A Diffusion Network Event History Estimator*

Preanalysis Plan

Jeffrey J. Harden†
Mark Brockway‡
Frederick J. Boehmke§
Bruce A. Desmarais¶
Scott LaCombe‖
Fridolin Linder**

March 31, 2019

Contents

*This preanalysis plan was deposited at the Political Science Registered Studies Dataverse on March 31, 2019.

†Associate Professor, Department of Political Science, University of Notre Dame, 2055 Jenkins Nanovic Halls, Notre Dame, IN 46556, jeff.harden@nd.edu.

‡Graduate Student, Department of Political Science, University of Notre Dame, 2060 Jenkins Nanovic Halls, Notre Dame, IN 46556, mark.d.brockway.4@nd.edu.

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

¶Associate Professor, Department of Political Science, Pennsylvania State University, 321 Pond Lab, University Park, PA 16802, bdesmarais@psu.edu.

‖Graduate Student, Department of Political Science, University of Iowa, 341 Schaeffer Hall, Iowa City, IA 52242, scott-lacombe@uiowa.edu.

**Graduate Student, Department of Political Science, Pennsylvania State University, 230 Pond Lab, University Park, PA 16802, fridolin.linder@psu.edu.

1 Introduction

This document describes plans for a methodological replication analysis designed to compare the conventional method for estimating models of public policy adoption—discrete-time event history analysis (EHA) and pooled event history analysis (PEHA)—with an alternative suite of methods that we refer to as network event history analysis (NEHA). These new methods are potentially beneficial to applied researchers because they incorporate estimation of parameters on covariates *and* latent diffusion networks that connect units, which Desmarais, Harden, and Boehmke (2015) demonstrate to be an important part of the policy adoption process (see also Boehmke et al. 2017; Boehmke et al. 2019).

Following the evaluation framework described in Harden, Sokhey, and Wilson (2019), this preanalysis plan is intended to document our goals prior to executing the replication analysis. Doing so guards against selection bias and overconfidence in new methods. While these methods are sufficiently general for use in a variety of substantive settings, our focus is on models of policy adoption in the American states.

2 Summaries of Methods

Here we summarize the methods we plan to compare in this replication analysis. These methods include approaches for analyzing single-policy diffusion studies and multiple-policy studies.

- **EHA (single-policy studies)**. Berry and Berry (1990) pioneered the use of discrete-time event history analysis of policy diffusion in their study of state lottery adoptions. The unit of analysis for this estimator is state-year and the outcome variable is coded 1 if a state adopted the policy in a given year and 0 otherwise. Additionally, states that adopt the policy drop out of the data in years after it adopted. The analyst can estimate parameters on state-level and time-varying covariates.
- **PEHA (multiple-policy studies)**. This method extends the logic of EHA (Kreitzer and Boehmke 2016, 122). It is a discrete-time event history estimator that models adoption of multiple policies by stacking the data from each policy and estimating the parameters in a

single model. The unit of analysis is policy-state-year, and the researcher estimates a single set of parameters that represent the average effects of the covariates on policy adoption across the range of policies. With sufficient policies included, researchers may include random effects and random coefficients to account for heterogeneity across policies (or units). PEHA allows researchers to test theories of policy diffusion by focusing on consistent patterns of diffusion over many policies rather than informally aggregating results one policy at a time.

- **NEHA (single-policy studies)**. This method employs a two-step procedure to include patterns of policy diffusion into a standard EHA model of adoption of one policy. Specifically, the model first uses the `NetInf` algorithm (Gomez-Rodriguez, Leskovec, and Krause 2010; Desmarais, Harden, and Boehmke 2015) to infer latent diffusion network ties from a large database of policies, such as Boehmke et al.'s (2019) State Policy and Innovation Database (SPID). Then information from this inferred network is incorporated into the EHA model with a covariate measuring the number or proportion of a given state's diffusion sources (as inferred by the network) who previously adopted the policy of interest.
- **NEHA (multiple-policy studies)**. This method incorporates inference regarding latent diffusion pathways into a standard PEHA model. It is a discrete-time event history estimator that models (1) unit-level effects on policy adoption, (2) dyadic effects on the tendency of units to emulate each other, and (3) residual dyadic ties between units (i.e., latent diffusion networks). NEHA includes functions of previous policy adoptions by other units (i.e., states) as covariates in the standard PEHA linear predictor. These latent "network effects" are constrained such that they must increase the likelihood of policy adoption. Additionally, it employs regularization to impose sparsity on the latent network effects, ensuring that the model does not estimate a parameter on many potential ties between units, which would be infeasible.

3 The Evaluation Framework

In Harden, Sokhey, and Wilson's (2019) evaluation framework the first step is to preregister the replication plan. This step involves (a) defining the relevant population of studies, (b) listing the target sample from this population, and (c) defining replication quantities of interest. We address each of these items below.

3.1 Defining the Population

The relevant population for this replication analysis is any political science or public policy study that employs EHA (for single policies) or PEHA (for multiple policies) to study policy adoption and/or policy diffusion between the American states. As noted above, this is a somewhat narrow definition in that the study of diffusion is quite broad, spanning American politics, comparative politics, and international relations (Graham, Shipan, and Volden 2013). Nonetheless, we focus on public policy diffusion in American state politics research because that subfield was instrumental in the development of EHA and PEHA and uses those methods frequently.

3.2 The Target Sample

To collect a sample of replication studies from this population, we first performed a search for peer-reviewed journal articles, books, and other scholarly sources that relied on data pertaining to diffusion of policies in the American states. First, we used Google scholar to search for sources using the following search terms:

```
(((((policy AND diffus*) OR (policy AND innovat*)) OR (policy AND
contagion)) OR (policy AND adopt*)) AND (Walker)).
```

In addition, we searched for sources that cited Walker (1969) under the assumption that any study of state policy diffusion would do so due to the fact that it is a seminal article in this literature. This strategy produced several thousand results from which several hundred candidate articles were identified that specifically related to policy diffusion in the American states. From these, 94 journal articles, books, and other sources were culled and analyzed for data fit and the availability of replication materials. Finally, we consulted our own knowledge of the literature and conferred with colleagues to identify other studies that we might have missed in this systematic search.

Overall, we acquired replication materials for *and* successfully replicated 16 EHA or PEHA models.[1] We report information on these studies in Table 1. Many of them report multiple model

[1] Our focus in this replication analysis is on monadic policy diffusion models, in which the unit of analysis is

Table 1: State Policy Diffusion Studies in the Planned Replication Analysis

Citation	Policy Data	Policy/Policy Area	Model	Specification
Berry and Berry (1990)	Single	Lotteries	Probit	Table 1, column 1 (406)
Boehmke (2005) [1]	Single	Captial punishment	Probit	Table 4.2 (85)
Boehmke (2005) [2]	Single	Indian gaming	Probit	Table 4.4 (89)
Shipan and Volden (2006) [1]	Single	Government building smoking bans	Logit	Table 2, column 2 (836)
Shipan and Volden (2006) [2]	Single	Restaurant smoking bans	Logit	Table 2, column 3 (836)
Shipan and Volden (2006) [3]	Single	Out-of-package cigarette sale restrictions	Logit	Table 2, column 4 (836)
Boehmke (2009*b*)	Single	Pain management policy	Logit	Table 2, column 5 (1135)
Hannah and Mallinson (2018)	Single	Medical marijuana laws	Logit	Table 2 (415)
Boehmke (2009*a*)	Multiple	Obesity policy (11)	Probit	Table 3, column 2 (243)
Makse and Volden (2011)	Multiple	Criminal justice (27)	Logit	Table 2, column 6 (118)
Desmarais et al. (2015)	Multiple	Several (152)	Logit	Appendix Table A.3 (22)
Hinkle (2015)	Multiple	Abortion (6)	Probit	Table A4 (1018)
Boushey (2016)	Multiple	Criminal justice (44)	Probit	Table 2, column 1 (206)
Kreitzer and Boehmke (2016)	Multiple	Abortion (29)	Logit	Table 1, column 1 (135)
Karch et al. (2016)	Multiple	Interstate compacts (79)	Logit	Table 2, column 1 (90)
Boehmke et al. (2017)	Multiple	Several (88)	Logit	Table 1 (294)

Note: Cell entries report data and specification information on the EHA (single policy) and PEHA (multiple policy) models we plan to replicate. For multiple-policy data, the number of policies is reported in parentheses in the policy area column.

specifications in their text. We plan to analyze all coefficient estimates in the model specifications listed in Table 1. We identified these specifications as the main regressions from the authors' discussion of their results.[2]

3.3 Quantities of Interest

We plan to focus on several replication quantities of interest (RQI) in our evaluation of these two different methods. The first RQI is the ratio of the absolute values of the EHA or PEHA coefficient estimates to the absolute NEHA estimates.[3] We refer to this quantity as the coefficient ratio (CR). It is equal to 1 if the EHA (or PEHA) and NEHA estimates are equivalent, greater than 1 if the EHA (PEHA) estimate is larger in magnitude, and less than 1 if the NEHA coefficient is

state-year or state-policy-year. Dyadic analyses also appear in the literature, but those models are beyond the scope of the NEHA estimator.

[2]If more than one model appeared to be the main one, we chose the first one presented in the text.

[3]We will compute this measure for coefficients on substantive covariates only because those are the parameters that both methods estimate.

larger in magnitude. This RQI will provide a measure of the difference in point estimates between the two types of estimation method. We will measure CR at the model-coefficient level; there will be a row in our final dataset for every coefficient estimated in a given model specification.

Second, we will compute the average marginal effect (AME) as an alternative to CR. This RQI is the change in the probability of policy adoption for a change from the minimum to the maximum value of a covariate, averaged over the observed values in the data (see Hanmer and Kalkan 2013). We plan to compute this measure because coefficient changes in binary dependent variable models can arise because the error variance is fixed (Long 1997). This process can lead to different coefficient estimates from variations in model specification even when neither one is biased or produces different substantive conclusions. We will compare our replication results using both CR and AME as our measure of difference between the two estimators.

Third, we will compute the ratio of the EHA or PEHA standard errors to the analagous NEHA standard errors.[4] We refer to this quantity as the standard error ratio (SER). It is equal to 1 if the EHA (or PEHA) and NEHA standard errrors are equivalent, greater than 1 if the EHA (PEHA) standard error is larger in magnitude, and less than 1 if the NEHA standard error is larger in magnitude. This RQI will provide a measure of the difference in uncertainty between the two types of estimation method. We will measure SER at the model-standard error level; there will be a row in our final dataset for every standard error estimated in a given model specification.

Fourth, we will compute a likelihood ratio test (LR) comparing the EHA or PEHA model to the NEHA model. The EHA or PEHA model is nested in the NEHA model, which makes the LR statistic a valid measure of model fit. This RQI will yield expectations for how often in applied research the NEHA model can be expected to outperform the convetional modeling approach. Additionally, it will allow us to evaluate whether the LR test's selection is associated with the magnitude of the differences in the other RQI. For instance, does the LR test choose NEHA more often when there are larger differences between NEHA and PEHA? This quantity will be measured at the model level; there will be a row in our final dataset for every model we replicate.

[4]We will match the standard error estimation methods in each of the original studies (e.g., robust and/or clustered standard errors).

Finally, we will monitor the statistical significance of individual coefficient estimates in each model. We will compute the ratio of proportion of estimates that are significant in the EHA or PEHA model to the same proportion in the NEHA model.[5] This RQI—statistical significance ratio (SSR)—will provide an assessment of statistical power in the two competing estimators. As with LR, the unit of analysis is the model with this RQI.

4 Statistical and Substantive Significance

We plan to conduct what Harden, Sokhey, and Wilson (2019) refer to as a full inference replication analysis (FIRA). We will use results from our replication analysis to make inferences about the differences between the two estimation methods.[6] This process will involve an assessment of statistical and substantive significance. Harden, Sokhey, and Wilson (2019) recommend declaring our criteria for assessing significance ahead of time. We present those criteria here. Regarding statistical significance, we plan to use the conventional $\alpha = 0.05$ (i.e., 95% confidence level) threshold for all hypothesis tests.

Harden, Sokhey, and Wilson (2019) recommend using Rainey's (2014) approach to assessing substantive significance. Specifically, we must choose a specific value for our RQI, denoted m, that defines the smallest substantively meaningful average RQI. This value represents how different, on average, the existing and new methods must be such that we believe it would be inadvisable for an applied researcher to ignore the NEHA method as a possible empirical strategy. This value will always be arbitrary to some degree, which is why we declare it in this preanalysis plan to ensure that we do not choose the value based on the results.

Specifically, for the CR, AME, SER, and SSR measures we select $m = \pm 0.10$. In other words, we declare a substantively meaningful difference to be one in which the EHA or PEHA coefficient estimate, marginal effect, standard error, or proportion of significant estimates is 10% larger or

[5]Using a proportion is necessary because the NEHA model will have more estimated parameters.

[6]Specifically, we will test a null hypothesis of no substantive difference between the two estimation methods by computing a posterior distribution for the mean of each RQI. As Harden, Sokhey, and Wilson (2019) explain, this computation employs a Bayesian bootstrap method that involves resampling the individual observations (coefficient ratios). Because the coefficients are clustered in statistical models, we will also consider bootstrapping at the cluster level (Cameron, Gelbach, and Miller 2008; Harden 2011, 2012).

10% smaller than the NEHA estimates of those same measures. With respect to the LR test, we will consider NEHA to be a substantively useful estimator if it is the better-fitting estimator for at least 25% of the models we replicate.

5 Updates

The original version of this plan was deposited on March 31, 2019. We plan to update it as needed over the course of the project. If we make any changes to our analysis after depositing this document, we will explain and justify those changes here.

References

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

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

Boehmke, Frederick J. 2009*a*. "Approaches to Modeling the Adoption and Diffusion of Policies with Multiple Components." *State Politics & Policy Quarterly* 9(2): 303–329.

Boehmke, Frederick J. 2009*b*. "Policy Emulation or Policy Convergence? Potential Ambiguities in the Dyadic Event History Approach to State Policy Emulation." *Journal of Politics* 71(3): 1125–1140.

Boehmke, Frederick J., Abigail Matthews Rury, Bruce A. Desmarais, and Jeffrey J. Harden. 2017. "The Seeds of Policy Change: Leveraging Diffusion to Disseminate Policy Innovations." *Journal of Health Politics, Policy and Law* 42(2): 285–307.

Boehmke, Frederick J., Mark Brockway, Bruce A. Desmarais, Jeffrey J. Harden, Scott LaCombe, Fridolin Linder, and Hanna Wallach. 2019. "A New Database for Inferring Public Policy Innovativeness and Diffusion Networks." Forthcoming, *Policy Studies Journal*.

Boushey, Graeme. 2016. "Targeted for Diffusion? How the Use and Acceptance of Stereotypes Shape the Diffusion of Criminal Justice Policy Innovations in the American States." *American Political Science Review* 110(1): 198–214.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors." *Review of Economics and Statistics* 90(3): 414–427.

Desmarais, Bruce A., Jeffrey J. Harden, and Frederick J. Boehmke. 2015. "Persistent Policy Pathways: Inferring Diffusion Networks in the American States." *American Political Science Review* 109(2): 392–406.

Gomez-Rodriguez, Manuel, Jure Leskovec, and Andreas Krause. 2010. Inferring Networks of Diffusion and Influence. In *The 16th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*.

Graham, Erin R., Charles R. Shipan, and Craig Volden. 2013. "The Diffusion of Policy Diffusion Research in Political Science." *British Journal of Political Science* 43(3): 673–701.

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

Hannah, A. Lee, and Daniel J. Mallinson. 2018. "Defiant Innovation: The Adoption of Medical Marijuana Laws in the American States." *Policy Studies Journal* 46(2): 402–423.

Harden, Jeffrey J. 2011. "A Bootstrap Method for Conducting Statistical Inference with Clustered Data." *State Politics & Policy Quarterly* 11(2): 223–246.

Harden, Jeffrey J. 2012. "Improving Statistical Inference with Clustered Data." *Statistics, Politics, and Policy* 3(1): 1–27.

Harden, Jeffrey J., Anand E. Sokhey, and Hannah Wilson. 2019. "Replications in Context: A Framework for Evaluating New Methods in Quantitative Political Science." *Political Analysis* 27(1): 119–125.

Hinkle, Rachael. 2015. "Into the Words: Using Statutory Text to Explore the Impact of Federal

Courts on State Policy Diffusion." *American Journal of Political Science* 59(4): 1002–1021.

Karch, Andrew, Sean C. Nicholson-Crotty, Neal D. Woods, and Ann O'M. Bowman. 2016. "Policy Diffusion and the Pro-innovation Bias." *Political Research Quarterly* 69(1): 83–95.

Kreitzer, Rebecca J., and Frederick J. Boehmke. 2016. "Modeling Heterogeneity in Pooled Event History Analysis." *State Politics & Policy Quarterly* 16(1): 121–141.

Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks, CA: Sage Publications.

Makse, Todd, and Craig Volden. 2011. "The Role of Policy Attributes in the Diffusion of Innovations." *Journal of Politics* 73(1): 108–124.

Rainey, Carlisle. 2014. "Arguing for a Negligible Effect." *American Journal of Political Science* 58(4): 1083–1091.

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.

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