Understanding the Past: Statistical Analysis of Causal Attribution

Teppei Yamamoto   Massachusetts Institute of Technology

Would the third-wave democracies have been democratized without prior modernization? What proportion of the past militarized disputes between nondemocracies would have been prevented had those dyads been democratic? Although political scientists often ask these questions of causal attribution, existing quantitative methods fail to address them. This article proposes an alternative statistical methodology based on the widely accepted counterfactual framework of causal inference. The contribution of this article is threefold. First, it clarifies differences between causal attribution and causal effects by specifying the type of research questions to which each quantity is relevant. Second, it provides a clear resolution of the long-standing methodological debate on "selection on the dependent variable." Third, the article derives new nonparametric identification results, showing that the complier probability of causal attribution can be identified using an instrumental variable. The proposed framework is illustrated via empirical examples from three subfields of political science.

Did George W. Bush win the 2000 election because of the butterfly ballot? Would the third-wave democracies have been democratized without prior modernization? How many of the past international conflicts would have been prevented had there been democratic institutions? These types of questions often arise when political scientists try to make inferences about past events. These questions are *causal*, instead of associational, because they explicitly refer to a comparison between factual and counterfactual states of the same unit of analysis. These questions are about *attribution*, instead of effects, because their primary concern is the extent to which the actual occurrence of events can be attributable to a suspected cause.

The goal of this article is to propose a methodological framework for analyzing causal attribution. Political scientists are often interested in whether a past event can be causally attributed to another event or preexisting condition. In other words, they often ask how likely it is that an observed event would not have happened if hypothetically another event or condition had been absent. In the rest of this article, I provide a set of methodological tools that can be used to statistically analyze this type of question based on empirical data.

For example, the proposed procedure estimates the proportion of the past interstate conflicts between nondemocracies that would not have happened if those states had been democratic (see the empirical section below). Despite the fact that such a quantity is often the main target of inference in security studies, it cannot be consistently estimated with methods that are currently used in the discipline. The standard methods for causal inference are instead aimed at inferences about causal effects, which represent the magnitude of changes in the outcome of interest due to a hypothetical policy intervention. While estimation of causal effects is relevant to know the likely effect of a policy, causal attribution should be the target of inference when one is interested in explaining events that happened in the past. Table 1 summarizes key differences between these two causal concepts, on which I elaborate more in the next section.

To see what I mean by causal attribution in more specific terms, imagine a political scientist who is

TABLE 1 Comparison of Causal Attribution and Causal Effect

	Causal Attribution	Causal Effect
Unit-level Definition	$Y_i(0) = 0$ for $X_i = Y_i = 1$	$Y_i(1) - Y_i(0)$
Common Estimand	$p_A \equiv \Pr(Y_i(0) = 0 \mid X_i = Y_i = 1)$	$\tau \equiv \Pr(Y_i(1) = 1) - \Pr(Y_i(0) = 1)$
Interpretation	Proportion of observed y prevented by the hypothetical absence of x	Change in the probability of y induced by the hypothetical manipulation of x
Application	Explain why past events happened	Predict the impact of policy intervention

Note: The table summarizes key differences between the two causal concepts. See the following section for more explanation on each item.

interested in the relationship between the past assassinations of autocrats and the democratic transitions that ensued. She might then ask, "How likely is it that these regime changes were caused by those assassinations?" Her immediate goal is not to estimate the effect of assassinating political leaders on the probability of democratic transition in general, but to provide a causal explanation to the specific democratic transitions that happened in the past. Given the fact that these transitions occurred after the assassinations of leaders, can she conclude that the former can be causally attributed to the latter? Although this example is purposely narrow and specific to convey the key intuition, the type of causal question being asked here is widespread in political science. In fact, as Fearon points out, political scientists "routinely evaluate causal hypotheses by discussing or simply referring to *counterfactual cases* in which a hypothesized causal factor is supposed to have been absent" (1991, 169, emphasis in original). In addition, as discussed below, the interest in explaining actual cases that happened in the past is often a primary motivation for "selection on the dependent variable," a controversial methodological practice (e.g., King, Keohane, and Verba 1994; Pape 2008).

This article contributes to the causal inference literature by deriving for the first time the exact conditions under which causal attribution can be identified with the aid of an instrumental variable. In the previous studies, it has been shown that the probability of causal attribution (defined formally in the next section) can be fully identified under the assumptions of exogeneity and monotone treatment effect (Pearl 1999; Rosenbaum 2001). However, these assumptions are often too strong to justify in observational studies, where researchers are especially likely to be interested in causal attribution. In the analytical section, I show that when the instrumental variable satisfies a standard set of properties, the probability of causal attribution can be identified for a meaningful subset of the population even without the exogenous treatment assumption.

In the empirical section, I apply the proposed framework to three problems in political science. First, I examine a recent study of the assassination of political leaders (Jones and Olken 2009) to estimate what proportions of the past regime changes can be attributed to the assassination of leaders. Next, I revisit the "democratic peace" debate in international relations and estimate the proportion of the past militarized interstate disputes that could have been prevented by democratic institutions (Oneal and Russett 2001). Finally, I use the variation in compulsory schooling laws across states as an instrument to estimate the extent to which the turnout of educated voters can be causally attributed to their education in each U.S. national election between 1978 and 1996 (Milligan, Moretti, and Oreopoulos 2004).

Definitions and Conceptual Issues

Causal inference is an important goal of social science. Scholars in various disciplines seek to uncover causal relationships underlying social phenomena. Although more basic, descriptive inferences are often of immediate concern in empirical research, political science theories usually go beyond pure description and make statements about how social entities and phenomena are causally related with each other. Some may also argue that prediction is the ultimate goal of any scientific discipline, but good predictive inference typically requires sufficient understanding of underlying causal relationships.

However, the meaning of something being "causal" is often not made explicit in applied studies. For example, what does it exactly mean that x causes y? Is it different from x having a positive causal effect on y? Indeed, there are numerous types of causal questions one can ask about a given set of variables, and one must choose appropriate methods to analyze different types of causal hypotheses. Below, I focus on one of these possible causal hypotheses which has been substantively important, but methodologically almost neglected, in the field of political science.

Quantity of Interest

A causal hypothesis in political science often takes the following simple form: "*x* was a cause of *y*," or equivalently, "*x* caused *y*." This type of question typically arises when a researcher seeks to explain the reason a past event happened. For example, in comparative politics, one long-standing question is why some third-world countries successfully became democracies after World War II while many others did not. To answer this question, some argue that a sufficient level of economic development was a major cause of democratization (e.g., Lipset 1959; Przeworski et al. 2000), while others maintain that cultural and religious changes played a more important causal role in some cases (Huntington 1991). In American politics, scholars are interested in whether a particular irregularity in the 2000 presidential election, such as a butterfly ballot (Wand et al. 2001) or miscounted overseas absentee ballots (Imai and King 2004), caused the victory of George W. Bush.

But what does this statement *exactly* mean? As demonstrated by the long history of philosophical literature (e.g., Hempel 1965; Hume 1739; Mackie 1965; Mill 1843), a cause can be defined in various ways. However, one widely accepted definition is the one based on counterfactuals, which has origins in both analytic philosophy (Lewis 2001) and statistics (Neyman 1923). Under this framework, causality is defined via counterfactual events, or *what would have happened if a certain event had been absent*. In recent years, this definition has become increasingly popular among both statisticians and applied researchers. Indeed, the introduction of counterfactuals to statistics has spawned various methods for both randomized experiments (e.g., Fisher 1971) and observational data (e.g., Rosenbaum 2002) to discover causal relationships. The empirical applications of these methods are too numerous to be listed here. Moreover, the advantages of the counterfactual account of causation are also widely recognized in the philosophical literature (e.g., Halpern and Pearl, 2005a, 2005b; Hitchcock and Woodward 2003; Menzies and Price 1993; Price 1991; Woodward 2003).

From the counterfactual viewpoint, *x* is said to have caused *y* when *y* would not have happened had *x* been absent. That is, an event was a cause of another event when the absence of the former would also have implied the absence of the latter. In fact, this is often precisely the way many social scientists conceptualize a causal relationship, either implicitly or explicitly. For example, Huntington states that "*if it were not for* the changes within the Catholic Church and the resulting actions of the Church against authoritarianism, fewer third wave transitions to democracy *would have occurred* and many

that did occur *would have occurred* later" (1991, 85, emphasis added). Similarly, Wand et al. summarize their main conclusion in the following way: "*Had* [Palm Beach County] used a ballot format in the presidential race that did not lead to systematic biased voting errors, our findings suggest that, other things equal, Al Gore *would have won* a majority of the officially certified votes in Florida" (2001, 803, emphasis added). Indeed, as Fearon (1991) pointed out, political scientists routinely engage in this type of counterfactual reasoning when they make causal claims about past phenomena.[1]

To formalize this notion of causality, I use the notation based on the potential outcomes model of causal inference (Holland 1986; Neyman 1923; Rubin 1974). Let $X_i \in \{0, 1\}$ be a binary variable representing either the presence ($X_i = 1$) or absence ($X_i = 0$) of a condition or an event for unit *i*. This variable will also be referred to as the *treatment*, following the convention. Next, I consider the following two binary variables denoting the potential outcomes for each unit: $Y_i(1) \in \{0, 1\}$, which represents whether the outcome, or the event of interest, would have occurred ($Y_i(1) = 1$) or would not have occurred ($Y_i(1) = 0$) with the presence of condition *x*, and $Y_i(0) \in \{0, 1\}$, which denotes the same event without the presence of *x*.[2] Then, based on this notation, it can be said that for unit *i* who actually experienced both *x* and *y* (i.e., $X_i = Y_i = 1$), *x* was a cause of *y* if and only if $Y_i(0) = 0$. It is important to note that the presence of *x* here is a necessary condition for the occurrence of *y* for *this particular unit i*, holding other preexisting conditions constant.[3]

Although identifying this type of cause for particular units is often the ultimate goal of empirical research, it is usually not possible without making very strong assumptions. This is primarily because, by definition, one can only observe the actual outcome, $Y_i = Y_i(X_i)$, for each unit and never observe the conterfactual outcome, $Y_i(1 - X_i)$. Thus, instead of seeking to provide a causal explanation for a particular unit, it is necessary to modify the causal question and focus on an alternative quantity of

---

[1] More precisely, Fearon made this remark in relation to small-N studies in comparative politics and international relations (1991, 169). However, the example just cited is a large-N quantitative study in American politics.

[2] This notation implicitly makes the stable unit treatment value assumption (SUTVA), which states that unit *i*'s outcome is only affected by the treatment of *i* itself and not by any other unit's treatment status (Cox 1958).

[3] In particular, this statement does not make any reference to other counterfactual worlds where conditions other than *x* were also different. In other words, it does not preclude the possibility that there exist some configurations of preexisting conditions that do not involve *x* but would still have led to *y*.

interest. Particularly natural is to examine the *probability of causal attribution* defined as

$$p_A = \Pr(Y_i(0) = 0 \mid X_i = 1, Y_i = 1), \qquad (1)$$

which was first introduced by Pearl (1999) and also studied by Rosenbaum (2001) in a different context.[4] Essentially, this quantity is a probabilistic version of the previous statement. That is, instead of asking *whether y* would not have happened had *x* not been present, this probability quantifies *how likely y* would not have occurred without *x* in the case where both *x* and *y* were actually present. One particularly intuitive interpretation of this quantity is to think of it as the expected proportion of the units for whom *x* was a cause of *y* among the subpopulation with $X_i = Y_i = 1$. In other words, p_A measures the proportion of the observed events that can be causally attributed to the hypothesized cause.

The rest of this article is focused on the question of how we can make valid inferences about the probability of causal attribution, as defined in equation (1), using observed data. In principle, the identification of p_A faces the same problem as that of a unit-level causation since the former quantity still refers to the counterfactual outcome. However, it can be shown that this probability can be identified by observed data under less restrictive sets of assumptions, as I discuss in the analytical section below.

Causal Attribution vs. Causal Effects

Before analyzing the problem of causal attribution, it is worthwhile to clarify the differences between this concept and *causal effects*. Attribution and effects are quite distinct causal concepts, though they are of course closely related[5] (Table 1). While causal attribution is defined at the unit level by whether the event *y* would still have occurred in the hypothetical absence of *x*, a causal effect is typically defined by the change in the value of Y_i induced by the hypothetical manipulation of X_i. That is, *x* is said to have a causal effect on *y* for unit *i* when $Y_i(1) - Y_i(0) \neq 0$. As further explained below, an important difference between

the causal effect and causal attribution is that the former is defined for any unit in the population regardless of the actual value of Y_i, whereas the latter can be conceptualized only for those who actually experienced both *x* and *y*, i.e., $X_i = Y_i = 1$.

Although political scientists often posit causal hypotheses in the form of causal attribution, they routinely use quantitative methods that do not have direct correspondence to these hypotheses. In particular, a vast majority of standard tools for causal inference are for the estimation of causal effects, instead of causal attribution. For example, a typical causal estimand in both experimental and observational studies is the average treatment effect (ATE) of *x* on outcome *y*, which is defined as

$$\begin{aligned}
\tau &= \mathbb{E}(Y_i(1) - Y_i(0)) \\
&= \Pr(Y_i(1) = 1) - \Pr(Y_i(0) = 1), \qquad (2)
\end{aligned}$$

where the second equality follows because Y_i is binary. This quantity represents the mean difference in the outcome variable under two scenarios, one of which is counterfactual for a given unit *i*. That is, it compares the value of the outcome variable that would be observed when the treatment took the value of 1 with the value of the outcome when X_i was fixed at 0. This difference represents a causal effect of X_i on Y_i because it measures the magnitude of the change in Y_i that would be induced by a hypothetical manipulation of X_i averaged over all units in the population. A closely related causal quantity is the average treatment effect on the treated (ATT), $\mathbb{E}(Y_i(1) - Y_i(0) \mid X_i = 1)$. By the same logic, this represents the causal effect of X_i on Y_i among those units who are actually assigned $X_i = 1$.

It is well known that the standard tools of causal inference, such as randomized experiments and covariate adjustments for observational data, are often well suited to estimating these causal effects. For example, the difference-in-means estimator is unbiased for the ATE in a randomized experiment (Rubin 1974), and in an observational study one can consistently estimate the ATE via a linear regression of Y_i on X_i and a set of pretreatment covariates when the strong ignorability assumption is satisfied (Freedman 2008; Rubin 1977). Moreover, the ATT is typically the estimand of a matching estimator, such as propensity score matching (Rosenbaum and Rubin 1983). Thus, when one uses any of these standard techniques to draw a causal conclusion, it is causal effects that are being estimated.

The differences between causal attribution and effects are not only mathematical or semantic, but also substantive and practical. The probability of causal attribution is defined as the conditional probability of the absence of the outcome in the hypothetical absence of the

---

[4] Pearl (1999) called this quantity the "probability of necessity," while Rosenbaum (2001) used the name of the "attributable effect." Prior to these two studies, a slightly different version of the same quantity was named the "probability of causation" by Robins and Greenland (1989). Here I use this terminology in order to emphasize the causal nature of the quantity as well as its close connection to the concept of attribution.

[5] For example, consider the alternative causal quantity $\mathbb{E}(Y_i(1) - Y_i(0) \mid X_i = Y_i = 1)$, which we may call the average treatment effect on the treated and successful. The probability of causal attribution is equal to one minus the second term in this expression. I exploit such mathematical relations in my nonparametric identification analysis below.

treatment ($Y_i(0) = 0$) given the actual presence of both ($X_i = Y_i = 1$). The target of inference here is thus the subpopulation to which the outcome of interest in fact occurred. This implies that the probability of causal attribution is of most interest when this subpopulation is clearly identified, i.e., when the event y already happened in the past. Thus, causal attribution is of particular relevance when one's interest is in explaining past phenomena.

In contrast, estimating the ATE corresponds to estimating the magnitude of the impact of a hypothetical intervention to the population, regardless of the actual outcome that would result without the intervention. Such an inference is most meaningful when one can think of implementing this type of intervention as a real possibility, as opposed to a pure thought experiment. One such example in social science is the use of a natural experiment or randomized field experiment when a researcher seeks to estimate the likely effect of implementing a particular policy, such as an increase in the minimum wage (Card and Krueger 1994) and the use of telephone calls as opposed to mailing in get-out-the-vote campaigns (Gerber and Green 2000). Thus, while causal attribution is useful for understanding the past, causal effects have more relevance to questions which involve the possibility of future interventions.

Implication for "Selection on the Dependent Variable"

The discussion in the previous paragraphs sheds light on the important methodological debate regarding whether one should "select on the dependent variable." When one's interest is in explaining past events, a common practice is to study only cases where these events actually happened. This type of research design is particularly prevalent in qualitative comparative studies (e.g., Porter 1998) but can also be found in large-n studies (e.g., Pape 2003). The current wisdom of political methodology strongly recommends against such a practice. Because studying only positive cases represents an extreme example of selection bias, the estimates of causal effects based on such samples will be seriously biased (Ashworth et al. 2008; Geddes 1990; King, Keohane, and Verba 1994). However, this advice is often received with bewilderment by applied researchers to whom the only "universe of a phenomenon" that is relevant is positive cases (Pape 2008).

From the proposed perspective, the motivation for selecting on the dependent variable stems from the researcher's interest in causal attribution. When one is interested in explaining why past events happened, the goal is to make an inference about whether these particular events would not have happened in the absence of the hypothesized cause. Focusing on these positive cases thus appears to be reasonable, as some researchers have argued (Collier, Mahoney, and Seawright 2004; Pape 2008). However, while it is perfectly reasonable to set one's *population of interest* based on the dependent variable, it does not imply that one can also *sample* based on the dependent variable. As will become clear in the next section, a representative sample from the entire population is still necessary even when one's direct interest is in the population of positive cases, because the negative cases must be used to make inferences about counterfactual outcomes of the positive cases. Thus, the standard recommendation against selection on the dependent variable is also valid for estimating causal attribution, even though the original argument was made for inferences about causal effects. This highlights the importance of making a clear distinction between estimands (i.e., quantities being estimated) and estimators (i.e., statistics used for estimation) in causal inference.

Analytical Results

The potential outcomes model of causal inference provides a useful methodological framework for analyzing causal attribution. However, this counterfactual nature of the proposed approach also leads to an important challenge. Under the current binary setup, a researcher can only identify either $Y_i(0)$ or $Y_i(1)$ as a realization of the potential outcome from the observed data for each i. Because of this "fundamental problem of causal inference" (Holland 1986), it is generally impossible to directly compute a unit-level causal quantity (e.g., the first row of Table 1) from the observed information.

Thus, an important question is how much we can learn about a causal quantity of interest from the observed data alone before making any statistical or functional form assumptions about the underlying causal process. Nonparametric identification analysis, as advocated by Manski (1995, 2007), seeks to answer this question by formally deriving the bounds within which the quantity of interest is guaranteed to be located.[6] By deriving these nonparametric bounds under various sets of assumptions, the identification analysis can formally characterize the role of each assumption in the inferential process. In this section, I use this approach to explore the sets of

---

[6] Recent applications of this approach in political science include Hanmer (2007), Ashworth et al. (2008), and Imai and Yamamoto (2010).

assumptions which allow for meaningful inferences about the probability of causal attribution.

Identification and Inference under Exogeneity

Before presenting the main analytical results of this article, I briefly review the existing identification results by Pearl (1999), which, among others, specify a set of assumptions for identifying the probability of causal attribution. The key result is that the same assumption that is sufficient for the identification of the ATE also *partially* identifies the probability of causal attribution, although an additional assumption is necessary for its full identification.

It is widely known that the ATE given in equation (2) and related causal effects (e.g., ATT) can be fully identified when the following assumption is satisfied.

Assumption 1 (Exogeneity)

$$\{Y_i(1), Y_i(0)\} \perp\!\!\!\perp X_i \mid W_i,$$

where W_i represents the set of confounders that affect both the treatment and outcome. Hereafter, conditioning on W_i is omitted for the sake of notational simplicity unless it needs to be made explicit. Assumption 1 is also known under various alternative names (e.g., ignorability, unconfoundedness, selection on the observables, no omitted variable bias). For our purpose, however, the question is whether this assumption is also sufficient to identify the probability of causal attribution. In a seminal article, Pearl (1999) shows that under Assumption 1 the probability of causal attribution can be identified only up to the following sharp (i.e., tightest possible) bounds.

Theorem 3.1 (Pearl, 1999). *Under Assumption 1, the probability of causal attribution has the following sharp upper and lower bounds.*

$$\max\left\{0, \frac{\Pr(Y_i = 1 \mid X_i = 1) - \Pr(Y_i = 1 \mid X_i = 0)}{\Pr(Y_i = 1 \mid X_i = 1)}\right\}$$

$$\leq p_A \leq \min\left\{1, \frac{\Pr(Y_i = 0 \mid X_i = 0)}{\Pr(Y_i = 1 \mid X_i = 1)}\right\}. \quad (3)$$

The bounds may or may not be informative (i.e., tighter than the original bounds, $[0, 1]$), depending on the strength of association between X_i and Y_i. One important observation is that the lower bound is informative (i.e., greater than zero) if and only if the ATE is also positive. This implies that when X_i has a positive causal effect on Y_i, there are at least some units in the population for whom the observed y can be causally attributed to the existence of x. In addition, the expression also makes it

clear that the standard estimate of the ATE is always less than the lower bound, revealing the danger of makings inferences about causal attribution based on empirical findings on causal effects.

Relatedly, Tian and Pearl (2000) derive sharp bounds on p_A for the situation where Assumption 1 does not hold in the population, but one can also obtain experimental data by randomly assigning X_i and combine them with the observational data. They show that the resulting bounds with the combined data are tighter than those with the experimental data alone. Researchers should thus use their bounds instead of those in equation (3) if such data are available. In practice, however, only in rare circumstances can one randomly manipulate the causal variable of interest, as well as obtaining observational data on the same population.[7] This article therefore focuses on the more realistic bounds in Theorem 3.1 where researchers must rely on control variables W_i to ensure the exogeneity of X_i.

To draw stronger conclusions about causal attribution, one needs to impose more stringent assumptions about the causal relationship between x and y. Pearl (1999) considers the assumption that the effect of x on y is nonnegative for every unit in the population, which can be stated as follows.

Assumption 2 (Monotone Treatment Effect)

$$Y_i(0) \leq Y_i(1) \quad \text{for all } i.$$

He then proves that Assumptions 1 and 2 are sufficient to identify the probability of causal attribution. The resulting expression for p_A is given in the following theorem.

Theorem 3.2 (Pearl, 1999). *Under Assumptions 1 and 2, p_A is identifiable and given by*

$$p_A = \frac{\Pr(Y_i = 1 \mid X_i = 1) - \Pr(Y_i = 1 \mid X_i = 0)}{\Pr(Y_i = 1 \mid X_i = 1)}. \quad (4)$$

Rosenbaum (2001) also derives the same expression under the same assumptions. This result implies that in order to fully identify the probability of causal attribution, one must make the rather strong assumption that x can only affect y in one direction for every unit, in addition to the exogeneity assumption. Thus, causal attribution is more difficult to identify than causal effects.

Nevertheless, it is noteworthy that the quantity in equation (4) coincides with the lower bound of

---

[7] Another possible situation is a natural experiment where the causal variable of interest is random in the observed data without manipulation (see the first example in the empirical section). Not surprisingly, the bounds in equation (3) and those of Tian and Pearl (2000) are identical in that case.

Theorem 3.1. This implies that, even when Assumption 2 is not entirely plausible, the estimated probability of causal attribution based on equation (4) is never greater than the true value of p_A (provided, of course, that Assumption 1 is satisfied) after taking into account sampling variability. This result is useful when one is specifically interested in the least possible proportion of the units for which the (hypothetical) absence of *x* would have prevented the (actual) occurrence of *y*. One example of such a situation is when one's aim is to test whether *x* was a necessary cause of *y* for the entire population, i.e., $p_A = 1$.

Estimation and statistical inference. To apply these identification results to empirical data, it is necessary to estimate the population quantities like equations (3) and (4) from a random sample. For empty W_i, estimation of these quantities is straightforward; one can first estimate the conditional distributions of Y_i given $X_i = x$ for $x = 0$, 1 using their sample analogues, and then substitute these estimates to obtain a consistent estimate of the whole expression.

When W_i is nonempty, the population bounds on p_A can be obtained by adding the conditioning on W_i to all the probabilities in each expression and taking the population average of the entire expression over the distribution of W_i.[8] Then, to estimate these bounds using a random sample, one can in principle compute the bounds for each stratum defined by W_i and then take the weighted average of these values across the strata. In practice, however, this fully nonparametric subclassification approach is often infeasible because W_i is high-dimensional and/or W_i contains continuous or near-continuous variables. Thus, one must typically rely on some semi-parametric or fully parametric estimation techniques. Fortunately, the expressions in Theorems 3.1 and 3.2 have a form particularly suited for such estimation procedures. Since all the components are conditional probabilities of Y_i given X_i and W_i, they can all be estimated by standard regression models, such as logit, probit, or linear probability models. It is also straightforward to use other balancing techniques, such as propensity score matching, subclassification, or weighting procedures. The empirical section of this article provides some examples.

The estimation of statistical uncertainty is slightly more complex. When the probability of causal attribution is point identified under Assumptions 1 and 2 and given by equation (4), one can estimate its variance via standard

---

[8] For example, the lower bound in equation (3) becomes $\max\{0, 1 - \int \Pr(Y_i = 1 \mid X_i = 0, W_i = w)/\Pr(Y_i = 1 \mid X_i = 1, W_i = w)dF_{W_i}(w)\}$ where $F_{W_i}(\cdot)$ represents the distribution function of W_i.

procedures, such as the Delta method and simulation-based approaches. For example, Cai and Kuroki (2006) derive an approximate variance of equation (4) based on the Delta method for the case where W_i is empty. In the empirical applications, I use confidence intervals based on bootstrap standard errors.

For the sharp bounds under Assumption 1 (i.e., equation 3), the computation of confidence intervals is more complicated because the true value of p_A can be located anywhere within the interval. The literature divides into two approaches, depending on whether to focus on the confidence regions that cover the entire bounds (e.g., Chernozhukov, Hong, and Tamer 2007; Imai and Soneji 2007) or those that contain the true value with a certain confidence level (see Tamer 2010, for a recent review). In the empirical applications in this article, I use the symmetric and uniformly convergent confidence interval by Imbens and Manski (2004).

Identification and Inference with an Instrumental Variable

The findings of Pearl (1999) show that the probability of causal attribution is identifiable under the assumptions of exogeneity (Assumption 1) and monotone treatment effect (Assumption 2). A natural question, then, is whether these assumptions are plausible in applied studies in political science. Unfortunately, both assumptions are often dubious in situations where researchers are interested in causal attribution. First, Assumption 1 requires either that the condition *x* be effectively random (in which case the conditioning set W_i would be empty) or that there be no omitted variable that affects both *x* and the outcome *y* (i.e., W_i includes all such variables). Since in typical observational studies *x* cannot be viewed as randomly assigned (but see the first empirical example, which is based on a natural experiment), and it is difficult, if not impossible, to collect sufficient information about suspected confounders, the exogeneity assumption is likely to be violated in many situations. Second, Assumption 2 requires that the researcher a priori knows that *x* can never affect *y* negatively for every unit in the population. Although this assumption may be reasonable in some situations, many applied researchers are likely to be uncomfortable about such a strong statement, particularly because the quantity of interest itself concerns the causal relationship between *x* and *y*.

Thus, an important question is whether one can make any progress without making these rather strong assumptions. In this section, I extend the findings by Pearl (1999) to the important case where exogene-

ity does not hold but there is an instrumental variable which satisfies a standard set of assumptions. In many fields of social science, using an instrumental variable has become a common identification strategy in observational studies where the exogeneity assumption is likely to be violated (Sovey and Green 2011). Here, I consider the three assumptions about the nature of a binary instrumental variable (IV), Z_i, that were originally introduced by Angrist, Imbens, and Rubin (1996). They prove that these assumptions have to be made in order to interpret standard IV estimators, such as the two-stage least squares (2SLS) estimator, as estimating causal effects (Angrist and Imbens 1995; Imbens and Angrist 1994). Below I show that these standard IV estimators can also be used to estimate causal attribution under the same set of identifying assumptions.[9]

Assumptions.   Following Angrist, Imbens, and Rubin (1996), I first introduce additional notation to incorporate an instrumental variable into our statistical framework. Let $X_i(z)$ be the potential treatment status for unit i when the unit receives z as the value of the instrument. Similarly, we denote by $Y_i(z, x)$ the potential outcome for unit i when the unit receives x and z for the treatment and instrumental variable. Then the observed value of the treatment and outcome can be written as $X_i = X_i(Z_i)$ and $Y_i = Y_i(Z_i, X_i(Z_i))$, respectively. By construction, Z_i is a variable that causally precedes both X_i and Y_i and potentially has causal effects on these variables.

With this notational framework, I can now state the three assumptions. First, I assume that the assignment of Z_i is exogenous with respect to the potential outcomes and the potential treatment status, possibly after controlling for observable pretreatment confounders. Formally, this assumption can be stated as

Assumption 3 (Exogeneity of the Instrument)

$$\{Y_i(0, X_i(0)), Y_i(1, X_i(1)), X_i(0), X_i(1)\} \perp\!\!\!\perp Z_i \mid W_i = w, \tag{5}$$

where it is also assumed that $\Pr(Z_i = z \mid W_i = w) > 0$ for all $z \in \{0, 1\}$ and $w \in \mathcal{W}$.

Second, I assume that Z_i can affect Y_i only through its effect on X_i. Formally, we write this assumption as

Assumption 4 (Exclusion Restriction)

$$Y_i(1, x) = Y_i(0, x) \text{ for any } x \in \{0, 1\} \text{ and } i. \tag{6}$$

This assumption says that the potential outcomes are invariant to changes in the value of Z_i as long as the treat-

[9] Of course, the plausibility of these assumptions also needs to be carefully evaluated in specific empirical contexts (Sovey and Green 2011). See the end of the empirical section for a discussion of alternative identification strategies.

ment X_i stays at the same level. Under this assumption, the potential outcome no longer directly depends on the value of Z_i and thus can be rewritten simply as $Y_i(x)$, $x \in \{0, 1\}$.

Finally, I assume that the instrument cannot affect the treatment status in the negative direction. Formally,

Assumption 5 (Monotone Instrument Effect on Treatment)

$$X_i(0) \leq X_i(1) \quad \text{for all } i. \tag{7}$$

Under this assumption, the instrument can only encourage the units to receive the treatment. Below I examine the consequences of these three assumptions for the identification of causal attribution.

Identification for compliers.   Under the current binary setup, the units can be categorized into four distinct causal types, or *principal strata* (Frangakis and Rubin 2002), based on the potential response of their treatment status to the instrumental variable. First, there are units whose treatment values are positively affected by the instrumental variable, i.e., $X_i(1) = 1$ and $X_i(0) = 0$, or more concisely $X_i(z) = z$ for $z \in \{0, 1\}$. These units are called *compliers* (Angrist, Imbens, and Rubin 1996). Second, there are units who would always be in the treatment condition regardless of the value of the instrument ($X_i(1) = X_i(0) = 1$) and also units who would always be in the control condition ($X_i(1) = X_i(0) = 0$). These types are called *always-takers* and *never-takers*, respectively. Finally, there are units, called *defiers*, whose treatment is negatively affected by the instrument ($X_i(1) = 0$ and $X_i(0) = 1$). By Assumption 5, there exists no defier in the case currently being considered.

Angrist, Imbens, and Rubin (1996) show that under Assumptions 3, 4, and 5, the average causal effect on compliers can be identified from the observed information. Here I focus on an analogous quantity of interest, the *complier probability of causal attribution*, which is defined as follows:

$$p_C = \Pr(Y_i(0) = 0 \mid X_i = 1, Y_i = 1, X_i(z) = z). \tag{8}$$

This quantity differs from the probability of causal attribution only in that it additionally conditions on the fact that unit i is a complier. Thus, p_C represents the proportion of the compliers to whom the hypothetical presence of x would have prevented the outcome y from happening. This implies that the conclusions based on this quantity may not generalize to the entire population because the compliers may be systematically different from the rest of the population with respect to their response to the treatment. Although it is an important limitation of the proposed methodology, the lack of external validity in

exchange for internal validity is a common feature of any identification strategy based on an instrumental variable.

It can be shown that under Assumptions 3, 4, and 5, the complier probability of causal attribution can be partially identified and the sharp bounds can be expressed in a closed form. The result is summarized in the following proposition.

Proposition 1 (Sharp Bounds on p_C with an Instrumental Variable). *Under Assumptions 3, 4, and 5, the sharp bounds on the complier probability of causal attribution can be given as*

$$\max\left\{0, \frac{\Pr(Y_i = 1 \mid Z_i = 1) - \Pr(Y_i = 1 \mid Z_i = 0)}{\Pr(Y_i = 1 \mid X_i = 1)\{\Pr(X_i = 1 \mid Z_i = 1) - \Pr(X_i = 1 \mid Z_i = 0)\}}\right\}$$
$$\leq p_C \leq \min\left\{1, \frac{\Pr(Y_i = 0 \mid X_i = 0)}{\Pr(Y_i = 1 \mid X_i = 1)}\right\}. \tag{9}$$

A proof is given in Appendix A.1. Proposition 1 shows that if an instrumental variable is available, it is possible to make a valid inference about the complier probability of causal attribution without either Assumption 1 or Assumption 2, although the probability is only identified up to the large-sample bounds given in equation (9). Analogously to Theorem 3.1, the lower bound is informative if and only if the complier average treatment effect (Angrist, Imbens, and Rubin 1996) is greater than zero, implying that a positive causal effect translates into some degree of causal attribution. The upper bound is identical to that of Theorem 3.1, even though Assumption 1 is now violated. Although Proposition 1 shows that the point identification is not possible without making additional assumptions, these bounds are often narrow enough to draw substantively interesting conclusions about causal attribution; the last empirical application of this article provides such an example.

Next, I consider the situation where in addition to Assumptions 3, 4, and 5 one can assume that the effect of the treatment on the outcome is also nonnegative, i.e., Assumption 2. The consequence of this additional assumption is substantial; in fact, one can now *fully* identify the complier probability of causal attribution. The result is given in the following proposition.

Proposition 2 (Identification of p_C with an Instrumental Variable). *Under Assumptions 2, 3, 4, and 5, the complier probability of causal attribution defined in equation (8) can be point identified and given by*

A proof is provided in Appendix A.2. Proposition 2 states that if one is willing to make the additional assumption of monotone treatment effect, i.e., Assumption 2, the proportion of the compliers for whom the outcome can be causally attributed to the treatment is point identified and thus can be consistently estimated from observed information. Here again, an analogy can be made between this proposition and Theorem 3.2; in both cases, adding Assumption 2 to the set of assumptions that identifies average causal effects is sufficient to further identify the probabilities of causal attribution.

Another interesting connection with earlier results is that p_C here is identified as the complier average treatment effect (Angrist, Imbens, and Rubin 1996) divided by $\Pr(Y_i = 1 \mid X_i = 1)$, whereas p_A in Theorem 3.2 is identified as the ATE divided by the same probability. Finally, equation (10) coincides with the sharp lower bound given in equation (9). This implies that even if Assumption 2 is in fact violated, the estimate based on Proposition 2 will *never overestimate* the true complier probability of causal attribution. Thus, by reporting the point estimates on the basis of equation (10), one is always conservative about the degree of causal attribution.

Identification for the population. So far, my discussion has focused on the group of compliers, a particular subset of the population from which an observed sample is drawn. As mentioned earlier, this implies that the substantive conclusions based on Propositions 1 and 2 may not be generally applicable for the entire population. Thus, an important question is whether one can learn anything about causal attribution for the population as a whole when an instrumental variable is available.

To address this question, I derive the nonparametric bounds for p_A, the original probability of causal attribution given in equation (1), under the same two sets of assumptions as Propositions 1 and 2. The results are summarized in the following propositions.

Proposition 3 (Sharp Bounds on p_A with an Instrumental Variable). *Under Assumptions 3, 4, and 5,*

$$p_C = \frac{\Pr(Y_i = 1 \mid Z_i = 1) - \Pr(Y_i = 1 \mid Z_i = 0)}{\Pr(Y_i = 1 \mid X_i = 1)\{\Pr(X_i = 1 \mid Z_i = 1) - \Pr(X_i = 1 \mid Z_i = 0)\}}. \tag{10}$$

the identification region of the probability of causal attribution is characterized by the following sharp bounds:

$$\max\left\{0, \frac{\{\Pr(Y_i = 1 \mid Z_i = 1) - \Pr(Y_i = 1 \mid Z_i = 0)\}\Pr(Z_i = 1)}{\Pr(Y_i = 1 \mid X_i = 1)\left\{\sum_{z=0}^{1}\Pr(X_i = 1 \mid Z_i = z)\Pr(Z_i = z)\right\}}\right\} \le p_A \le 1. \qquad (11)$$

Proposition 4 (Sharp Bounds on p_A with an Instrument and Monotonicity). *Under Assumptions 2, 3, 4, and 5, the identification region of the probability of causal attribution is characterized by the following sharp bounds:*

$$\frac{\{\Pr(Y_i = 1 \mid Z_i = 1) - \Pr(Y_i = 1 \mid Z_i = 0)\}\Pr(Z_i = 1)}{\Pr(Y_i = 1 \mid X_i = 1)\left\{\sum_{z=0}^{1}\Pr(X_i = 1 \mid Z_i = z)\Pr(Z_i = z)\right\}} \le p_A \le$$

$$\min\left\{1, \frac{\{\Pr(Y_i = 1 \mid Z_i = 1) - \Pr(Y_i = 1 \mid Z_i = 0)\}\Pr(Z_i = 1) + \Pr(Y_i = 1 \mid X_i = 1)\Pr(X_i = 1 \mid Z_i = 0)}{\Pr(Y_i = 1 \mid X_i = 1)\left\{\sum_{z=0}^{1}\Pr(X_i = 1 \mid Z_i = z)\Pr(Z_i = z)\right\}}\right\}. \qquad (12)$$

These two propositions formally characterize the identification regions of the probability of causal attribution under the alternative sets of assumptions. Proofs are given in Appendix A.3. Proposition 3 shows that observed information is completely uninformative about the upper bound of p_A even when there is an instrumental variable; we can never exclude the possibility that the hypothetical absence of x would have prevented the outcome for all units with $X_i = Y_i = 1$. On the other hand, the lower bound may or may not be informative. Proposition 4 shows that the lower bound of p_A is always informative (i.e., greater than zero) when Assumption 2 is additionally made, whereas the upper bound may or may not be less than one.

In practice, these bounds are often too wide to be meaningful in substantive terms. The empirical study in the next section provides such an example. Nevertheless, even this kind of limited information about the probability of causal attribution may be useful for answering particular research questions. For example, suppose one's interest is in testing whether x was a *necessary* cause of y, i.e., whether y would *never* have happened had x been absent. Then it suffices to know that p_A is statistically indistinguishable from one, and thus the information about its upper bound based on equation (12) may prove to be enough to reject this hypothesis. Here again, specifying one's causal hypothesis of interest is essential for causal inference.

Estimation and statistical inference. To estimate these large sample bounds from an observed sample, one can

essentially follow the same strategies as those discussed in the previous section. The first step is to estimate the conditional distributions of X_i given Z_i and also Y_i given X_i, possibly also conditional on observable covariates W_i, just as one would do in a standard IV estimation procedure. A typical example is to use 2SLS, where in the first stage X_i is linearly regressed on $\{Z_i, W_i\}$ and in the second stage Y_i is regressed on the fitted values from the first regression as well as W_i to correct the bias. Alternatively, one can use binary response models such as logit or probit to model these conditional distributions.

Next, it is also necessary to estimate $\Pr(Y_i = 1 \mid Z_i = z)$, which can be achieved by either combining the estimates of $\Pr(Y_i = 1 \mid X_i = x)$ and $\Pr(X_i = x \mid Z_i = z)$ for $x, z \in \{0, 1\}$ from the previous step or by fitting a reduced form regression of Y_i on Z_i (and possibly also on W_i and averaging over the empirical distribution of W_i; Tan 2006). The marginal distribution of Z_i, which is also needed to compute the quantities of interest, can be estimated simply by its sample analogue. Finally, these estimated quantities can be substituted into the corresponding probabilities in equations (9), (10), (11), and (12) to produce plug-in estimators of the desired quantities. For uncertainty estimates, one can use the bootstrap approach as in the previous case. I apply these procedures to an empirical example on voter turnout and education at the end of the next section.

Empirical Applications

In this section, I apply the proposed methods for the identification of causal attribution to empirical problems from various subfields of political science. First, I examine a natural experiment where successful assassinations of political leaders are considered to have occurred

randomly (Jones and Olken 2009). Next, I investigate a typical observational study from the so-called democratic peace literature (Oneal and Russett 2001), where regression adjustments are used to control for possible observable confounders. Finally, I apply the new identification results with an instrumental variable in the previous section to revisit the study of education and voter turnout by Milligan, Moretti, and Oreopoulos (2004).

Assassination of Political Leaders

Political leadership is a contentious but somewhat understudied topic in political science. Some scholars have emphasized the roles national leaders can play in determining institutional and policy outcomes in various contexts (e.g., Allison 1971; Greenstein 2004). To study the effect of leadership, researchers face a serious endogeneity problem; that is, while changes in leadership may affect the various societal characteristics, these changes themselves are often consequences of changes in the society. Jones and Olken (2009) use a novel approach to partially cope with this major methodological challenge. They argue that because whether an assassination attempt on a political leader will succeed or fail is essentially randomly determined, they can exploit this inherent randomness as a natural experiment to isolate the effect of leadership on various outcomes. To estimate the causal effects of successful assassinations, they collect data on all publicly reported assassination attempts for all national leaders since 1875. They then analyze these data to see whether a successful assassination increases the probability of regime transition and the intensity of conflicts.

In this section, I follow the identification strategy of Jones and Olken (2009) to estimate the probability of causal attribution (p_A) for regime transition. This alternative quantity is of particular interest given the context of the study. The original authors estimate the average causal *effects* of successful assassination by employing both parametric and nonparametric specifications. In this context, these causal estimands correspond to the expected difference in regime type (i.e., democracy or autocracy) that would realize between hypothetical successful assassinations and unsuccessful ones. While estimating such a quantity would be useful for predicting the impact of assassinating a leader, it may not directly interest political scientists or historians who are primarily concerned with explaining why past regime transitions happened. As discussed earlier, the probability of causal *attribution* addresses this question because it represents the proportion of those actual regime transitions that would not have happened without successful assassinations.

To estimate p_A for successful assassination and regime transition, I employ the estimation procedure explained in the previous section. More specifically, I first estimate the conditional probability of regime transition as a function of the successful assassination indicator, using the same parametric model as the original study.[10] Then the identification regions of p_A are calculated for different sets of identification assumptions based on the estimated coefficients as well as equations (3) and (4). Finally, confidence intervals are computed using the nonparametric bootstrap with 2,000 resamples and the procedure of Imbens and Manski (2004). The results are presented in Figure 1.

In each panel of Figure 1, the point estimate of p_A under both exogeneity (Assumption 1) and monotone treatment effect (Assumption 2) is presented on the left, while the sharp bounds under Assumption 1 alone are given on the right, both along with the 95% confidence intervals. The left panel shows the estimates for all types of regime changes pooled together, democratic transitions, and autocratic transitions (from left to right). The results suggest that under the assumption that the success or failure of an assassination attempt is essentially random (Assumption 1), at least 63%, and possibly all, of the regime transitions following the assassination of national leaders would not have occurred if the leader had survived the attempt. The 95% confidence interval for this estimate is relatively wide but does not cover zero ([0.272, 1]), implying that a significant fraction of these past regime changes would have been prevented had the assassinations been unsuccessful. In addition, if it is also assumed that successful assassination cannot prevent the regime transition that would otherwise have happened, p_A is point identified and equals the previous lower bound (0.630). The 95% confidence interval is still relatively wide ([0.204, 1]) but again does not cover zero.[11] Moreover, the estimated p_A dramatically increases when I focus on the democratic

---

[10] Specifically, I use the linear probability model containing weapon type and the number of serious attempts as covariates. The coefficients are then estimated via ordinary least squares. Because the estimates are obtained from the same statistical model, any difference between the original results and mine is solely due to the difference of the causal quantity of interest. I also conducted the estimation using the probit model with the same set of covariates to address the problem in the original specification that predicted probabilities went outside of the interval [0, 1] for some resamples. The results were qualitatively similar to the ones reported here.

[11] Readers may wonder why the confidence intervals of the bounds appear to be wider than those of the point estimates in this and other examples. This is because p_A is restricted to be in the range of the unit interval, as well as because the bounds confidence intervals cover the true value located somewhere inside the bounds with fixed probability instead of the entire bounds. See Imbens and Manski (2004) for the convergence properties of their confidence intervals.

FIGURE 1 Assassination of Political Leaders and Regime Change
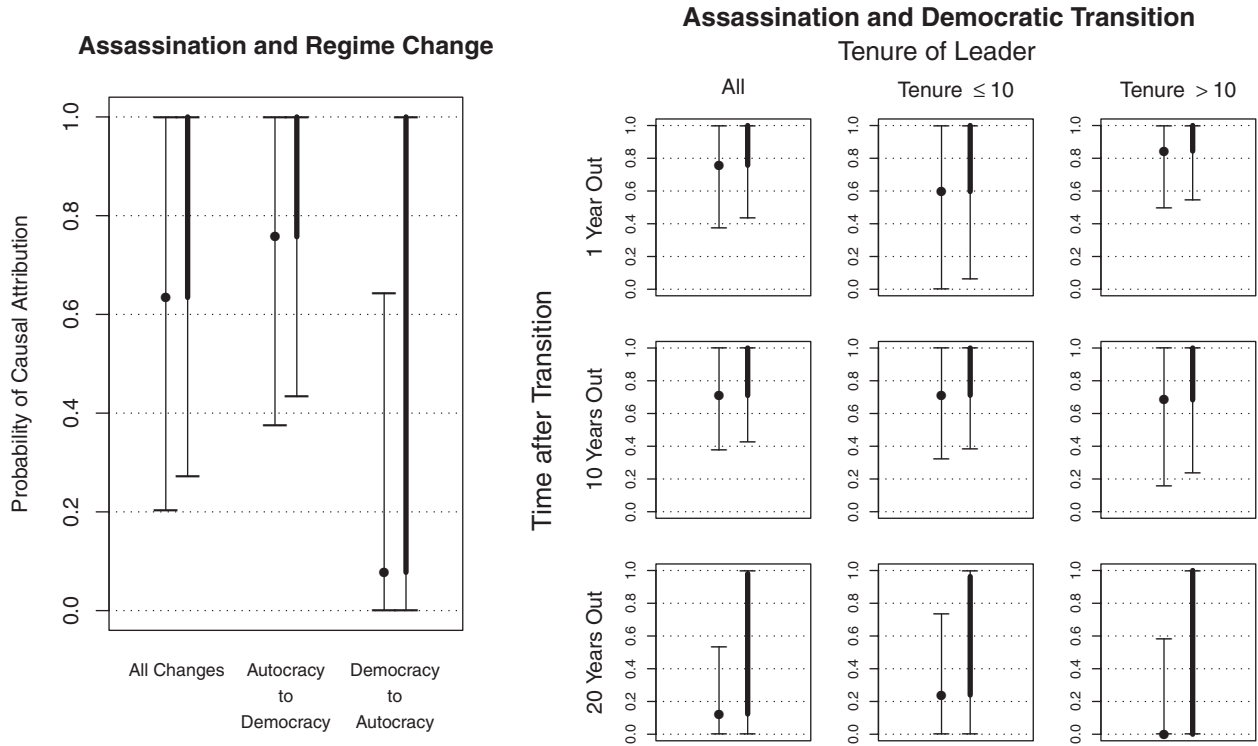
Note: The pair of vertical lines in each plot represents the identification regions for the probability of causal attribution (p_A) under Assumptions 1 and 2 (solid circle on the left) and under Assumption 1 alone (thick bar on the right) along with their 95% confidence intervals (thin solid lines). The left panel shows the estimated probabilities that the regime changes observed after the assassinations of political leaders would not have happened had the assassination attempts failed. The right panel breaks down the estimated p_A for successful assassination and democratic transition by the tenure of the assassinated leader and by the time since the assassination. For a full description of the coding scheme for these groupings, see Jones and Olken (2009). The results suggest that more than 75% of these democratic transitions can be causally attributed to the assassinations. The proportion is particularly large for the short-run moves to democracy following the assassination of long-tenured autocrats.

transitions. The lower bound under Assumption 1 (or equivalently the point estimate under Assumptions 1 and 2) is estimated to be 0.757, with the 95% confidence interval of [0.433, 1] ([0.376, 1]). Thus, a large majority of the past democratic transitions can be causally attributed to the preceding assassinations of the autocrats.

To analyze the past democratic transitions in greater detail, the nine plots on the right decompose them on the basis of the tenure of the assassinated leaders and the duration of the new regimes. The results show that the probability of causal attribution is particularly high for the short-run moves to democracy following the assassination of long-tenured autocrats (upper-right panel). Indeed, at least 84% of these democratic transitions would not have happened if the assassination attempts had failed. However, the same cannot be said for relatively new autocrats who were assassinated with fewer than 10 years of tenure (upper middle panel), as the confidence intervals essentially cover zero ([0, 1] for the point estimate and [0.062, 1] for the bounds). Interestingly, for

the democracies whose former autocratic leaders were assassinated 10 years ago (middle row), p_A is estimated to be roughly the same regardless of tenure. Finally, the bounds and confidence intervals are both wide and uninformative in the three panels on the bottom; this indicates, not surprisingly, that it is rather difficult to trace the cause of established democracies back to the assassinations that occurred 20 or more years ago.

Democracy and International Conflict

A major question in the literature of international security is whether democratic institutions are preventive of militarized international disputes (MIDs). Although this causal hypothesis has been posited in various specific forms, including sufficiency (Russett 1995) and causal effects (e.g., Maoz and Russett 1993), an interesting question is what proportion of the past MIDs could have been prevented by democratic institutions. In a famous

quantitative study, Oneal and Russett (2001) used data on all state dyads that existed between the period of 1886–1992 to show that democracy has a strong negative effect on the probability of the onset of a MID after controlling for various covariates, contrary to the claim by Green, Kim, and Yoon (2001).

By regressing MID onsets on democracy and other control variables and computing the marginal effects of democracy, Oneal and Russett (2001) and Green, Kim, and Yoon (2001) are implicitly making an inference about the causal effect of democracy on the probability of MID onset. As discussed earlier, this corresponds to the causal question of what would have been the impact of hypothetical intervention to the level of democracy at each time point of the past period covered by this dataset. While this quantity may well be of substantive interest, an alternative, perhaps more directly relevant, question is what percentage of the actually observed MIDs was causally attributable to the absence of democratic institutions. Thus, an interesting analysis is to estimate p_A using this dataset. Throughout the analysis, I maintain the exogeneity assumption as in the original study.[12]

Figure 2 reports the estimated probability of causal attribution under the exogeneity and monotone treatment effect assumptions (Assumptions 1 and 2) as well as just under exogeneity (Assumption 1). Since democracy here is measured by the Polity score, which is a discrete scale ranging from −10 to 10, I used values from 0 to 10 as the threshold for a dyad to be defined as democratic and then calculated the probability of causal attribution for each of these coding schemes. Based on the whole period covered by the dataset (1886 to 1992, top panel), the probability of causal attribution under Assumptions 1 and 2 (or equivalently, the lower bound of this quantity under Assumption 1 alone) almost monotonically increases as we change the threshold for democracy from low to high, with a possible exception of the thresholds at

6 and 7. That is, when the comparison is between dyads with a Polity score less than zero and those greater than or equal to zero, the estimated probability is at least 0.482 with the 95% confidence interval of [0.323, 0.641]. This means that about 48.2% of the observed MIDs among the dyads in the first group would not have happened had those dyads switched to the latter category. Furthermore, this probability sharply increases as we shift the reference of comparison to more restrictive sets of democracies. Indeed, when the threshold is set at the standard defined by the Polity IV Project (6 or higher), the estimated probability of causal attribution becomes about 0.676 with the confidence interval of [0.580, 0.772], and the probability jumps to as high as about 0.768 with the confidence interval of [0.648, 0.888] when only the most developed democracies are used as the comparison group. This implies that, holding every pretreatment condition constant, about 76.8% of the past military disputes (excluding the ones fought between two fully developed democracies) would have been prevented if the states engaged in these disputes had been full democracies. This is indeed a high proportion, even compared to the original estimate of the causal effect by Oneal and Russett (2001) based on the identical model.

One point of contention between Green, Kim, and Yoon (2001) and Oneal and Russett (2001) is whether the effect of democracy had changed after 1951. The results presented in the bottom panel of Figure 2 weakly support this hypothesis, albeit in a nuanced way. When the threshold for democracy is set at relatively low levels (0 to 4), the estimated probability of causal attribution is slightly higher when we focus on the period after 1951, although the confidence intervals for these estimates are rather wide. For example, the difference is as large as 0.116 when the threshold is set at one. That is, for this particular definition of democratic dyads, the MIDs involving nondemocracies became more preventable after 1951 by a margin of 11.6 percentage points. This finding is interesting and also consistent with the earlier results by Farber and Gowa (1995). However, this gap effectively vanishes when the threshold is set at levels higher than five.

The analysis has so far led to remarkable conclusions, suggesting that many of the past MIDs would have been prevented by democratic institutions. However, one should be careful about these results since they all rest on the crucial assumption of exogeneity (Assumption 1) and also monotone treatment effect (Assumption 2) if one interprets them as point estimates rather than the lower bounds of the probability of causal attribution. Unlike the previous example, which exploited the inherent randomness in the success or failure of assassination attempts, the current example relies on a more tenuous assumption

---

[12] I use the same estimation procedure as in the previous example. Namely, I estimate the regression model that is identical to the one used in the original study (i.e., the first model in Table 1, 473, which uses the generalized estimating equation with the AR(1) working correlation matrix). The result is then used to construct the estimates of the identification regions. The confidence intervals are obtained via the block bootstrap clustered at the dyad level with 1,000 resamples, using normal approximation and Imbens and Manski's formula. Although I followed the original procedure to facilitate comparison, the model contains a likely specification error for the purpose of causal inference. That is, the covariates include posttreatment confounders such as alliance and capability ratio, which will induce bias in the regression estimates of the causal quantities. Thus, I also estimated an alternative model which only contains the variables that are unlikely to be affected by the level of democracy in a given year (i.e., contiguity, log distance, and lagged dependent variable). The results were qualitatively very similar to the ones reported in the main text.

FIGURE 2 Proportion of the Past MID Onsets Attributable to the Lack of Democracy
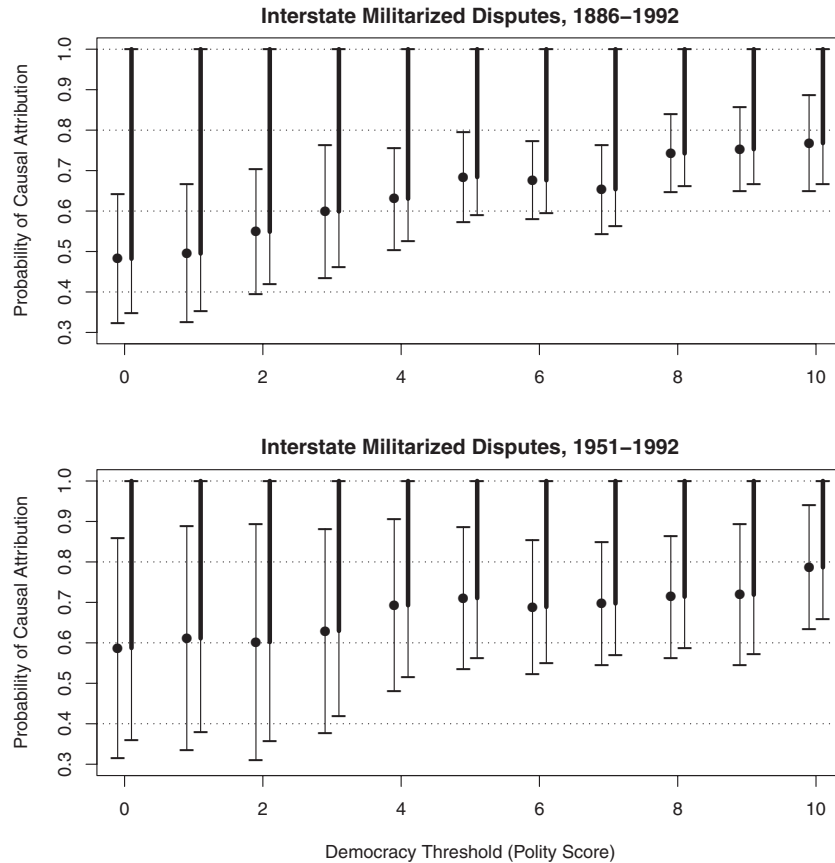
Note: In each panel, the horizontal axis represents the values of the Polity score used as the threshold for a dyad to be coded as democratic. The results in the top panel are based on the full dataset while the bottom panel is restricted to the postwar period. For the meaning of the graph components, see the caption of Figure 1. The results suggest that the probability of causal attribution is roughly monotonically increasing as a function of the threshold for both time periods.

that controlling for the observable covariates included in the model is sufficient to consider the level of democracy to be effectively randomly assigned by nature. This assumption, unfortunately, is difficult to justify. In fact, the original issue raised by Green, Kim, and Yoon (2001) was precisely the possible existence of unobservable confounders in this dataset. In the next example, I consider the situation where one can address this problem with the aid of an instrumental variable.

Education and Voter Turnout

Education has been identified as an important predictor of voter turnout by numerous empirical studies (e.g., Wolfinger and Rosenstone 1980), but whether the ob-

served strong association between education and turnout represents a true causal relationship remains a contentious issue (e.g., Brady, Verba, and Schlozman 1995). Various unmeasured confounders, such as unobservable individual characteristics and environmental factors in which survey respondents are raised, can both increase their educational attainment and probability of voting. To cope with this endogeneity problem, Milligan, Moretti, and Oreopoulos (2004) employ an identification strategy based on instrumental variables.[13] Following the approach of Acemoglu and Angrist (2000), Milligan, Moretti, and Oreopoulos use the variation in compulsory

[13] A recent study by Sondheimer and Green (2010) takes a similar approach and uses past field experiments and a quasi-experiment as instruments. Their focus is also to estimate the causal effect of education on voter turnout.

schooling laws across the United States as an exogenous encouragement for the citizens to receive different levels of education. They then estimate the causal effects of receiving a relatively high level of education (defined as high school graduation or higher) on various types of civic participation, including voting in the national election held in the year of the survey.[14]

While using an instrumental variable is a reasonable strategy to address the endogeneity problem, focusing on causal *effects* may not be entirely justified depending on the type of causal question one is interested in. For example, one empirical puzzle in the political participation literature is why voter turnout has largely stayed the same (or even declined slightly) in the United States since the 1970s despite the increase in educational levels of the electorate (Abramson and Aldrich 1982; Cassel and Luskin 1988). Because the goal underlying this question is to explain the past levels of turnout, it may be more reasonable to focus on causal *attribution* and estimate the extent to which citizens' actual behavior is attributable to their educational levels. More specifically, how likely is it that those who voted would have abstained if they had been less educated? Conversely, what is the probability that the uneducated nonvoters in the past elections would have voted had they been better educated?

Here, I apply the new identification results developed in the previous section to answer these questions. I estimate the complier probabilities of causal attribution (p_C) for education and turnout for the subpopulation of actual voters and nonvoters, using the state compulsory attendance rules as an instrument. To estimate this quantity, I followed the procedures used in the original study,[15] except that I used the estimated model parameters to compute the point and bound estimates of p_A and p_C (given in Propositions 1, 2, 3, and 4) using the procedure described in the previous section. The results are shown in Figure 3.

The left two panels present the estimated population (left two bars) and complier (right two bars) probabilities of causal attribution, along with their 95% normal confidence intervals based on the nonparametric bootstrap with 1,000 resamples (and Imbens and Manski's formula

for bounds). Within each pair of vertical bars, the left bar corresponds to the estimate under Assumptions 2, 3, 4, and 5 (Propositions 4 and 2) while the right bar represents the bounds under Assumptions 3, 4, and 5 (Propositions 3 and 1). The first panel shows the probability of attribution for the subpopulation of educated voters, i.e., the proportion of those who would not have voted if they had been less educated. The second panel, in contrast, shows the probability of attribution for uneducated nonvoters, i.e., the probability with which they would have voted if they had been better educated. Finally, for the purpose of comparison, I also include the estimated complier average treatment *effect* in the rightmost panel. This is the quantity typically reported by empirical researchers when 2SLS is used to estimate the causal effect of a binary treatment (Angrist, Imbens, and Rubin 1996).

Figure 3 shows that when the treatment is not exogenous, inference about the population probability of causal attribution is difficult even with the aid of an instrumental variable. The estimated bounds for p_A are almost uninformative for the educated voters under either set of assumptions. However, there is some evidence that the probability is positive for the uneducated nonvoters. The estimated bounds for this subpopulation are [0.049, 0.969] ([0.049, 1] without Assumption 2), and their 95% confidence interval does not include zero ([0.046, 0.971]; [0.046, 1] without Assumption 2). This implies that at least about 4.9% of the uneducated nonvoters would have voted if given education better than high school.

Inference about compliers, however, is substantially easier. The complier probability of causal attribution represents the degree to which voting (nonvoting) can be causally attributed to education among the educated (uneducated) voters whose education levels would be positively affected by compulsory attendance laws.[16] For these groups, there is more evidence that turnout is attributable to education; the estimated p_C for the educated compliers who voted is equal to 0.262 under Assumptions 2, 3, 4, and 5 (with the 95% confidence interval of [0.047, 0.478]) and within the bounds of [0.262, 0.703] under Assumptions 3, 4, and 5 (with the confidence interval of [0.081, 0.775]). This implies that at least about 26.2% of the educated complier voters would not have voted had they not received high education.

Moreover, these estimates are somewhat larger for the uneducated compliers who actually did not vote; their p_C is 0.349 ([0.349, 0.898] without Assumption 5) with the

---

[14] Milligan, Moretti, and Oreopoulos (2004) provide evidence from both the National Election Studies and Current Population Surveys (CPS). This article focuses on the latter.

[15] More specifically, I use 2SLS where the high school graduation indicator is regressed on the instrument and covariates (black, female, state of birth, and quartic function of age) in the first stage and then turnout is regressed on the fitted values from the first stage and the covariates. Since the proposed method relies on a binary instrument, I dichotomize the compulsory schooling variable using 10-year compulsory attendance as the cutpoint.

[16] More specifically, this group represents the voters with (without) a high school diploma who would not have (would have) graduated from high school if their state of residence had had less (more) restrictive compulsory attendance laws.

FIGURE 3 Probability of Causal Attribution and Local Average Treatment Effect of Education and Turnout
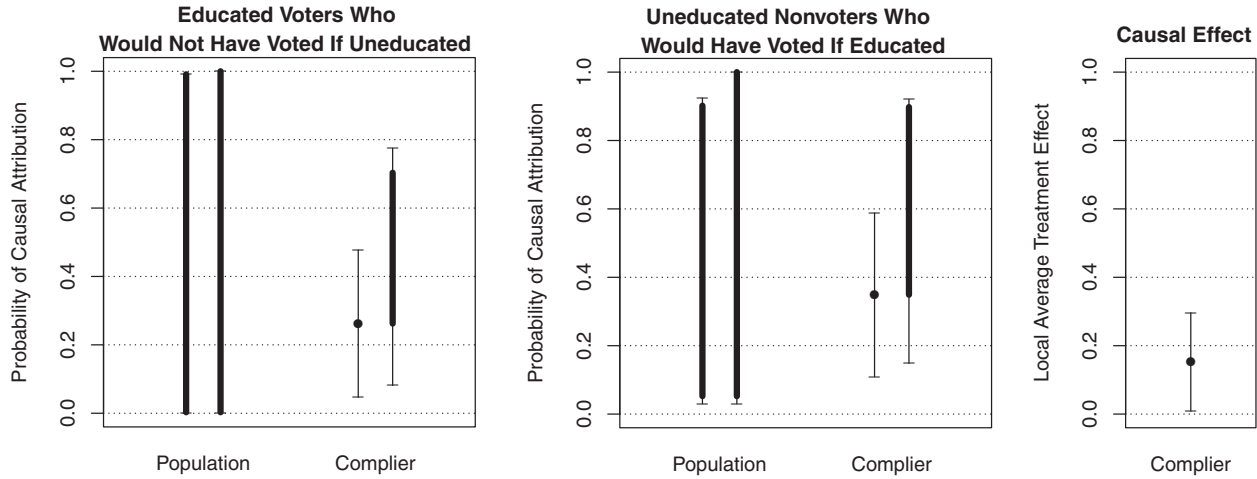
Note: The left two panels show the estimated sharp bounds on the probabilities of causal attribution for the population (p_A, left two vertical bars) and for the compliers (p_C, right two vertical bars) along with their 95% confidence intervals. The first panel shows the proportions of the educated voters who would not have voted if they had been less educated, while the second panel shows the probabilities with which the uneducated nonvoters would have voted if they had been better educated. There is little evidence that the complier probability of attribution differs between the educated voters and the uneducated nonvoters, although it is clearly statistically different from zero for both subpopulations. The estimated complier average treatment *effect* (the rightmost panel), however, is only marginally significant.

confidence interval of [0.110, 0.588] ([0.148, 0.920]), implying that at least about 34.9% of them would have voted if they had graduated from high school. However, these estimates are not statistically significantly different from those for the educated complier voters. This is substantively interesting because, a priori, the educated voters and uneducated nonvoters have different characteristics and their counterfactual turnout behavior may well be much different. Finally, the estimated complier average treatment effect is only 0.152. The standard 2SLS causal estimate is thus misleading for inference about causal attribution. Indeed, because this estimate is only marginally statistically significant ([0.009, 0.296]), one may incorrectly conclude that turnout is not causally attributable to education while in fact it is on the basis of the proposed framework.

A potential problem in the above analysis is that the compulsory schooling variable may not be an ideal instrument, as previously pointed out by Tenn (2007). If the identification assumptions are violated, the estimates shown in Figure 3 will be biased and invalid. More generally, the availability of good instrumental variables is often a major obstacle in causal inference with observational data. Researchers may find it helpful to supplement the IV-based analysis with the more traditional approach based on conditional exogeneity (i.e., Assumption 1) in such a situation. For example, Kam and Palmer (2008)

collect a large number of covariates that may confound the relationship between education and turnout, such as respondents' preadult experiences and parental background, and use propensity score matching to estimate the causal effect of education on participation. When such extensive information about pretreatment characteristics is available, the exogeneity-based approach may provide a viable alternative to the IV approach, especially if coupled with estimation procedures robust to functional form misspecifications (e.g., matching). Those multivariate estimation procedures can be easily incorporated for analyzing causal attribution, as illustrated through the other empirical applications in this section.

Concluding Remarks

In this article, I proposed an alternative approach to causal inference in political science. The framework starts from a specific type of causal question commonly asked in the discipline—what proportion of the past events can be attributed to a suspected cause—and defines a quantity of interest that is directly tailored for this type of question. This contrasts with the other methods for causal inference currently employed in social sciences, which are more appropriate for estimating the effect of

hypothetical intervention to a causal relationship. The discussion of the differences between these two types of causal quantities also sheds light on a methodological debate about selecting on the dependent variable.

Building on the previous results about the identification of this quantity under the exogeneity assumption, I derived a set of new analytical results for the situation where exogeneity is violated but an instrumental variable is available. Because the exogeneity assumption is often difficult to justify in observational studies, the existing results are limited in their applicability to many empirical problems of interest. In contrast, the results in this article expand the scope of this framework to at least some of these empirical questions. Both the previous and new results were illustrated by three empirical examples from different subfields of political science.

In closing, I emphasize again that the proposed framework should not be considered as a replacement to standard methods for the estimation of causal effects. Rather, the central message of this article is that one should carefully consider the quantity of interest in light of the causal question at hand before diving into such standard analyses. Causal inference is an intricate problem which requires more than a blind application of single methodology, be it regression, matching, experiment, or combinations of those. This article provides a guideline for choosing an appropriate quantity of interest for one's empirical question, as well as a set of methods for analyzing causal attribution.

A Mathematical Appendix
A.1 Proof of Proposition 1

First, note that under Assumption 4, the complier probability of causal attribution (p_C) can be expressed with respect to the population proportions of principal strata as

$$p_C = \Pr(Y_i(0) = 0 \mid Y_i(1) = 1, X_i(1) = 1, X_i(0) = 0)$$
$$= \frac{\pi_{1010}}{\pi_{1010} + \pi_{1110}}, \qquad (13)$$

where $\pi_{y_1 y_0 x_1 x_0} = \Pr(Y_i(1) = y_1, Y_i(0) = y_0, X_i(1) = x_1, X_i(0) = x_0)$. Under Assumptions 3, 4, and 5, the denominator of this quantity is identifiable from the observed information because

$$\pi_{1010} + \pi_{1110} = \Pr(Y_i(1) = 1, X_i(1) = 1, X_i(0) = 0)$$
$$= \Pr(Y_i(1) = 1, X_i(1) = 1, X_i(0) = 1)$$
$$+ \Pr(Y_i(1) = 1, X_i(1) = 1, X_i(0) = 0)$$
$$- \Pr(Y_i(1) = 1, X_i(1) = 1, X_i(0) = 1)$$

$$= \Pr(Y_i(1) = 1, X_i(1) = 1)$$
$$- \Pr(Y_i(1) = 1, X_i(0) = 1)$$
$$= \mathbb{E}\{\mathbb{E}(Y_i(1) X_i(1) \mid W_i)\}$$
$$- \mathbb{E}\{\mathbb{E}(Y_i(1) X_i(0) \mid W_i)\}$$
$$= \mathbb{E}\{\mathbb{E}(Y_i(1) X_i(1) \mid Z_i = 1, W_i)\}$$
$$- \mathbb{E}\{\mathbb{E}(Y_i(1) X_i(0) \mid Z_i = 0, W_i)\}$$
$$= \Pr(Y_i = 1, X_i = 1 \mid Z_i = 1)$$
$$- \Pr(Y_i = 1, X_i = 1 \mid Z_i = 0),$$

where the expectations are with respect to the distribution of W_i. The third equality holds by Assumption 5, the fourth by the law of iterated expectations, the fifth by Assumption 3, and the last by Assumption 4 and the law of iterated expectations. Thus, the sharp bounds on equation 13 can be obtained by maximizing and minimizing the numerator, π_{1010}, under the constraints implied by the maintained assumptions. First, Assumption 3 (and the law of iterated expectations) implies the following linear constraints on the possible values of this quantity:

$$\Pr(X_i = 1, Y_i = y \mid Z_i = 1) = \sum_{y_0=0}^{1} \sum_{x_0=0}^{1} \pi_{y y_0 1 x_0},$$

$$\Pr(X_i = 0, Y_i = y \mid Z_i = 1) = \sum_{y_1=0}^{1} \sum_{x_0=0}^{1} \pi_{y_1 y 0 x_0},$$

$$\Pr(X_i = 1, Y_i = y \mid Z_i = 0) = \sum_{y_0=0}^{1} \sum_{x_1=0}^{1} \pi_{y y_0 x_1 1},$$

$$\Pr(X_i = 0, Y_i = y \mid Z_i = 0) = \sum_{y_1=0}^{1} \sum_{x_1=0}^{1} \pi_{y_1 y x_1 0},$$

for $y \in \{0, 1\}$. Next, Assumption 5 implies $\pi_{y_1 y_0 01} = 0$ for all $y_1, y_0 \in \{0, 1\}$, because these strata represent the units for which the instrument negatively affects the treatment. Finally, the principal strata probabilities must add up to one, i.e., $\sum_{y_1} \sum_{y_0} \sum_{x_1} \sum_{x_0} \pi_{y_1 y_0 x_1 x_0} = 1$. Solving the linear programming problem via a standard algorithm (Balke and Pearl 1997) yields the following sharp bounds on π_{1010},

$$\max\{0, \ \Pr(Y_i = 1 \mid Z_i = 1) - \Pr(Y_i = 1 \mid Z_i = 0)\}$$
$$\leq \pi_{1010} \leq$$
$$\min\{1, \ \Pr(Y_i = 0, X_i = 0 \mid Z_i = 0)$$
$$- \Pr(Y_i = 0, X_i = 0 \mid Z_i = 1)\}.$$

By substituting and rewriting the above, we obtain the sharp bounds on p_C given in Proposition 1. $\qquad \square$

A.2 Proof of Proposition 2

First, note that we have the identical expression for p_C given in equation 13 and the same set of linear constraints as in the proof of Proposition 1 by Assumptions 3, 4, and 5. In addition, Assumption 2 implies that $\pi_{01x_1x_0} = 0$ for all $x_1, x_0 \in \{0, 1\}$. These constraints imply that $\pi_{1110} = \Pr(X_i = 0, Y_i = 1 \mid Z_i = 0) - \Pr(X_i = 0, Y_i = 1 \mid Z_i = 1)$ and $\pi_{1010} = \Pr(Y_i = 1 \mid Z_i = 1) - \Pr(Y_i = 1 \mid Z_i = 0)$. Substituting these to equation (13) and rewriting yields the expression in Proposition 2. $\qquad\square$

A.3 Proof of Propositions 3 and 4

The proof follows the same linear programming procedure as in the proof of Proposition 1, except that the objective function is now

$$
\begin{aligned}
p_A = {} & \frac{\Pr(Z_i = 1)}{\Pr(X_i = Y_i = 1)}(\pi_{1010} + \pi_{1011}) \\
& + \frac{\Pr(Z_i = 0)}{\Pr(X_i = Y_i = 1)}(\pi_{1001} + \pi_{1011}).
\end{aligned}
\tag{14}
$$

Maximizing and minimizing this quantity using the corresponding sets of assumptions yields the expressions in Propositions 3 and 4. $\qquad\square$

References

Abramson, Paul R., and John H. Aldrich. 1982. "The Decline of Electoral Participation in America." *Americal Political Science Review* 76(3): 502–21.

Acemoglu, Daron, and Joshua Angrist. 2000. "How Large Are Human Capital Externalities? Evidence from Compulsory Schooling Laws." *NBER Macroeconomics Annual* 15(1): 9–59.

Allison, Graham. 1971. *Essence of Decision: Explaining the Cuban Missile Crisis*. New York: Little, Brown.

Angrist, Joshua D., and Guido W. Imbens. 1995. "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity." *Journal of the American Statistical Association* 90(430): 431–42.

Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91(434): 444–55.

Ashworth, Scott, Joshua Clinton, Adam Meirowitz, and Kristopher W. Ramsay. 2008. "Design, Inference, and the Strategic Logic of Suicide Terrorism." *American Political Science Review* 102(2): 269–73.

Balke, Alexander, and Judea Pearl. 1997. "Bounds on Treatment Effects from Studies with Imperfect Compliance." *Journal of the American Statistical Association* 92(439): 1171–76.

Brady, Henry E., Sidney Verba, and Kay Lehman Schlozman. 1995. "Beyond SES: A Resource Model of Political Participation." *American Political Science Review* 89(2): 271–94.

Cai, Zhihong, and Manabu Kuroki. 2006. "Variance Estimators for Three 'Probabilities of Causation.'" *Risk Analysis* 25(6): 1611–20.

Card, David, and Alan B. Krueger. 1994. "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania." *American Economic Review* 84(4): 772–93.

Cassel, Carol A., and Robert C. Luskin. 1988. "Simple Explanations of Turnout Decline." *American Political Science Review* 82(4): 1321–30.

Chernozhukov, Victor, Han Hong, and Elie E. Tamer. 2007. "Estimation and Confidence Regions for Parameter Sets in Econometric Models." *Econometrica* 75(5): 1243–84.

Collier, David, James Mahoney, and Jason Seawright. 2004. "Claiming Too Much: Warnings about Selection Bias." In *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, ed. Henry E. Brady and David Collier. Lanham, MD: Rowman and Littlefield, 85–102.

Cox, Donald R. 1958. *Planning of Experiments*. New York: John Wiley & Sons.

Farber, Henry S., and Joanne Gowa. 1995. "Polities and Peace." *International Security* 20(2): 123–46.

Fearon, James D. 1991. "Counterfactuals and Hypothesis Testing in Political Science." *World Politics* 43(2): 169–95.

Fisher, Ronald A. 1971. *The Design of Experiments*. 9th ed. New York: Macmillan.

Frangakis, Constantine E., and Donald B. Rubin. 2002. "Principal Stratification in Causal Inference." *Biometrics* 58(1): 21–29.

Freedman, David A. 2008. "On Regression Adjustments in Experiments with Several Treatments." *Annals of Applied Statistics* 2(1): 176–96.

Geddes, Barbara. 1990. "How the Cases You Choose Affect the Answers You Get: Selection Bias in Comparative Politics." *Political Analysis* 2(1): 131–50.

Gerber, Alan S., and Donald P. Green. 2000. "The Effects of Canvassing, Telephone Calls, and Direct Mail on Voter Turnout: A Field Experiment." *American Political Science Review* 94(3): 653–63.

Green, Donald P., Soo Yeon Kim, and David H. Yoon. 2001. "Dirty Pool." *International Organization* 55(2): 441–68.

Greenstein, Fred I. 2004. *The Presidential Difference: Leadership Style from FDR to George W. Bush*. Princeton, NJ: Princeton University Press.

Halpern, Joseph Y., and Judea Pearl. 2005a. "Causes and Explanations: A Structural-Model Approach—Part I: Causes." *British Journal of Philosophy of Science* 56(4): 843–87.

Halpern, Joseph Y., and Judea Pearl. 2005b. "Causes and Explanations: A Structural-Model Approach—Part II: Explanations." *British Journal of Philosophy of Science* 56(4): 889–911.

Hanmer, Michael J. 2007. "An Alternative Approach to Estimating Who Is Most Likely to Respond to Changes in Registration Laws." *Political Behavior* 29(1): 1–30.

Hempel, Carl G. 1965. *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. New York: Free Press.

Hitchcock, Christopher, and James F. Woodward. 2003. "Explanatory Generalizations, Part I: A Counterfactual Account." *Nôus* 37(1): 1–24.

Holland, Paul W. 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association* 81(396): 945–60.

Hume, David. 1739. *A Treatise of Human Nature*. Oxford: Oxford University Press.

Huntington, Samuel P. 1991. *The Third Wave: Democratization in the Late 20th Century*. Oklahoma City: University of Oklahoma Press.

Imai, Kosuke, and Gary King. 2004. "Did Illegal Overseas Absentee Ballots Decide the 2000 U.S. Presidential Election?" *Perspectives on Politics* 2(3): 537–49.

Imai, Kosuke, and Samir Soneji. 2007. "On the Estimation of Disability-Free Life Expectancy: Sullivan's Method and Its Extension." *Journal of the American Statistical Association* 102(480): 1199–1211.

Imai, Kosuke, and Teppei Yamamoto. 2010. "Causal Inference with Differential Measurement Error: Nonparametric Identification and Sensitivity Analysis." *American Journal of Political Science* 54(2): 543–60.

Imbens, Guido W., and Charles F. Manski. 2004. "Confidence Intervals for Partially Identified Parameters." *Econometrica* 72(6): 1845–57.

Imbens, Guido W., and Joshua D. Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62(2): 467–75.

Jones, Benjamin F., and Benjamin A. Olken. 2009. "Hit or Miss? The Effect of Assassinations on Institutions and War." *American Economic Journal: Macroeconomics* 1(2): 55–87.

Kam, Cindy D., and Carl L. Palmer. 2008. "Reconsidering the Effets of Education on Political Participation." *Journal of Politics* 70(3): 612–31.

King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton, NJ: Princeton University Press.

Lewis, David K. 2001. *Counterfactuals*. 2nd ed. New York: Wiley-Blackwell.

Lipset, Seymour M. 1959. "Some Social Requisites of Democracy: Economic Development and Political Legitimacy." *American Political Science Review* 53(1): 69–105.

Mackie, John L. 1965. "Causes and Conditions." *American Philosophical Quarterly* 2(4): 245–64.

Manski, Charles F. 1995. *Identification Problems in the Social Sciences*. Cambridge, MA: Harvard University Press.

Manski, Charles F. 2007. *Identification for Prediction and Decision*. Cambridge, MA: Harvard University Press.

Maoz, Zeev, and Bruce Russett. 1993. "Normative and Structural Causes of Democratic Peace, 1946–1986." *American Political Science Review* 87(3): 624–38.

Menzies, Peter, and Huw Price. 1993. "Causation as a Secondary Quality." *British Journal for the Philosophy of Science* 44(2): 187–203.

Mill, John Stuart. 1843. *A System of Logic*. London: John Parker.

Milligan, Kevin, Enrico Moretti, and Philip Oreopoulos. 2004. "Does Education Improve Citizenship? Evindence from the United States and the United Kingdom." *Journal of Public Economics* 88(9–10): 1667–95.

Neyman, J. 1923. "On the Application of Probability Theory to Agricultural Experiments: Essay on Principles, Section 9. (Translated in 1990)." *Statistical Science* 5(4): 465–80.

Oneal, John R., and Bruce Russett. 2001. "Clear and Clean: The Fixed Effects of the Liberal Peace." *International Organization* 55(2): 469–85.

Pape, Robert A. 2003. "The Strategic Logic of Suicide Terrorism." *American Political Science Review* 97(3): 343–61.

Pape, Robert A. 2008. "Methods and Findings in the Study of Suicide Terrorism." *American Political Science Review* 102(2): 275–77.

Pearl, Judea. 1999. "Probabilities of Causation: Three Counterfactual Interpretations and Their Identification." *Synthese* 121(1/2): 93–149.

Porter, Michael E. 1998. *The Competitive Advantage of Nations: With a New Introduction*. New York: Free Press.

Price, Huw. 1991. "Agency and Probabilistic Causality." *British Journal for the Philosophy of Science* 42(2): 157–76.

Przeworski, Adam, Michael E. Alvarez, Jose Antonio Cheibub, and Fernando Limongi. 2000. *Democracy and Development: Political Institutions and Well-Being in the World, 1950–1990*. Cambridge: Cambridge University Press.

Robins, James, and Sander Greenland. 1989. "The Probability of Causation Under a Stochastic Model for Individual Risk." *Biometrics* 45(4): 1125–38.

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

Rosenbaum, Paul R. 2001. "Effects Attributable to Treatment: Inference in Experiments and Observational Studies with a Discrete Pivot." *Biometrika* 88(1): 219–31.

Rosenbaum, Paul R. 2002. *Observational Studies*. 2nd ed. New York: Springer-Verlag.

Rubin, Donald B. 1974. "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology* 66(5): 688–701.

Rubin, Donald B. 1977. "Assignment to Treatment Group on the Basis of a Covariate." *Journal of Educational and Behavioral Statistics* 2(1): 1–26.

Russett, Bruce. 1995. "The Democratic Peace: 'And Yet It Moves.'" *International Security* 19(4): 164–75.

Sondheimer, Rachel Milstein, and Donald Green. 2010. "Using Experiments to Estimate the Effects of Education on Voter Turnout." *American Journal of Political Science* 54(1): 174–89.

Sovey, Allison J., and Donald P. Green. 2011. "Instrumental Variables Estimation in Political Science: A Readers' Guide." *American Journal of Political Science* 55(1): 188–200.

Tamer, Elie E. 2010. "Partial Identification in Econometrics." *Annual Review of Economics* 2: 167–95.

Tan, Zhiqiang. 2006. "Regression and Weighting Methods for Causal Inference Using Instrumental Variables." *Journal of the American Statistical Association* 101(476): 1607–18.

Tenn, Steven. 2007. "The Effect of Education on Voter Turnout." *Political Analysis* 15(4): 446–64.

Tian, Jin, and Judea Pearl. 2000. "Probabilities of Causation: Bounds and Identification." *Annals of Mathematics and Artificial Intelligence* 28(1–4): 287–313.

Wand, Jonathan N., Kenneth W. Shotts, Jasjeet S. Sekhon, Walter R. Mebane Jr., Michael C. Herron, and Henry E. Brady.

2001. "The Butterfly Did It: The Aberrant Vote for Buchanan in Palm Beach County, Florida." *American Political Science Review* 95(4): 793–810.

Wolfinger, Raymond E., and Steven J. Rosenstone. 1980. *Who Votes?* New Haven, CT: Yale University Press.

Woodward, James F. 2003. *Making Things Happen: A Theory of Causal Explanation.* Oxford: Oxford University Press.

Yamamoto, Teppei. 2011. "Replication Data for: Understanding the Past: Statistical Analysis of Causal Attribution." hdl:1902.1/16248. The Dataverse Network.