounding states that have passed laws, the more that public opinion and group support becomes a prominent topic, with the likely explanation being that public support increases—and thus is discussed more frequently—as more surrounding states have passed similar laws.

Finally, the third and fourth figures show negative relationships between between the spatial lag and the topic prevalence. For Sales Restrictions, which often focus on the access that youths have to getting cigarettes (e.g., laws preventing restaurants from having vending machines that sell cigarettes), the negative slope can be interpreted as evidence that these sorts of restrictions have become less controversial over time. The figure for Legislative Action tells a somewhat similar story. During early adoptions, there was a fair amount of controversy over whether the legislature should be passing such laws, and what form these laws should take. Over time this controversy lessened, leading to far fewer debates and discussions being framed in terms of whether it was appropriate for the legislature to take such actions.

We hasten to add that at this point we need to be cautious and not read too much into these figures—we are, after all, simply showing correlations and then constructing explanations that are consistent with the correlations that we see. Still, several findings emerge from these figures.

First, there are clearly different types of policy perceptions that emerge in the media coverage of the debates and discussions about whether to adopt new policies. We have identified five that correspond well to previous discussions of the ways in which debates over anti-smoking laws are framed, relating to health, business interests, sales restrictions, legislative action, and public opinion. Second, there is variation across these perceptions we have identified, in terms of prevalence. Business concerns end up being the most prevalent, while at the other end, perceptions of the issue that revolve around sales restrictions are at the lower end. Third, there is also variation within the topics. As already pointed out,
some of these frames decrease in prevalence as a function of earlier adoptions; others increase; and one other remains fairly flat. Even within perceptions that show similar patterns, differences emerge—the increase for Groups and Voters, for example, is far less than that for Business, Customers, and Workers, which shows a more dramatic increase. The more central and general point, though, is that the STM allows us to model topic prevalence as a function of policies in other states, which puts these policy perceptions directly in the overall diffusion process.

5 Conclusion

Policy diffusion is a multi-stage process, but most research has been limited to an examination of only two of these stages—the initial adoption (or adoptions) in some set of states, and then whether future adoptions are influenced by these earlier adoptions. We argue that an intermediary stage is of crucial importance, both because it is affected by earlier adoptions and because it can affect later adoptions. More specifically, it is during this intermediary stage—the second stage of the diffusion process—that states form specific perceptions of policies. These perceptions can plausibly influence the likelihood of adoption, but our interest in this paper is on examining these perceptions themselves. What perceptions exist? Do these perceptions vary over time? And most importantly, are these perceptions a function of earlier adoptions elsewhere? To the extent that these perceptions are a function of earlier adoptions, we should recognize them as a critical part of the overall diffusion process.

Our analysis provides a first step toward better understanding how policy perceptions can diffuse—or more accurately, how the first stage of the diffusion process, in which other states adopt policies, can influence the next stage, in which policy perceptions are formed. We have put forward a preliminary analysis of the diffusion of the perception of smoking bans in US states based on a structural topic model of over 52,000 paragraphs in thirty newspapers, showing that there is variation in the incidence of these perceptions, as well as connections between these perceptions and the prevalence of prior adoptions in neighboring states.

Of course, much work remains to be done. In terms of data, the first step is to complete our newspaper sample. We also are in the process of collecting newspaper articles regarding the consideration of anti-smoking policies in Swiss cantons, which will provide for a useful comparison with the US states. We also are continuing to work on the topic models themselves, including improving the classifier used to weed out irrelevant texts, the use of Named Entity Recognition tools to identify states and cities...
in the texts, and a more careful consideration of time dependence. In addition, obtaining estimates of whether the newspaper coverage was positive or negative will allow us to ascertain whether not only the frame, but the nature of the frame, varies in response to earlier adoptions. We also can look to see whether perceptions vary based on adoptions by non-neighboring states, for instance using the “diffusion pathways” identified by Desmarais, Harden and Boehmke (2015). And finally, we can use our data to examine a range of other important questions—whether perceptions change within a state after it adopts a policy, for example, or how the relative prevalence of frames within a state changes over time. For now, however, our preliminary analysis has established a foothold for the usefulness of structural topic models and support for the idea that policy perceptions are an important part of the diffusion process.
References


A1 Coverage of smoking bans in individual newspapers
A2 Structural Topic Models

A2.1 Ten topics, smoking bans in restaurants
A2.2 Twenty topics, smoking bans in restaurants
A2.3  Ten topics, smoking bans in seven areas
A2.4 Twenty topics, smoking bans in seven areas