Approaches to Modeling the Adoption and Diffusion of Policies with Multiple Components

Frederick J. Boehmke

University of Iowa
Department of Political Science
341 Schaeffer Hall
Iowa City, IA 52242

April 10, 2008

Contact: frederick-boehmke@uiowa.edu. Boehmke is Associate Professor of Political Science at the University of Iowa. Most the work on this paper was conducted during a stint as a Robert Wood Johnson Scholar in Health Policy Research at the University of Michigan; the support of the Robert Wood Johnson Foundation is gratefully acknowledged. I thank Megan Sussek for her assistance gathering the state obesity legislation; I thank Corrine McConnaughy and Chuck Shipan for valuable comments and suggestions.

ABSTRACT

Scholars have begun to move beyond the dichotomous dependent variable — indicating whether a state adopts a policy or not in a given year — usually employed in event history analysis. In particular, they have devoted increasing attention to the components of policies that states adopt. In this paper I discuss a variety of estimators that have been employed to analyze the adoption and modification of policies with multiple components, including various forms of event history analysis, OLS, and event count models. Each of these estimators has its strengths and may be appropriate in certain circumstances; in the majority of cases, however, some version of event history analysis for multiple or repeat failures is likely to be preferred. With various modifications, the researcher can estimate models that treat each component as distinct, pool these models to leverage commonalities across components, or treat the components as identical parts of the same process. The different approaches are illustrated by studying state adoption of various obesity-related policies.

1 Introduction

For the last sixteen years the standard approach to modeling state policy innovation has been event history analysis (EHA).[1] This approach has proven to be extremely flexible and has allowed scholars to answer a variety of important research questions. The methodology has become so widespread that cataloging the breadth of its applications would be a sizable task in its own right (see Karch (2007) for a recent review). To some degree, however, the field of political science has reached a point of diminishing marginal returns from the standard EHA model, by which I mean a logit or probit analysis of state policy adoption measured with a dichotomous variable observed annually and generating a sequence of zeros followed by a one in the year the event occurs.

At the same time, scholars have begun to make advancements in many directions by pushing against the confines of the standard EHA approach, at least three of which involve policies with multiple components. First, Boehmke and Witmer's (2004) study of Indian gaming adoptions tries to separate policy diffusion resulting from economic competition or social learning by modeling the number of Indian gaming compacts agreed to in a state. Second, scholars have found it difficult to fit some policies into the dichotomous requirements of the standard EHA. While some policies, like whether a state has a lottery (e.g., Berry and Berry 1990), are easily approximated with a discrete indicator, others are not. For example, does a state have good pain-management policy if it only allows three of the eight measures advocated by the Pain Policy Studies Group (Imhof 2006)? Third, studying policy emulation requires more nuanced information about the contents of policy. Examining changes in individual components of a policy helps us understand their evolution as states revise and expand policies to reflect shifting needs among their citizens, to respond to technological changes, or to incorporate successful advances developed in other states (e.g., Volden 2006).

One of the consequences of these many issues is that scholars have begun to move beyond the dichotomous dependent variable — indicating whether a state simply adopts a policy or not in a given year — usually employed in EHA. One way that they have done this is by focusing on the multiple components that a given policy encompasses. These components may be complementary or alternative ways to accomplish a given policy objective; states may adopt none, some or all of them in a given year and may adopt additional components in later years. Examples include the nine components of state medical malpractice reform studied by Yackee (2005) and the three components of anti-smoking policy studied by Shipan and Volden (2006, 2007).[2]

Considering the details of policy adoption by studying the presence or level of specific components has both methodological and substantive advantages. Substantively, their consideration allows us a richer understanding of the heterogeneity of state policy and reflects the variation in the methods through which states seek to accomplish their objectives. Methodologically, their consideration can add information to our models and allows us to test previously unexplored theories. By considering multiple components of a single policy, scholars can leverage their similarities to gain both theoretical and empirical traction, while also allowing some of the richness of that policy to emerge within a parsimonious empirical model.

Yet the selection of an appropriate estimator for these models has been somewhat trickier. Recent studies use a mélange of methods, including more advanced forms of EHA, event count methods, ordered models, and linear regression. And while the preferred model will vary across applications given the structure of the data and the question(s) being asked, it seems appropriate to consider the various approaches at researchers' disposal. Most of these approaches correspond to relatively straightforward applications of more advanced EHA techniques; no "new" estimators are proposed here, only (re)conceptualizations of existing techniques for policies with multiple components. In particular, one can accomplish quite a bit by pooling separate EHA models for each component into a single model and then making modifications to account for various forms of heterogeneity.

2 Policies With Multiple Components

In this section I provide some clarification of what I mean by a policy component and provide some examples to illustrate the discussion. An important issue to address up front is whether it is best to view a set of components as pieces of a single policy or whether they are, in fact, each separate policies. If the latter is true, then analysis can proceed as usual with the appropriate measure of each component studied separately. If not, then the researcher must make some decision about how to combine the information about a state's particular array of adopted components. The desire to do this will likely be a function of both the nature of the policy at hand and the researcher's question. For example, in most cases one would probably analyze state gaming policies separately: does a state have a lottery; does it allow commercial casino-style gaming; does it allow Indian gaming? While each of these is a part of state gaming policy, they are not necessarily alternate parts of a single policy decision. On the other hand, the decision about whether to negotiate agreements with Indian nations to allow parimutuel wagering, card games, games of chance or house-banked games may be seen as part of the same overall decision-making process. Even ignoring the possibility that there is some natural tendency for certain policy decisions to be viewed as components of a single policy rather than as distinct ones, it is clear that researchers are more frequently making this distinction.

Given the willingness to consider the different components of a single policy, what do those components look like? To put a little precision on the discussion, I introduce some notation. In a standard EHA the analysis focuses on explaining whether a state adopts a given policy. Let Y_{it} be an indicator variable for whether state i adopts policy Y in year t. In a standard EHA, states are included in the sample until they adopt the policy, after which they are excluded. This status is summarized by the risk set at time t: $i \in R(t)$ means that state i has not adopted the policy by time t and is therefore included in the analysis. In a repeated failures context, states are not precluded from experiencing future events and may remain in the risk set.[3]

Now assume that the policy has multiple components; here I will assume that those components are each dichotomous. In particular, assume that there are K distinct, binary components. The dependent variable and the risk set must be indexed by component, so let Y_{ikt} indicate whether state i adopts component k ($1 \leq k \leq K$) at time t and let $i \in R_k(t)$ indicate that state i is at risk of adopting that component at time t. Because these are distinct components, a state does not have to adopt all of them at the same time. In fact, a state can adopt none, all, or any combination of the K components in a given year. One can allow for single or repeated failures for each component.

Some of the features I have put on the components are restrictive and may not be realistic in a given policy area. For example, some of the components may not be dichotomous. Volden's (2006) six components of state Children's Health Insurance Program plans include some that he measures dichotomously, such as whether there is a monthly premium or a co-pay; others that are nominal, such as whether the plan is run under Medicaid, run separately, or as a combination of the two; and some that are continuous, such as the maximum income for eligibility. Policies that involve such rich components will be even more difficult to analyze in any way save separately.

An additional concern is whether there is strong interdependence among some of the components. This could happen in two ways. Some of the components may be complementary in the sense that adopting one essentially guarantees or requires adopting the other. If this complementarity is strong enough, it may make sense to view them as a single component, especially if viewing them separately adds no more information than viewing them together. On the other hand, some of the components may be substitutes — they both accomplish the same goal, but implementing one makes the other superfluous. Finally, some of the components may be in opposition to each other, meaning that the implementation of two or more is mutually contradictory. For example, some states have passed laws protecting fast food retailers from obesity-related lawsuits while others have promoted awareness, treatment and physical fitness. In practice, of course, few components will fall precisely into these extreme categories and may exhibit one or more of them to varying degrees. Further, it may be up to the researcher to determine whether components fall into these categories by definition or whether they appear to do so because of a specific pattern of implementation across states.

Lastly, there is the question of exactly how many components there are within a given policy. In some cases this emerges naturally from the theoretical question being asked, in others it may emerge from the structure of the policy in question. Yet often there will be a subjective part in its determination. One option is to rely on the decisions made by policymakers or policy experts when they consider the policy in question; Many research groups or advocacy groups explicitly break policies into multiple components when comparing them across states.

Certainly these various issues indicate the potential complexity of considering the components of a single policy. Yet they also reflect the potential richness of the policies that we have studied (and may help explain why others have not been studied) and have often treated as dichotomous. While it may be readily apparent when a state has adopted a lottery, it is harder to say exactly when it has adopted a pain-management policy: when it has one of the eight components? at least four of them? or all of them? Considering multiple components together may allow researchers to leverage the information provided by variation across states in how they adopt policies while also reflecting the underlying objective to arrive at a relatively concise explanation of policy in a given area.

3 Potential Modeling Strategies

In this section I discuss a number of ways through which one could attempt to analyze the adoption of a policy with components of the nature described above. I consider three broad approaches to modeling the data, including using distinct models for each component, pooling the components into a single model, and treating the components as part of the same event structure rather than as distinct. Each of the methods has different strengths and weaknesses; some may be more appropriate for certain substantive or theoretical questions than others. In general, though, researchers will likely often prefer an approach that lies somewhere between the first and second, which accounts for both commonalities and differences in the factors that influence adoption across components; deviations from this approach will be driven by appropriate theoretical or policy considerations.

As an example of situations under which a research may prefer different models, consider the different questions one might wish to study about Indian gaming policy. First, one may wish to understand when states first negotiate gaming agreements with any tribes — i.e., when does the state adopt Indian gaming (Boehmke 2005)? Second, one may wish to study with how many tribes the state negotiates agreements (Boehmke and Witmer 2004). Third, one could study the types of games allowed under those compacts. Fourth, one may wish to study the expansion of Indian gaming either by looking at expansion in the number of tribes with agreements or at expansion in the types of games or number of slots allowed under existing agreements. Each of these questions is valid and each potentially dictates a different empirical approach. In the rest of this section I review the relationship between these questions and various approaches researchers have employed.

3.1 Timing to Adoption

The first approach is to work within the EHA approach and modify the standard implementation to adapt to different questions. There are a variety of ways one can study the adoption process in this framework; I will focus on three general classes. The first consists of separate analysis of each component. The second involves pooling these analyses into one model that simultaneously estimates adoption of each component; this approach takes the state-policy-year as the level of analysis. The third treats the components as identical and considers adoption as part of the same process; this approach takes the state-year as the level of analysis. Each of these approaches allows one to draw on the considerable strengths of EHA, including dealing with common issues like right-censoring and duration dependence (see Beck, Katz and Tucker 1998) and they can each be performed using any variety of EHA models, including discrete ones like logit or probit, or continuous ones like the Weibull and Cox models (though I focus on discrete versions here).[4]

3.1.1 Separate Analysis of Each Component

In the first approach, one would merely estimate K separate EHA models — one for each component; this is equivalent to stacking K different data sets on top of each other.[5] Note that this does not necessarily rule out accounting for activity on the other components in each equation, however, as one can include variables measuring the status and characteristics of a state's activity on the other $K-1$ policy areas; for example, one could test whether diffusion pressures operate across components by including separate diffusion variables for each component. This approach has the advantage of allowing full flexibility in modeling each component — one could conceivably enter completely different explanatory variables in each equation, for example. The trade-off is in the form of parsimony and ignoring the potential commonalities of the components of the policy in question.

3.1.2 Pooled Analysis of All Components

In the second approach, the K EHA equations are pooled into a single equation (that is, the data for the K components are stacked as in the previous approach, but now all the parameters are estimated in a single model).[6] At one extreme, one could assume that the coefficients are the same for each component and estimate a single one for each independent variable. At the other extreme, one could allow distinct coefficients for each variable for each component, effectively estimating the separate models with one command. The first approach emphasizes parsimony, with only M parameter estimates (i.e., if there are the same $M - 1$ variables and a constant in each equation), while the latter maximizes uniqueness, with $K \times M$ estimates. The best approach will usually lie in between these extremes.

To continue the diffusion example, one could use this approach to test whether adoptions of the same component in contiguous states have the same effect for all components by including one diffusion variable or, alternatively, whether diffusion works differently for each component by including separate diffusion measures for each component.[7] Note that in a data set with observations at the state-year-component level, and in a single failure approach (for each component), states drop out of the risk set for a given component once they have adopted it, but they remain in the analysis if they have not yet adopted all K components. In a repeated failures context, they may remain in the risk set if additional adoptions of specific components are possible.

There are a number of assumptions that one may wish to relax in this approach. Besides the possibility of allowing the coefficients for specific explanatory variables to be different for some components, one could allow the baseline hazard rates to vary across components (i.e., stratifying by component). This would be relevant if some components are adopted at a slower rate than others. Allowing this additional form of flexibility is as simple as adding indicator variables for $K-1$ of the components.[8] Further, one can model various patterns of adoption over time. For example, states that have already adopted three components may be more likely to adopt the remaining components than states that have adopted none of them. This can be accounted for by including indicators for the event number (i.e., stratifying by event) or by including a running tally of the total number of components already adopted (various alternatives can be employed to allow for possibly nonlinear effects).[9]

There are, of course, advantages and disadvantages to pooling the components and assuming consistency of coefficients. Advantages include parsimony and the ability to pool information and get more precise estimates of the coefficients by increasing the number of observations. Pooling can be particularly advantageous when relatively few states adopt many of the components. This situation makes it difficult to get good parameter estimates on a component-by-component basis, and can often make it necessary to omit desired explanatory variables from the analysis to avoid perfectly predicting many observations. Pooling components can alleviate many of these problems. With respect to standard errors, pooling creates additional possibilities for clustering: by state (the usual approach), by component, or by component and state.[10]

The disadvantage occurs when the homogeneity assumption is violated, in which case pooling with constant coefficients would result in biased estimates. This could be problematic, for example, if the effect of a variable is positive for one component and negative for a second component, potentially leading to an estimated coefficient near zero. Note that this is not a consequence of pooling the estimating of the components, but of failing to properly allow the effect of the relevant variable to differ across components. For this reason, it will often make sense to start with separate models and then compare the results to identify variables whose coefficients could be pooled into one estimate, for example by conducting explicit tests for pooling those coefficients. Alternatively, one may start with the pooled model and then relax this assumption for specific variables. With at least $K \times M$ different coefficients in the separate approach, researchers may let theoretical concerns and substantive expertise of the policy area help guide them in determining the appropriate level of parsimony.

3.1.3 Combined Analysis of All Components

While the previous approach is analogous to an elapsed time interpretation of the data (i.e., each state is at risk of each event from $t = 0$), the third EHA approach I discuss treats each component as part of the same process. That is, one can think of modeling activity within a policy area by studying whether a state adopts any component in a given year, rather than modeling each component distinctly. This is analogous to a gap time model, which looks at the duration of time between each event.[11] For example, one could ask whether a state adopts any of three anti-smoking policy components in a given year rather than whether it adopts a specific component (see, e.g., Shipan and Volden 2006, 2007). After it adopted its first component, it is still at risk of adopting each of the other components, leading to the repeated failures interpretation; each component is assumed to be an indicator of activity on the broader policy area. The dependent variable for this approach is therefore an indicator for whether there is any policy activity in a given year: $Y_{it} = \max_k \{Y_{ikt} : i \in \cup_k R_k(t)\}$. States are in the risk set as long as they are still at risk of adopting at least one of the K components (though one could allow for repeated adoption of individual components, which would keep states in the risk set indefinitely).[12]

This approach has the drawback that it ignores the uniqueness of each component and makes it difficult to include component-specific variables or coefficients, including explicitly modeling complementarity or substitutability across components. Yet is has at least one advantage relative to a standard EHA, in that by focusing on multiple components rather than the policy as a whole, it retains states in the analysis after initial adoption. This makes it possible to study policy adoption and modification with one model. Since this approach treats the adoption of different components as repeated failures, one may wish to allow for different baseline hazards or covariate effects for the first, second, third events, etc. This is accomplished in the same fashion as discussed in the previous section by including variables for different strata, except here each previously adopted component is considered to be an occurrence of the same event. Conceptually, this approach differs from the pooled approach by assuming that states are at risk of adopting a second component only after they have adopted the first component; in the pooled approach, states are at risk of adopting any component at any time.

3.2 Number of Components Adopted

One common feature of the previous methods discussed is that they each model a dichotomous outcome. The pooled elapsed-time approach is the only one that allows multiple failures to occur in the same year. The repeated events approach only analyzes whether at least one component is adopted in a given year, but ignores information about the number of events adopted that year. This information can be valuable, especially if adoption of multiple components in the same year is fairly common.

One way to incorporate this information is to model the number of components adopted in each year. Thus one would analyze $Y_{it} = \sum_k Y_{ikt}$ for $i \in R_k(t)$.[13] The dependent variable therefore changes from dichotomous to count, and different methods are required.[14] Common ways to model such dependent variables are linear regression, ordered logit or probit, and event count models. For a variety of reasons, event count models such as Poisson regression and the negative binomial model are preferred in such cases, particularly when the counts tend to be small (see, e.g., King (1989a); King (1989b); Maddala (1983); or Cameron and Trivedi (1998) for more information on event count models). In the analysis of state policy adoption one would almost certainly prefer the negative binomial to the Poisson model since the latter assumes that the rate of occurrence of events is constant within time periods (years, in this case). Since states that adopt multiple components in the same year are likely to do so at the same time, this assumption probably does not hold.

One issue that arises in an event count framework is how to deal with states that have adopted some components but not all of them. Assuming that states can only adopt each component once limits the count of components that can be adopted in a given year to the number that have not yet been adopted. This maximum possible count is referred to as exposure and can be accounted for by including the log of the number of unadopted components as an independent variable[15]

Assuming a repeated events context would not limit the total number of potential events. To allow the rate of adoption to depend on previous activity, one may wish to include a tally of or indicators for the number of events that have already occurred.

4 Illustration and Application of Different Approaches

In this section I compare the results obtained using the estimators discussed in the previous section. It is important to remember that the different models correspond to different questions, so one should not expect *à priori* that the results will be the same across estimators. Further, results will also differ across models that assume greater homogeneity of the effects of variables when that assumption is not valid. For the comparison in this paper I use data from state obesity policy.

4.1 Components of State Obesity Legislation

Rising obesity rates in the U.S. — about one third of Americans are overweight (i.e., with a body mass index between 25 and 30) and another third are obese (BMI of at least 30) — have put increasing pressure on policymakers to deal with the associated costs.[16] In the last decade, every state saw an increase in obesity rates, with an average increase of 40% (Mokdad et al.1999). With no action at the Federal level as of 2004, advocats of government action have been forced to rely on the courts or to press state or local officials for action (Kersh and Morone 2005).

As a result, I focus on efforts by state lawmakers to address the obesity issue by studying legislation passed in the fifty states between 1998 and 2005. States have attempted to address the issue in a variety of ways. For example, many states have banned the sale of sugary drinks during certain school hours; others have increased physical education requirements; still others have passed legislation requiring insurers to cover certain types of treatments for obesity (including gastric bypass surgery). At the same time, other states have responded by passing so-called "cheeseburger bills" that protect the fast food industry from consumer lawsuits. Thus there has been a wide range of legislative activity in the states over the past decade to address obesity, with some legislatures seeking to aid the public and others protecting businesses from litigation.

[Table 1 Here.]

To obtain a more complete picture of state activity on this issue, we gathered data from a number of sources.[17] After reading through the summaries of these bills we coded them based on the different issues that they addressed and then coded these different issues into eleven primary components. Table 1 lists these eleven components and shows the number of adopted bills that address each component per year (additional information on each policy area is contained in an online appendix). Note that since bills may address multiple areas at once, there are more components adopted than there are bills enacted. Further, states may pass more than one piece of legislation in each area, which explains how some components are adopted more than fifty times. The four areas with the most activity are Committee/Research (76 times), Physical Education (67), Awareness (65), and School Nutrition (59); at the other end of the spectrum are Insurance (11) and Labeling (3). These data are used to construct various models to explain state obesity policy and to illustrate many of the issues discussed in the previous section.

4.2 Control Variables

The models contain a number of variables intended to explain state-level adoptions of obesity legislation components, most of which are common in the state policy adoption and diffusion literature. Political variables include state-level ideology (Erikson, Wright and McIver 1993), divided government, unified Republican government, and legislative professionalism in 1990 (King 2000). Measures of innovativeness include total state population and real income per capita.[18] I include additional variables to help measure the demand for public policy solutions to the obesity epidemic by calculating the percentage of a state's population that is overweight or obese.[19] The results presented are not intended to be definitive explanations of obesity legislation adoption, but are used to facilitate model comparison. That said, the results tend to be fairly robust to the inclusion of other variables.

In order to illustrate the differences between the methodological approaches, I focus on interstate diffusion pressures. In the context of state obesity policy, most components would be expected to diffuse based more on social learning pressures rather than economic pressures, but components such as insurance funding and medical treatment could lead to interstate competition for business. Diffusion is generally measured by the number of neighboring states that have adopted the policy in question (see, e.g., Berry and Berry 1990). When studying policies with multiple components, however, the measure of diffusion may differ depending on the level of analysis or the research question. When studying each component separately, one would likely measure diffusion pressures with the number of neighboring states that have adopted that component, ignoring neighbors' adoptions of other components. When pooling the analysis, one could retain this measure, use a measure of the total number of all components adopted by neighboring states, or both. FInally, when aggregating the data to the state-year, one could measure the number of neighboring states that have adopted at least one component or the total number of components adopted in neighboring states. Additional variations could be employed as desired.

4.3 Analyzed Separately

The first analysis I present compares separate EHA analyses of the adoption of ten components (Labeling did not have enough activity to generate meaningful statistical results – it is therefore not analyzed on its own). This approach ignores any commonality or cross-component influences and involves estimating ten separate event history models. This may be useful in the context of obesity legislation since some laws protect consumers while others protect the fast food industry; further, some of the laws are perhaps mere grandstanding (e.g., "Walking Wednesday," which encourages parents and children to walk to school together) while others have very specific policy bite. For the diffusion variable I measure the number of contiguous neighbors that have adopted the same provision before the current year. Here I report the results for a repeated events framework, and omit the results for the time until the first event, which were broadly similar. To account for different baseline hazards across events, I include an indicator for whether a state has previously passed legislation for the component in question.

[Table 2 Here.]

Overall, the results in Table 2 indicate some consistency across components in the factors that influence adoption. For example, states with unified Republican governments are consistently less likely to adopt any of the provisions, with significant coefficients for five of the ten areas; larger states tend to be significantly more likely to adopt many components as well. On the other hand, some differences do emerge. For example, percent overweight has a positive effect on Awareness legislation but a negative effect on Liability legislation; liberal ideology has a positive effect on nutrition legislation but a negative effect on transportation/infrastructure legislation. This serves to highlight the fact that adoption of policy components may be driven by different factors. Further, note that there is little evidence of event dependence, in the sense that states that have already adopted a component are not significantly more or less likely to pass more legislation on that component (the one exception is nutrition legislation). With respect to policy diffusion, note that while seven of the ten coefficients are positive, the Transportation component produces the lone significant finding with a negative coefficient. Based on this analysis, then, one would not conclude that diffusion pressures play a significant role. Note that since many of the components experience relatively little activity (the majority are adopted no more than eighteen times), analyzing the components separately may not produce the most informative comparison since fluctuations in coefficients may emerge somewhat randomly.

4.4 Pooled Analyses

In this section I report the results of variations on two pooled analyses: one that examines the time until the first event for each component and one that allows for repeated adoption of each component. Observations are therefore at the state-year-component level. In its simplest form, this assumes that the coefficients for each of the eleven issues areas are the same across the separate models. Yet this assumption can be relaxed in a number of ways, some of which I highlight here. First, I include fixed effects for each component. This allows the baseline rate of adoption to vary across components, which seems likely given that some components are adopted by most states while others, such as restaurant labeling, are only adopted by a handful. I also study patterns of adoption across components by including ten indicator variables for whether a state has adopted each of the other provisions.[20] These indicators allow me to test, for example, whether states that have adopted liability laws are less likely to adopt other components.

[Table 3 Here.]

The results for these more parsimonious models are presented in Table 3. These results show more consistency in the effects of the various independent variables: five of them have significant effects in every model. The percentage of a state's population that is overweight has a positive and significant effect on adoption. Both unified Republican government and divided government have negative and significant effects, while total population is always positive and significant. Also, states with more professionalized legislature are less likely to adopt obesity-related legislation, with significant coefficients in three models (and nearly so in a fourth). Compared to the previous analysis by component, some of these results are not surprising — population had a positive effect in nine of ten components and was significant for five of those nine. Others may be more so: divided government was negative for seven components, but significant in only one.

In terms of diffusion pressures, the results depend on whether the model accounts for component heterogeneity. The two models that pool the components but do not include fixed effects to allow for different baseline hazard rates for each component produce positive and highly significant coefficients for adoptions in neighboring states. But the two models that add component fixed effects produce smaller, insignificant coefficients. Not all methods of accounting for heterogeneity affect the results, though, as the final model shows that adding fixed effects for whether a state has adopted the other components still results in strong evidence for cross-state diffusion.

Examining the fixed effects in more detail indicates that they are statistically significant: a χ^2 test of their joint effect strongly rejects the null hypothesis that they should be excluded. While these coefficients are not shown, the results for the component fixed effects indicate that compared to the baseline category of awareness legislation, all of the other categories have significantly lower baseline hazard rates, with the exceptions of Creating Committees, Physical Education and School Nutrition, which are generally negative though not significant. The results for lagged adoption of other components indicate that the adoption of labeling laws significantly decreases while the adoption of nutrition laws significantly increases future adoptions of other components. Finally, the repeated events models indicate that a another form of heterogeneity matters: when a state has already adopted a provision, it is significantly more likely to pass additional legislation on that provision.

4.5 Combined Analysis for Same Event

In this section I present results from models that treat the different components as identical events. That is, the unit of analysis shifts to the state-year level and the adoption of any component is considered as an obesity legislation event. In the event history models, the dependent variable is coded as a one in any year in which a state adopts at least one component. In the negative binomial and OLS models for the number of events, the dependent variable is the number of components adopted. In the obesity legislation case, about two-thirds of the observations with at least one event have more than one component adopted. This suggests that a substantial amount of information may be lost when this count is reduced to an indicator.[21]

[Table 4 Here.]

These results are presented in Table 4. Some familiar patterns emerge. The percentage of a state's population that is overweight has a positive and significant effect in two of the models, and nearly significant effects in the other two. Unified Republican government and poorer states have fewer adoptions. Aggregating the pooled analysis to the state-year level results in less information, however, as evidenced by the current lack of significance for divided government and legislative professionalism. The models of the number of events appear to recapture some of this information, as divided government significantly decreases the number of components adopted. Finally, note that none of these models provides any evidence of cross-state diffusion pressures.

5 Discussion

This paper discusses various models available for studying the adoption of policies with multiple components. While there are often theoretical and statistical advantages to considering the richness of a policy through its multiple components, doing so also presents methodological challenges. In light of the variety of approaches that have been applied in studies of policies with multiple components to date, the goal of this paper has been to present various approaches, discuss the different questions they address, some of their strengths and weaknesses, and new methodological issues that arise in the study of multiple components.

The most important issue that emerges is the researcher's desire to achieve parsimony while respecting heterogeneity across components. The balance between these two will help determine the preferred amount of pooling. If any pooling is done, however, the researcher needs to evaluate the appropriateness of the associated assumptions. For example, some policy components may be subject to economic competition pressures across states, while others may not (e.g., the number and type of slot machines allowed for Native American casinos versus whether or not the state has a revenue-sharing fund to spread profits to all tribes). In this situation one may wish to estimate two coefficients for diffusion: one for components that involve interstate competition and one for those that do not. One could then test whether the two types of components have the predicted effects.

When allowing for cross-component heterogeneity, at a very minimum one should consider controlling for differences in the baseline hazard rates across components and across repeated failures. As evidenced in the analysis of state obesity policy, allowing for component-specific differences can have important consequences for the conclusions reached about theoretically relevant variables, such as policy diffusion. One can allow for these forms of heterogeneity with fixed or random effects at the component level; a more advanced treatment could allow for both state- and component-specific effects. In addition, one should also pay attention to the estimates of the standard errors. In addition to possibly clustering errors by state, one could cluster them by component or by component and state.

While I wish to underscore the lack of a one-method-fits-all approach to these models given the different theoretical and substantive nuances that may arise, some approaches may be more appropriate more often. In general, then, one may prefer to start with a separate analysis of each component and then look for opportunities to move towards a pooled analysis that emphasizes as much parsimony as possible. One may then consider moving to a combined model that treats each component as part of the same process in cases in which there is not much evidence of heterogeneity, since these models limit the opportunity to model it. Along the way, the researcher should use statistical tests, theoretical foundations, and substantive knowledge as guidance.

Notes

[1] I use the publication of Berry and Berry's (1990) article as the inception of the EHA era in political science.

[2] Alternatively, some studies construct policy indices that take a linear combination of a state's values for the different components (e.g., Yu 2005; Jacoby and Schneider 2001). Since these studies move well beyond event history approaches, I do not discuss them here.

[3] Readers seeking additional information and discussion of event history model, and duration models in general, are referred to Allison(1984) and Box-Steffensmeier and Jones (2004).

[4] See Jones and Branton (2005) for a discussion of the use of continuous-time duration models for studying state policy adoptions. I do not discuss these estimators in this paper, but all of the discrete EHA models I estimate could just as easily be estimated using a continuous-time approach.

[5] In Stata, one can go from state-year observations with variables for each component to a data set with state-year-component observations using the `reshape` command. To estimate, one would merely focus on the observations for the relevant component: `logit y x_1 ...x_n if component==k`, where k identifies the component.

[6] Alternatively, one could apply the approach proposed by Wei, Lin, and Weissfeld (1989) to model repeated failures as competing risks. Because the policy components being adopted are believed to be potentially different in the approach that I outline here, it is appropriate to assume that a state is at risk of each event (adopting a specific component) in each period for which it is in the risk set for that component.

[7] With two components, one would specify `logit y x_1 ...x_n diff` for the first and `logit y x_1 ...x_n diff_1 diff_2` for the second specification (where `diff_1` is zero for observations for component 2 and vice versa and `diff=diff_1 + diff_2`; a likelihood ratio test would

help distinguish between the two.

[8]This can be accomplished manually or automatically using the `xi` command in Stata: `xi: logit y x_1 ...x_n i.component`.

[9]Interested readers are referred to Box-Steffensmeier and Zorn (2002) for a detailed discussion of these issues and other important considerations that arise in the context of repeated events duration analysis.

[10]A convenient way to cluster by component and state is to create a variable that uniquely identifies the intersection of the two. With the fifty state states indexed by i and the components by k, one could merely create a variable equal to $50 \times i + k$ (assuming there are fewer than fifty components).

[11]Note that under the assumption that adoption of any component represents the same event, this model is essentially the gap-time model for repeated events proposed by Prentice, Williams and Peterson (1981).

[12]Note that a special case of this approach is modeling the time until the first component is adopted: a state is in the risk set as long as none of the components is adopted and once at least one is adopted, it is dropped from the analysis (e.g., $R(t) = \cap_k R_k(t)$). Essentially, one would estimate the effects on timing until first activity.

[13]In some applications states may enact multiple pieces of legislation for a single component in the same year, which would merely require allowing Y_{ikt} to be integer-valued.

[14]See Alt, King, and Signorino (2001) for a discussion of moving between different levels of aggregation for event data. Here, one aggregates from the state-policy-year level to the state-year level.

[15]If the coefficient is restricted to be one, this is equivalent to estimating the proportion of remaining components adopted in a given year. See King (1989b) or Maddala (1983) for more information on including the logarithm of exposure as an independent variable.

[16]See Flegal, Carroll, Ogden, and Johnson (2002) for definitions and rates. Studies suggest that the costs of obesity-related conditions account for 7-10% of health care costs in the U.S. (Mokdad et al. 1999; Kersh and Morone 2005).

[17]These included reports and databases from the National Conference of State Legislatures, reports sponsored by the Robert Wood Johnson Foundation and put together by Health Policy Tracking Service, and two databases available through the Center for Disease Control: State Obesity-Related Legislation and Nutrition and Physical Activity. Legislation was cross-checked and in some cases compared to the listings available from state legislature's websites; in a few cases discrepancies between the various sources could not be resolved.

[18]These variables are culled from the U.S. Census and the Statistical Abstract of the States, various years.

[19]These two variables are constructed according the definitions mentioned previously and are calculated using data from the Behavioral Risk Factor Surveillance System's 1998-2005 studies.

[20]I was only able to include these indicators in the repeated events model. Also, I set the value of the variable to zero for observations corresponding to the same component since I already have a variable for the presence of the current component.

[21]The number of events ranges from zero to seventeen, with a mean of one. There are 147 cases out of 400 with a positive number of events, and the conditional mean number of events is 2.75.

References

Allison, Paul D. 1984. *Event History Analysis: Regression for Longitudinal Event Data.* Sage University Paper Series on Quantitative Applications in the Social Sciences, series no. 07-046. Beverly Hills and London: Sage Publications.

Alt, James E.; Gary King; and Curtis Signorino. 2001. "Aggregation Among Binary, Count, and Duration Models: Estimating the Same Quantities from Different Levels of Data." *Political Analysis* 9 (1): 21-44.

Beck, Nathaniel, Jonathan N. and Richard Tucker. 1998. "Taking Time Seriously: Time-Series-Cross-Section Analysis with a Binary Dependent Variable." *American Journal of Political Science* 42(4): 1260-1288.

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

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

Boehmke, Frederick J. and Richard Witmer. 2004. "Disentangling Diffusion: The Effects of Social Learning and Economic Competition on State Policy Innovation and Expansion." *Political Research Quarterly* 57: 39-51.

Box-Steffensmeier, Janet M. and Bradford D. Jones. 2004. Event History Modeling: A Guide for Social Sciences. Cambridge: Cambridge University Press.

Box-Steffensmeier, Janet M., and Christopher Zorn. 2002. "Duration Models for Repeated Events." *Journal of Politics* 64(4): 1069-94.

Cameron, A. Colin and Pravin K. Trivedi. 1998. *Regression Analysis of Count Data.* Cambridge: Cambridge University Press.

Council of State Governments. Various years. *The Book of the States.* Lexington: Council of State Governments.

Erikson, Robert S., Gerald C. Wright, and John P. McIver. 1993. *Statehouse Democracy: Public Opinion and Policy in the American States*. Cambridge: Cambridge University Press.

Flegal, Katherine M., Margaret D. Carroll, Cynthia L. Ogden, and Clifford L. Johnson. 2002. "Prevalence and Trends in Obesity Among US Adults, 1999-2000." *Journal of the American Medical Association* 288(14): 1723-27.

Imhof, Sara. 2006. *Promoting a "Good Death": An Investigation of the Determinants that Explain State Medical Board Adoptions of Pain Management-Enhancing Policies in the United States.* Ph.D. dissertation, University of Iowa.

Jacoby, William G.; Saundra K. Schneider. 2001. "Variability in State Policy Priorities: An Empirical Analysis." *The Journal of Politics* 63(2): 544-568.

Jones, Bradford S. and Regina P. Branton. 2005. "Beyond Logit and Probit: Cox Duration Models for State Policy Adoption." *State Politics and Policy Quarterly* 5(4): 420-443.

Karch, Andrew. 2007. "Emerging Issues and Future Directions in State Policy Diffusion Research." *State Politics and Policy Quarterly* 7 (1): 54-80.

Kersh, Rogan and James A. Morone. 2005. "Obesity, Courts, and the New Politics of Public Health." *Journal of Health Politics, Policy and Law* 30(5): 839-868.

King, Gary. 1989a. "Statistical Models for Political Science Event Counts: Bias in Conventional Procedures and Evidence for The Exponential Poisson Regression Model." *American Journal of Political Science* 32: 838-863.

King, Gary. 1989b. "Variance Specification in Event Count Models: From Restrictive Assumptions to a Generalized Estimator." *American Journal of Political Science* 33: 762-784.

King, James D. 2000. "Changes in Professionalism in U.S. State Legislatures." *Legislative Studies Quarterly* 25(2): 327-343.

Maddala, G.G. 1983. *Limited Dependent and Qualitative Variables in Econometrics.* Cambridge: Cambridge University Press.

Mokdad, Ali H., Mary K. Serdula, William H. Dietz, Barbara A. Bowman, James S. Marks, and Jeffrey P. Koplan. 1999. "The Spread of the Obesity Epidemic in the United States, 1991-1998." *Journal of the American Medical Association* 282(16): 1519-22.

Prentice, R. L.; B. J. Williams; A. V. Peterson. 1981. "On the Regression Analysis of Multivariate Failure Time Data." *Biometrika* 68(2): 373-379.

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. 2007. "Local Antismoking Laws and the Mechanisms of Policy Diffusion." Working paper.

U.S. Bureau of the Census. Various years. *Statistical Abstract of the United States.* Washington: US Government Printing Office.

Volden, Craig. 2006. "States as Policy Laboratories: Emulating Success in the Children's Health Insurance Program." *American Journal of Political Science* 50: 294-312.

Wei, L. J.; D. Y. Lin; and L. Weissfeld. 1989. "Regression Analysis of Multivariate Incomplete Failure Time Data by Modeling Marginal Distributions." *Journal of the American Statistical Association* 84(408): 1065-73.

Yackee, Susan Webb. 2005. "Coalitions in Conflict: Competing Interests in State Medical Malpractice Laws." Paper Presented at the State Politics and Policy Conference.

Yu, Eric C. H. 2005. "The Impact of Welfare Reforms on Welfare Caseload." Paper Presented at the State Politics and Policy Conference.

Table 1: Adoption of Obesity Legislation Components by Year

	1998	1999	2000	2001	2002	2003	2004	2005	Total	# States
Awareness	1	2	0	4	8	20	14	16	65	29
Committee/Research	1	2	2	16	4	26	12	13	76	31
Health Education	0	0	0	0	0	7	6	16	29	18
Insurance	0	1	0	1	1	2	2	4	11	9
Labeling	0	0	0	0	0	1	1	1	3	3
Liability	0	0	0	0	0	2	14	8	24	22
Medical Treatment	0	0	2	1	0	4	3	8	18	11
Nutrition/Wellness	0	0	0	2	2	5	8	11	28	13
Physical Education	0	0	1	4	5	14	16	27	67	31
School Nutrition	0	0	0	2	8	14	10	25	59	32
Transportation	0	0	0	4	2	3	9	7	25	16
Total	2	5	5	34	30	98	95	136	405	215

Source: author’s data, compiled from various sources. See text for more information.

Table 2: Probit Event History Analyses of Repeated Policy Adoptions, by Component

	Aware.	Comm.	Health Ed.	Insur.	Liab.	Treat.	Nutr.	P.E.	Sch. Nutr.	Transp.
% Overweight BMI	6.468**	2.028	9.809	1.258	−22.816**	−2.993	−0.000	1.675	3.383	0.246
	(2.348)	(2.871)	(13.061)	(4.285)	(11.632)	(3.441)	(3.152)	(2.244)	(2.308)	(3.061)
% Obese BMI	−2.720	−1.154	−3.300	2.718	−9.649	1.849	2.669	−4.252	−4.200	2.685
	(3.749)	(4.246)	(6.665)	(8.345)	(7.255)	(6.649)	(5.795)	(3.690)	(4.011)	(4.733)
Liberal Ideology	−0.006	−0.017	−0.039	−0.036	−0.034	−0.034	0.040*	−0.018	0.024	0.090**
	(0.016)	(0.021)	(0.033)	(0.025)	(0.028)	(0.025)	(0.023)	(0.016)	(0.017)	(0.032)
Unified Republican Government	−0.812**	−1.372**	−0.533	0.288	0.454	−0.387	−0.914*	−0.691*	−0.787**	−0.449
	(0.389)	(0.332)	(0.457)	(0.451)	(0.565)	(0.377)	(0.526)	(0.381)	(0.350)	(0.462)
Divided Government	−0.359	−0.284	−0.035	0.080	0.233	−0.070	−0.390	0.001	−0.347	−0.950**
	(0.250)	(0.176)	(0.389)	(0.439)	(0.491)	(0.241)	(0.282)	(0.310)	(0.294)	(0.257)
Total Population	0.051**	0.050**	0.057	−0.006	0.049	0.040**	0.021	0.093**	0.050**	0.050
	(0.022)	(0.018)	(0.035)	(0.031)	(0.031)	(0.017)	(0.034)	(0.022)	(0.023)	(0.040)
Personal Income	−0.048	−0.038	−0.008	0.079*	−0.017	0.007	−0.026	−0.019	−0.033	0.021
	(0.030)	(0.031)	(0.043)	(0.043)	(0.046)	(0.051)	(0.036)	(0.035)	(0.029)	(0.030)
Legislative Professionalism	0.685	0.262	−0.492	0.743	−2.574	0.281	−1.002	−3.054**	−1.413	−2.031
	(0.843)	(0.837)	(1.454)	(1.216)	(1.804)	(0.991)	(1.464)	(0.967)	(0.900)	(1.617)
Neighbors' Adoptions	0.021	0.070	−0.057	0.290	0.116	−0.330	0.021	0.002	0.102	−0.821**
	(0.097)	(0.081)	(0.235)	(0.316)	(0.230)	(0.432)	(0.217)	(0.111)	(0.091)	(0.360)
Time	0.497**	0.448**	1.998**	−0.038	11.177**	0.173	0.509	0.648**	0.590**	0.819**
	(0.252)	(0.185)	(0.965)	(0.215)	(4.641)	(0.288)	(0.321)	(0.285)	(0.281)	(0.410)
Time Squared	−0.034	−0.034	−0.125	0.016	−0.864**	0.019	−0.033	−0.020	−0.019	−0.044
	(0.029)	(0.023)	(0.091)	(0.029)	(0.383)	(0.035)	(0.035)	(0.032)	(0.030)	(0.047)
Has this Provision	0.017	0.018	−0.014	−0.044		0.268	0.790**	0.063	−0.176	−0.118
	(0.238)	(0.225)	(0.492)	(0.485)		(0.382)	(0.291)	(0.266)	(0.259)	(0.615)
Constant	−3.053**	−1.834	−11.773*	−6.439**	−25.905*	−3.465*	−2.031	−2.751**	−2.094**	−3.490**
	(1.155)	(1.205)	(6.201)	(1.360)	(14.349)	(1.901)	(1.307)	(1.153)	(1.024)	(1.242)

Notes: N=384 (except liability, for which N=369). ** indicates $p \leq .05$; * $p \leq .10$. Standard errors calculated by clustering on state.

Table 3: Pooled Event History Analyses of Obesity Legislation Adoptions

	First Event		Repeated Events		
% Overweight BMI	1.953**	1.864*	2.314**	1.991**	2.635 * *
	(0.889)	(1.029)	(0.809)	(0.915)	(0.880)
% Obese BMI	−1.761	−1.629	−2.370	−1.808	−3.022*
	(1.551)	(1.841)	(1.446)	(1.662)	(1.597)
Liberal Ideology	−0.006	−0.007	−0.003	−0.004	−0.004
	(0.009)	(0.010)	(0.007)	(0.009)	(0.009)
Unified Republican Gov't	−0.521**	−0.618**	−0.549**	−0.612**	−0.545 * *
	(0.162)	(0.178)	(0.146)	(0.159)	(0.159)
Divided Government	−0.246**	−0.273**	−0.216**	−0.226**	−0.185*
	(0.115)	(0.118)	(0.107)	(0.109)	(0.108)
Total Population	0.038**	0.046**	0.040**	0.047**	0.040 * *
	(0.008)	(0.009)	(0.009)	(0.009)	(0.010)
Personal Income	−0.008	−0.014	−0.011	−0.014	−0.014
	(0.014)	(0.016)	(0.013)	(0.015)	(0.013)
Legislative Professionalism	−0.672*	−0.779*	−0.658	−0.776*	−0.494
	(0.385)	(0.405)	(0.404)	(0.408)	(0.461)
Neighbors' Adoptions	0.217**	0.036	0.165**	0.002	0.167 * *
	(0.036)	(0.045)	(0.034)	(0.043)	(0.034)
Time	0.326**	0.334**	0.357**	0.361**	0.349 * *
	(0.116)	(0.120)	(0.122)	(0.126)	(0.125)
Time Squared	−0.014	−0.006	−0.017	−0.009	−0.014
	(0.012)	(0.013)	(0.013)	(0.013)	(0.013)
Has this Provision			0.378**	0.172*	0.331 * *
			(0.095)	(0.104)	(0.146)
Constant	−2.736**	−2.263**	−2.680**	−2.348**	−2.677 * *
	(0.471)	(0.535)	(0.449)	(0.502)	(0.444)
Fixed Effects by:	*None*	*Comp.*	*None*	*Comp.*	*Other Comp.*
χ^2 statistic (FE = 0)		78.59		113.33	34.60
p value		0.00		0.00	0.00
Observations	3873		4224		

Notes: Fixed effects' coefficients not reported. ** indicates $p \leq .05$; * $p \leq .10$.
Standard errors calculated by clustering on state.

Table 4: Analysis of Obesity Legislation Components Adopted per Year

	Discrete - EHA		Total Number	
	First	**Repeated**	**Neg. Bin.**	**OLS**
% Overweight BMI	4.752	$4.030*$	$4.706**$	2.694
	(2.935)	(2.177)	(2.091)	(1.950)
% Obese BMI	-6.974	-5.406	-4.693	-4.310
	(5.231)	(3.765)	(3.628)	(3.459)
Liberal Ideology	$-0.041**$	-0.013	-0.021	-0.005
	(0.019)	(0.016)	(0.017)	(0.014)
Unified Republican Gov't	$-1.528**$	$-1.224**$	$-1.233**$	$-1.026**$
	(0.454)	(0.335)	(0.329)	(0.341)
Divided Government	-0.464	-0.276	$-0.452*$	$-0.608*$
	(0.304)	(0.219)	(0.240)	(0.344)
Total Population	$0.065**$	$0.073**$	$0.075**$	$0.102**$
	(0.029)	(0.028)	(0.017)	(0.028)
Personal Income	-0.012	$-0.045*$	-0.029	$-0.041*$
	(0.028)	(0.026)	(0.028)	(0.024)
Legislative Professionalism	-0.637	-1.059	-0.648	0.106
	(1.059)	(1.141)	(0.778)	(1.052)
Neighbors' Adoptions	-0.074	-0.049	0.033	0.017
	(0.111)	(0.070)	(0.059)	(0.121)
Time	$0.618**$	$0.599**$	$0.945**$	0.064
	(0.232)	(0.185)	(0.280)	(0.165)
Time Squared	-0.013	-0.014	$-0.053*$	$0.041*$
	(0.031)	(0.021)	(0.029)	(0.020)
Has any Provision		0.173	$0.341*$	$0.543*$
		(0.195)	(0.190)	(0.294)
Constant	$-2.470**$	-1.263	$-3.302**$	0.939
	(1.037)	(0.916)	(1.029)	(0.627)
Observations	245		384	

Notes: ** indicates $p \leq .05$; * $p \leq .10$. Standard errors calculated by clustering on state.

A Components of Obesity Policy

1. *Awareness*: promote awareness of obesity as an issue and ways to deal with it. May create health awareness days, weeks, etc.; government advertising campaigns; promote BMI measurement; creating BMI report cards for students.

2. *Committee/Research*: create or continue a government commission, committee or task force, either at the state or local level, to study or create policy solutions.

3. *Health Education*: promote or require health education in schools.

4. *Insurance*: address coverage for obesity-related treatment or for treatment of related problems. Includes private and Medicaid coverage.

5. *Labeling*: restaurant food nutrition labeling.

6. *Liability*: provide immunity from liability for food retailers relating to individuals' consumption-related weight gain or associated conditions.

7. *Medical Treatment*: restrict or permit medical treatments for obesity.

8. *Nutrition/Wellness*: promote public nutrition, fitness and wellness. Includes nutritious grocers acts and farm to market acts.

9. *Physical Education*: relating to physical education requirements in schools or school athletic programs.

10. *School Nutrition*: promote nutritious meals in schools, school meal programs, vending machine restrictions, contracts with food and beverage vendors, farm to school policies, and school/community gardens.

11. *Transportation*: improve infrastructure for walking/jogging/cycling, safe routes to school acts, promote or facilitate personal locomotion.