Causal Inference in Policy Diffusion Studies: Sensitivity Analysis for Unmeasured Confounding*

Naoki Egami†

First Version: July 6, 2020
This Version: December 5, 2021

Abstract

Understanding how policies diffuse from one government to others is a central goal in many subfields of political science. In the policy diffusion literature where most studies rely on observational data, the necessity of explicitly addressing causal problems is high for two reasons. First, unmeasured confounding is common due to spatial and network dependence across governments. Second, scholars are interested in a wide range of causal effects, yet their different required assumptions are often unrecognized. In this paper, we first clarify identification assumptions and situate the widely-used event history analysis within a causal framework. Our main contribution is to propose a general sensitivity analysis, with which researchers can quantify the robustness of their causal conclusions to unmeasured confounding. Its implementation only requires a simple adjustment to the standard event history analysis. We illustrate the proposed approach using a diffusion study of abortion policies.

*The companion R package cEHA will be made available for implementing our methods.

†Assistant Professor, Department of Political Science, Columbia University, New York NY 10027. Email: naoki.egami@columbia.edu, URL: https://naokiegami.com

1 Introduction

How policies diffuse across governments has been a fundamental question in political science for decades. In American politics, researchers have examined the state-level diffusion process of policies, such as those on education, lotteries, smoking, and criminal justice (Walker, 1969; Gray, 1973; Berry and Berry, 1990; Shipan and Volden, 2006). Scholars of international relations have explored the diffusion dynamics of liberal economic policies, labor rights, and norms (Simmons and Elkins, 2004; Hyde, 2011). Diffusion studies of democracies and welfare policies have a long history in comparative politics (Collier and Messick, 1975; Starr, 1991). A recent review by Graham et al. (2013) found more than 800 articles about policy diffusion from the past 50 years across all subfields of political science.

There have been numerous methodological developments to analyze policy adoption patterns. The most influential approach has been event history analysis (EHA), also known as survival analysis in statistics, and duration analysis in economics (e.g., Beck, 1999; Box-Steffensmeier and Jones, 2004). Berry and Berry (1990) first adapted it to the policy diffusion literature, and since then, it has been a default method for three decades. Recent important extensions include pooled event history models (Shipan and Volden, 2006; Boehmke, 2009) that analyze multiple policies together, a method for dyadic data (e.g., Gilardi, 2010), and models for latent diffusion patterns (Desmarais et al., 2015).

Despite these numerous successful methodological developments in the policy diffusion literature, issues of causality have received less attention to date. However, the necessity of explicitly addressing causal problems is particularly high in policy diffusion studies for two reasons. First, well-known concerns about unmeasured confounding are even more severe than in usual observational studies because policy adoptions are theorized and empirically found to be interdependent between governments across space and networks. This interdependence is exactly the quantity of interest to policy diffusion scholars, and yet, at the same time, it makes unmeasured confounding significantly more complex due to spatial and network dependence (Galton, 1889; Fowler et al., 2011). Second, policy diffusion scholars examine a wide range of causal effects, including effects of time-varying and time-invariant variables, but the common implementation of EHA often fails to recognize different causal assumptions required for each causal quantity. This has led to a mismatch between causal quantities that analysts aim to estimate and causal assumptions that they implicitly make when applying EHA, which results in biased estimates.

To improve causal inference in policy diffusion studies, we first propose a causal inference framework by defining causal effects relevant to policy diffusion studies and clarifying their identification assumptions. We situate the widely-used EHA within this causal framework (Sec-

tion 2).

Our main contribution is to propose a new sensitivity analysis method, with which researchers can quantify the robustness of their causal estimates to unmeasured confounding (Section 3). The central benefit of the sensitivity analysis is that researchers can transparently examine how causal estimates would change depending on various scenarios of unmeasured confounding rather than focusing only on a binary question of whether there exists any unmeasured confounding or not. This is particularly important in policy diffusion studies as it is difficult to eliminate all concerns of unmeasured confounding. To avoid the subjective choice of sensitivity parameters, we also propose the evidence-value (E-value) as a single-value summary of the robustness to unmeasured confounding.

A key advantage of the proposed sensitivity analysis method is that we can apply it to various causal estimands relevant in policy diffusion studies, without changing how to interpret results from the sensitivity analysis. After we introduce our proposed method by focusing on the most common estimand — the causal effect of time-varying treatments on the immediate future outcome, we extend the proposed sensitivity analysis to long-term causal effects (Section 3.4) and causal effects of spatial/network diffusion variables (e.g., the number or proportion of prior adopters in contiguous spatial neighbors) (Section 3.5). In Section 4, we extend our proposed sensitivity analysis to the time-invariant treatment variable.

To facilitate the use of the proposed methods, we offer a companion R package cEHA that implements all the methods we discuss in this paper, including event history analysis and sensitivity analysis methods. We summarize our practical recommendations for making causal inference in policy diffusion studies in Section 5. We then illustrate our proposed approach by applying it to a diffusion study of abortion policies (Kreitzer, 2015) in Section 7. We also provide an additional empirical application based on Karch et al. (2016) and simulation studies in the Appendix to further illustrate the use of our proposed approach.

This paper builds on the large literature of event history analysis (e.g., Berry and Berry, 1990; Beck, 1999; Box-Steffensmeier and Jones, 2004; Shipan and Volden, 2006; Gilardi, 2010; Desmarais et al., 2015). Existing sensitivity analysis methods for event history analysis have applied ideas of the confounding ratio by Brumback et al. (2004) to a Cox proportional hazards model (e.g., Klungsøyr et al., 2009; Kasza et al., 2015). These approaches require researchers to specify a parametric sensitivity model of how unmeasured confounding changes over time and across policies. This is challenging in policy diffusion studies where unmeasured confounding is often heterogeneous over time and across policies. In contrast, our sensitivity analysis relies on a nonparametric approach, as described in Section 3, and thus we can avoid parametric

assumptions without sacrificing intuitive interpretation of sensitivity parameters. We follow the naming convention of the E-value by VanderWeele and Ding (2017), but the proposed E-value is distinct from their proposal because the underlying sensitivity analysis is completely different. Finally, while several sensitivity analysis methods are particularly popular in political science (e.g., Imai et al., 2010; Blackwell, 2014; Cinelli and Hazlett, 2020), they are developed for application areas different from policy diffusion studies. In contrast, our sensitivity analysis is specifically designed to accommodate methodological challenges in policy diffusion studies. We provide further discussions on relationships with existing approaches in Section 6.

2 Event History Analysis in A Causal Framework

We begin with the basic setup for making causal inference in policy diffusion studies and then discuss event history analysis (EHA), which is the most popular approach for analyzing policy diffusion. This section provides a key background for developing sensitivity analysis in the next section.

2.1 Setup

At its most basic, policy diffusion analysis starts with panel data on policy adoptions that measure the timing when governments adopt policies. When focusing on a single policy, each observation is a pair {government, time}, and the outcome variable is whether a given government adopts the policy at or before the specified time period. It is now common to examine multiple policies together, known as pooled event history analysis (Boehmke, 2009; Makse and Volden, 2011). In this case, each observation is a combination {government, time, policy}, and the outcome variable is whether a given government adopts a specific policy at or before each time period. For instance, Kreitzer (2015) examines the adoption of 29 anti-abortion policies by the American states from 1973 to 2013.

In general, we consider K policies and n units over T time periods. We then define Y_{itk} to be the outcome for unit i at time t for policy k, which represents the policy adoption status — whether unit i adopted policy k at or before time t (Beck, 1999; Box-Steffensmeier and Jones, 2004). This event history setting has an important feature; when unit i adopted policy k at time t, we know the outcome variable takes the value of one at or after time $t+1$, that is, when $Y_{itk} = 1$, $Y_{it'k} = 1$ for $t' \geq t+1$. For example, when we find State 1 adopted Policy A in year 2002, then we know the outcome variable of a pair {State 1, Policy A} remains one after year 2002 (see Figure 1). Note that in practice, researchers sometimes recode the outcome variables for years after the policy adoption as missing (`NA`) and remove them from analyses. This specific strategy is equivalent to our general setting under conventional event history analysis. However, a setup

	Year						
	1997	1998	1999	2000	2001	2002	2003
State 1, Policy A	0	0	0	0	0	1	1
State 2, Policy A	0	0	0	0	1	1	1
State 3, Policy A	0	0	0	0	0	0	0
State 1, Policy B	0	0	0	1	1	1	1
State 2, Policy B	0	1	1	1	1	1	1
State 3, Policy B	0	0	1	1	1	1	1

Figure 1: Example of Event History Data. *Note:* Event history data encodes when each unit adopts a specific policy. The first row shows that State 1 adopted Policy A in 2002, and the third row shows that State 3 did not adopt Policy A by the end of 2003. Following the widely-used pooled event history analysis, we consider non-competing events; that is, the same unit can adopt multiple policies (e.g., State 2 adopted Policy A in 2001 and Policy B in 1998).

we use here, which is more common in the statistics literature, is essential when we extend conventional EHA in Sections 3 and 4.

The baseline time period $t = 0$ is defined for each policy k such that no unit adopts the policy before the baseline time period. This is already common in most applications because researchers often start the period of observations with the year in which the first state adopts a policy innovation.[1] Finally, we follow traditions in the widely-used pooled event history analysis (Boehmke, 2009; Makse and Volden, 2011) in which multiple policies are not "competing events" in the survival analysis terminology, meaning that adopting one policy does not prevent units from adopting other policies; units can potentially adopt all K policies (see also Figure 1).

We define A_{itk} to be a treatment variable of unit i measured at time t for policy k. For concreteness, we focus on a *time-varying* treatment here, but we extend our results to a time-invariant treatment variable in Section 4, which requires some methodological modifications. For example, $A_{itk} = 1$ could represent whether there exists at least one interest group advocating for policy k in year t at state i. This section also naturally accommodates cases when the time-varying treatment variable does not vary across policies A_{it}, such as the partisanship of governors — whether a state governor is a Democrat or not in year t.

[1]When there are units that have adopted policies before the time periods measured in the study, we cannot estimate causal effects or even associations for such units without making stringent assumptions. To make reliable causal inference, we define causal estimands only for units that have not adopted policies before the study period.

We use the potential outcomes framework (Imbens and Rubin, 2015) to define causal effects. In particular, we define $Y_{itk}(A_{itk} = a)$ to be the potential outcome variable of unit i at time t for policy k that would be realized if the unit receives the treatment $A_{itk} = a$. For example, the potential outcome $Y_{itk}(A_{itk} = 1)$ could represent whether state i would adopt policy k at or before year t if there had been at least one interest group for policy k at state i in year t. The potential outcome under control $Y_{itk}(A_{itk} = 0)$ is also similarly defined. Importantly, because of the event history setting, we know that $Y_{itk}(A_{itk} = a) = 1$ for those who adopted policies before the administration of treatment at time t. This implies no causal effect of the treatment A_{itk} for already-adopters with $Y_{i,t-1,k} = 1$.

2.2 Causal Quantities of Interest

We now define a causal effect of the time-varying treatment A_{itk} as the difference in the average potential policy adoption rates among potential adopters. As we see below, this is closely related to the standard EHA model. Formally, we define the average treatment effect (ATE) of the time-varying treatment τ_{kt} as follows. For notational simplicity, we use $A_{itk} = 1$ and $A_{itk} = 0$ to denote a treatment group and a control group, but the same framework applies to categorical and continuous treatments with appropriate notational changes.

$$\tau_{tk} \;\coloneqq\; \Pr\{Y_{itk}(A_{itk} = 1) = 1 \mid Y_{i,t-1,k} = 0\} \;-\; \Pr\{Y_{itk}(A_{itk} = 0) = 1 \mid Y_{i,t-1,k} = 0\}, \qquad (1)$$

which quantifies the total causal effect of the time-varying treatment at time t on whether policy k was adopted at time t among potential adopters. For example, this could represent the average causal effect of having interest groups in year t on the policy adoption rate in year t. We will extend our results to long-term causal effects in Section 3.4.

Finally, while the ATE, τ_{tk}, is defined with specific timing of the treatment administration t and policy k, we can also define the time-average ATE and the policy- and time-average ATE (we provide their formal definitions in Appendix A.1). Because identification and estimation methods for the ATE are directly applicable to these other relevant causal estimands, our formal discussion will focus on the ATE without loss of generality.

2.3 Causal Assumption

What assumptions are required to identify the ATE? We need to assume that observed covariates include all relevant confounders at each time period for potential adopters (Robins, 1986; Blackwell and Glynn, 2018). In other words, a treatment variable should be as-if random at each time period given observed time-varying covariates and the history of outcomes.

Assumption 1 (Sequential Ignorability Among Potential-Adopters)

$$\{Y_{itk}(A_{itk}=1), Y_{itk}(A_{itk}=0)\} \perp\!\!\!\perp A_{itk} \mid Y_{i,t-1,k}=0, \mathbf{X}_{itk}. \tag{2}$$

This requires that the potential outcomes at time t are independent of the time-varying treatment at time t given the observed covariates $\mathbf{X}_{itk}$ for potential adopters with $Y_{i,t-1,k}=0$. For example, this assumption requires that analysts adjust for all confounders affecting policy adoption in year t and the treatment.

Under this sequential ignorability and the standard overlap assumption, we can identify the ATE with the usual identification formula, which we detail in Appendix A.2.

2.4 Standard EHA

Event history analysis (EHA) is by far the most widely-used estimation strategy in the policy diffusion literature. In particular, EHA is used to model the policy adoption outcome Y_{itk} with the treatment variable A_{itk} and some covariates $\mathbf{X}_{itk}$. In particular, standard EHA estimates the following conditional policy adoption rate (Berry and Berry, 1990; Beck, 1999; Box-Steffensmeier and Jones, 2004).

$$\Pr(Y_{itk}=1 \mid A_{itk}, \mathbf{X}_{itk}, Y_{i,t-1,k}=0) = g\{h(t) + A_{itk}\beta + \mathbf{X}_{itk}^{\top}\gamma\}, \tag{3}$$

where $g(\cdot)$ is a binary outcome model, such as logistic and probit models, $h(t)$ represents some flexible function of time-trends, such as third-degree polynomials or splines, and pre-treatment covariates $\mathbf{X}_{itk}$ can include both time-varying and time-invariant predictors as well as fixed and random effects. The central feature is that EHA models the policy adoption at time t only among potential new adopters at time t, i.e., those who have not yet adopted policies at time $t-1$ with $Y_{i,t-1,k}=0$. This is because units who have adopted policies at time $t-1$ have outcomes equal to one at time t by definition, and thus, they provide no additional information. In the survival analysis terminology, this conditional probability is called the discrete-time hazard rate.

Under the sequential ignorability and overlap assumptions, we can use this standard EHA to estimate the ATE. Formally,

$$\begin{aligned}\widehat{\tau}_{tk} \;=\; \widehat{\mathbb{E}}_{\mathbf{X}|\mathsf{PA}_t}\Big\{&\widehat{\Pr}(Y_{itk}=1 \mid A_{itk}=1, \mathbf{X}_{itk}=\mathbf{x}, Y_{i,t-1,k}=0) \\ &- \widehat{\Pr}(Y_{itk}=1 \mid A_{itk}=0, \mathbf{X}_{itk}=\mathbf{x}, Y_{i,t-1,k}=0)\Big\}\end{aligned} \tag{4}$$

where $\widehat{\mathbb{E}}_{\mathbf{X}|\mathsf{PA}_t}(\cdot)$ denotes the empirical average over the distribution of $\mathbf{X}_{itk}$ among potential adopters $Y_{i,t-1,k}=0$. In practice, we first compute the change in predicted probabilities from the estimated EHA model for all potential adopters at time t for policy k, and then we can simply average them to obtain an estimated ATE at time t for policy k.

We compute standard errors clustered at the level of treatment assignment. For example, in our empirical application based on Kreitzer (2015), we cluster standard errors at the state level. Specifically, researchers can estimate the standard EHA using the usual cluster robust standard errors or using nonparametric block bootstrap. We provide technical details about the estimation algorithm in Appendix C.

It is important to clarify that, while we focus on the standard EHA model in this section as it is the most popular approach in practice, the proposed framework and sensitivity analysis methods in this paper can be combined with any methods of estimation (e.g., multi-level EHA (Kreitzer and Boehmke, 2016)) and standard error calculation (e.g., cluster standard errors and block bootstrap). This is because our results are about nonparametric causal identification, and thus, our approach can be combined with existing models as long as they estimate the conditional policy adoption rate $\Pr(Y_{itk} = 1 \mid A_{itk}, \mathbf{X}_{itk}, Y_{i,t-1,k} = 0)$.

3 Sensitivity Analysis for Unmeasured Confounding

We emphasize that, like any other causal inference methods, the EHA estimator critically depends on causal assumptions, in particular, the sequential ignorability. In this section, we consider potential violations of the sequential ignorability assumption and develop a new sensitivity analysis method, with which researchers can assess the potential influence of unmeasured confounding on causal conclusions.

3.1 Setup

To conduct sensitivity analyses, we have to quantify the extent to which the required causal assumption — the sequential ignorability (Assumption 1) — is violated. Following a large literature in causal inference (e.g., Brumback et al., 2004; Blackwell, 2014), we consider a confounding ratio (C-ratio) to measure violations of the sequential ignorability. Intuitively, it quantifies how much more likely units in the treatment group are to adopt policies than the control group due to unmeasured confounding. Formally, the confounding ratio is the ratio of the potential policy adoption rate at time t for those under treatment $A_{itk} = 1$ and under control $A_{itk} = 0$.

$$\text{C-ratio} = \frac{\Pr\{Y_{itk}(A_{itk} = 0) = 1 \mid A_{itk} = 1, \mathbf{X}_{itk}, Y_{i,t-1,k} = 0\}}{\Pr\{Y_{itk}(A_{itk} = 0) = 1 \mid A_{itk} = 0, \mathbf{X}_{itk}, Y_{i,t-1,k} = 0\}}. \tag{5}$$

When the sequential ignorability holds, treatments are as-if random conditional on observed covariates, and thus, both treatment and control groups have the same potential policy adoption rates. In this case, the confounding ratio is equal to one. In contrast, when there exists unmeasured confounding, the treatment and control groups have different policy adoption rates, even in the absence of treatment administration. In this sense, the deviation of the confounding

ratio from one is a formal measure of how much the required sequential ignorability assumption is violated.

Existing methods based on the confounding ratio approach (e.g., Brumback et al., 2004; Blackwell, 2014) posit some parametric models for this confounding ratio, e.g., as simple as, $\mathsf{C\text{-}ratio} = \alpha$, and compute the ATE under different scenarios of unmeasured confounding by changing parameter α. While this is simple to use, this parametric approach is not suitable for policy diffusion studies in which the policy adoption rates often differ significantly over time and policies, which in turn strongly suggests unmeasured confounding also varies across time and policies.

We instead propose a nonparametric approach by incorporating the confounding ratio into the classical inequality-based sensitivity analysis framework, which has been studied separately (Cornfield et al., 1959; VanderWeele and Ding, 2017). The main idea of this approach is to consider the worst-case unmeasured confounding and derive bounds for the ATE. The central benefit is that we can avoid specific parametric assumptions on unmeasured confounding, which is crucial in policy diffusion studies where unmeasured confounding is complex and heterogeneous over time and across policies.

In particular, we define sensitivity parameter Γ to represent the worst-case unmeasured confounding; the confounding ratio is at most $\Gamma+1$. The proposed sensitivity analysis is to consider different values of Γ and to obtain bias-corrected estimates of the ATE for each scenario. By investigating different values of the sensitivity parameter, researchers can quantify the robustness of causal findings to violations of causal assumptions. Formally, we define sensitivity parameter $\Gamma(\geq 0)$ as the worst-case confounding ratio. For $a \in \{0,1\}$, $\mathbf{x} \in \mathcal{X}$, and $t \in \{1,\ldots,T_k\}$ and $k \in \{1,\ldots,K\}$,

$$\frac{1}{\Gamma+1} \;\leq\; \frac{\Pr\{Y_{itk}(A_{itk}=a)=1 \mid A_{itk}=1, \mathbf{X}_{itk}=\mathbf{x}, Y_{i,t-1,k}=0\}}{\Pr\{Y_{itk}(A_{itk}=a)=1 \mid A_{itk}=0, \mathbf{X}_{itk}=\mathbf{x}, Y_{i,t-1,k}=0\}} \;\leq\; \Gamma+1, \tag{6}$$

where the middle term is a generalization of the $\mathsf{C\text{-}ratio}$ in equation (5). The lower bound $1/(\Gamma+1)$ covers scenarios in which the treatment group is less likely to adopt policies than the control group due to unmeasured confounding. Therefore, it encompasses both cases when the treatment group has a higher potential policy adoption rate and cases when the control group has a higher rate. When we choose $\Gamma = 0.2$, we take into account unmeasured confounding that makes one group more likely to adopt policies than the other group up to 20%. In general, sensitivity parameter Γ represents unmeasured confounding that makes one group more likely to adopt policies than the other group up to $\Gamma \times 100\%$. By choosing a larger value of Γ, analysts can account for stronger unmeasured confounding. In a special case of $\Gamma = 0$, the confounding ratio

is equal to one, which corresponds to no unmeasured confounding. Therefore, the sensitivity analysis contains sequential ignorability as a special case.

3.2 Sensitivity Analysis

How can we use sensitivity parameter Γ to estimate the ATE while accounting for unmeasured confounding? We can directly combine sensitivity parameter Γ and the standard EHA estimator. For each Γ, we can compute a bound for the ATE, and this bound is sharp in the sense that there always exists unmeasured confounding that requires the maximum correction and attain the lower and upper bounds. Formally, we can write the lower bound as follows.

$$\begin{aligned}
\widehat{\mathbb{E}}_{\mathbf{X}|\mathsf{PA}_t}\Bigg[& \underbrace{\widehat{\Pr}(Y_{itk}=1 \mid A_{itk}=1, \mathbf{X}_{itk}, Y_{i,t-1,k}=0)}_{\text{Predicted Probability for the treatment group}} \\
& \times \underbrace{\left\{ \widehat{\Pr}(A_{itk}=1 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) + \frac{\widehat{\Pr}(A_{itk}=0 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0)}{\Gamma+1} \right\}}_{\text{Bias-correction term for the treatment group}} \\
& - \underbrace{\widehat{\Pr}(Y_{itk}=1 \mid A_{itk}=0, \mathbf{X}_{itk}, Y_{i,t-1,k}=0)}_{\text{Predicted Probability for the control group}} \\
& \times \underbrace{\left\{ (\Gamma+1) \times \widehat{\Pr}(A_{itk}=1 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) + \widehat{\Pr}(A_{itk}=0 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \right\}}_{\text{Bias-correction term for the control group}} \Bigg], \quad (7)
\end{aligned}$$

We provide the proof and an estimator corresponding to the upper bound in Appendix B.1.

The key insight is that the standard EHA estimator is adjusted by the bias-correction terms that include sensitivity parameter Γ. For the treatment group, the bias-correction term is less than one because $\Gamma \geq 0$, and for the control group, the corresponding term is larger than one, which together makes the estimate smaller than the original EHA estimator. We can easily see that bias-correction is larger when Γ is larger. When $\Gamma = 0$, this estimator naturally reduces to the conventional EHA estimator under sequential ignorability (equation (4)). To implement this estimator, researchers just need to weight estimates from the standard EHA estimator with bias-correction terms, that is, $\widehat{\Pr}(A_{itk}=1 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) + \widehat{\Pr}(A_{itk}=0 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0)/(\Gamma+1)$ for the treatment group, and $(\Gamma+1) \times \widehat{\Pr}(A_{itk}=1 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) + \widehat{\Pr}(A_{itk}=0 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0)$ for the control group. We also extend the sensitivity analysis to a continuous treatment variable in Appendix B.2.

Sensitivity analysis is to estimate causal quantities of interest with a range of plausible values of sensitivity parameters Γ and check the robustness of causal findings to various scenarios of unmeasured confounding. A natural question would be about how to select sensitivity parameter Γ in practice. Several points are worth clarifying. First, as emphasized earlier, we defined Γ

to have a straightforward interpretation; unmeasured confounding that makes one group more likely to adopt policies than the other group up to $\Gamma \times 100\%$ (e.g., $\Gamma = 0.2$ means 20%). Thus, it is relatively easy to incorporate substantive knowledge when choosing sensitivity parameters, compared to conventional sensitivity analysis methods. While the exact choice should depend on substantive contexts, we use $\Gamma = \{0.05, 0.10, 0.20, 0.50, 1.00\}$ as default values, which would cover small to large unmeasured confounding scenarios in many applications (see also simulation studies in Appendix E that evaluate this choice). Second, to avoid the subjective choice of sensitivity parameters, we next propose the *evidence-value* as a single-value summary of the robustness to unmeasured confounding.

3.3 E-value

Some researchers might worry that the choice of sensitivity parameters can be subjective and difficult to justify in some applications. To address this concern, we develop an additional method that is useful when analysts want to avoid specifying sensitivity parameter Γ a priori. In particular, researchers can compute the *evidence-value* (E-value) that would explain away the causal effect of interest.[2] The E-value is defined as the minimum sensitivity parameter that would flip the sign of the original EHA estimate after accounting for unmeasured confounding. Suppose the original EHA estimate is positive, and the E-value is 0.3. This means that causal estimates are still positive even after accounting for unmeasured confounding that makes the confounding ratio up to 30%, but the original finding of the positive causal effect is not robust to unmeasured confounding stronger than that. Researchers can report the E-value as a single-value summary of the robustness to unmeasured confounding; the larger value means stronger evidence for causal conclusions — the sign of the original EHA estimate does not change even after accounting for large unmeasured confounding.

Formally, we can estimate the E-value as follows.

$$\text{E-value} = \begin{cases} \dfrac{\widehat{Q}(1,1) - \widehat{Q}(0,0)}{2\widehat{Q}(0,1)} + \sqrt{\left(\dfrac{\widehat{Q}(1,1) - \widehat{Q}(0,0)}{2\widehat{Q}(0,1)}\right)^2 + \dfrac{\widehat{Q}(1,0)}{\widehat{Q}(0,1)}} - 1 & (\text{when } \widehat{\tau}_{kt} \geq 0) \\ \dfrac{\widehat{Q}(0,0) - \widehat{Q}(1,1)}{2\widehat{Q}(1,0)} + \sqrt{\left(\dfrac{\widehat{Q}(0,0) - \widehat{Q}(1,1)}{2\widehat{Q}(1,0)}\right)^2 + \dfrac{\widehat{Q}(0,1)}{\widehat{Q}(1,0)}} - 1 & (\text{when } \widehat{\tau}_{kt} < 0) \end{cases} \tag{8}$$

[2]We follow the naming convention of the "E-value" by VanderWeele and Ding (2017). However, underlying sensitivity analyses are different, and thus, our proposed E-value is distinct from the existing E-value, as we clarify more in Section 6.

where, for $a_1, a_2 \in \{0, 1\}$,

$$\widehat{Q}(a_1, a_2) = \widehat{\mathbb{E}}_{\mathbf{X}|\mathsf{PA}_t}\left\{\widehat{\Pr}(Y_{itk} = 1 \mid A_{itk} = a_1, \mathbf{X}_{itk}, Y_{i,t-1,k} = 0)\widehat{\Pr}(A_{itk} = a_2 \mid \mathbf{X}_{itk}, Y_{i,t-1,k} = 0)\right\}.$$

We provide the proof in Appendix B.4. Each component of the E-value is already estimated when computing sensitivity analysis estimates, so analysts can compute the E-value without additional computational cost.

Some existing sensitivity analysis methods provide "benchmarks" by using some observed confounders as if they were unobserved confounders. However, Cinelli and Hazlett (2020) recently show that such analyses can be misleading because (a) the effect of observed confounders is also confounded when unmeasured confounding exists, and (b) it requires an additional stringent assumption, such as independence of observed and unobserved confounders.

Instead, in this paper, we designed the E-value such that it has a natural substantive meaning and researchers can use domain knowledge to understand it easily. When the estimated E-value is $\hat{e}$, it means that an estimated causal effect is robust to unmeasured confounding that makes one group more likely to adopt policies than the other group up to $\hat{e} \times 100\%$ (e.g., $\hat{e} = 0.2$ means 20%). Most importantly, this E-value is defined based on the policy adoption rates — the main outcome of interest, so researchers can reason about the E-value as they discuss substantive importance of any causal effects (e.g., whether the effect of 3 percentage points is "large" depends on applications, but researchers can judge its substantive relevance in each context). We also offer two empirical applications (Section 7 and Appendix D) and simulation studies (Appendix E) to show the range of E-values that are common in applications.

Finally, we emphasize that the sensitivity analysis is not a panacea for unmeasured confounding — the fundamental challenge in observational causal inference. It is possible that the true unmeasured confounding might be larger than researchers expect, and in this case, sensitivity analysis can still understate the potential influence of unmeasured confounding and overestimate causal effects. But in such scenarios, naturally, no methodologies can estimate causal effects without additional stringent assumptions. The central benefit of the sensitivity analysis is that researchers can transparently quantify the robustness of causal findings to various scenarios of unmeasured confounding rather than focusing only on a binary question of whether there exists any unmeasured confounding or not.

3.4 Long-term Causal Effects

In standard event history analysis, researchers often focus only on outcome Y_{itk}, which is measured immediately after the treatment assignment at time t. However, it is of great substantive importance to capture a longer-term cumulative causal effect on $Y_{i,t+s,k}$ where $s \geq 0$ because

many treatment variables important in the policy diffusion literature can take time to realize causal effects. The goal here is to estimate causal effects on the outcome measured at s time periods after the administration of treatment at time t (see Appendix A.2 for its formal definition). In Appendix B.1, we show that our proposed sensitivity analysis methods can be extended to this long-term causal effect, and we also illustrate its use in Section 7.

3.5 Spatial/Network diffusion variables as Treatments

In the policy diffusion literature, one of the most important treatment variables of interest is a spatial/network diffusion variable, which is defined as a function of spatial or network neighbors' policy adoption. For example, in American politics, one of the most well-studied diffusion hypotheses is a regional diffusion hypothesis (Berry and Berry, 1990; Graham et al., 2013), in which researchers study effects of the number (or proportion) of prior adopters in each state's spatial contiguous neighbors. When studying network-diffusion, researchers often define a diffusion variable as a weighted average of network peers' policies.

Importantly, it is well known that the risk of unmeasured confounding is considered to be even starker when examining spatial/network diffusion variables as treatments (Galton, 1889; Fowler et al., 2011). In the policy diffusion literature, Volden et al. (2008) use formal models to show that policy adoptions can be correlated across states even when there is no learning and independent governments only learn from their own experiences. We provide further discussions in Appendix A.4.

The difficulty of causal diffusion analysis should be reflected in how to select the plausible values of sensitivity parameter Γ. Researchers have to take into account stronger unmeasured confounding, that is, larger values of Γ than those considered in the unit-level treatment cases, a.k.a, "internal" factors. For example, causal diffusion effects have been found to be overestimated by a large amount, for example, by 300 – 700% (e.g., Eckles and Bakshy, 2017). In such cases, even the value of $\Gamma = 1$ might still understate the potential magnitude of omitted variable bias. Therefore, as default values, we consider $\Gamma = \{0.05, 0.30, 1.00, 1.50, 2.00\}$, which covers from 5% to 200% unmeasured confounding. Researchers who worry about the subjectivity of the choice of sensitivity parameters can rely on the E-value, as clarified in Section 3.3.

4 Extension to Time-Invariant Treatments

Scholars of policy diffusion are often interested in understanding causal effects of the time-invariant treatments as well. In this section, we clarify a different causal assumption required for identification of causal effects of time-invariant treatments. We then extend the proposed sensitivity analysis to this case.

4.1 Causal Assumption and Post-Treatment Bias

One common example of the time-invariant treatment is a political institution. For instance, researchers have examined causal effects of electoral systems, such as majoritarian and proportional representation systems. In some cases, analysts have to rely on time-invariant treatments due to data constraints, even though measures can be theoretically time-varying. For example, Kreitzer (2015)'s main hypothesis is about how state-level public opinions about abortion affect the adoption of abortion policies, and she uses the state-of-the-art measures of abortion-specific state opinions, which is time-invariant.

While time-invariant and time-varying treatments have been similarly analyzed in practice, they require different causal assumptions. Importantly, researchers have to be careful about post-treatment bias when examining the causal effect of time-invariant treatments.

We define the time-invariant treatment to be B_{ik} such that its difference from the time-varying treatment A_{itk} is clear. Then, we define the average treatment effect (ATE) of the time-invariant treatment as follows.[3]

$$\delta_k(s) \;\; := \;\; \Pr\{Y_{isk}(B_{ik}=1)=1\} \;-\; \Pr\{Y_{isk}(B_{ik}=0)=1\}, \tag{9}$$

which quantifies the causal effect of the time-invariant treatment B_{ik} on whether policy k was adopted by time s. One key difference from the time-varying treatment case is that all units are potential adopters of interest because the time-invariant treatment is administered at the baseline time period. This causal quantity is naturally the long-term causal effect because researchers are often interested in the causal effects not only at $t=1$ (right after implementing the time-invariant treatment at the baseline $t=0$) but also at other time points $s \geq 1$.

We require a different causal assumption for identifying the ATE of the time-invariant treatment (equation (9)). We need to assume a time-invariant treatment variable is as-if random given observed covariates at the *baseline* time period. As well-known in the literature (Robins, 1986), this is, in general, stronger than the sequential ignorability required for time-varying treatments.

Assumption 2 (Baseline Ignorability)

$$\{Y_{isk}(B_{ik}=1), Y_{isk}(B_{ik}=0)\} \perp\!\!\!\perp B_{ik} \mid \mathbf{X}_{ik}. \tag{10}$$

It requires that the potential outcomes at time s are independent of the time-invariant treatment B_{ik} given the observed covariates $\mathbf{X}_{ik}$ that are time-invariant or are measured at the baseline.

[3]Here, again for notational simplicity, we use $B_{ik}=1$ and $B_{ik}=0$ to denote a treatment group and a control group, but the same framework applies to categorical and continuous treatments with appropriate notational changes.

Under the baseline ignorability and overlap assumptions, we can identify the ATE of the time-invariant treatment (equation (9)). See Appendix A.3 for details of the identification formula.

Post-Treatment Bias in Standard EHA. The most important aspect of this baseline ignorability is that we should not adjust for previous outcomes Y_{itk} for $t \in \{0, \ldots, s-1\}$ in order to avoid post-treatment bias, which arises when we adjust for variables that are affected by the treatment. This point is critical because standard EHA model (equation (3)) would estimate $\Pr(Y_{isk} = 1 \mid Y_{i,s-1,k} = 0, B_{ik}, \mathbf{X}_{ik})$ — it includes the treatment B_{ik}, controls $\mathbf{X}_{ik}$, and most importantly, the previous outcome $Y_{i,s-1,k} = 0$, which is the post-treatment variable that is affected by the time-invariant treatment B_{ik} (Hernán, 2010). This implies that, in the standard EHA model (equation (3)), coefficients of the time-invariant treatments or change in predicted probabilities based on such coefficients do not have causal interpretation due to this inherent post-treatment bias. As mentioned in Section 2.1, removing units who have already adopted policies from data is equivalent to conditioning on the past outcome $Y_{i,s-1,k} = 0$, and therefore, analysts can suffer from the same post-treatment bias even if they do not explicitly include the previous outcomes in regression models. Unfortunately, this critical issue of the inherent post-treatment bias in standard EHA has received little attention in the policy diffusion studies.

While consequences are similar, we note that the underlying reason for this post-treatment bias is different from the one extensively studied in the panel causal inference literature (e.g., Blackwell and Glynn, 2018). In the panel causal inference literature, researchers have to be careful about time-varying confounders due to causal feedback between outcomes, treatments and confounders over time. In the event history analysis, even without any causal feedback, standard EHA with the time-invariant treatment suffers from post-treatment bias because the hazard rate $\Pr(Y_{isk} = 1 \mid Y_{i,s-1,k} = 0, B_{ik}, \mathbf{X}_{ik})$ incorporates the post-treatment variable $Y_{i,s-1,k}$ by definition. Therefore, a solution to the post-treatment bias is also different. Indeed, even though this inherent post-treatment bias in standard EHA (Hernán, 2010) can cause significant biases, it is relatively straightforward to resolve when we use an alternative estimator instead of standard EHA estimator.

4.2 Estimation

Under the baseline ignorability and the overlap, we can estimate the ATE of the time-invariant treatment (equation (9)) using the difference in predicted probabilities.

$$\widehat{\delta}_k(s) \;=\; \widehat{\mathbb{E}}_{\mathbf{X}} \left\{ \widehat{\Pr}(Y_{isk} = 1 \mid B_{ik} = 1, \mathbf{X}_{ik} = \mathbf{x}) \;-\; \widehat{\Pr}(Y_{isk} = 1 \mid B_{ik} = 0, \mathbf{X}_{ik} = \mathbf{x}) \right\} \quad (11)$$

where $\widehat{\mathbb{E}}_{\mathbf{X}}$ denotes the empirical average over the distribution of $\mathbf{X}_{ik}$. To estimate these predicted probabilities, researchers can use a binary outcome model $m(\cdot)$, e.g., a logistic or probit model.

$$\Pr(Y_{isk} = 1 \mid B_{ik}, \mathbf{X}_{ik}) = m(\alpha_s + B_{ik}\beta_B + \mathbf{X}_{ik}^{\top}\gamma_B), \tag{12}$$

where the outcome variable is the policy adoption status at time s, the treatment is the time-invariant variable B_{ik} measured at the baseline, and other control variables $\mathbf{X}_{ik}$ should also be measured at the baseline.

Most importantly, we do not condition on $Y_{i,s-1,k} = 1$. In practice, researchers often recode the outcome variables for years after the policy adoption as missing (`NA`) and remove them from analyses. However, this specific strategy is only valid for standard EHA, which is biased for time-invariant treatment. Therefore, when estimating the causal effects of time-invariant treatments, researchers should not remove units that have already adopted policies. Instead, we should keep them by setting their outcomes to be one. Then, fitting a binary outcome model that regresses the outcome variable at time s on the time-invariant treatment B_{ik} and time-invariant control variables $\mathbf{X}_{ik}$ will be consistent to the ATE under the baseline ignorability assumption and correct model specification.

4.3 Sensitivity Analysis

Importantly, estimation of the ATE of the time-invariant treatment is also subject to the risk of unmeasured confounding. We can extend our proposed sensitivity analysis to the time-invariant treatment. First, we generalize the definition of the sensitivity parameter Γ as follows.

$$\frac{1}{\Gamma + 1} \;\leq\; \frac{\Pr\{Y_{isk}(B_{ik} = b) = 1 \mid B_{ik} = 1, \mathbf{X}_{ik} = \mathbf{x}\}}{\Pr\{Y_{isk}(B_{ik} = b) = 1 \mid B_{ik} = 0, \mathbf{X}_{ik} = \mathbf{x}\}} \;\leq\; \Gamma + 1. \tag{13}$$

The substantive meaning is similar to the one given in equation (6). For each choice of Γ, researchers can obtain both lower and upper bounds for the ATE of the time-invariant treatment (see Appendix B.3 for the exact expressions).

We can also compute the E-value for the ATE of the time-invariant treatment using the same formula (equation (8)) by redefining Q function as follows.

$$\widehat{Q}(b_1, b_2) = \widehat{\mathbb{E}}_{\mathbf{X}}\left\{\widehat{\Pr}(Y_{isk} = 1 \mid B_{ik} = b_1, \mathbf{X}_{ik})\widehat{\Pr}(B_{ik} = b_2 \mid \mathbf{X}_{ik})\right\}.$$

Therefore, while technical details differ, researchers can conduct sensitivity analysis and compute the E-value for time-invariant treatments as they can for the time-varying treatment.

5 Practical Considerations

In this section, we summarize our practical recommendations. In particular, there are three important steps for making causal inference for policy diffusion studies.

	Time-varying Treatments		Time-Invariant Treatments
	Instantaneous effects	Long-term effects	
Assumption	Sequential Ignorability	Sequential Ignorability	Baseline Ignorability
Estimator	Standard EHA	EHA with long-term outcomes	Binary outcome model with long-term outcomes
Sensitivity Analysis	Bounds and E-value	Bounds and E-value	Bounds and E-value

Table 1: Summary of Proposed Methods.

Step 1: Specify Causal Quantities of Interest. Specification of causal quantities of interest is important in any causal analysis. However, it is even more critical in causal analysis of policy diffusion since there are many relevant causal quantities of interest researchers can estimate. Most importantly, researchers should clarify (1) the type of their treatment variable — a time-varying treatment or a time-invariant treatment, and (2) the type of causal effects — an instantaneous causal effect or a long-term causal effect — for the time-varying treatment.

Step 2: Specify Appropriate Models for Each Causal Quantity of Interest. Once researchers specify causal quantities of interest, they should specify an appropriate model for each causal quantity of interest. As we summarized in Table 1, even without any unmeasured confounding, researchers have to use different estimators for different causal quantities of interest. Fortunately, however, these estimators are similar and simple to implement, and our companion R package implements all of them.

Most importantly in practice, this means that researchers cannot simply fit one EHA model and interpret coefficients of different time-varying and time-invariant covariates causally. Especially when researchers are interested in long-term causal effects of time-varying treatments or causal effects of time-invariant treatments, they have to use alternative estimators — EHA with long-term outcomes (Section 3.4) and binary outcome models with long-term outcomes (equation (12)), respectively — because standard EHA will be biased.

Step 3: Evaluate Underlying Causal Assumptions Using Sensitivity Analysis. After estimating causal effects, researchers should carefully evaluate potential violations of underlying causal assumptions. The main contribution of this paper is to provide interpretable sensitivity analysis methods applicable to all the estimands.

Researchers can examine how bounds on causal effects change depending on the strength of potential unmeasured confounding, which we illustrate in Section 7. If researchers are worried

about the subjectivity of the choice of sensitivity parameters, they can report the E-value as a single summary statistic, which quantifies the minimum amount of unmeasured confounding that will change the sign of the original causal estimate.

6 Relationships with Existing Approaches

6.1 Sensitivity Analysis

The literature of sensitivity analysis is large, dating back to Cornfield et al. (1959), and has been essential in the literature of event history analysis and survival analysis (e.g., Andrea et al., 2001). In particular, Klungsøyr et al. (2009) and Kasza et al. (2015) apply ideas of the confounding ratio by Brumback et al. (2004) to a Cox proportional hazards model. However, like Brumback et al. (2004), their approach requires a specific parametric model for the confounding ratio, which means that researchers have to specify how unmeasured confounding changes over time and across policies in order to use their sensitivity analysis. This is particularly challenging in policy diffusion studies where unmeasured confounding is often heterogeneous over time and across policies. In contrast, our sensitivity analysis relies on a nonparametric inequality-based approach, as described in Section 3.1, and thus we can avoid parametric assumptions without sacrificing intuitive interpretation of sensitivity parameters.

Recently, VanderWeele and Ding (2017) developed a general inequality-based sensitivity analysis, which can be applied to event history analysis in principle. Unlike ours, their sensitivity analysis asks researchers to specify association that unmeasured confounders have with both the outcome and the treatment. This approach is powerful when researchers can explicitly name unmeasured confounders and hypothesize their relationship with the outcome and treatment. This could be challenging in policy diffusion studies where many potential unmeasured confounders can affect policy adoption. In contrast, we rely on the confounding ratio, and thus, researchers can focus on just one parameter — how much more likely units in the treatment group are to adopt policies than the control group due to unmeasured confounding. This is why the proposed E-value is distinct from VanderWeele and Ding (2017), even though we follow their naming convention.

Finally, we clarify relationships with sensitivity analysis methods widely used in political science. While details vary, all of them are developed for application areas different from policy diffusion studies. First, Imai et al. (2010) propose a widely used sensitivity analysis method for a different causal estimand, the average causal mediation effect. Second, Cinelli and Hazlett (2020) recently extended the omitted variable bias framework, but it is designed for OLS and is not directly applicable to the event history data. Finally, a popular approach by Blackwell (2014)

and Brumback et al. (2004) is powerful when researchers can specify a parametric model for a confounding function, but, as clarified above, its correct specification is especially challenging in policy diffusion studies where unmeasured confounding is heterogeneous over time and across policies.

6.2 Approaches using Alternative Causal Assumptions

While EHA estimators rely on ignorability assumptions, there exist other approaches based on alternative causal assumptions. First, the fixed effects regression estimator is one of the most popular methods in panel data analysis. Recently, Imai and Kim (2019) show that the fixed effects regression estimator assumes the absence of causal feedback between the treatment and outcome variable over time. However, in policy diffusion studies, this assumption is often untenable; for example, policy makers adopt policies responding to public opinions, and policies also affect public opinions. This is partly why EHA estimators based on ignorability assumptions have been by far the most popular approach in policy diffusion studies.

Second, synthetic control methods (Abadie et al., 2010) and their recent extensions (e.g., Xu, 2017) estimate a weighted average of control units to approximate the outcome trend of a treated unit. Interestingly, in policy diffusion studies, due to the event history data structure, we always compare only potential adopters who have exactly the same outcome history, i.e., the outcome variable is zero up to the treatment assignment timing. Therefore, EHA naturally inherits the main advantage of synthetic control type methods without estimating weights.

7 Application: Diffusion of Anti-Abortion Policies

Kreitzer (2015) studied the diffusion of anti-abortion policies across the American states by analyzing event history data of anti-abortion policies from 1973 to 2013.[4] She used standard EHA to test a number of influential hypotheses in the policy diffusion literature. Using this study, we show how the proposed approach can improve causal inference for a variety of substantive questions common in policy diffusion studies. See Appendix D for an additional empirical application and Appendix E for simulation studies.

Kreitzer (2015) considered three main hypotheses. While they all view the adoption of anti-abortion policies as the outcome variable, they differ in the main treatment variables of interest. The first is about the effects of partisanship of governors. Provided that abortion is a highly partisan issue nowadays, Kreitzer (2015) examined how the party control of the state

[4]Kreitzer (2015) also examines pro-abortion policies and other important hypotheses. See the original paper for details. We use the publicly available replication data from Kreitzer and Boehmke (2016), which examines the same set of anti-abortion policies.

government plays a role in shaping state abortion policies. For this partisanship hypothesis, the treatment variable was whether a state governor is a Democrat or not, which is a time-varying treatment variable. The second is one of the oldest and the most well-studied questions — the regional diffusion hypothesis (Berry and Berry, 1990; Graham et al., 2013), which argues a U.S. state is more likely to adopt a policy if its neighboring states have already done so. Even though empirical results are sometimes mixed, it is so fundamental in the literature that almost all policy diffusion models include this diffusion variable, including Kreitzer (2015). For this regional diffusion hypothesis, the time-varying treatment variable is the number of prior adopters in states' contiguous neighbors.

Finally, the original author was interested in how state-level public opinions on abortion influence the policy adoption rate in each state. Kreitzer (2015) relies on measures of abortion-specific state opinions from Norrander (2001), which ranges from 1 to 5, with higher values representing more conservative opinions. Because Norrander (2001) averages over related survey questions from 1988, 1990, and 1992 Senate National Election Studies to create this measure, this state public opinion variable is a time-invariant treatment.

The outcome variable is the policy adoption status Y_{itk} representing whether state i adopts policy k at or before year t. We include fifty states $i \in \{1, \ldots, 50\}$, and the total number of anti-abortion policies is $K = 26$ after removing three policies that only have three years of data. To understand long-term effects on policy adoption, we study not only causal effects on the immediate future outcome but also on longer-term outcomes up to three years relative to the timing of the treatment $s \in \{0, 1, 2, 3\}$. For each hypothesis, we closely follow the original analysis to select control variables. For the partisanship hypothesis, we use the same pre-treatment covariates used in the original paper and add policy-fixed effects and third-degree polynomials of time trends. For the regional diffusion hypothesis, we use the same pre-treatment covariates used in the original paper and add policy-fixed effects and third-degree polynomials of time trends. We also explicitly adjust for the lagged treatment variable by exploiting the staggered adoption design of this diffusion treatment variable (i.e., the number of prior adopters never decreases over time). Therefore, we can compare treatment and control groups that have exactly the same outcome and treatment histories. For the public opinion hypothesis, standard EHA used in the original analysis includes post-treatment variables as controls. We instead use a binary outcome model (equation (11)) that only incorporates the baseline measures of control variables to avoid post-treatment bias. All standard errors are clustered at the state level.

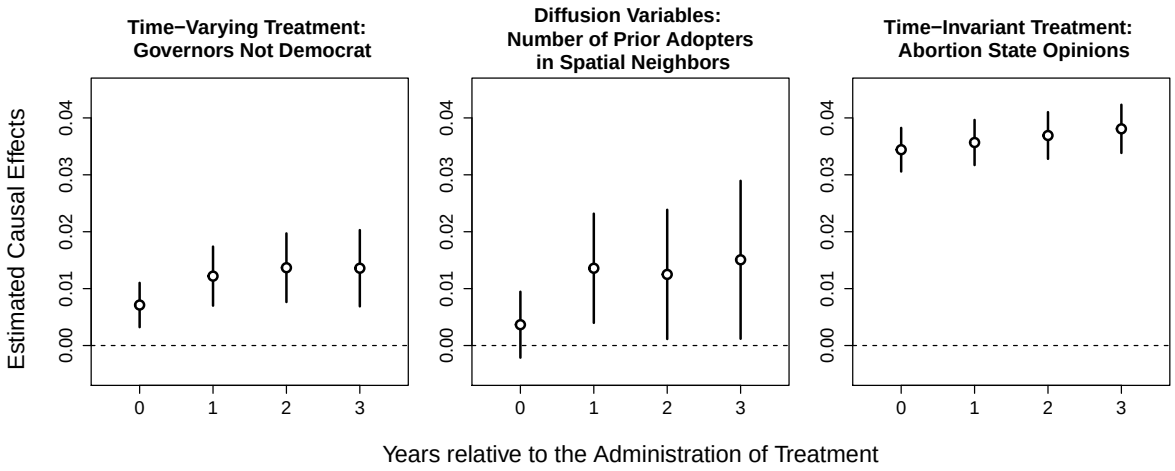

Figure 2: Causal Estimates for Three Hypotheses in Kreitzer (2015). *Note:* We report causal estimates with their 95% confidence intervals.

7.1 Estimating Causal Effects Under Ignorability Assumptions

First, we estimate the ATEs under the ignorability assumptions. We consider sensitivity analysis in the subsequent section. We report the policy- and time-average causal effects as a single-value causal summary for each hypothesis. When useful, researchers can also study the ATEs for each policy and time separately, using the same method. Figure 2 presents the results for each hypothesis. In the left panel, we report estimates of the ATE of the partisanship of state governors. We estimate the causal effect of having non-Democratic governors relative to Democratic governors such that the partisanship hypothesis predicts the positive ATE. The ATE on the immediate future outcome ($s = 0$) is 0.71 percentage points (95% CI = $[0.32, 1.10]$), which is substantively meaningful given the overall policy adoption rate is small. Interestingly, the longer-term ATEs are much larger. For example, after three years of the treatment, the ATE is as large as 1.36 percentage points (95% CI = $[0.69, 2.03]$).

Second, we also find some causal effects from spatial neighbors. We present the causal effects of changing the number of prior adopters in contiguous neighbors from 0 to 1. Because the overall policy adoption rate is small in this data, about 77% of observations take one of these two treatment values. Again, researchers can choose any causal contrasts that are relevant to their substantive questions, such as from 0% to 50% of states' spatial neighbors. The ATE on the immediate future outcome is 0.37 percentage points (95% CI = $[-0.21, 0.95]$). After three years of the treatment assignment, the ATE is as large as 1.51 percentage points (95% CI = $[0.12, 2.89]$), which is substantively large and statistically significant. Even though the policy diffusion literature has analyzed the regional diffusion hypothesis for more than three decades, they have almost exclusively focused only on a short-term diffusion effect. However, as social

learning theories naturally imply, causal diffusion of policies can take time. It is thus critical to examine causal effects not only on short-term outcomes but also on longer-term outcomes.

Finally, in the right panel, we present estimates of the ATE of the state-level public opinions about abortion. We focus on the causal effects of changing the state-level public opinion from its 40th percentile to its 60th percentile, with higher values representing more conservative opinions. An estimate of the ATE on the immediate future outcome is 3.44 percentage points (95% CI = $[3.06, 3.82]$), and the long-term ATE after three years of the treatment is 3.81 percentage points (95% CI = $[3.38, 4.23]$).

7.2 Sensitivity Analysis for Unmeasured Confounding

In practice, it is always critical to assess the robustness of causal conclusions to unmeasured confounding. We now show how to use the proposed sensitivity analysis to do so.

Figure 3 reports the results of the sensitivity analysis for the three causal estimands. For all causal estimands, we conduct the sensitivity analysis for the long-term causal effects measured three years after the treatment administration ($s = 3$). Researchers can do similar sensitivity analyses for any outcome measures $s \geq 0$. The blue thick bars present bounds for the ATE and thin bars show the 95% confidence interval of the bounds. On the top left corners, we also report the E-values.

In the left panel, we consider the long-term ATE of the partisanship of state governors and investigate sensitivity parameters $\Gamma = \{0.00, 0.05, 0.10, 0.20, 0.50, 1.00\}$. As a reference, we also report $\Gamma = 0$, which corresponds to a scenario of no unmeasured confounding and reproduces estimates from Figure 2. Importantly, the ATE is robust to a small to moderate amount of unmeasured confounding, such as $\Gamma = 0.10$, which represents unmeasured confounding that makes one group more likely to adopt policies than the other group up to 10%. The E-value is estimated to be 0.17. This means that the estimated ATE can be explained away if there exists unmeasured confounding that makes one group (treatment or control) 17% more likely to adopt policies than the other group.

Second, we consider the regional diffusion hypothesis reported in the middle panel. As we emphasized in Section 3.5, identification of causal diffusion effects is particularly difficult, which should be reflected in the choice of sensitivity parameters. Following discussions in Section 3.5, we use $\Gamma = \{0.00, 0.05, 0.30, 1.00, 1.50, 2.00\}$ where $\Gamma = 0$ is again reported as a reference point. Even though the ATE estimate under the spatial version of sequential ignorability ($\Gamma = 0$) is statistically significant, results are not as robust to unmeasured confounding as we might expect from the original results. The E-value is estimated to be 0.11.

The final panel presents results for the ATE of the state-level public opinions about abortion.

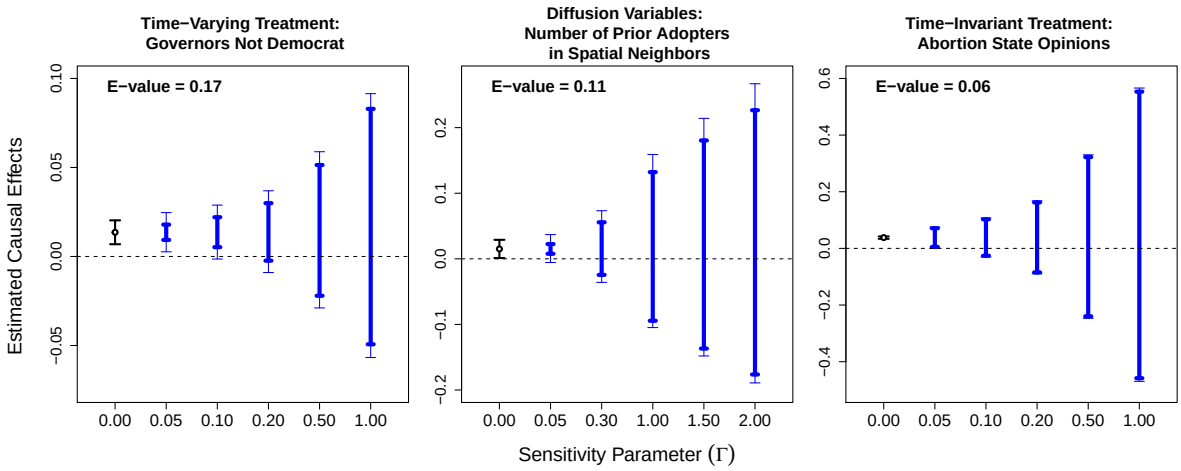

Figure 3: Sensitivity Analysis for Three Hypotheses in Kreitzer (2015). *Note:* We report causal estimates from the proposed sensitivity analysis with their 95% confidence intervals.

The results are less robust to unmeasured confounding than the other two results. The E-value is estimated to be 0.06, which means that the point estimate becomes zero as long as we account for unmeasured confounding that makes one group more likely to adopt policies than the other group by 6%. Importantly, even if the original ATE estimate is highly significant under the assumption of no unmeasured confounding, it can still be sensitive to unmeasured confounding, as in this case. Therefore, researchers have to explicitly conduct sensitivity analyses to check the causal robustness to unmeasured confounding.

These results together highlight the importance of causal inference and sensitivity analyses. Applying the sensitivity analyses, researchers can go beyond a binary debate about whether causal assumptions hold or not and explicitly quantify the robustness to unmeasured confounding.

8 Concluding Remarks

Explaining what factors influence governments' decisions on policy adoption is fundamental to all subfields of political science. Even though the policy diffusion literature has made significant progress both in terms of substantive findings and methodologies, issues of causality have received little attention to date. In this paper, we propose a coherent causal framework for policy diffusion studies and show that researchers can study a wide variety of causal estimands for analyzing policy adoptions. Recognizing the difficulty of eliminating concerns of unmeasured confounding, we develop a general sensitivity analysis approach, with which analysts can assess the robustness of causal conclusions to unmeasured confounding. To avoid the subjective choice of sensitivity parameters, we also propose the E-value as a single-value summary of the robust-

ness to unmeasured confounding. To facilitate the use of the proposed approach, we provide a companion R package cEHA that implements all the methods we discuss in this paper. We hope this article helps researchers make causal inference in policy diffusion studies by examining underlying assumptions, estimating causal effects, and explicitly accounting for potential unmeasured confounding.

References

Abadie, Alberto, Alexis Diamond and Jens Hainmueller. 2010. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *Journal of the American Statistical Association* 105(490):493–505.

Andrea, Rotnitzky, Daniel Scharfstein, Ting-Li Su and James Robins. 2001. "Methods for Conducting Sensitivity Analysis of Trials with Potentially Nonignorable Competing Causes of Censoring." *Biometrics* 57(1):103–113.

Beck, Nathaniel. 1999. Modelling Space and Time: The Event History Approach. In *Research Strategies in Social Science*. Oxford University Press.

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.

Blackwell, Matthew. 2014. "A Selection Bias Approach To Sensitivity Analysis For Causal Effects." *Political Analysis* 22(2):169–182.

Blackwell, Matthew and Adam N Glynn. 2018. "How to Make Causal Inferences With Time-Series Cross-Sectional Data Under Selection On Observables." *American Political Science Review* 112(4):1067–1082.

Boehmke, Frederick J. 2009. "Approaches to Modeling the Adoption and Diffusion of Policies with Multiple Components." *State Politics & Policy Quarterly* 9(2):229–252.

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

Brumback, Babette A, Miguel A Hernán, Sebastien JPA Haneuse and James M Robins. 2004. "Sensitivity Analyses for Unmeasured Confounding Assuming A Marginal Structural Model for Repeated Measures." *Statistics in Medicine* 23(5):749–767.

Cinelli, Carlos and Chad Hazlett. 2020. "Making Sense of Sensitivity: Extending Omitted Variable Bias." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 82(1):39–67.

Collier, David and Richard E Messick. 1975. "Prerequisites versus Diffusion: Testing Alternative Explanations of Social Security Adoption." *American Political Science Review* 69(4):1299–1315.

Cornfield, Jerome, William Haenszel, E Cuyler Hammond, Abraham M Lilienfeld, Michael B Shimkin and Ernst L Wynder. 1959. "Smoking and Lung Cancer: Recent Evidence and A Discussion of Some Questions." *Journal of the National Cancer Institute* 22(1):173–203.

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.

Eckles, Dean and Eytan Bakshy. 2017. "Bias and High-Dimensional Adjustment in Observational Studies of Peer Effects." *arXiv preprint arXiv:1706.04692* .

Fowler, James H, Michael T Heaney, David W Nickerson, John F Padgett and Betsy Sinclair. 2011. "Causality in Political Networks." *American Politics Research* 39(2):437–480.

Galton, Francis. 1889. "'Comments' on E.B. Tylor 'On a Method of Investigating the Development of Institutions: Applied to Laws of Marriage and Descent'." *Journal of the Royal Anthropological Institute* 18:245–256, 261–269.

Gilardi, Fabrizio. 2010. "Who Learns from What in Policy Diffusion Processes?" *American Journal of Political Science* 54(3):650–666.

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(03):673–701.

Gray, Virginia. 1973. "Innovation in the States: A Diffusion Study." *American Political Science Review* 67(4):1174–1185.

Hernán, Miguel A. 2010. "The Hazards of Hazard Ratios." *Epidemiology* 21(1):13.

Hyde, Susan D. 2011. "Catch Us If You Can: Election Monitoring and International Norm Diffusion." *American Journal of Political Science* 55(2):356–369.

Imai, Kosuke and In Song Kim. 2019. "When Should We Use Unit Fixed Effects Regression Models for Causal Inference with Longitudinal Data?" *American Journal of Political Science* 63(2):467–490.

Imai, Kosuke, Luke Keele and Teppei Yamamoto. 2010. "Identification, Inference and Sensitivity Analysis for Causal Mediation Effects." *Statistical Science* 25(1):51–71.

Imbens, Guido W and Donald B Rubin. 2015. *Causal Inference in Statistics, Social, and Biomedical Sciences*. Cambridge University Press.

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.

Kasza, Jessica, Kevan R Polkinghorne, Mark R Marshall, Stephen P McDonald and Rory Wolfe. 2015. "Clustering and Residual Confounding in the Application of Marginal Structural Models: Dialysis Modality, Vascular Access, and Mortality." *American Journal of Epidemiology* 182(6):535–543.

Klungsøyr, Ole, Joe Sexton, Inger Sandanger and Jan F Nygard. 2009. "Sensitivity Analysis for Unmeasured Confounding in a Marginal Structural Cox Proportional Hazards Model." *Lifetime Data Analysis* 15(2):278–294.

Kreitzer, Rebecca J. 2015. "Politics and Morality in State Abortion Policy." *State Politics & Policy Quarterly* 15(1):41–66.

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

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

Norrander, Barbara. 2001. "Measuring State Public Opinion With the Senate National Election Study." *State Politics & Policy Quarterly* 1(1):111–125.

Robins, James. 1986. "A New Approach To Causal Inference In Mortality Studies With A Sustained Exposure Period — Application to Control of the Healthy Worker Survivor Effect." *Mathematical modelling* 7(9-12):1393–1512.

Shipan, Charles R and Craig Volden. 2006. "Bottom-up Federalism: The Diffusion of Antismoking Policies from US Cities to States." *American journal of political science* 50(4):825–843.

Simmons, Beth A and Zachary Elkins. 2004. "The Globalization of Liberalization: Policy Diffusion in the International Political Economy." *American Political Science Review* pp. 171–189.

Starr, Harvey. 1991. "Democratic Dominoes: Diffusion Approaches to the Spread of Democracy in the International System." *Journal of Conflict Resolution* 35(2):356–381.

VanderWeele, Tyler J and Peng Ding. 2017. "Sensitivity Analysis in Observational Research: Introducing the E-value." *Annals of Internal Medicine* 167(4):268–274.

Volden, Craig, Michael M Ting and Daniel P Carpenter. 2008. "A Formal Model of Learning and Policy Diffusion." *American Political Science Review* 102(3):319–332.

Walker, Jack L. 1969. "The Diffusion of Innovations Among the American States." *American political science review* 63(03):880–899.

Xu, Yiqing. 2017. "Generalized Synthetic Control Method: Causal Inference with Interactive Fixed Effects Models." *Political Analysis* 25(1):57–76.

Online Appendix

Causal Inference in Policy Diffusion Studies:
Sensitivity Analysis for Unmeasured Confounding

Table of Contents

A Identification of Causal Estimands

A.1 Definitions of Causal Effects

Here, we formally define the time-average ATE and the policy- and time-average ATE for completeness.

$$\tau_k \coloneqq \sum_{t=1}^{T_k} w_{kt}\tau_{tk} \quad \text{where} \quad w_{tk} = \frac{\sum_{i=1}^{n}(1 - Y_{i,t-1,k})}{\sum_{t'=1}^{T_k}\sum_{i=1}^{n}(1 - Y_{i,t'-1,k})} \tag{A.2}$$

$$\tau \coloneqq \frac{1}{K}\sum_{k=1}^{K}\tau_k, \tag{A.3}$$

where T_k denotes the number of time periods for policy k. The former takes a weighted average of the ATE over the timing of the treatment administration with weights w_{tk} proportional to the number of potential adopters in each time period. The later averages the time-average ATE over policies.

Because identification and estimation methods for the ATE are directly applicable to these other relevant causal estimands, our formal discussion will focus on the ATE without loss of generality.

A.2 Time-Varying Treatments

We consider a general setting of long-term causal effects as identification of the ATE on the immediate future outcome is a special case.

To study causal effects on long-term outcomes, we define the main outcome of interest to be $Y_{i,t+s,k}$, the policy adoption status at time $t+s$, which represents whether unit i adopts policy k at or before time $t+s$. The goal is to estimate causal effects on the outcome measured at s time periods after the administration of treatment at time t. In standard event history analysis, researchers often focus only on the immediate future outcome Y_{itk} where $s = 0$. However, it is of great substantive importance to capture a longer-term cumulative causal effect on $Y_{i,t+s,k}$ because many treatment variables important in the policy diffusion literature can take time to realize causal effects. If we only focus on effects on the immediate future outcomes, we might miss important treatment variables that affect policy adoption processes in the long-run.

Formally, the long-term ATE can be defined as,

$$\tau_{kt}(s) \coloneqq \Pr\{Y_{i,t+s,k}(A_{itk} = 1) = 1 \mid Y_{i,t-1,k} = 0\} \;-\; \Pr\{Y_{i,t+s,k}(A_{itk} = 0) = 1 \mid Y_{i,t-1,k} = 0\},$$

which quantifies the total causal effect of the time-varying treatment at time t on whether policy k was adopted by time $t+s$ among potential adopters. For example, this could represent the average causal effect of having interest groups in year t on the policy adoption rate by year $t+s$. By changing s, analysts can study different long-term causal effects on policy adoptions.

Remark. The long-term ATE is different from the lagged causal effects studied in marginal structural models (e.g., Robins et al., 2000; Blackwell and Glynn, 2018), which quantifies the causal effect of treatment history. In contrast, the long-term ATE is the total causal effect of the time-varying treatment at time t on the outcome at time $t+s$. This is the causal quantity we can estimate if we can hypothetically randomize the time-varying treatment at time t. Therefore,

this causal quantity inherits from a randomized experiment the simple interpretation as the total causal effect. □

We need to make a general version of the sequential ignorability assumption.

$$\{Y_{i,t+s,k}(A_{itk}=1), Y_{i,t+s,k}(A_{itk}=0)\} \perp\!\!\!\perp A_{itk} \mid Y_{i,t-1,k}=0, \mathbf{X}_{itk}. \tag{A.4}$$

This is a general version of the sequential ignorability assumption we presented in Assumption 1. The overlap assumption requires

$$0 < \Pr(A_{itk}=1 \mid Y_{i,t-1,k}=0, \mathbf{X}_{itk}=\mathbf{x}) < 1 \tag{A.5}$$

for all $\mathbf{x} \in \mathcal{X}$ where $\mathcal{X}$ is the support of $\mathbf{X}_{itk}$ given $Y_{i,t-1,k}=0$.

Under the general version of sequential ignorability and overlap assumptions, we can derive the standard identification formula (Rosenbaum and Rubin, 1983; Robins, 1986; Pearl, 1995) as follows.

$$\begin{aligned}
& \Pr\{Y_{i,t+s,k}(A_{itk}=1)=1 \mid Y_{i,t-1,k}=0\} \;-\; \Pr\{Y_{i,t+s,k}(A_{itk}=0)=1 \mid Y_{i,t-1,k}=0\} \\
= & \int \Big\{ \Pr\{Y_{i,t+s,k}(A_{itk}=1)=1 \mid Y_{i,t-1,k}=0, \mathbf{X}_{itk}\} \\
& \qquad - \;\Pr\{Y_{i,t+s,k}(A_{itk}=0)=1 \mid Y_{i,t-1,k}=0, \mathbf{X}_{itk}\} \Big\} dP(\mathbf{X}_{itk} \mid Y_{i,t-1,k}=0) \\
= & \int \Big\{ \Pr\{Y_{i,t+s,k}(A_{itk}=1)=1 \mid A_{itk}=1, Y_{i,t-1,k}=0, \mathbf{X}_{itk}\} \\
& \qquad - \;\Pr\{Y_{i,t+s,k}(A_{itk}=0)=1 \mid A_{itk}=0, Y_{i,t-1,k}=0, \mathbf{X}_{itk}\} \Big\} dP(\mathbf{X}_{itk} \mid Y_{i,t-1,k}=0) \\
= & \int \Big\{ \Pr(Y_{i,t+s,k}=1 \mid A_{itk}=1, Y_{i,t-1,k}=0, \mathbf{X}_{itk}) \\
& \qquad - \;\Pr(Y_{i,t+s,k}=1 \mid A_{itk}=0, Y_{i,t-1,k}=0, \mathbf{X}_{itk}) \Big\} dP(\mathbf{X}_{itk} \mid Y_{i,t-1,k}=0),
\end{aligned}$$

where the first equality follows from the rule of conditional expectation, the second from sequential ignorability, and the third from consistency of the potential outcomes.

A.3 Time-Invariant Treatments

For the time-invariant treatment, we can similarly identify the ATE under the baseline ignorability (Assumption 2) and the corresponding overlap assumption, which requires that

$$0 < \Pr(B_{ik}=1 \mid \mathbf{X}_{ik}=\mathbf{x}) < 1 \tag{A.6}$$

for all $\mathbf{x} \in \mathcal{X}$ where $\mathcal{X}$ is the support of $\mathbf{X}_{ik}$. We can derive the standard identification formula as follows.

$$\begin{aligned}
& \Pr\{Y_{isk}(B_{ik}=1)=1\} \;-\; \Pr\{Y_{isk}(B_{ik}=0)=1\} \\
= & \int \Big\{ \Pr\{Y_{isk}(B_{ik}=1)=1 \mid \mathbf{X}_{ik}\} - \;\Pr\{Y_{isk}(B_{ik}=0)=1 \mid \mathbf{X}_{ik}\} \Big\} dP(\mathbf{X}_{ik}) \\
= & \int \Big\{ \Pr\{Y_{isk}(B_{ik}=1)=1 \mid B_{ik}=1, \mathbf{X}_{ik}\} - \;\Pr\{Y_{isk}(B_{ik}=0)=1 \mid B_{ik}=0, \mathbf{X}_{ik}\} \Big\} dP(\mathbf{X}_{ik})
\end{aligned}$$

$$= \int \Big\{ \Pr(Y_{isk} = 1 \mid B_{ik} = 1, \mathbf{X}_{ik}) - \Pr(Y_{isk} = 1 \mid B_{ik} = 0, \mathbf{X}_{ik}) \Big\} dP(\mathbf{X}_{ik}),$$

where the first equality follows from the rule of conditional expectation, the second from baseline ignorability, and the third from consistency of the potential outcomes.

A.4 Difficulty of Identification for Spatial/Network Diffusion Variables as Treatments

The risk of unmeasured confounding is considered to be even more stark when we examine spatial/network diffusion variables as treatments. The difficulty of identifying causal diffusion effects has been known for a long time (Galton, 1889). This methodological challenge is so fundamental that similar cautions have been repeated regularly across disciplines, such as Manski (1993); Shalizi and Thomas (2011); Fowler et al. (2011); Ogburn (2018); Pang and Liu (2020). In the policy diffusion literature, Volden et al. (2008) use formal models to show that policy adoptions can be correlated across states even when there is no learning and independent governments only learn from their own experiences.

Formally, this well-known methodological challenge is about pervasive concerns over a spatial/network version of the ignorability assumption required for causal diffusion analysis. In particular, the identification of causal diffusion effects requires that observed covariates include all relevant confounders that affect the policy adoption rate and the diffusion treatment variable, which captures the spatial and network environments of each unit.

Why do many scholars explicitly and separately debate about this assumption in diffusion studies, even though the issue of unmeasured confounding is classical and the most common problem even in standard observational studies? This is because the spatial/network ignorability assumption implies the absence of two well-known types of biases, *contextual confounding* and *homophily bias*, that are difficult to adjust for in the policy diffusion studies where units are connected across space and networks.

Contextual confounding — the primary focus of the spatial diffusion literature — can exist when units share some unobserved contextual factors or when unobserved confounders are correlated across space or networks. For example, in many applications, the probability of policy adoption might be affected by interest groups (Balla, 2001; Volden et al., 2008), whose behaviors and influences are often spatially correlated. In this case, unless researchers can adjust for all relevant local interest groups, they might observe spatial clusters of policy adoptions even without any causal policy diffusion. Graham et al. (2013) provide another example by writing, "neighbourhood adoption patterns that appear to be a function of interdependence instead may occur because these neighbouring polities face similar policy problems at about the same time" (p.694), which is formally a problem of contextual confounding.

In the policy diffusion studies, researchers are often interested in diffusion from "peers," such as trade partners (Greenhill et al., 2009) and sociocultural peers (Simmons and Elkins, 2004) in international relations, and states with similar political ideology (e.g., Volden, 2006) or similar ethnic diversity (e.g., Desmarais et al., 2015) in American politics. Another well-known type of bias — *homophily bias* — arises unless we can adjust for all the similarities between peers. For example, scholars of international relations have examined diffusion effects from trade-partners (e.g., Greenhill et al., 2009). In those studies, countries do not choose their trade relationships randomly, and therefore, trade partners can have correlated policy choices without

any causal diffusion if there are unmeasured confounders affecting both trade partnerships and their policies. Thus, to account for homophily bias, researchers have to consider why two countries become trade partners in addition to what affects each country's policy adoption.

B Sensitivity Analysis

B.1 Time-Varying Treatments

We begin by deriving the lower bound of the ATE. To compute the lower bound for the ATE, we need to derive the lower bound for

$$\Pr\{Y_{i,t+s,k}(A_{itk}=1) \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0\}, \tag{A.7}$$

and the upper bound for

$$\Pr\{Y_{i,t+s,k}(A_{itk}=0) \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0\}. \tag{A.8}$$

Using the sensitivity parameter Γ, we now derive the lower bound for the first quantity.

$$\begin{aligned}
& \Pr\{Y_{i,t+s,k}(A_{itk}=1) \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0\} \\
= \;& \Pr\{Y_{i,t+s,k}(A_{itk}=1) \mid A_{itk}=1, \mathbf{X}_{itk}, Y_{i,t-1,k}=0\} \Pr(A_{itk}=1 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \\
& + \Pr\{Y_{i,t+s,k}(A_{itk}=1) \mid A_{itk}=0, \mathbf{X}_{itk}, Y_{i,t-1,k}=0\} \Pr(A_{itk}=0 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \\
= \;& \Pr\{Y_{i,t+s,k} \mid A_{itk}=1, \mathbf{X}_{itk}, Y_{i,t-1,k}=0\} \Pr(A_{itk}=1 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \\
& + \Pr\{Y_{i,t+s,k}(A_{itk}=1) \mid A_{itk}=0, \mathbf{X}_{itk}, Y_{i,t-1,k}=0\} \Pr(A_{itk}=0 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \\
\geq \;& \Pr(Y_{i,t+s,k} \mid A_{itk}=1, \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \Pr(A_{itk}=1 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \\
& + \left\{ \frac{1}{\Gamma+1} \Pr(Y_{i,t+s,k} \mid A_{itk}=1, \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \right\} \Pr(A_{itk}=0 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \\
= \;& \Pr(Y_{i,t+s,k} \mid A_{itk}=1, \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \\
& \qquad \times \left\{ \Pr(A_{itk}=1 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) + \frac{1}{\Gamma+1} \Pr\{A_{itk}=0 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \right\},
\end{aligned}$$

where the third inequality follows from the definition of the sensitivity parameter. Similarly, we can derive the upper bound for the second quantity.

$$\begin{aligned}
& \Pr\{Y_{i,t+s,k}(A_{itk}=0) \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0\} \\
= \;& \Pr\{Y_{i,t+s,k}(A_{itk}=0) \mid A_{itk}=1, \mathbf{X}_{itk}, Y_{i,t-1,k}=0\} \Pr(A_{itk}=1 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \\
& + \Pr\{Y_{i,t+s,k}(A_{itk}=0) \mid A_{itk}=0, \mathbf{X}_{itk}, Y_{i,t-1,k}=0\} \Pr(A_{itk}=0 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \\
= \;& \Pr\{Y_{i,t+s,k}(A_{itk}=0) \mid A_{itk}=1, \mathbf{X}_{itk}, Y_{i,t-1,k}=0\} \Pr(A_{itk}=1 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \\
& + \Pr(Y_{i,t+s,k} \mid A_{itk}=0, \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \Pr(A_{itk}=0 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \\
\leq \;& \left\{ (\Gamma+1) \times \Pr(Y_{i,t+s,k} \mid A_{itk}=0, \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \right\} \Pr(A_{itk}=1 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \\
& + \Pr(Y_{i,t+s,k} \mid A_{itk}=1, \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \Pr(A_{itk}=0 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \\
= \;& \Pr(Y_{i,t+s,k} \mid A_{itk}=0, \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \\
& \qquad \times \left\{ (\Gamma+1) \times \Pr(A_{itk}=1 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) + \Pr(A_{itk}=0 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \right\}.
\end{aligned}$$

Taken together,

$$
\begin{aligned}
& \Pr\{Y_{i,t+s,k}(A_{itk}=1) \mid \mathbf{X}_{itk}, Y_{i,t+j-1,k}=0\} - \Pr\{Y_{i,t+s,k}(A_{itk}=0) \mid \mathbf{X}_{itk}, Y_{i,t+j-1,k}=0\} \\
\geq\ & \Pr(Y_{i,t+s,k} \mid A_{itk}=1, \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \\
& \qquad \times \left\{ \Pr(A_{itk}=1 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) + \frac{1}{\Gamma+1} \Pr(A_{itk}=0 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \right\} \\
& - \Pr(Y_{i,t+s,k} \mid A_{itk}=0, \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \\
& \qquad \times \left\{ (\Gamma+1) \times \Pr(A_{itk}=1 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) + \Pr[A_{itk}=0 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \right\}.
\end{aligned}
$$

By marginalizing over $\mathbf{X}_{itk}$ with $dP(\mathbf{X}_{itk} \mid Y_{i,t-1,k}=0)$, we obtain the desired result (equation (7)).

For the upper bound of the ATE, we need to derive the upper bound for equation (A.7) and the lower bound for equation (A.8). Using the sensitivity parameter Γ_A, we now derive the upper bound for equation (A.7) as follows.

$$
\begin{aligned}
& \Pr\{Y_{i,t+s,k}(A_{itk}=1) \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0\} \\
=\ & \Pr\{Y_{i,t+s,k}(A_{itk}=1) \mid A_{itk}=1, \mathbf{X}_{itk}, Y_{i,t-1,k}=0\} \Pr(A_{itk}=1 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \\
& + \Pr\{Y_{i,t+s,k}(A_{itk}=1) \mid A_{itk}=0, \mathbf{X}_{itk}, Y_{i,t-1,k}=0\} \Pr(A_{itk}=0 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \\
=\ & \Pr\{Y_{i,t+s,k} \mid A_{itk}=1, \mathbf{X}_{itk}, Y_{i,t-1,k}=0\} \Pr(A_{itk}=1 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \\
& + \Pr\{Y_{i,t+s,k}(A_{itk}=1) \mid A_{itk}=0, \mathbf{X}_{itk}, Y_{i,t-1,k}=0\} \Pr(A_{itk}=0 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \\
\leq\ & \Pr(Y_{i,t+s,k} \mid A_{itk}=1, \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \Pr(A_{itk}=1 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \\
& + \left\{ (\Gamma+1) \times \Pr(Y_{i,t+s,k} \mid A_{itk}=1, \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \right\} \Pr(A_{itk}=0 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \\
=\ & \Pr(Y_{i,t+s,k} \mid A_{itk}=1, \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \\
& \qquad \times \left\{ \Pr(A_{itk}=1 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) + (\Gamma+1) \times \Pr\{A_{itk}=0 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \right\},
\end{aligned}
$$

Similarly, we can derive the lower bound for the second quantity.

$$
\begin{aligned}
& \Pr\{Y_{i,t+s,k}(A_{itk}=0) \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0\} \\
=\ & \Pr\{Y_{i,t+s,k}(A_{itk}=0) \mid A_{itk}=1, \mathbf{X}_{itk}, Y_{i,t-1,k}=0\} \Pr(A_{itk}=1 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \\
& + \Pr\{Y_{i,t+s,k}(A_{itk}=0) \mid A_{itk}=0, \mathbf{X}_{itk}, Y_{i,t-1,k}=0\} \Pr(A_{itk}=0 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \\
=\ & \Pr\{Y_{i,t+s,k}(A_{itk}=0) \mid A_{itk}=1, \mathbf{X}_{itk}, Y_{i,t-1,k}=0\} \Pr(A_{itk}=1 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \\
& + \Pr(Y_{i,t+s,k} \mid A_{itk}=0, \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \Pr(A_{itk}=0 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \\
\geq\ & \left\{ \frac{1}{\Gamma+1} \times \Pr(Y_{i,t+s,k} \mid A_{itk}=0, \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \right\} \Pr(A_{itk}=1 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \\
& + \Pr(Y_{i,t+s,k} \mid A_{itk}=1, \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \Pr(A_{itk}=0 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \\
=\ & \Pr(Y_{i,t+s,k} \mid A_{itk}=0, \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \\
& \qquad \times \left\{ \frac{1}{\Gamma+1} \times \Pr(A_{itk}=1 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) + \Pr(A_{itk}=0 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \right\}.
\end{aligned}
$$

Taken together,

$$
\Pr\{Y_{i,t+s,k}(A_{itk}=1) \mid \mathbf{X}_{itk}, Y_{i,t+j-1,k}=0\} - \Pr\{Y_{i,t+s,k}(A_{itk}=0) \mid \mathbf{X}_{itk}, Y_{i,t+j-1,k}=0\}
$$

$$
\begin{aligned}
\leq\ & \Pr(Y_{i,t+s,k} \mid A_{itk}=1, \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \\
& \qquad \times \Big\{ \Pr(A_{itk}=1 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) + (\Gamma+1) \times \Pr(A_{itk}=0 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \Big\} \\
& - \Pr(Y_{i,t+s,k} \mid A_{itk}=0, \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \\
& \qquad \times \Big\{ \frac{1}{\Gamma+1} \times \Pr(A_{itk}=1 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) + \Pr[A_{itk}=0 \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \Big\}.
\end{aligned}
$$

By marginalizing over $\mathbf{X}_{itk}$ with $dP(\mathbf{X}_{itk} \mid Y_{i,t-1,k}=0)$, we obtain the upper bound. □

B.2 Continuous Treatments

Here, we generalize the sensitivity analysis to continuous time-varying treatments. The same proof applies to continuous time-invariant treatments, and thus, their proofs are omitted.

For continuous treatments, we define the long-term ATE as follows.

$$
\tau_{kt}(s) \coloneqq \Pr\{Y_{i,t+s,k}(A_{itk}=a_1)=1 \mid Y_{i,t-1,k}=0\} \ - \ \Pr\{Y_{i,t+s,k}(A_{itk}=a_0)=1 \mid Y_{i,t-1,k}=0\},
$$

where $\{a_1, a_0\}$ are two constants researchers specify. Then, we define sensitivity parameter Γ as follows. For $a^* \in \{a_1, a_0\}$ and $\tilde{a} \in \mathcal{A}$ where $\mathcal{A}$ is the support of A_{itk},

$$
\frac{1}{\Gamma+1} \leq \frac{\Pr\{Y_{i,t+s,k}(A_{itk}=a^*)=1 \mid A_{itk}=\tilde{a}, \mathbf{X}_{itk}=\mathbf{x}, Y_{i,t-1,k}=0\}}{\Pr\{Y_{i,t+s,k}(A_{itk}=a^*)=1 \mid A_{itk}=a^*, \mathbf{X}_{itk}=\mathbf{x}, Y_{i,t-1,k}=0\}} \leq \Gamma+1.
$$

To compute the lower bound for the ATE, we need to derive the lower bound for

$$
\Pr\{Y_{i,t+s,k}(A_{itk}=a_1) \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0\}, \tag{A.9}
$$

and the upper bound for

$$
\Pr\{Y_{i,t+s,k}(A_{itk}=a_0) \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0\}. \tag{A.10}
$$

Using the sensitivity parameter Γ, we now derive the lower bound for the first quantity.

$$
\begin{aligned}
& \Pr\{Y_{i,t+s,k}(A_{itk}=a_1) \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0\} \\
= & \int_{\mathcal{A}} \Pr\{Y_{i,t+s,k}(A_{itk}=a_1) \mid A_{itk}=a, \mathbf{X}_{itk}, Y_{i,t-1,k}=0\} f(A_{itk}=a \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \\
\geq & \int_{\mathcal{A}} \Big\{ \frac{1}{\Gamma+1} \Pr\{Y_{i,t+s,k}(A_{itk}=a_1) \mid A_{itk}=a_1, \mathbf{X}_{itk}, Y_{i,t-1,k}=0\} \Big\} f(A_{itk}=a \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \\
= & \frac{1}{\Gamma+1} \Pr(Y_{i,t+s,k} \mid A_{itk}=a_1, \mathbf{X}_{itk}, Y_{i,t-1,k}=0),
\end{aligned}
$$

where $f(A_{itk}=a \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0)$ is the conditional probability density function of A_{itk} conditional on $\mathbf{X}_{itk}$ and $Y_{i,t-1,k}=0$. Similarly, we can derive the upper bound for the second quantity.

$$
\begin{aligned}
& \Pr\{Y_{i,t+s,k}(A_{itk}=a_0) \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0\} \\
= & \int_{\mathcal{A}} \Pr\{Y_{i,t+s,k}(A_{itk}=a_0) \mid A_{itk}=a, \mathbf{X}_{itk}, Y_{i,t-1,k}=0\} f(A_{itk}=a \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0) \\
\leq & \int_{\mathcal{A}} \Big\{ (\Gamma+1) \times \Pr\{Y_{i,t+s,k}(A_{itk}=a_0) \mid A_{itk}=a_0, \mathbf{X}_{itk}, Y_{i,t-1,k}=0\} \Big\} f(A_{itk}=a \mid \mathbf{X}_{itk}, Y_{i,t-1,k}=0)
\end{aligned}
$$

$$= \quad (\Gamma + 1) \times \Pr(Y_{i,t+s,k} \mid A_{itk} = a_0, \mathbf{X}_{itk}, Y_{i,t-1,k} = 0).$$

Taken together,

$$\begin{aligned}
& \Pr\{Y_{i,t+s,k}(A_{itk} = a_1) \mid Y_{i,t-1,k} = 0\} - \Pr\{Y_{i,t+s,k}(A_{itk} = a_0) \mid Y_{i,t-1,k} = 0\} \\
= \quad & \int \Bigg[\Pr\{Y_{i,t+s,k}(A_{itk} = a_1) \mid Y_{i,t-1,k} = 0, \mathbf{X}_{itk}\} \\
& \qquad - \Pr\{Y_{i,t+s,k}(A_{itk} = a_0) \mid Y_{i,t-1,k} = 0, \mathbf{X}_{itk}\} \Bigg] dP(\mathbf{X}_{itk} \mid Y_{i,t-1,k} = 0) \\
\geq \quad & \int \Bigg[\frac{1}{\Gamma + 1} \Pr(Y_{i,t+s,k} \mid A_{itk} = a_1, \mathbf{X}_{itk}, Y_{i,t-1,k} = 0) \\
& \qquad - (\Gamma + 1) \times \Pr(Y_{i,t+s,k} \mid A_{itk} = a_0, \mathbf{X}_{itk}, Y_{i,t-1,k} = 0) \Bigg] dP(\mathbf{X}_{itk} \mid Y_{i,t-1,k} = 0).
\end{aligned}$$

□

B.3 Time-Invariant Treatments

To compute the lower bound for the ATE, we need to derive the lower bound for

$$\Pr\{Y_{isk}(B_{ik} = 1) \mid \mathbf{X}_{ik}\}, \tag{A.11}$$

and the upper bound for

$$\Pr\{Y_{isk}(B_{ik} = 0) \mid \mathbf{X}_{ik}\}. \tag{A.12}$$

Using the sensitivity parameter Γ, we now derive the lower bound for the first quantity.

$$\begin{aligned}
& \Pr\{Y_{isk}(B_{ik} = 1) \mid \mathbf{X}_{ik}\} \\
= \quad & \Pr\{Y_{isk}(B_{ik} = 1) \mid B_{ik} = 1, \mathbf{X}_{ik}\} \Pr(B_{ik} = 1 \mid \mathbf{X}_{ik}) + \Pr\{Y_{isk}(B_{ik} = 1) \mid B_{ik} = 0, \mathbf{X}_{ik}\} \Pr(B_{ik} = 0 \mid \mathbf{X}_{ik}) \\
= \quad & \Pr(Y_{isk} \mid B_{ik} = 1, \mathbf{X}_{ik}) \Pr(B_{ik} = 1 \mid \mathbf{X}_{ik}) + \Pr\{Y_{isk}(B_{ik} = 1) \mid B_{ik} = 0, \mathbf{X}_{ik}) \Pr(B_{ik} = 0 \mid \mathbf{X}_{ik}) \\
\geq \quad & \Pr(Y_{isk} \mid B_{ik} = 1, \mathbf{X}_{ik}) \Pr(B_{ik} = 1 \mid \mathbf{X}_{ik}) + \left\{ \frac{1}{\Gamma + 1} \Pr(Y_{isk} \mid B_{ik} = 1, \mathbf{X}_{ik}) \right\} \Pr(B_{ik} = 0 \mid \mathbf{X}_{ik}) \\
= \quad & \Pr(Y_{isk} \mid B_{ik} = 1, \mathbf{X}_{ik}) \times \left\{ \Pr(B_{ik} = 1 \mid \mathbf{X}_{ik}) + \frac{1}{\Gamma + 1} \Pr(B_{ik} = 0 \mid \mathbf{X}_{ik}) \right\},
\end{aligned}$$

where the third inequality follows from the sensitivity parameter defined in equation (13). Similarly, we can derive the upper bound for the second quantity.

$$\begin{aligned}
& \Pr\{Y_{isk}(B_{ik} = 0) \mid \mathbf{X}_{ik}\} \\
= \quad & \Pr\{Y_{isk}(B_{ik} = 0) \mid B_{ik} = 1, \mathbf{X}_{ik}\} \Pr(B_{ik} = 1 \mid \mathbf{X}_{ik}) + \Pr\{Y_{isk}(B_{ik} = 0) \mid B_{ik} = 0, \mathbf{X}_{ik}\} \Pr(B_{ik} = 0 \mid \mathbf{X}_{ik}) \\
= \quad & \Pr\{Y_{isk}(B_{itk} = 0) \mid B_{ik} = 1, \mathbf{X}_{ik}\} \Pr(B_{ik} = 1 \mid \mathbf{X}_{ik}) + \Pr(Y_{isk} \mid B_{ik} = 0, \mathbf{X}_{ik}) \Pr(B_{ik} = 0 \mid \mathbf{X}_{ik}) \\
\leq \quad & \left\{ (\Gamma + 1) \times \Pr(Y_{isk} \mid B_{ik} = 0, \mathbf{X}_{ik}) \right\} \Pr(B_{ik} = 1 \mid \mathbf{X}_{ik}) + \Pr(Y_{isk} \mid B_{ik} = 0, \mathbf{X}_{ik}) \Pr(B_{ik} = 0 \mid \mathbf{X}_{ik}) \\
= \quad & \Pr(Y_{isk} \mid B_{ik} = 0, \mathbf{X}_{ik}) \times \left\{ (\Gamma + 1) \times \Pr(B_{ik} = 1 \mid \mathbf{X}_{ik}) + \Pr(B_{ik} = 0 \mid \mathbf{X}_{ik}) \right\}.
\end{aligned}$$

Taken together,

$$\Pr\{Y_{isk}(B_{itk} = 1) \mid \mathbf{X}_{ik}\} - \Pr\{Y_{isk}(B_{itk} = 0) \mid \mathbf{X}_{ik}\}$$

$$
\begin{aligned}
&\geq \Pr(Y_{isk} \mid B_{ik} = 1, \mathbf{X}_{ik}) \times \left\{ \Pr(B_{ik} = 1 \mid \mathbf{X}_{ik}) + \frac{1}{\Gamma + 1} \Pr(B_{ik} = 0 \mid \mathbf{X}_{ik}) \right\} \\
&\quad - \Pr(Y_{isk} \mid B_{ik} = 0, \mathbf{X}_{ik}) \times \left\{ (\Gamma + 1) \times \Pr(B_{ik} = 1 \mid \mathbf{X}_{ik}) + \Pr(B_{ik} = 0 \mid \mathbf{X}_{ik}) \right\}.
\end{aligned}
$$

By marginalizing over $\mathbf{X}_{itk}$ with $dP(\mathbf{X}_{ik})$, we obtain the desired result. □

B.4 E-value

B.4.1 Binary Treatments

We first define a general Q function as follows. For the time-varying treatment,

$$
\begin{aligned}
&Q(a_1, a_2) \\
= \;& \mathbb{E}_{\mathbf{X} \mid \mathsf{PA}_t} \left\{ \Pr(Y_{i,t+s,k} = 1 \mid A_{itk} = a_1, \mathbf{X}_{itk}, Y_{i,t-1,k} = 0) \Pr(A_{itk} = a_2 \mid \mathbf{X}_{itk}, Y_{i,t-1,k} = 0) \right\},
\end{aligned}
$$

where $\mathbb{E}_{\mathbf{X} \mid \mathsf{PA}_t}(\cdot)$ denotes the empirical average over the distribution of $\mathbf{X}_{itk}$ among potential adopters $Y_{i,t-1,k} = 0$.

For the time-invariant treatment,

$$
Q(b_1, b_2) = \mathbb{E}_{\mathbf{X}} \left\{ \Pr(Y_{isk} = 1 \mid B_{ik} = b_1, \mathbf{X}_{ik}) \Pr(B_{ik} = b_2 \mid \mathbf{X}_{ik}) \right\}.
$$

where $\mathbb{E}_{\mathbf{X}}(\cdot)$ denotes the empirical average over the distribution of $\mathbf{X}_{itk}$.

Given this general notation, we can write the lower bound for each corresponding causal quantity as

$$
\big\{ Q(1,1) + \frac{1}{\Gamma + 1} Q(1,0) \big\} - \big\{ (\Gamma + 1) \times Q(0,1) + Q(0,0) \big\}.
$$

Therefore, when the original estimate is positive, the E-value is Γ with which the lower bound becomes zero. By solving the following equality,

$$
\big\{ Q(1,1) + \frac{1}{\Gamma + 1} Q(1,0) \big\} - \big\{ (\Gamma + 1) \times Q(0,1) + Q(0,0) \big\} = 0,
$$

we obtain

$$
\text{E-value} = \frac{\widehat{Q}(1,1) - \widehat{Q}(0,0)}{2\widehat{Q}(0,1)} + \sqrt{\left(\frac{\widehat{Q}(1,1) - \widehat{Q}(0,0)}{2\widehat{Q}(0,1)} \right)^2 + \frac{\widehat{Q}(1,0)}{\widehat{Q}(0,1)}} - 1
$$

Similarly, we can write the upper bound for each corresponding causal quantity as

$$
\big\{ Q(1,1) + (\Gamma + 1) \times Q(1,0) \big\} - \left\{ \frac{1}{\Gamma + 1} \times Q(0,1) + Q(0,0) \right\}.
$$

Therefore, when the original estimate is negative, the E-value is Γ with which the upper bound becomes zero. By solving the following equality,

$$
\big\{ Q(1,1) + (\Gamma + 1) \times Q(1,0) \big\} - \left\{ \frac{1}{\Gamma + 1} \times Q(0,1) + Q(0,0) \right\} = 0,
$$

we obtain

$$
\text{E-value} = \frac{\widehat{Q}(0,0) - \widehat{Q}(1,1)}{2\widehat{Q}(1,0)} + \sqrt{\left(\frac{\widehat{Q}(0,0) - \widehat{Q}(1,1)}{2\widehat{Q}(1,0)} \right)^2 + \frac{\widehat{Q}(0,1)}{\widehat{Q}(1,0)}} - 1.
$$

B.4.2 Continuous Treatments

We now re-define a general Q function as follows. For the time-varying treatment,

$$Q(a) = \mathbb{E}_{\mathbf{X}|\mathsf{PA}_t}\Bigg\{\Pr(Y_{i,t+s,k} = 1 \mid A_{itk} = a, \mathbf{X}_{itk}, Y_{i,t-1,k} = 0)\Bigg\},$$

where $\mathbb{E}_{\mathbf{X}|\mathsf{PA}_t}(\cdot)$ denotes the empirical average over the distribution of $\mathbf{X}_{itk}$ among potential adopters $Y_{i,t-1,k} = 0$.

For the time-invariant treatment,

$$Q(b) = \mathbb{E}_{\mathbf{X}}\Bigg\{\Pr(Y_{isk} = 1 \mid B_{ik} = b, \mathbf{X}_{ik})\Bigg\}.$$

where $\mathbb{E}_{\mathbf{X}}(\cdot)$ denotes the empirical average over the distribution of $\mathbf{X}_{itk}$.

Given this general notation, we can write the lower bound for each corresponding causal quantity as

$$\frac{1}{\Gamma + 1} Q(a_1) - (\Gamma + 1) \times Q(a_0).$$

Therefore, when the original estimate is positive, the E-value is Γ with which the lower bound becomes zero. By solving the following equality,

$$\frac{1}{\Gamma + 1} Q(a_1) - (\Gamma + 1) \times Q(a_0) = 0,$$

we obtain

$$\text{E-value} \quad = \quad \sqrt{\frac{Q(a_1)}{Q(a_0)}} - 1.$$

We can write the upper bound for each corresponding causal quantity as

$$(\Gamma + 1) \times Q(a_1) - \frac{1}{\Gamma + 1} \times Q(a_0).$$

Therefore, when the original estimate is negative, the E-value is Γ with which the upper bound becomes zero. By solving the following equality,

$$(\Gamma + 1) \times Q(a_1) - \frac{1}{\Gamma + 1} \times Q(a_0) = 0,$$

we obtain

$$\text{E-value} \quad = \quad \sqrt{\frac{Q(a_0)}{Q(a_1)}} - 1,$$

which completes the proof. □

C Estimation of Point Estimates and Standard Errors

We first describe a general approach to compute point estimates and standard errors for the ATE. In Section C.2, we provide a concrete example based on the standard EHA.

C.1 General Algorithm

Suppose we define parameters of interest in our model to be θ. Then, we define a general function ψ_{tk} such that $\tau_{tk} = \psi_{tk}(\theta)$. To obtain a point estimate of τ_{tk}, we use $\tau_{tk} = \psi_{tk}(\widehat{\theta})$. To obtain standard errors of $\widehat{\tau}_{tk}$, we have two approaches.

1. Cluster robust standard errors. We first estimate parameters of interest in our model θ using cluster robust standard errors. Then, sample B copies of θ from its sampling distribution $\mathcal{N}(\widehat{\theta}, \widehat{\text{Var}}(\widehat{\theta}))$ where $\widehat{\text{Var}}(\widehat{\theta})$ is estimated by cluster robust standard errors. We denote each copy by $\theta^{(b)}$ where $b \in \{1, \ldots, B\}$. This approach is called a quasi-Bayesian approach (King et al., 2000; Imai et al., 2010).

2. Nonparametric Block Bootstrap. We estimate the sampling distribution of parameters of interest in our model θ using nonparametric block bootstrap. In particular, we repeat nonparametric block bootstrap B times and denote each bootstrap estimate by $\theta^{(b)}$ where $b \in \{1, \ldots, B\}$.

For each $b \in \{1, \ldots, B\}$, we compute $\tau_{tk}^{(b)} = \psi_{tk}(\theta^{(b)})$. Finally, we compute the point estimate of τ_{tk} and its uncertainty estimates from the distribution of $\{\tau_{tk}^{(b)}\}_{b=1}^{B}$. For example, the sample mean and the sample standard deviation of the distribution can be used as the point estimate of τ_{kt} and its standard error, whereas percentiles of this distribution can serve as confidence intervals for τ_{kt}. To estimate the time-average ATE, we can use $\tau_k^{(b)} = \sum_{t=1}^{T_k} w_{kt}\psi_{tk}(\theta^{(b)})$ where $w_{tk} = \sum_{i=1}^{n}(1 - Y_{i,t-1,k}) / \sum_{t'=1}^{T_k}\sum_{i=1}^{n}(1 - Y_{i,t'-1,k})$. To estimate the policy- and time-average ATE, we can use $\tau^{(b)} = \frac{1}{K}\sum_{k=1}^{K}\tau_k^{(b)}$.

C.2 Standard EHA

Here we focus on the standard EHA model as a concrete example. Suppose we use the following standard EHA model.

$$\Pr(Y_{itk} = 1 \mid A_{itk}, \mathbf{X}_{itk}, Y_{i,t-1,k} = 0) = g(\alpha_1 t + \alpha_2 t^2 + \alpha_3 t^3 + A_{itk}\beta + \mathbf{X}_{itk}^{\top}\gamma),$$

where we include third-degree polynomials as time-trends and $g(\cdot)$ is a binary outcome model, such as logistic and probit models. Then, parameters of the model is defined as $\theta = (\alpha_1, \alpha_2, \alpha_3, \beta, \gamma)$. Then, an estimator for the ATE is given by

$$\begin{aligned}
&\widehat{\tau}_{tk} \\
&= \widehat{\mathbb{E}}_{\mathbf{X}|\mathsf{PA}_t}\left\{\widehat{\Pr}(Y_{itk} = 1 \mid A_{itk} = 1, \mathbf{X}_{itk} = \mathbf{x}, Y_{i,t-1,k} = 0) \;-\; \widehat{\Pr}(Y_{itk} = 1 \mid A_{itk} = 0, \mathbf{X}_{itk} = \mathbf{x}, Y_{i,t-1,k} = 0)\right\} \\
&= \widehat{\mathbb{E}}_{\mathbf{X}|\mathsf{PA}_t}\left\{g(\widehat{\alpha}_1 t + \widehat{\alpha}_2 t^2 + \widehat{\alpha}_3 t^3 + \widehat{\beta} + \mathbf{X}_{itk}^{\top}\widehat{\gamma}) - \; g(\widehat{\alpha}_1 t + \widehat{\alpha}_2 t^2 + \widehat{\alpha}_3 t^3 + \mathbf{X}_{itk}^{\top}\widehat{\gamma})\right\} \\
&= \frac{1}{n_{kt}}\sum_{i=1}^{n}\mathbf{1}\{Y_{i,t-1,k} = 0\}\left\{g(\widehat{\alpha}_1 t + \widehat{\alpha}_2 t^2 + \widehat{\alpha}_3 t^3 + \widehat{\beta} + \mathbf{X}_{itk}^{\top}\widehat{\gamma}) - \; g(\widehat{\alpha}_1 t + \widehat{\alpha}_2 t^2 + \widehat{\alpha}_3 t^3 + \mathbf{X}_{itk}^{\top}\widehat{\gamma})\right\}
\end{aligned}$$

where $n_{kt} = \sum_{i=1}^{n}\mathbf{1}\{Y_{i,t-1,k} = 0\}$.

Therefore, we can compute

$$\begin{aligned}
&\tau_{tk}^{(b)} \\
&= \frac{1}{n_{kt}}\sum_{i=1}^{n}\mathbf{1}\{Y_{i,t-1,k} = 0\}\{g(\alpha_1^{(b)} t + \alpha_2^{(b)} t^2 + \alpha_3^{(b)} t^3 + \beta^{(b)} + \mathbf{X}_{itk}^{\top}\gamma^{(b)})
\end{aligned}$$

$$- g(\alpha_1^{(b)} t + \alpha_2^{(b)} t^2 + \alpha_3^{(b)} t^3 + \mathbf{X}_{itk}^\top \gamma^{(b)})\}$$

where $\theta^{(b)} = (\alpha_1^{(b)}, \alpha_2^{(b)}, \alpha_3^{(b)}, \beta^{(b)}, \gamma^{(b)})$ are computed based on cluster standard errors or non-parametric block bootstrap.

D Empirical Application: Karch et al. (2016)

In this section, we use Karch et al. (2016) to illustrate our proposed approach. While Karch et al. (2016) consider a variety of hypotheses, this section focuses on two main hypotheses — (1) the causal effect of the partisanship of state governors and (2) the regional diffusion hypothesis — that are shared with Kreitzer (2015), which we analyzed in the main text. In this way, it is easier to compare results, while we should emphasize that these two papers examine different policies.

The outcome variable is the policy adoption status Y_{itk} representing whether state i adopts policy k at or before year t. We include fifty states $i \in \{1, \ldots, 50\}$, and the total number of interstate compacts is $K = 43$. To understand long-term effects on policy adoption, we study not only causal effects on the immediate future outcome but also on longer-term outcomes up to three years relative to the timing of the treatment $s \in \{0, 1, 2, 3\}$. For each hypothesis, we closely follow the original analysis to select control variables. For the partisanship hypothesis, we use the same pre-treatment covariates used in the original paper and add policy-fixed effects and third-degree polynomials of time trends. For the regional diffusion hypothesis, we use the same pre-treatment covariates used in the original paper and add policy-fixed effects and third-degree polynomials of time trends.

D.1 Estimating Causal Effects Under Sequential Ignorability Assumptions

First, we estimate the ATEs under the sequential ignorability assumption. We consider sensitivity analysis in the next section. We report the policy- and time-average causal effects as a single-value causal summary for each hypothesis. Figure A1 presents the results for each hypothesis. In the left panel, we report estimates of the ATE of the partisanship of state governors. We estimate the causal effect of having Republican governors relative to non-Republican governors. The ATE on the immediate future outcome ($s = 0$) is 0.08 percentage points (95% CI = $[-0.17, 0.34]$). After three years of the treatment, the ATE is -0.02 percentage points (95% CI = $[-0.45, 0.41]$).

Second, we find some causal effects from spatial neighbors. We present the causal effects of changing the standardized number of prior adopters in spatial contiguous neighbors (computed in the original paper) from its 20th percentile to its 80th percentile. The ATE on the immediate future outcome is 1.15 percentage points (95% CI = $[0.85, 1.45]$). After three years of the treatment assignment, the ATE is as large as 1.53 percentage points (95% CI = $[1.07, 2.00]$), which is substantively large and statistically significant.

D.2 Sensitivity Analysis for Unmeasured Confounding

In practice, it is important to assess the robustness of causal conclusions to unmeasured confounding. We now show how to use the proposed sensitivity analysis to do so.

Figure A2 reports the results of the sensitivity analysis for the two causal estimands. For all causal estimands, we conduct the sensitivity analysis for the long-term causal effects measured at three years after the treatment administration ($s = 3$). The blue thick bars present bounds

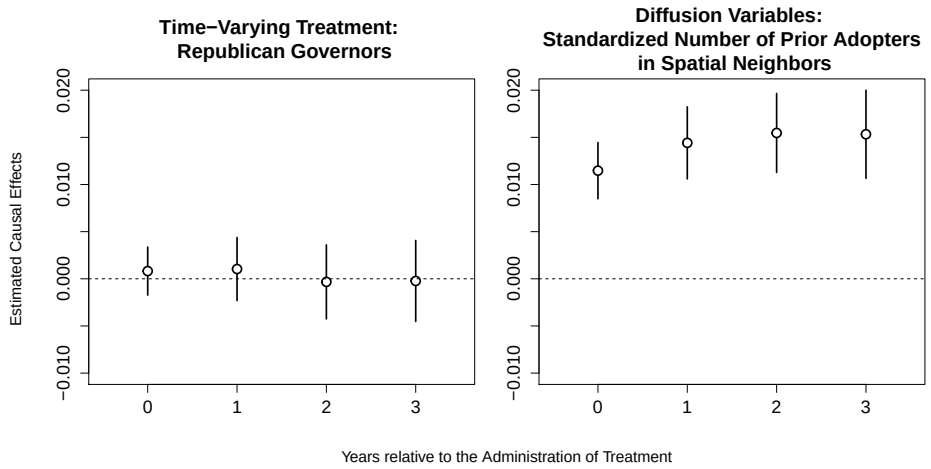

Figure A1: Causal Estimates for Two Hypotheses. *Note:* We report causal estimates with their 95% confidence intervals.

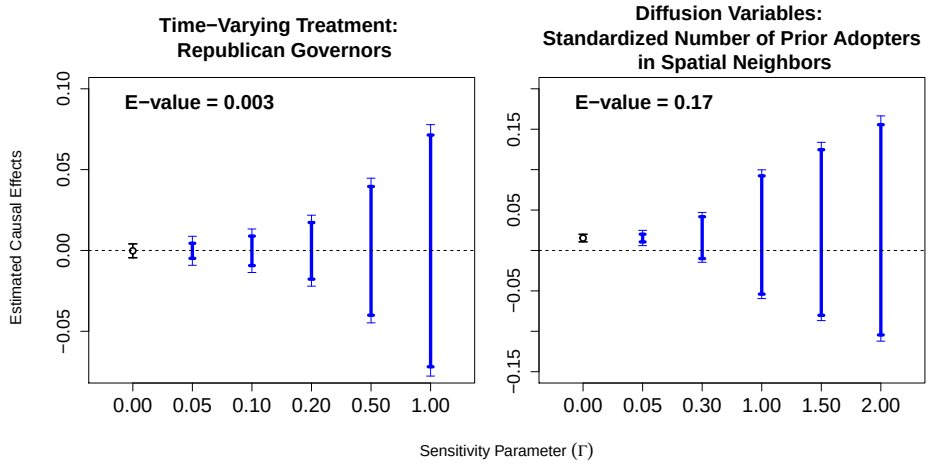

Figure A2: Sensitivity Analysis for Two Hypotheses. *Note:* We report causal estimates from the proposed sensitivity analysis with their 95% confidence intervals.

for the ATE and thin bars show the 95% confidence interval of the bounds. On the top left corners, we also report the E-values.

In the left panel, we consider the long-term ATE of the partisanship of state governors and investigate sensitivity parameters $\Gamma = \{0.00, 0.05, 0.10, 0.20, 0.50, 1.00\}$. As a reference, we also report $\Gamma = 0$, which corresponds to a scenario of no unmeasured confounding and reproduces estimates from Figure A1. As we expect from the original estimate, the ATE is not robust to a small amount of unmeasured confounding, such as $\Gamma = 0.05$, which represents unmeasured confounding that makes one group more likely to adopt policies than the other group up to 5%. The E-value is estimated to be as small as 0.003. This means that the estimated ATE can be explained away if there exists unmeasured confounding that makes one group (treatment or control) 0.3% more likely to adopt policies than the other group.

Second, we consider the regional diffusion hypothesis reported in the right panel. As we emphasized in Section 3.5, identification of causal diffusion effects is particularly difficult, which

should be reflected in the choice of sensitivity parameters. Following discussions in Section 3.5, we use $\Gamma = \{0.00, 0.05, 0.30, 1.00, 1.50, 2.00\}$ where $\Gamma = 0$ is again reported as a reference point. Even though the ATE estimate under the spatial version of sequential ignorability ($\Gamma = 0$) is statistically significant, results are not as robust to unmeasured confounding as we might expect from the original results. The E-value is estimated to be 0.17. This means that the estimated ATE can be explained away if there exists unmeasured confounding that makes one group (treatment or control) 17% more likely to adopt policies than the other group. This E-value is similar to the E-value for the regional diffusion hypothesis tested in Section 7, even though they examine different sets of policies and different control variables.

E Simulation Studies

In this section, we use simulations to investigate the finite sample performance of the proposed approach. After introducing the setup of the simulation, we first investigate scenarios in which the sequential ignorability assumption holds. We verify that our proposed estimators are consistent and confidence intervals have good coverage rates. We then examine scenarios in which the sequential ignorability assumption is violated. After showing EHA estimators relying on the sequential ignorability assumption can suffer from significant bias, we show bounds from our proposed sensitivity analysis cover the true causal effect even with unmeasured confounding.

Importantly, this simulation suggests that our default sensitivity parameters $\Gamma = \{0.00, 0.05, 0.20, 0.50, 1.00\}$ effectively capture moderate (30% $\sim$ 50% of the true causal effect) to large bias (60% $\sim$ 80% of the true causal effect).

E.1 Setup

For each unit $i \in \{1, \ldots, n\}$, time $t \in \{1, \ldots, 10\}$, and policy $k \in \{1, \ldots, 10\}$, we simulate data with the following data-generating mechanism.

- Observed covariates: For $\ell \in \{1, 2, 3\}$,

$$X_{itk\ell} = \rho_\ell X_{i,t-1,k,\ell} + \epsilon_{itk\ell}$$

  where $(\rho_1, \rho_2, \rho_3) = (0.8, 0.6, 0.4)$, $\epsilon_{itk\ell} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$. We simulate observed covariates at the first period as follows. $X_{i,1,k,\ell} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu_\ell, 1)$ with $(\mu_1, \mu_2, \mu_3) = (2, 1, 0)$.

- Confounder:

$$U_{ik} \overset{\text{i.i.d.}}{\sim} \text{Bernoulli}(0.4)$$

- Treatment Assignment: $A_{itk} \overset{\text{i.i.d.}}{\sim} \text{Bernoulli}(\Pr(A_{itk} = 1 \mid U_{ik}))$ where

$$\Pr(A_{itk} = 1 \mid U_{ik}) = 0.1 + 0.4 \times U_{ik}.$$

- Outcome model: For a given choice of lead $s \in \{0, 3\}$, $Y_{i,t+s,k} \overset{\text{i.i.d.}}{\sim} \text{Bernoulli}(\Pr(Y_{i,t+s,k} = 1 \mid A_{itk}, \mathbf{X}_{itk}, U_{ik}, Y_{i,t-1,k} = 0))$ where

$$\Pr(Y_{i,t+s,k} = 1 \mid A_{itk}, \mathbf{X}_{itk}, U_{ik}, Y_{i,t-1,k} = 0) = \text{logit}^{-1}(\alpha_k + \mathbf{X}_{itk}^\top \beta_X + A_{itk}\beta_A + U_{ik}\beta_{UY})$$

  where $\beta_X = (0.5, -0.5, 1)$ and $\beta_A = 1$. α_k is a policy-specific fixed effect, and $\alpha_k \overset{\text{i.i.d.}}{\sim} \text{Uniform}(-8, -5)$. β_{UY} captures the effect of confounder U on the outcome.

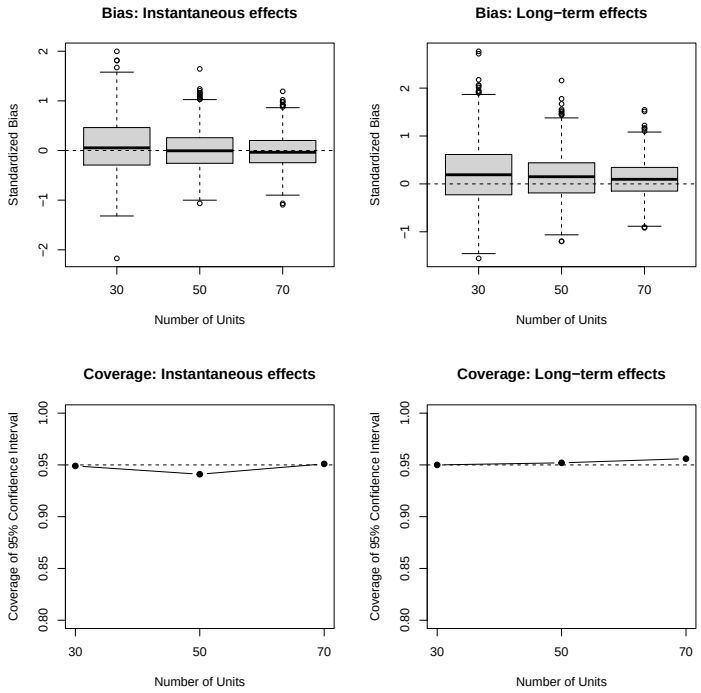

Figure A3: Bias and Coverage of EHA Estimators under Sequential Ignorability.

Note: When $s = 0$, the causal estimand is the instantaneous effect, and this corresponds to the standard EHA. When $s = 3$, the causal estimand is the long-term causal effect (i.e., the total causal effect on the adoption rate three years after the treatment assignment), and it corresponds to the EHA with long-term outcomes (see Appendix A.2).

Below, we vary the number of units n, whether we estimate the instantaneous effect s, and the strength of confounding β_{UY}. We generate 1000 simulations in each scenario to evaluate estimators in terms of bias and coverage of 95% confidence intervals. We standardize bias by the true causal effect to ease interpretation.

E.2 When Sequential Ignorability Holds

We first consider cases where the sequential ignorability assumption holds. In particular, we adjust for both $\mathbf{X}_{itk}$ and U_{ik}. In this scenario, as we do not have unmeasured confounding, we fix $\beta_{UY} = 0.5$. Results are reported in Figure A3. As we expect from causal identification results, the EHA estimators discussed in the paper have small bias and great coverages.

E.3 When Sequential Ignorability is Violated

We now consider cases where the sequential ignorability assumption is violated. In particular, we adjust for only $\mathbf{X}_{itk}$ and we omit U_{ik}. To consider different levels of unmeasured confounding, we use $\beta_{UY} = 0.5$, which we call "Moderate Bias" case, and $\beta_{UY} = 1$, which we call "Strong Bias" case. To investigate the performance of our proposed sensitivity analysis, we estimate bounds with four default sensitivity parameters $\Gamma = \{0.05, 0.20, 0.50, 1.00\}$. As we emphasize in the paper, the exact choice of sensitivity parameters should be based on domain knowledge, but we will show that this default choice of sensitivity parameters covers a wide range of unmeasured

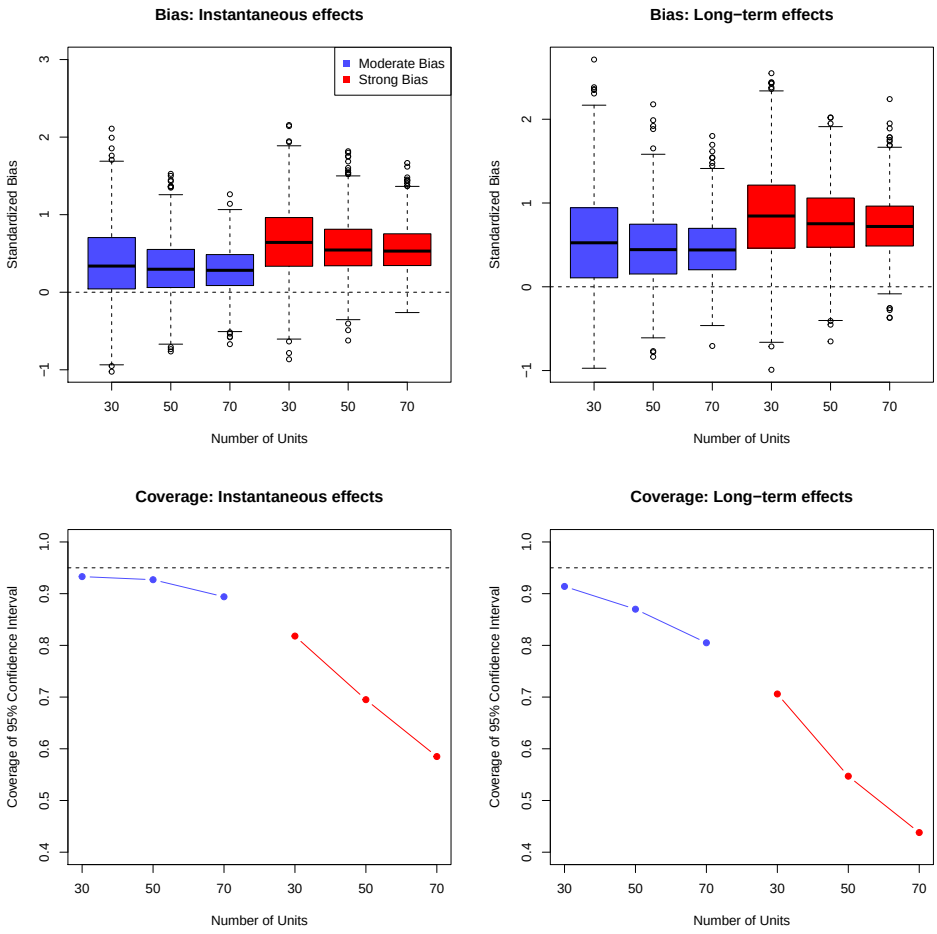

Figure A4: Bias and Coverage of EHA Estimators when Sequential Ignorability is Violated. *Note:* We use blue to represent "Moderate Bias" cases, and red to represent "Strong Bias" cases.

confounding.

First, we verify that the EHA estimators are biased and confidence intervals have poor coverages when the sequential ignorability is violated. Results are reported in Figure A4. As we expect from causal identification results, the EHA estimators discussed in the paper have large bias and confidence intervals fail to achieve 95% coverage rate. In the "Moderate Bias" case (i.e., $\beta_{UY} = 0.5$), the bias is about 30% of the true causal effect when considering the instantaneous causal effect ($s = 0$; the first column in Figure A4), and the bias is about 50% of the true causal effect when considering the long-term causal effect ($s = 3$; the second column in Figure A4). In the "Strong Bias" case (i.e., $\beta_{UY} = 1$), the bias is as large as 60% of the true causal effect when considering the instantaneous causal effect ($s = 0$; the first column in Figure A4), and the bias is about 80% of the true causal effect when considering the long-term causal effect ($s = 3$; the second column in Figure A4).

Finally, we consider whether proposed bounds from our sensitivity analysis can cover the true causal effect. Results are reported in Figure A5. For "Moderate Bias" cases, this simulation shows researchers have to use $\Gamma \geq 0.20$ to have 95% coverage rates. This aligns with our

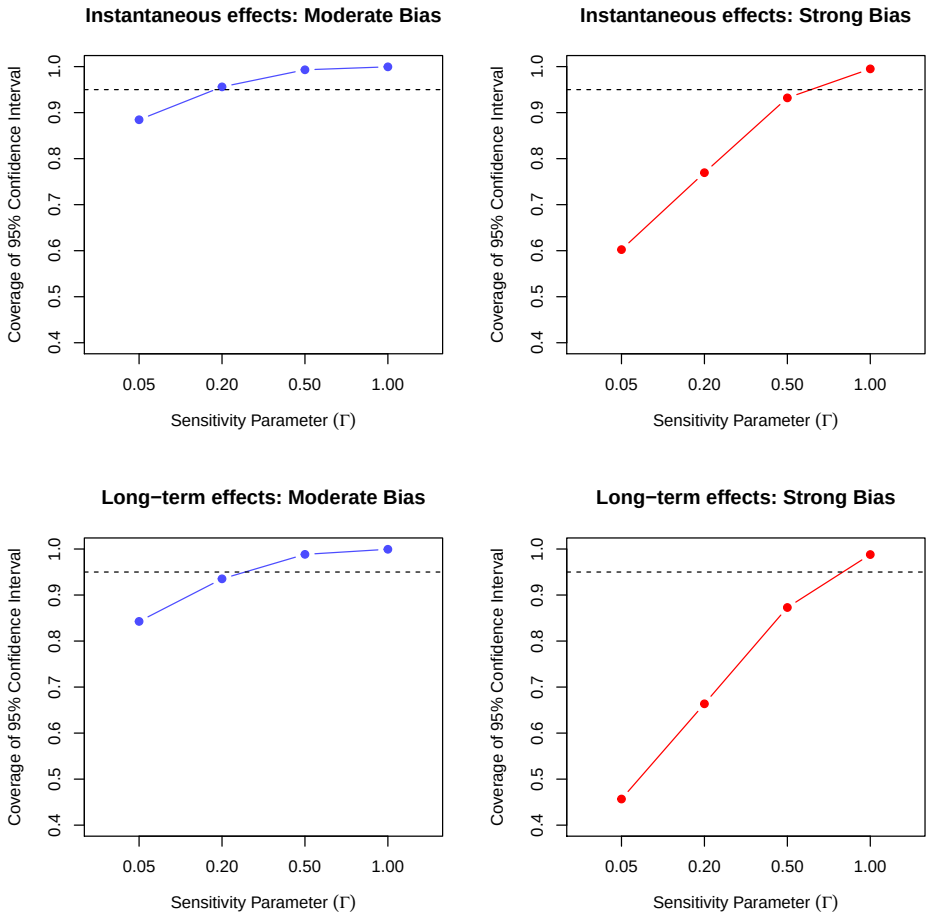

Figure A5: Coverage of Proposed Bounds from Sensitivity Analysis with different values of sensitivity parameters. *Note:* We use blue to represent "Moderate Bias" cases, and red to represent "Strong Bias" cases.

argument in the main text that we should in general be worried about moderate size bias, and $\Gamma = 0.20$ is a common level of confounding. This suggests that, when the estimated E-value is larger than 0.20, causal estimates are robust to "Moderate Bias" in this simulation.

To account for "Strong Bias" cases, this simulation shows researchers have to use $\Gamma \geq 0.50$ or even larger to have 95% coverage rates. When researchers are worried about a large amount of unmeasured confounding, they should take into account Γ as large as 0.50 and 1.00. This suggests that, in order to have the causal estimates robust to "Strong Bias" in this simulation, the estimated E-value needs to be as larger as 0.50 or even larger.

Importantly, this simulation suggests that our default sensitivity parameters $\Gamma = \{0.00, 0.05, 0.20, 0.50, 1.00\}$ in fact capture moderate ($30\% \sim 50\%$ of the true causal effect) to large bias ($60\% \sim 80\%$ of the true causal effect).

We want to re-emphasize here that the sensitivity analysis is not a panacea for unmeasured confounding — the fundamental challenge in observational causal inference. It is possible that the true unmeasured confounding might be larger than researchers expect. The central benefit of the sensitivity analysis is that researchers can transparently quantify the robustness of causal

findings to various scenarios of unmeasured confounding rather than focusing only on a binary question of whether there exists any unmeasured confounding or another binary question of whether estimates are robust to unmeasured confounding or not.

References

Balla, Steven J. 2001. "Interstate Professional Associations and The Diffusion of Policy Innovations." *American Politics Research* 29(3):221–245.

Blackwell, Matthew and Adam N Glynn. 2018. "How to Make Causal Inferences With Time-Series Cross-Sectional Data Under Selection On Observables." *American Political Science Review* 112(4):1067–1082.

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.

Fowler, James H, Michael T Heaney, David W Nickerson, John F Padgett and Betsy Sinclair. 2011. "Causality in Political Networks." *American Politics Research* 39(2):437–480.

Galton, Francis. 1889. "'Comments' on E.B. Tylor 'On a Method of Investigating the Development of Institutions: Applied to Laws of Marriage and Descent'." *Journal of the Royal Anthropological Institute* 18:245–256, 261–269.

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(03):673–701.

Greenhill, Brian, Layna Mosley and Aseem Prakash. 2009. "Trade-based Diffusion of Labor Rights: A Panel Study, 1986–2002." *American Political Science Review* 103(4):669–690.

Imai, Kosuke, Luke Keele and Dustin Tingley. 2010. "A General Approach to Causal Mediation Analysis." *Psychological Methods* 15(4):309–334.

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.

King, Gary, Michael Tomz and Jason Wittenberg. 2000. "Making the Most of Statistical Analyses: Improving Interpretation and Presentation." *American Journal of Political Science* 44(2):347–361.

Kreitzer, Rebecca J. 2015. "Politics and Morality in State Abortion Policy." *State Politics & Policy Quarterly* 15(1):41–66.

Manski, Charles F. 1993. "Identification of Endogenous Social Effects: The Reflection Problem." *The Review of Economic Studies* 60(3):531–542.

Ogburn, Elizabeth L. 2018. Challenges to Estimating Contagion Effects From Observational Data. In *Complex Spreading Phenomena in Social Systems*. Springer pp. 47–64.

Pang, Xun and Licheng Liu. 2020. "A Bayesian Method for Modeling Dynamic Network Influence With TSCS Data." *Presented at the 2020 Annual Meeting of the Society for Political Methodology* .

Pearl, Judea. 1995. "Causal Diagrams for Empirical Research." *Biometrika* 82(4):669–688.

Robins, James. 1986. "A New Approach To Causal Inference In Mortality Studies With A Sustained Exposure Period — Application to Control of the Healthy Worker Survivor Effect." *Mathematical modelling* 7(9-12):1393–1512.

Robins, James M, Miguel Angel Hernan and Babette Brumback. 2000. "Marginal Structural Models and Causal Inference in Epidemiology." *Epidemiology* 11(5):550–560.

Rosenbaum, Paul R and Donald B Rubin. 1983. "The Central Role of The Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70(1):41–55.

Shalizi, Cosma Rohilla and Andrew C Thomas. 2011. "Homophily and Contagion are Generically Confounded in Observational Social Network Studies." *Sociological Methods & Research* 40(2):211–239.

Simmons, Beth A and Zachary Elkins. 2004. "The Globalization of Liberalization: Policy Diffusion in the International Political Economy." *American Political Science Review* pp. 171–189.

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

Volden, Craig, Michael M Ting and Daniel P Carpenter. 2008. "A Formal Model of Learning and Policy Diffusion." *American Political Science Review* 102(3):319–332.