The Perils of the All Cause Model*

Luke Keele† Randolph T. Stevenson‡

December 8, 2014

Abstract

One of the most common identification strategies in political science is selection on observables. Under this strategy, analysts assume that they observed enough covariates to make treatment status as-if random. Adjustments are then made for observed confounders through statistical methods such as regression or matching. Under adjustment methods such as matching or inverse probability weighting, coefficients for control variables are treated as nuisance parameters and are not directly estimated. This is in direct contrast to regression approaches where estimated parameters are observed for all covariates. Analysts often find it tempting to give a causal interpretation to all the parameters in such regression models, which is not possible under the controls as nuisance parameter approach. In this paper, we illustrate the dangers of treating all the parameters in a regression model as causal parameters. Using Directed Acyclic Graphs, we show how even if some effects are identified in a regression model, many estimated parameters do not represent causal effects or may be direct effects. The general recommendation is for analysts to attempt to identify a single effect and limit interpretation of models to that effect.

*I would like to thank Felix Elwert for first showing me how DAGs can be useful in this context.

†Associate Professor, Department of Political Science, 211 Pond Lab, Penn State University, University Park, PA 19130 Email: ljk20@psu.edu.

‡Professor, Department of Political Science, Rice University, P.O. Box 1892, MS-24, Houston, TX 77251, Email: stevenso@rice.edu

1 Introduction

Statistics and statistical modeling form an important and widely used component of research in political science. While there is some debate about the validity of a statistical approach to the study of politics, a cursory glance at most journals show that the use of statistics is widespread. The use of statistics in political science is generally for three main purposes: description, prediction and causality. That is, analysts often use statistics to describe empirical phenomena. Far less frequently, statistic models are used to predict political outcomes. Here, the focus tends to be on the prediction of election outcomes and militarized conflict. Finally, statistical models are used to draw inferences about causal relationships. While each use of statistical models has merit, quite often the focus of statistical models is drawing causal inferences.

Over the last fifteen years, political science has been gripped by an identification revolution that has focused on the assumptions needed to identify causal effects with statistical models. The new focus on identification has produced some calls for whole sale change in the practice of research design and statistical analysis in the field. In general, the argument is that research designs in political science should focus on randomized experiments if possible and natural experiments where randomization is not possible (Gerber, Green and Kaplan 2004; Dunning 2009; Sekhon 2009; Keele and Minozzi 2012). Of course, while this experimental approach is far more widespread than it once was, many analysts still rely on assumptions about statistical specification to identify causal effects. Here, the goal is to make units similar on observed characteristic such that valid counterfactual comparisons are possible.

Here we outline a series of interpretation pitfalls that can arise when control parameters are given a causal interpretation. We contrast two different method of statistical adjustment. In the first approach, which we deem the regression approach, all covariates in the models produce some statistical summary that can be given an interpretation under the principle of ceteris paribus. In the second approach, we call this the treatment effect approach, the effects of control variables are treated as nuisance parameters and left unspecified. Here, while there are many covariates in the

model, only a single parameter is produced and interpreted. The motivation behind the treatment effect approach is to reduce the possibility of misspecfication due to an incorrect functional form. We highlight an additional advantage of this approach. Often under the regression approach, multiple covariates are interpreted as causal parameters. Using Directed Acyclic Graphs, we show how the practice of interpreting control covariate parameters as causal parameters can lead to nonsensical results. We demonstrate how controlling for a variable does not imply that the effect of the control on the outcome is causal. As such it can be dangerous to give the parameters from control variable causal interpretations. We demonstrate how such nonsensical interpretations can arise using two cases studies: one from the literature on the democratic peace and one from the literature on vote choice.

2 The All Cause Model

We begin with a discussion of how analysts might approach the interpretation of estimated parameters in a multivariate regression models. We do this through a set of hypothetical statistical models that one might use in some common political science applications. We start with a brief overview of concepts in the study of causal relationship with statistical models.

In many statistical analyses, the goal is to establish whether there is a causal relationship between two variables which we denote as D and Y with D being the cause and Y the outcome. We say that there is an association between D and Y if distribution of Y varies across levels of D. If this does not hold, then D and Y are said to be independent. We say there is a causal relationship between D and Y in a population of units if and only if there is at least one unit for which changing or intervening on D will change Y. How are association and causation related? Intuition tells us that associations are the result of causal relationships. That is if D causes Y this will produce an association between D and Y. The great difficulty is that while we can readily observe associations between D and Y, there are forces other than causality such as confounding and selection which can induce an association between D and Y. Moreover, association is nondirectional. An association might imply that Y causes D instead of D causing

Y.

An alternative way to discuss whether a statistical estimate is causal or not is to discuss it in terms of identification. Data only provide us with estimates of associations. These observed associations contain some unknown mix of causal and non-causal (spurious) components. A causal effect is said to "identified" when we have removed all non-causal components from the observed association (Pearl 1995). Resolution of the problem requires an *identification strategy*: a set of assumptions that warrant inferences based on observable quantities. Any research design, at least implicitly, adopts an identification strategy. Identification assumptions thus bridge theoretical and observable quantities. When identification assumptions hold, our estimate of the causal parameter is said to be is identified. See Keele (2014) for a detailed overview of identification and identification strategies. One identification strategy is the randomized experiment. I now present a hypothetical application to discuss the interpretation of estimated parameters in statistical models when the goal is causal inference.

We being with an application where we assume the analysts has collected data from a field experiment on get-out-the-vote techniques. Let's assume that in this field experiment the researcher was interested in whether door-to-door canvassing has a causal effect on turnout. We also assume that the treatment was randomized[1] When the experiment has concluded with data collected, the analyst might estimate the following regression model:

$$Turnout = \beta_0 + \beta_1 Canvas + \beta_2 Age + \beta_3 Income + \beta_4 Education + \beta_5 Female$$

In this model, $Canvas$ represents an indicator for whether a subject was exposed to the door-to-door canvassing treatment. Given that the treatment was randomized, one would be justified interpreting $\hat{\beta}_1$ as the causal effect of the treatment on turnout. How should one interpret the other parameters in the model? It is well understood that those parameters do *not* represent causal effects and should not be given a causal interpretation. Gerber and Green (2012) provide a

[1]We also assume that treatment spillovers did not occurs such that the stable unit treatment value assumption (SUTVA) (Rubin 1986) holds.

full overview of the purpose of covariates in experiments, but briefly since we have not randomized quantities like age and income, we cannot give those estimate parameters causal interpretation despite their inclusion in the regression model. In this model, there is only one parameter that represents a causal effect, β_1, the rest are in some sense nuisance parameters that we cannot give a causal interpretation to since the are not randomized.

Now let's turn to a different example. In this example, we are interested in estimate the causal effect of election day registration (EDR) on turnout. In this setting, randomization is infeasible, so we must use observational data to study this question. The most common approach to identification of causal effects with observational data is to make a specification assumption. That is we must assume that our statistical model is "correctly" specified. Under this approach, analysts collect all known confounders and use a statistical estimator to make treated and control comparable while the treatment effect is estimated. One statement of this assumption is that of "selection on observables" (Barnow, Cain and Goldberger 1980),[2] to emphasize that this assumption requires *all* covariates that predict both the outcome *and* the treatment.

Let's assume that the researcher estimates the following model to estimate the causal effect of EDR on turnout:

$$Turnout = \beta_0 + \beta_1 EDR + \beta_2 Age + \beta_3 Income + \beta_4 Education + \beta_5 Female$$

For the moment, let's assume that this is the correct specification in that the effect of EDR on turnout is identified. While this may strike many readers as implausible, this point is orthogonal to a larger point about interpretation. So if the identification assumptions hold, we may interpret $\hat{\beta}_1$ as the causal effect of EDR on turnout. How should we interpret the other coefficients in this model? Should we treat these parameter like those from the randomized experiment as nuisance parameters? Perhaps while these covariates are needed for identification, we should not interpret the coefficients associated with these covariates? Some statistical methods of adjustment do exactly that. We next review an the nuisance parameter approach to statistical estimation

[2]Other names for this assumption include "conditional ignorability" and "ignorable treatment assignment."

2.1 The Nuisance Parameter Approach

Under the selection on observables assumption, the analysts needs to apply a statistical method to adjust for observed confounders. By far the most common statistical technique applied to data across the social sciences is that of regression. Here, we use the word regression to describe a wide range of statistical techniques that impose very different functional form assumptions. In this essay, we group under the term regression not just the use of least squares but also generalized linear models which includes various types of logit and probit models as well as count models like negative binomial regression. We also include nonparametric regression models as well. Here we group together anything that produces interpretable output for all right hand side covariates. For example, generalized additive models, allow the analyst to specify that continuous right hand side predictors be modeled nonparametrically to allow for arbitrary nonlinear partial associations (Hastie and Tibshirani 1990; Keele 2008). However, for any right hand side covariate with a nonparametric fit one can produce a plot that allows for basic interpretation of the partial association between the covariate and the outcome. The method of kernel regularized least squares (KRLS), which relaxes the additivity assumption from generalized additive models, allows analysts to back out interpretable partial associations for all right hand side covariates (Hainmueller and Hazlett 2013).

There are two primary alternatives to regression models for statistical adjustment: weighting and matching. The method of matching is the more well known than weighting methods. See Sekhon (2009) for an overview of matching estimators in political science. Under weighting methods, propensity scores are used to weight observations to correct for selection bias. The methods often are referred to as inverse probability weighting (IPW) techniques. Weighting methods include both marginal structural models and a class of models known as "doubly robust." With marginal structural models, after the estimation of propensity scores, a second stage regression model is estimate with only the treatment indicator on the right hand side of the model. This second stage regression is weighted by the propensity score estimates (Robins 1999; Robins, Hernan and Brumback 2000; Robins and Rotnitzky 2004). One popular alternative implementation

of weighting methods is not only weight by propensity scores in the outcome model, but to also include covariates in the outcome model as well. The hope is that if either the propensity score model or the outcome model is correctly specified then the estimate of the treatment effect will be consistent. Such methods are said to be "doubly robust" (Robins 1994; Kang, Schafer et al. 2007). See Glynn and Quinn (2010) for an introduction to such methods in political science. Typically under the doubly robust approach, interpretable parameters are not produced for any of the covariates included in the outcome model. There are, of course exceptions. Ho et al. (2007) propose a double robust estimator through matching and regression. Under this approach, interpretable parameters are produced for all right hand side variables.

Usually, the contrast drawn between methods such as matching and weighting and regression is that former methods are nonparametric methods of adjustment while regression is parametric. That is true when one compares matching to linear regression estimated with least squares, but methods like KRLS are also a fully nonparametric form of adjustment. The other difference, which we highlight here, is that under the nuisance parameter approach is that the parameters for covariates that are included as confounders are not estimated. Under both matching and most weighting methods, the parameters for confounders are treated as nuisance parameters and cannot be interpreted since they are not actually directly estimated. In a standard matching analysis, while one might match on many covariates only a single parameter for the treatment is produced. Here, we group statistical methods that only produce a treatment effect estimate the nuisance parameter approach.

Is the nuisance parameter approach to causal estimation just an accident of these methods or actually a feature? As we will argue below, it is a feature. Under the nuisance parameter approach, a single covariate is designated as the treatment, and the analyst only gives that estimate a causal interpretation. When every covariate has an estimated parameter, analysts are often tempted to interpret all these estimates as causal effects. As we outline, below, while a covariate may be a necessary control that needs to be adjusted for, it may be the case that one cannot provide any interpretation for the estimated parameter. To illustrate how this can happen,

we first define some concepts and introduce directed acyclic graphs (DAGs).

3 DAGs and Types of Causal Effects

In an analysis of causal effects, we can broadly define three types of effects: total, direct, and indirect. Here, I use the potential outcomes framework (see, e.g. Rubin 1974), which often referred to as the Rubin Causal Model to define these three types of effects (Holland 1986). In the potential outcomes model, each individual has two *potential outcomes* but only one *actual outcome*. Potential outcomes represent individual behavior in the presence and the absence of a treatment, and the actual outcome depends on the treatment actually occurs. We denote a binary treatment status with $D \in \{0, 1\}$. We use values of 0 and 1 refer to the values that D might take. While the treatment can take on many values, we focus on the binary case for clarity.[3] The potential outcomes are $Y(D) \in \{0, 1\}$, and the actual outcome is is a function of treatment assignment and potential outcomes such that $Y = DY(1) + (1 - D)Y(0)$.

The potential outcomes model formalizes the idea that the individual-level causal effect of a law is an unobservable counterfactual quantity since we define the individual level causal effect as $Y(1) - Y(0)$ (Holland 1986). Next, we define M as a post-treatment variable affected by D. Typically M is referred to as a mediator. Let $M(D)$ denote the potential mediator, the value of the mediator that would be realized under the treatment status d. We define the average causal effect which is equivalent to the total effect of D on Y as

$$\tau \equiv Y(1, M(1)) - Y(0, M(0)). \tag{1}$$

In the examples from the last section, the regression models were providing estimates of this quantity. The total effect can be defined in terms of a mediator, but the total effect holds the mediator constant and does not reveal any information about mediating variables.

[3] We also omit a subscript i that would indicate that these are individual-level variables.

Using this notation, the indirect effect is defined as

$$\delta(d) \equiv Y(d, M(1)) - Y(d, M(0)), \tag{2}$$

for $d = 0, 1$. In this definition, the causal mediation effect represents the indirect effects of the treatment on the outcome through the mediating variable (Pearl 2001; Robins 2003). The indirect effect essentially asks the following counterfactual question: What change would occur to the outcome if the mediator changes from what would be realized under the treatment condition, i.e., $M(1)$, to what would be observed under the control condition, i.e., $M(0)$, while holding the treatment status at d? Although $Y(d, M_i(d))$ is observable for units with $D = d$, $Y(d, M(1-d))$ can never be observed for any unit. Seelmai et al. (2011) for a detailed discussion of why researchers are often interested in indirect effects.

Next, we define the direct effect of the treatment for each unit as

$$\zeta(d) \equiv Y(1, M(d)) - Y(0, M(d)), \tag{3}$$

for $d = 0, 1$.[4] While the indirect effect represents the effect of the treatment through M, the direct effect represents the effect of the treatment through all other possible causal mechanisms. The total effect of the treatment, τ, can be decomposed into the indirect and direct effects in the following manner, $\tau \equiv Y(1, M(1)) - Y(0, M(0)) = \frac{1}{2}\sum_{d=0}^{1}\{\delta(d) + \zeta(d)\}$, where we simply average over the two treatment levels.[5]

Importantly, under selection on observables, we can only identify the total effect. Identification of direct and indirect effects require an additional assumption. Typically analysts use an assumption known as sequential ignorability, which rules out confounding between M and Y

[4] Pearl (2001) calls $\zeta_i(t)$ as *natural direct effects* to distinguish them from *controlled direct effects* of the treatment.

[5] In addition, if direct and causal mediation effects do not vary as functions of treatment status (i.e., $\delta = \delta(1) = \delta(0)$ and $\zeta = \zeta(1) = \zeta(0)$, the assumption often called the no-interaction assumption), then the total effect is the simple sum of the mediation and direct effects, i.e., $\tau = \delta + \zeta$. The total effect is equivalent to the unit-level causal effect of D.

(Imai et al. 2011). Sequential ignorability, like selection on observables, assumes we observe all confounders and is equally untestable. Why is the distinction between total, direct, and indirect effects important? The estimated parameters in a single regression equation may represent either total effects or direct effects. Later we show that when analysts interpret all the estimated parameters in a regression model they may mistake a direct effect for a total effect. Next, we provide a brief review of DAGs to illustrate how analysts understand what types of casual effects are identifiable given the causal structure they think holds.

3.1 DAGs

One tool that is sometimes used in the literature on causal inference is that of causal graphs or DAGs. DAGs are graphical representations of nonparametric structural equation models (Pearl 2009). DAGS are often useful for reasoning about causal structure, since they allow us to formalize identification assumptions for total, indirect, and direct effects. From a given DAG, a researcher can derive which effects are identified and the assumptions needed for identification. This is particularly useful with the selection on observables identification strategy since one can derive what covariates are needed to identify a specific causal effect. Here, we provide a brief overview of DAGs and conclude with an illustration of how DAGs can be used to perform an identification analysis under selection on observables.

In a DAG, a node represents a covariate or vector of covariates. We represent measured variables that we can condition on with a solid circle and unobserved variables with an open circle. When one variable causes another, we draw an arrow between them. Arrows also imply that a variable that precedes another variable has temporal precedence. If there is no arrow that means we are willing to assume that there is no causal relationship between two nodes, which corresponds to an exclusion restriction. Arrows only flow in one direction hence the name acyclic. From a given graph, we can derive which paths are nonparametrically identified, derive expected conditional independencies, and identify which variables or sets of variables are necessary for identification via adjustment.

To provide an overview of some useful constructs that can be encoded in a DAG, we start with Figure 1 which is the simplest possible DAG. This DAG implies that D causes Y, but it also encodes the strong assumption that there are no other covariates that are common causes of both D and Y. The key piece of information in this DAG is that since there are no other variables that also affect Y, this implies that there are no other causes of Y that also cause D. A DAG of this type could be used to represent the causal effect in a randomized experiment since randomization implies that D and Y have no common causes. One key way we encode assumptions into DAGs is by omitting arrows, since that implies that certain causal effects are thought be absent.

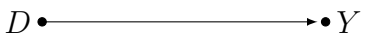

Figure 1: DAG with no common causes for D and Y.

In an observational study, one key threat to identification is the presence of common causes of D and Y. In Figure 2, D causes Y, but both D and Y are common causes of L. If we ignore L or L is unobserved the total effect of D on Y will not be identified. Under the selection on observables identification strategy, analysts attempt to identify all possible common causes of D and Y. If the analyst can condition on all common causes of D and Y identification of the total effect is possible.

Thus far, we have assumed that D has only a total effect on Y. The causal effect of D does not flow through any mediators. Figure 3 is a DAG that represents a causal mechanism. In this DAG, the treatment D has a total effect, that can be decomposed into a direct and indirect component. All three effects are identified in this DAG, since there are no variables, observed or unobserved, that are common causes of D, M, or Y. While Figure 3 is a simple DAG, it represents a subtle but important point. In this DAG, since there are no common causes of D and Y the total effect is identified. Since there are no common causes of M and Y, the indirect effect is identified. It is often the case, that we can identify the total effect but not the direct and indirect effects. If D were as-if random that is enough to identify the total effect but would not allow us to identify the indirect and direct effects. It is only by ruling out common causes for

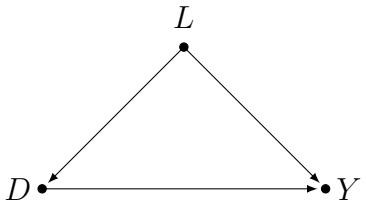

Figure 2: Identification of the Total Effect of T on Y via Selection on Observables

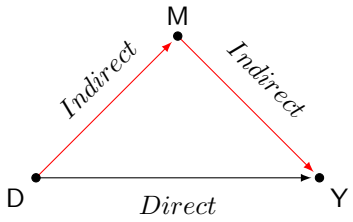

Figure 3: The Direct and Indirect Components of the Total Effect of T on Y.

M and Y that this is possible.

These are simple applications of DAGs that while useful as an introduction do little to illustrate the general utility of DAGs. How might we use DAGs to study applied questions in political science? The basic procedure for using a DAG is as follows. Based on theoretical and scientific knowledge, the analyst writes down a DAG that reflects the causal structure he or she thinks is true. Then based on the DAG drawn by the analyst, one can derive which effects (total, direct, or indirect) are identified from that causal structure, identify the set of covariates that one needs to condition on to identify effects, and derive a set of testable independencies, which can help validate the causal structure in the DAG. Here, all stated identification results are conditional on the validity of the DAG (Pearl 1995). The idea is for researchers to use the DAG to clearly convey what they think is the causal structure derived from theory; this can then be critiqued by other scholars and defended or revised by the researcher.

This process is best communicated through an example. Rogowski (2014) tests whether ideological conflict has a causal effect on voter turnout. He is primarily interested in the total effect of ideological conflict on turnout, but also explores some possible mechanisms. We use his article as an example how one might use a DAG to structure the analysis. We present a somewhat stylized version of his argument to simplify the illustration. In Figure 4, we write down in a DAG

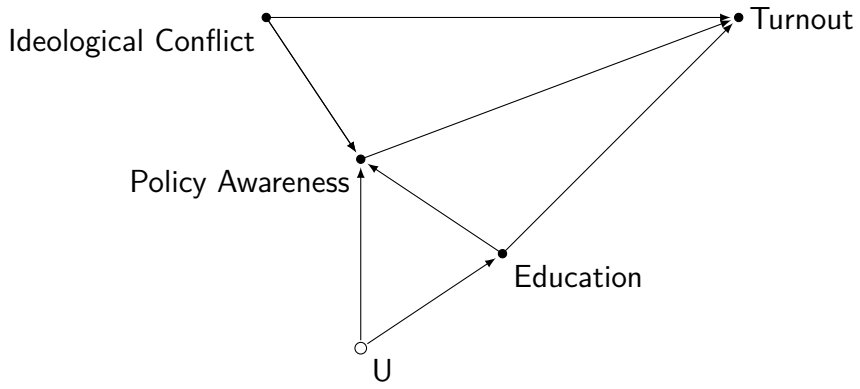

Figure 4: One possible DAG for a model of ideological conflict and voter turnout.

one possible causal structure for this application based on our interpretation of his theory. In this DAG, ideological conflict has both a direct effect on turnout as well as an indirect effect on turnout through policy awareness. This mechanism is one proposed and explored in the original article. Here, we propose that education has a direct and indirect effect on turnout as well. That is education has a causal effect on both turnout and policy awareness. Importantly, this DAG assumes that the policy awareness and ideological conflict do not share any unobservable common causes but education and policy awareness do. This implies that ideological conflict is as-if randomly determined or we might say it is completely exogenous. Unless we can randomize ideological conflict or have a natural experiment, this would be a strong and dubious assumption. Moreover education and turnout also do not share any unobservable common causes, which also implies that education is as-if randomly assigned. While both are strong assumptions, both are implied in the original article. Based on this DAG, what can we learn about identification of the causal effect of ideological turnout?

From this DAG, we can deduce the following implications. First, so long as we condition on education and do not condition on policy awareness the total effect of ideological conflict on turnout will be identified. Importantly, as we noted above, this identification result depends on the validity of the in the DAG in Figure 4. If, for example, education and ideological conflict share a common unobserved cause, this effect would no longer be identified, as the DAG in Figure 4 is no longer valid. In this DAG, we can also identify the direct effects of both policy awareness and

education, but we cannot identify the total effects of either covariate.

Beyond deriving identifiability results, we can also derive testable implications. For example, the DAG also implies that ideological conflict and education should be independent. This serves as a testable implication of this DAG. In the data, we should find that measures of education and ideological conflict are independent of each other. If they are not independent, that serves as empirical evidence that the causal model represented by this DAG is wrong. In this way, DAGs can be used to structure causal inferences. They allow the analysts to understand what it is possible to identify from a particular causal model. Here for example, while we can identify the total effect of ideological conflict, but we can only identify the direct effects of education and policy awareness. We now use DAGs in two different empirical contexts to demonstrate how misplaced interpretation of estimated parameters in regression models can lead incorrect causal inferences.

4 All Cause Models

We now use DAGS in two applications to demonstrate the dangers of interpreting all the estimates from regression models as causal effects. We show that while some effects might be identified, many of the estimates from a regression model are either direct effects or the estimates do not represent effects at all. The key point is that analysts generally need to focus on identification of a single causal effect. The estimates for covariates that are designated as controls should not be interpreted, since that covariate may be a confounder, and we need to condition on that covariate, but the estimated parameter for that covariate may not have a causal interpretation. We start with an example from the literature on the democratic peace.

4.1 The Democratic Peace

For the first empirical application, we use the democratic peace from international relations. In international relations, much research focuses on the factors that promote or inhibit militarized conflict. While a large literature has developed on this topic, much of the focus is on whether democracies are less likely to engage in conflict with other democracies and whether trading

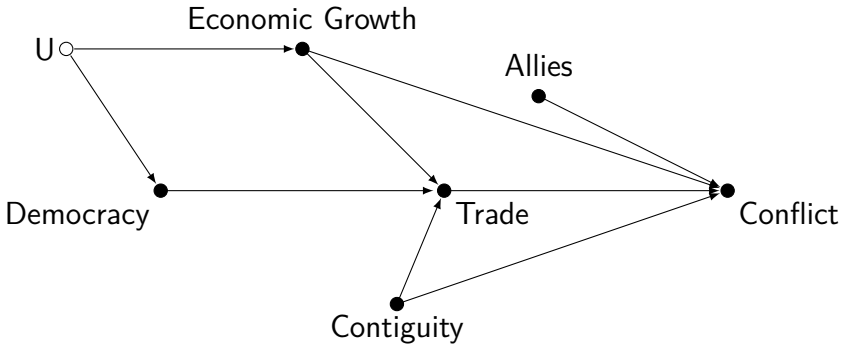

Figure 5: One possible DAG for the effect of ideological conflict on voter turnout.

partners are less likely to engage in conflict with each other. Many researchers have provided evidence for these two propositions (Russett 1990, 1993; Maoz and Russett 1992, 1993; Oneal et al. 1996; Oneal and Russett 1997). In general, it has been found that even controlling for a variety of confounding factors, the foundations for the "Democratic Peace" findings remain intact, suggesting that democratic governance is the best guarantor of peace. The seven factors that are often identified as being important in this literature are: (1) the level of *democracy* in the dyad as measured by a scale of Polity III scores, (2) *economic growth*, (3) the presence of an *alliance* between the two nations in the dyad, (4) geographical *contiguity* in the dyad, (5) the level of intradyadic *trade*.

Figure 5 contains *one* DAG that represents the possible causal structure for interstate conflict. We want to emphasize that a large number of alternative causal models might also be true. This is just one model that strikes us as plausible. We also want to assume that the main causal of effect of interest is the effect of democracy on conflict. The other covariates represent possible confounders, and identification is via the selection on observables. In this DAG, allies and contiguity are exogenous while economic growth, democracy and contiguity are all causes of trade. Under selection on observables, an analyst may attempt to estimate the causal effect of democracy on conflict using the following regression model:

$$Conflict = \beta_0 + \beta_1 Trade + \beta_2 Contiguity + \beta_3 Allies + \beta_4 Democracy + \beta_5 EconomicCapacity. \tag{4}$$

What happens if we treat the model in Equation 4 as an all cause model and attempt to interpret each coefficient in the model. Specifically, what are the consequences of this practice if Figure 5 represents the true causal structure? Using graphical identification techniques, we can deduce exactly with effects are identifiable from the DAG in Figure 5. Below is a summary of what is identifiable assuming the causal structure in Figure 5 is true:

- β_1 Total Effect of Trade is Identified
- β_2 Direct Effect of Contiguity is Identified
- β_3 Total Effect of Allies is Identified
- β_4 Effect of Democracy is Unidentified
- β_5 Direct Effect of Economic Capacity is Identified.

First, if one attempted to fully interpret the parameters in Equation 4 only two of the five parameters represent total causal effects. In this model, the total effect of allies is identified as is the total effect of trade. The effect of trade is identifiable only if we condition on economic growth, contiguity, and democracy. Under this DAG, we need not condition on allies since we are interested in the effect of democracy on conflict, and allies does not aid in identifiability of that effect. Next, if we attempt to interpret the effects of contiguity and economic capacity, the estimated regression parameters only represent direct effects and do not represent total effects. For both covariates, it is not possible to identify the total effect. Finally, the estimated regression parameter for democracy does not represent a causal effect of any kind. While democracy is necessary to identify the effect of trade on conflict, its casual effect is unidentifiable. This represents an important point about specifications. While democracy is necessary as a control variable for the total effect of trade, we cannot give its estimated parameter any causal interpretation. This demonstrates the hazards of attempting to treat all the parameters in a regression model as causal. Unless one understands the causal structure, regression coefficients can represent any number of quantities. I turn to one more example.

4.2 Vote Choice

The topic of vote choice is one where political scientists have long been interested in drawing causal inferences. Early work proposed the "funnel of causality" as one causal model that represented the vote choice process. In general, vote choice in American politics is a function of economic performance, party loyalty, qualities of the candidates, and the where the candidates stand on the issues. It is also well understood that characteristics such as sex and gender also play a role. In Figure 6 we present a DAG that represents the possible causal structure for vote choice. As before, this DAG represents one possible causal model for vote choice, and, for the moment, we assume it represents the true causal structure. One might imagine a much more complex DAG is more realistic, however, we think this DAG captures the essentials of what is understood about vote choice. One could quite easily redraw the DAG such that only the effects of race and gender are identifiable, leaving aside whether statistical associations for nonmanipulable covariates such as these should be given a causal interpretation. In fact, for vote choice, it might be true that causal effects are generally unidentifiable. For now, let's assume Figure 6 is the true causal model.

Let's next assume that we estimate the following regression model:

$$Vote = \beta_0 + \beta_1 PartyId + \beta_2 Female + \beta_3 CandidateQualities + \beta_4 Issues + \beta_5 Race + \beta_6 Economy \quad (5)$$

In the regression model, vote choice is function of all the observed covariates in the DAG. Under standard advice about complete specifications, one would hesitate to leave any of these covariates out of the right hand side of the model. In models of this type, it rarely the case that one focuses on identification of just one of the covaraites in the DAG. From the DAG, we can again derive which effects are identifiable. The following effects are identifiable.

- β_1 Direct Effect of Party Identification on Vote
- β_2 Total Effect of Female on Vote

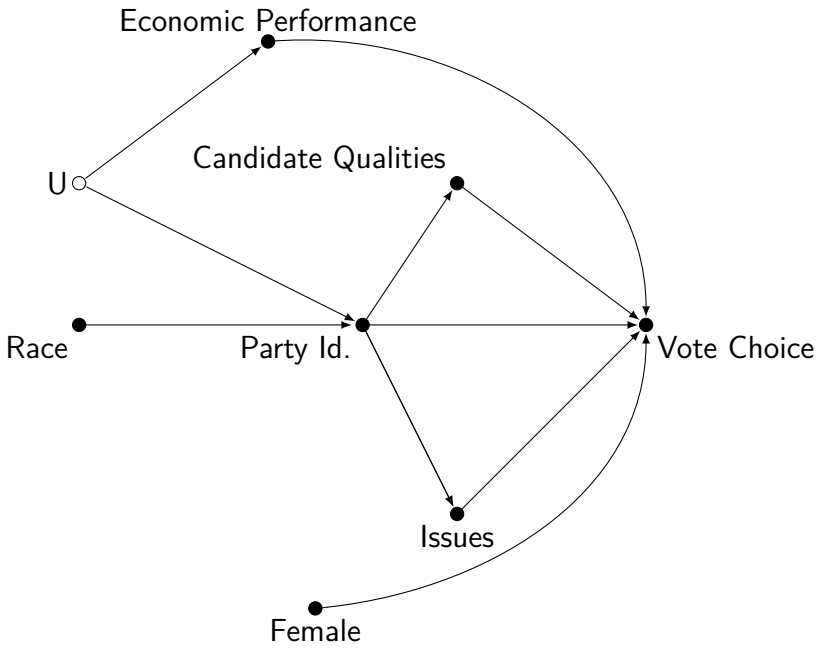

Figure 6: One possible DAG for a model of vote choice.

- β_3 Total Effect of Candidate characteristics on Vote
- β_4 Total Effect of Issues on Vote
- β_5 Total Effect Race on Vote
- β_6 Economic Performance is Unidentified

Of the six estimated parameters, we can identify four causal effects. In this regression model, the estimated parameters for race, sex, candidate characteristics, and issues are all identified. For candidate characteristic and issues, identifiability is contingent on the assumption that these covariates do not share unobserved common causes with party identification or economic performance. For party identification, the estimated parameter represent the direct effect but not the total effect. Finally, the estimated parameter for the effect of economic performance on vote choice does not represent a causal effect.

5 Discussion

We have outlined how while it may be necessary to condition on a covariate, the effect of that covariate itself may not be identifiable. Given this fact, it can be ill-advised to attempt to interpret all the parameters in regression models. Some methods of statistical adjustment only attempt to provide an estimate for a single causal parameter while treating the parameters for all other covariates as nuisance parameters. We contrasted these methods with regression methods that provide estimates for all covariates included in the model. Strictly speaking having estimates for all the covariates included in the model does not itself threaten the identification of causal effects. The danger is if analysts interpret all the parameters from such models as causal effects. Unless one clearly understands the causal structure, such parameters can represent either direct effects or may not be causal effects at all. This fact suggests that analysts should seek to clearly understand which effects they seek to identify whether a covariate is simply a control variable or instead a causal variable. While the use of DAGs is not necessary, they can be helpful to clarify what analyst thinks is the causal structure. The difficulty with DAGs is that it can be hard to reduce a complex causal model into graphical form. Moreover, DAGs often reveal that nothing may be identifiable. One could certainly argue that for both the model of interstate conflict and vote choice, there is no realistic causal model in which effects are identifiable. In general, this note serves as a notice that estimated parameters are not always what they appear. The perils of the all-cause model are great unless care is taken.

References

Barnow, B.S., G.G. Cain and A.S. Goldberger. 1980. Issues in the Analysis of Selectivity Bias. In *Evaluation Studies*, ed. E. Stromsdorfer and G. Farkas. Vol. 5 San Francisco, CA: Sage.

Dunning, Thad. 2009. "Improving Causal Inference: Strengths and Limitations of Natural Experiments." *Political Research Quarterly* 61(2):282–293.

Gerber, Alan S. and Donald P. Green. 2012. *Field Experiments: Design, Analysis, and Interpretation*. New York, NY: Norton.

Gerber, Alan S., Donald P. Green and Edward H. Kaplan. 2004. The Illusion of Learning From Observational Research. In *Problems and Methods in the Study of Politics*, ed. Ian Shapiro, Rogers Smith and Tarek Massoud. New York: Cambridge University Press pp. 251–273.

Glynn, Adam N and Kevin M Quinn. 2010. "An introduction to the augmented inverse propensity weighted estimator." *Political Analysis* 18(1):36–56.

Hainmueller, Jens and Chad Hazlett. 2013. "Kernel Regularized Least Squares: Reducing Misspecification Bias with a Flexible and Interpretable Machine Learning Approach." *Political Analysis* p. mpt019.

Hastie, T.J. and R.J. Tibshirani. 1990. *Generalized Additive Models*. London: Chapman and Hall.

Ho, Daniel E, Kosuke Imai, Gary King and Elizabeth A Stuart. 2007. "Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference." *Political analysis* 15(3):199–236.

Holland, Paul W. 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association* 81(396):945–960.

Imai, Kosuke, Luke Keele, Dustin Tingley and Teppei Yamamoto. 2011. "Unpacking the Black Box of Causality: Learning about Causal Mechanisms from Experimental and Observational Studies." *American Political Science Review* 105(4):765–789.

Kang, Joseph DY, Joseph L Schafer et al. 2007. "Demystifying double robustness: A comparison of alternative strategies for estimating a population mean from incomplete data." *Statistical science* 22(4):523–539.

Keele, Luke. 2008. *Semiparametric Regression for the Social Sciences*. Chichester, UK: Wiley and Sons.

Keele, Luke J. 2014. "The Statistics of Causal Inference."

Keele, Luke J. and William Minozzi. 2012. "How Much is Minnesota Like Wisconsin? Assumptions and Counterfactuals in Causal Inference with Observational Data." *Political Analysis* 21(2):193–216.

Maoz, Zeev and Bruce Russett. 1992. "Alliance, Contiguity, Wealth, and Political Stability: Is the Lack of Conflict Among Democracies A Statistical Artifact?" *International Interactions* 17(3):245–267.

Maoz, Zeev and Bruce Russett. 1993. "Normative and Structural Causes of Democratic Peace, 1946-1986." *American Political Science Review* 87(3):639–56.

Oneal, John R. and Bruce Russett. 1997. "The Classial Liberals Were Right: Democracy, Interdependence, and Conflict, 1950-1985." *International Studies Quarterly* 41:267–294.

Oneal, John R., Frances H. Oneal, Zeev Maoz and Bruce Russett. 1996. "The Liberal Peace: Interdependence, Democracy, and International Conflict, 1950-1985." *Journal of Peace Research* 33:11–28.

Pearl, Judea. 1995. "Causal Diagrams for Empirical Research." *Biometrika* 82(4):669–710.

Pearl, Judea. 2001. Direct and indirect effects. In *Proceedings of the seventeenth conference on uncertainty in artificial intelligence*. Morgan Kaufmann Publishers Inc. pp. 411–420.

Pearl, Judea. 2009. *Causality: Models, Reasoning, and Inference*. 2nd ed. New York: Cambridge University Press.

Robins, James M. 1994. "Correcting for Non-compliance in Randomized Trials Using Structural Nested Mean Models." *Communications in Statistics-Theory and Methods* 23(8):2379–2412.

Robins, James M. 1999. Marginal structural models versus structural nested models as tools for causal inference. In *Statistical Methods in Epidemiology: The Environment and Clinical Trials*, ed. Halloran E. and D. Berry. New York, NY: Springer-Verlag p. 95134.

Robins, James M. and Andrea Rotnitzky. 2004. "Estimation of Treatment Effects in Randomised Trials with Non-Compliance and a Dichotomous Outcome Using Structural Mean Models." *Biometrika* 91(4):763–783.

Robins, James M, Miguel Angel Hernan and Babette Brumback. 2000. "Marginal structural models and causal inference in epidemiology." *Epidemiology* 11(5):550–560.

Robins, J.M. 2003. Semantics of Causal DAG Models and the Identification of Direct and Indirect Effects. In *Highly Structured Stochastic Systems*, ed. P.J. Green, N.L. Hjort and S. Richardson. Oxford: Oxford University Press pp. 70–81.

Rogowski, Jon C. 2014. "Electoral Choice, Ideological Conflict, and Political Participation." *American Journal of Political Science* 58(2):479–494.

Rubin, Donald B. 1974. "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology* 6:688–701.

Rubin, Donald B. 1986. "Which Ifs Have Causal Answers." *Journal of the American Statistical Association* 81(396):961–962.

Russett, Bruce. 1990. Economic Decline, Electoral Pressure, and the Initiation of International Conflict. In *Prisoners of War?: Nation-States in the Modern Era*, ed. Charles Gochman and Ned Allan Sabrosky. Lexington, MA: Lexington Books.

Russett, Bruce. 1993. *Grasping the Democratic Peace*. Princeton, NJ: Princeton University Press.

Sekhon, Jasjeet S. 2009. "Opiates for the Matches: Matching Methods for Causal Inference." *Annual Review of Political Science* 12:487–508.