A Framework for Dynamic Causal Inference in Political Science

Matthew Blackwell   University of Rochester

Dynamic strategies are an essential part of politics. In the context of campaigns, for example, candidates continuously recalibrate their campaign strategy in response to polls and opponent actions. Traditional causal inference methods, however, assume that these dynamic decisions are made all at once, an assumption that forces a choice between omitted variable bias and posttreatment bias. Thus, these kinds of "single-shot" causal inference methods are inappropriate for dynamic processes like campaigns. I resolve this dilemma by adapting methods from biostatistics, thereby presenting a holistic framework for dynamic causal inference. I then use this method to estimate the effectiveness of an inherently dynamic process: a candidate's decision to "go negative." Drawing on U.S. statewide elections (2000–2006), I find, in contrast to the previous literature and alternative methods, that negative advertising is an effective strategy for nonincumbents. I also describe a set of diagnostic tools and an approach to sensitivity analysis.

What candidate would plan all of their rallies, write all of their speeches, and film all of their advertisements at the beginning of a campaign, then sit back and watch them unfold until Election Day? Clearly this is absurd, and yet it is the only setup that the usual ways of making causal inferences allows us to study. While political science has seen enormous growth in attention to causal inference over the past decade, these advances have heavily focused on snapshots where the dynamic nature of politics is crammed into a single point in time. As political science finds itself with a growing number of motion pictures—panel data, time-series cross-sectional data—a tension has emerged between substance and method. Indeed, applied to dynamic data, the best practices of *single-shot* causal inference methods provide conflicting advice and fail to alleviate omitted variable or posttreatment bias.

This article focuses on a specific dynamic process: negative advertising in 176 U.S. Senate and gubernatorial elections from 2000 until 2006. Candidates in these races change their tone over the course of the campaign, reacting to their current environment. A single-shot causal-inference method compares campaigns that are similar on a host of pre-election variables in order to eliminate omitted variable bias. While this is often the best approach with single-shot data, such an approach ignores the fundamentally dynamic nature of campaigns: races that become close over the course of the campaign

are more likely to go negative than those that are safe. Attempting to correct for this dynamic selection by controlling for polls leads to posttreatment bias since earlier campaign tone influences polling. The inappropriate application of single-shot causal inference therefore leaves scholars between a rock and a hard place, steeped in bias with either approach. This dilemma is not limited to negative advertising or campaigns—every field of political science has a variable of interest that evolves over time.

This article solves this dilemma by presenting a framework for dynamic causal inference and a set of tools, developed in biostatistics and epidemiology (Robins, Hernán, and Brumback 2000), to estimate dynamic causal effects. These tools directly model dynamic selection and overcome the above problems of single-shot causal inference. Actions (such as campaign tone) are allowed to vary over time along with any confounding covariates (such as polling). Thus, we can study the effects of the *action history* (candidate's tone across the entire campaign) as opposed to a single action (simply "going negative").

To estimate dynamic causal effects, this article applies inverse probability of treatment weighting (IPTW) to a class of semiparametric models called marginal structural models (MSM). These models dictate the form of the relationship between the large number of possible action histories and the outcome and serve to reduce the number of causal parameters. The dynamic causal effects

---

are encoded as parameters of the MSM and, under certain assumptions, IPTW estimates them free of the biases inherent in single-shot methods. With this approach, each unit is weighted by the inverse of the estimated probability of its observed action history. This weighting creates a pseudosample where dynamic selection is eliminated, circumventing the dilemmas posed by single-shot causal inference. Since these methods require strong assumptions, this article also develops a novel diagnostic tool, the history-adjusted balance, and describes a sensitivity-analysis framework to address potential causes of concern.

Once I correct for the biases due to time, I find that negative advertising is an effective strategy for Democratic nonincumbents. This stands in contrast to the previous literature on negative advertising, which, according to Lau, Sigelman, and Rovner (2007, 1176), "does not bear out the idea that negative advertising is an effective means of winning votes." The previous approaches to estimating the effectiveness of negative advertising relied on single-shot methods, and when I apply these methods to the present data, I find similar noneffects. These single-shot results are worrisome since, as I show below, polls in the middle of a campaign are one of the most important predictors of the decision to go negative. This crucial selection issue is inherently dynamic and has been largely ignored by previous work on this topic.

The article proceeds as follows. The second section describes how dynamic causal inference extends single-shot methods. The third section introduces marginal structural models and the assumptions that they use. The following section describes a weighting approach to estimating dynamic causal effects, and the fifth section applies the techniques to estimating the effectiveness of "going negative" in campaigns. The sixth section discusses useful diagnostics and a sensitivity-analysis framework for marginal structural models. The last section concludes with directions for future research.

Single-Shot and Dynamic Causal Inference

The goal of a single-shot causal inference approach is to estimate the effect of a single *action* on an outcome at a single point in time.[1] With the example of campaigns, we might be interested in the effect of a Democratic candidate running a negative campaign or a positive one on his or her share of the two-party vote. There are many situations in political science, including campaigns, where actions

---

[1] *Action* here is synonymous with *treatment*, a more common term in the causal inference literature.

evolve over time and react to the current state of affairs. In this case, a campaign can "go negative" at multiple points over the course of the campaign. Perhaps a candidate attacks early, before her opponent has a footing, or perhaps she runs negative ads late, responding to smear tactics. These two situations, as far apart as they are, would both register as "going negative" in a single-shot model since they ignore time and implicitly assume that all actions occur at once. This is an acceptable framework for many problems because actions really do occur once. When actions unfold over time, however, the incorporation of time and its implications become necessary.

Actions versus Action Sequences

Dynamic causal inference, in contrast, allows the actions to vary over time. In this framework, we investigate the effect of an *action sequence* on the outcome of interest. In this framework, we have sequences such as (positive$_1$, negative$_2$), where a candidate stays positive in the first part of the campaign and then goes negative later. We might ask how this differs from initiating negativity earlier in the race: (negative$_1$, negative$_2$). This framework has two primary advantages over single-shot methods in dynamic situations. First, the comparison of action sequences naturally handles a richer set of causal questions, which include both the presence and timing of actions. As Pierson (2000) points out, *when* an action occurs is often as important as *if* it occurs at all. Second, and more important, this framework clarifies and resolves certain dilemmas posed by single-shot methods.

Confounders in a Single-Shot World

In order to separate a causal effect from a mere association, we must be confident that the observed correlations are not due to some other variable. In political science, we call this assumption *no omitted variables,* and it is made, implicitly or explicitly, in almost all empirical research in political science. It states that we have measured and appropriately controlled for any variable that could potentially cause bias in our causal estimates. We call this bias *confounding* and the variables that cause it *confounders.* For instance, if we were to run a regression of the Democratic vote share in Senate elections on a measure of Democratic negativity, we would also want to control for variables that might cause a correlation between negativity and vote shares. In this case, the incumbency status of the Democrat would be a confounder because incumbents are less likely to go negative and also more likely to win

FIGURE 1 **Directed Acyclic Graphs Showing Single-Shot and Dynamic Causal Inference Frameworks**

(a) Single-shot causal inference.

(b) Dynamic causal inference.
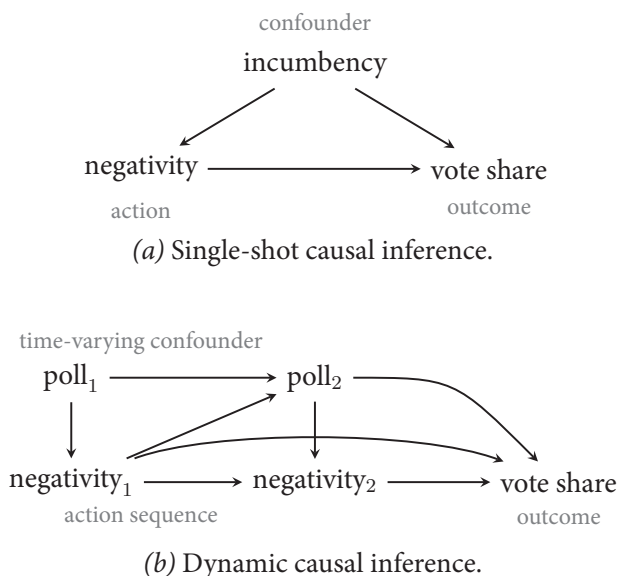
Note: Each arrow represents a causal relationship.

an election. Figure 1a shows a graphical representation of the causal relationship between confounders, actions, and the outcome. We would want to include confounders like these in our analysis, be it using a linear model, a generalized linear model, or a matching estimator.

How do we choose which of our variables are confounders? A common definition is this: a confounder is any variable that (a) is correlated with the outcome, (b) causes, or shares a common cause with, the action, and (c) is not affected by the action. Thus, in any regression or matching estimation of a causal effect, we would want to control for or match on any *preaction variables* in the sense that they are causally prior to the action of interest. In negative advertising, these variables would either affect or be correlated with the decision to go negative, but never be affected by the decision to go negative. We avoid controlling for *postaction* variables because doing so can induce bias in estimating the causal effect. This is known in the causal inference literature as *posttreatment bias* (Ho et al. 2006).

There are two related sources of posttreatment bias. First, conditioning on postaction variables can "block" part of the action's overall effect. For instance, suppose a researcher controlled for polling results from the day of the election when attempting to estimate the effect of incumbency. This will understate the effect of incumbency since most of the effect flows through the standing of candidates late in the race. Second, conditioning on a

postaction variable can induce selection bias even when no bias exists absent the conditioning. For instance, suppose at the start of a campaign we randomly assigned high and low budgets to different Democratic candidates for Senate. If we condition on the polls sometime during the campaign, we can seriously bias our estimates of the effect of campaign budgets. Those leading Democrats who had high budgets are likely to differ strongly from leaders with small budgets. For example, if higher budgets help a candidate, then those low-budget leaders are actually much stronger candidates than the high-budget leaders, since they were able to lead in the polls without the additional funding. Thus, comparing high- and low-budget leaders would give a misleading estimate of the causal effect of campaign finance, even though it was randomly assigned.

The Problem of Time-Varying Confounders

When we force an inherently dynamic situation into a single-shot framework, the above discussion of confounders and posttreatment bias becomes muddled. Take negative advertising, for example: how should we treat polling data from over the course of the campaign? It is surely *preaction* in the sense that polling affects the decision to go negative, and it is correlated with the outcome, the election results. At the same time, polling is *postaction* since it is affected by negativity earlier in the race. Polling is an example of a *time-varying confounder*, which is a confounder that both affects future treatment and is affected by past treatment.

The single-shot advice to include preaction confounders and exclude postaction variables appears to recommend both courses of action in this situation, leaving a researcher without a palatable solution. In fact, both of these approaches will bias causal estimates, albeit in different ways. In the above hypothetical regression of Democratic vote share on Democratic negativity, we could omit polling data from the regression on the grounds that it is posttreatment, yet this would lead to omitted variable bias. Note in Figure 1b that polling in period 2 affects negativity in period 2, perhaps because candidates who are trailing are more likely to go negative. If we exclude polling from our regression (or matching analysis), it might seem that negativity is a bad strategy even though this is wholly due to candidates going negative when they are in trouble. Thus, we must include polling in our analyses. Doing so, however, also biases our estimates, since the polls in period 2 are partially a result of negativity in period 1. For instance, a candidate who stays positive early and whose polls decline might have done better if she had gone negative early. If we control for polling in period 2, we block that part of the effect in our analysis and introduce

posttreatment bias into our estimates. Either approach, ignoring or including polling, will lead to some form of bias. These problems with time-varying confounders cannot be solved by single-shot methods even if we assume, as I do below, no omitted variables in each time period. They are fundamental to situations where actions unfold over time. The assumptions and methods presented here represent a solution to this dilemma under the weakest possible assumptions on the causal structure of the data.[2]

A Weighting Approach to Dynamic Causal Inference

A key characteristic of single-shot methods, such as regression and matching, is that they provide no way of removing omitted variable bias due to time-varying confounders without inducing posttreatment bias. These estimators divide up the data into comparable subsets and estimate causal effects within these subsets. The overall effect is simply a combination of these *stratum-specific* effects. This broad approach is called *stratification,* and it breaks down in dynamic settings as described above: stratifying removes omitted variable bias but induces posttreatment bias for time-varying confounders.

An alternative to stratification estimators are *inverse probability of treatment weighting* (IPTW) estimators, which reweight the data to alleviate the omitted variable bias.[3] To see how these weights work, note that, in negative campaigning, certain strategies are used more often than others: candidates tend to go negative when they are trailing and stay positive when they are leading. In Figure 1b, we represent this in the arrows from polling to negativity. This, of course, causes confounding. To remove this time-varying confounding, we can give less weight to common strategies so that, in the reweighted data, all strategies have the same weight: as many trailers go negative as stay positive. Thus, in the reweighted data, the action sequences are balanced across time-varying confounders, and there is no omitted variable bias. Crucially, Robins (1999) shows that IPTW estimators do not introduce post-

treatment bias because they avoid stratifying the outcome by time-varying confounders.

Alternatives to IPTW estimators for dynamic causal inference include structural nested models, structural equation modeling, synthetic control methods, and principal stratification, but each of these methods has a disadvantage when compared to the weighting approach. The IPTW estimator described below is far more general than the latter three approaches, while being less model dependent than structural nested models. Robins (2000) points out that while these structural nested models can estimate more flexible causal quantities of interest than IPTW, they also require models for each time-varying confounder in the data.[4] While this weighting approach is less flexible than structural nested models, it is much more flexible than other related methods. Structural equation modeling requires a constant effects assumption in order to estimate dynamic causal effects.[5] Synthetic control methods for comparative case studies, for example, focus on a single intervention for each unit and thus limit the number of possible estimable quantities (Abadie, Diamond, and Hainmueller 2010). Principal stratification (Frangakis and Rubin 2002) can recover causal-effect estimates when a posttreatment variable defines the available sample, such as censoring by death. Frangakis and Rubin (2002) note, however, that this approach is more appropriate for nonmanipulable posttreatment variables. When the relevant posttreatment variable is manipulable and truly part of the treatment, as is the case here, principal stratification needlessly restricts the quantities of interest under investigation.

A Framework for Dynamic Causal Inference

To show how IPTW can estimate causal effects in a dynamic setting, it is useful to extend the single-shot causal-inference framework to explicitly include time.[6] Suppose i indexes the campaign, with $i = 1, \ldots, N$. Let t denote the week of the campaign, taking possible values $1, \ldots, T$, where T is the final week before Election Day. We refer

---

[2]We could change our quantity of interest to avoid this problem, but we would be restricting our analysis to the effect of the last action. Furthermore, the assumptions used to identify this quantity of interest would likely be as strong as the assumptions used to identify the full set of causal quantities.

[3]IPTW estimators have a long history in statistics, beginning with the Horvitz-Thompson estimator (Horvitz and Thompson 1952), which has been applied to many problems outside of causal inference, including survey sampling. For an introduction to these estimators for causal inference in political science, see Glynn and Quinn (2010).

[4]For a more detailed discussion of the advantages and disadvantages of IPTW approaches versus g-estimation and structural nested models, see Robins (2000).

[5]See Glynn (2012) for a description of this problem in the context of mediation in linear structural equation models.

[6]The following section rests heavily on the potential outcomes-based model of causal inference championed by Rubin (1978) and extended to dynamic settings by Robins (1986, 1997).

to $t = 1$ as the "baseline" time period; it is the time period before the campaign begins, assumed to be the first week after the primary. In each period, campaigns can either go negative, denoted $A_{it} = 1$, or remain positive, $A_{it} = 0$.

Campaigns face a rapidly evolving environment. To account for this, let X_{it} represent the characteristics of the campaign in week t that affect the Democrat's decision to go negative in week t. This would include recent polling or Republican negativity in the previous weeks. This definition assumes that the decision to go negative occurs "after" the variables in X_{it}, so that they are preaction for week t.[7] Instead of containing all variables occurring at time t, the set of covariates describes the information setting for the action decision at time t. Simply put, X_{it} is the most recent set of variables that could possibly affect A_{it}.[8] The baseline covariates, X_{i1}, include background information that remains static over the course of the study. For campaigns, these could be perceived competitiveness of the election, number of ads shown in the primary, incumbency status, or challenger quality. The choice of relevant covariates of course depends on the outcome, Y, which in this case is the Democratic percent of the two-party vote.

Dynamic settings require references to the *history* of a variable. A history is the set of all instances of that variable up to some point in time. In this example, it may be the sequence of campaign tone or poll results in each week. Underlines indicate the history of a variable, so that $\underline{A}_{it}$ would be the negativity up through time t: $\underline{A}_{it} \equiv (A_{i1}, A_{i2}, \ldots, A_{it})$. The covariate history, $\underline{X}_{it}$, is defined similarly. One possible realization of $\underline{A}_{it}$ is $\underline{a}_t \equiv (a_1, \ldots, a_t)$, where each a_t can take the values 0 or 1. Furthermore, let $\underline{A}_i = \underline{A}_{iT}$ be the sequence of negativity over the course of the entire campaign. Let $\underline{a}$ be a representative campaign tone history and $\underline{\mathcal{A}}$ as the set of all possible values of $\underline{a}$; that is, all the possible ways a candidate could go negative over the course of the campaign. Let $\underline{X}_i$, $\underline{x}_t$, and $\underline{x}$ be defined similarly for the covariate history.

Each possible negativity sequence, $\underline{a}$, has an associated potential electoral outcome. Let $Y_i(\underline{a})$ be the Democratic percent of the two-party vote if we forced candidate i to implement the campaign $\underline{a}$. Note that there are 2^T pos-

sible sequences $\underline{a}$. As before, any individual candidate can experience at most one of these potential outcomes, which is the one associated with her observed action history. The rest of the potential outcomes will be counterfactual; they are *what would have happened* if the unit had followed a different sequence. Suppose campaigns only lasted two weeks. In this world, $Y_i(0, 1)$ would be the Democratic vote share if candidate i were to remain positive in week one and go negative in week two. To complete the definition of the potential outcomes, we connect them to the observed outcomes, Y_i. When some unit is observed to have followed action sequence $\underline{a}$, then we observe the potential outcome for that sequence, or $Y_i(\underline{a}) = Y_i$ when $\underline{A}_i = \underline{a}$.[9]

We say that $\underline{X}_t$ contains a time-varying confounder if it (1) affects the election outcome, (2) affects future negativity, and (3) is affected by past negativity. In estimating the effect of Democrats going negative, the advertising tone of the Republican would be a time-dependent confounder. Democrats are more likely to go negative if their opponent has gone negative, and their opponent's actions are likely related to the outcome. Note that X_t could include past values of Y, in which case the lagged dependent variable would be a time-dependent confounder.

The Assumptions

The goal of dynamic causal inference is to estimate the means of the potential outcomes under various action sequences. These are population-based quantities of interest: what would happen if *every* Democrat remained positive? In the sample, however, the candidates who actually went negative always might be different from those who did not. Thus, the sample of units who followed the strategy would be an unrepresentative sample of the potential outcome under that strategy. In order to rid our analysis of the above selection problems, we must be able to identify and measure all possible confounders.

Assumption 1 (Sequential Ignorability). *For any action sequences $\underline{a}$, covariate history $\underline{X}_t$, and time t, if $\underline{A}_{t-1} = \underline{a}_{t-1}$, then $Y(\underline{a}) \perp\!\!\!\perp A_t | \underline{X}_t, \underline{A}_{t-1} = \underline{a}_{t-1}$.*

Here, $B \perp\!\!\!\perp C | D$ means that B is independent of C, conditional on D, and notation for units has been

---

[7] The causal ordering here is notationally arbitrary as its reversal would require only a change in subscript. More crucially, researchers must determine what information is pre- and postaction in a given period for the substantive question at hand.

[8] Note that these variables are possibly affected by past treatment, but I suppress the potential outcome notation for the covariates for exposition as it adds no consequences for the present discussion.

[9] This is commonly referred to as the consistency assumption in the epidemiology literature (Robins 1997). It implicitly assumes what Rubin (1978) refers to as the stable unit treatment value assumption, or SUTVA. Recent work has further formalized and clarified this assumption (Cole and Frangakis 2009; Pearl 2010; VanderWeele 2009).

FIGURE 2 Directed Acyclic Graphs Representing Different Assumptions about Sequential Ignorability, Where U Is an Unobserved Variable
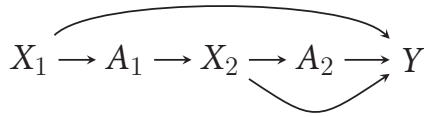
(a) Sequential ignorability holds.
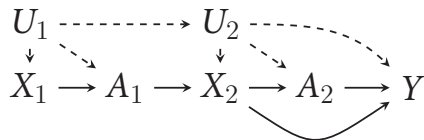
(b) Sequential ignorability fails to hold.

suppressed. The assumption of sequential ignorability extends the conditional ignorability assumption to time-varying actions. It states that action decision at time t is independent of the potential outcomes, conditional on the covariate and action histories up to that point. That is, conditional on the past, those who go negative are similar to those who stay positive. Figure 2a shows a causal directed acyclic graph (DAG) in which sequential ignorability holds, while Figure 2b shows a situation where the assumption fails to hold due to an omitted variable U. If decisions are made by a coin flip, then clearly this assumption will hold. If units act based on the covariate history, however, then it will fail to hold unless the analyst can observe all of those covariates. For instance, the assumption would be violated if campaigns made the decision to go negative based on polling data, but the analyst did not have access to that polling data. The goal for researchers, then, is to collect all the covariates that might influence the decision to go negative in some week. While this is a daunting task in an observational study, it is no harder than satisfying conditional ignorability in the single-shot case, and the sixth section shows how to relax the assumption in a sensitivity analysis.

Finally, in order to compare the various action sequences, each must have some positive probability of occurring. It is nonsensical to estimate the effect of a sequence that could never occur.

Assumption 2 (Positivity). *For any sequences* $\underline{a}_t = (\underline{a}_{t-1}, a_t)$ *and* $\underline{x}_t$, *and time t, if* $\Pr(\underline{A}_{t-1} = \underline{a}_{t-1}, \underline{X}_t = \underline{x}_t) > 0$, *then* $\Pr(A_t = a_t | \underline{X}_t = \underline{x}_t, \underline{A}_{t-1} = \underline{a}_{t-1}) > 0$.

This assumption outlines the types of strategies we can study. Positivity can break down when some sequences fail to occur in the actual data even though they are theoretically possible. In negative advertising, for instance, candidates with extremely safe seats never go negative, even though nothing is stopping them from doing so. Unfortunately, we will be unable to estimate the effect of going negative for these candidates. These empirical violations of positivity are closely related to the assumption of *common support* often invoked in the matching literature. The sixth section discusses these practical problems with positivity and how to restrict the analysis to the common support.

Marginal Structural Models for the Potential Outcomes

In the single-shot approach, estimating a causal effect only involves two quantities, one corresponding to each action: $E[Y_i(1)]$ and $E[Y_i(0)]$. In dynamic causal inference, there is one potential outcome for each action sequence. A key consequence is that even with a small number of time periods, there will be an overwhelming number of possible action sequences. With two potential outcomes, we can nonparametrically estimate the mean outcome in the treated and control groups by taking sample means. With just 10 periods, however, there would be 1,024 possible action sequences, making it unlikely that there will be even one unit following any particular sequence. Thus, the nonparametric approach of single-shot methods will be useless here.

To overcome this curse of dimensionality, we can use a parametric model to relate the action sequences to the potential outcomes. That is, we will suppose that "similar" action sequences should have "similar" potential outcomes. Imposing this structure on the problem reduces the dimensionality of the problem at the expense of possible model misspecification. Robins, Hernán, and Brumback (2000) introduced a parsimonious class of semiparametric models for this problem called marginal structural models (MSM). In this class of models, we assume a parametric form for the mean of the potential outcome

$$E[Y(\underline{a})] = g(\underline{a}; \beta), \qquad (1)$$

while leaving the rest of the distribution of $Y(\underline{a})$ unspecified. This model is semiparametric in the sense that it

leaves unrestricted the relationship between the outcome and the covariates, though it does require a model of the relationship between the actions and the covariates.

The function g defines our assumptions about which action sequences should have similar potential outcomes. We may have, for instance,

$$g(\underline{a}; \beta) = \beta_0 + \beta_1 c(\underline{a}), \qquad (2)$$

where $c(\underline{a}) = \Sigma_{t=0}^T a_t$ is the cumulative action. This model assumes that units with the same number of total periods acted should have similar potential outcomes, with β_1 as the causal effect of an additional period of the action. In the context of negative campaigning, β_1 is the effect of an additional week of negativity. An assumption here is that going negative for the first five weeks of the campaign is the same as going negative for the last five weeks of the campaign. Depending on the application, this might be a more or less plausible assumption, and, in general, these types of modeling assumptions will always produce some amount of bias. The greater flexibility we allow for $g(\underline{a}; \beta)$, however, the more variable our estimates become. The substance of the problem and the amount of data on hand will determine what model makes sense for the potential outcomes.

Supposing that Equation (2) was the correct model for the potential outcomes, we want to estimate its causal parameters. One approach would be to estimate

$$E[Y|\underline{A} = \underline{a}] = \gamma_0 + \gamma_1 c(\underline{a}), \qquad (3)$$

which omits any covariates X_t and simply regresses the outcome on the observed action. This approach replaces the potential outcomes $Y(\underline{a})$ with the observed outcomes Y, holding the model fixed. If X_t affects the action and the outcome, however, the associational parameter, γ_1, will not equal the causal parameter, β_1, due to omitted variable bias. That is, differences in the observed outcomes could be due to difference in the covariate history, not the action sequence. We could instead condition on X_t by estimating

$$E[Y|\underline{A} = \underline{a}, X_t] = \delta_0 + \delta_1 c(\underline{a}) + \delta_2 X_t, \qquad (4)$$

either through a regression that includes X_t or a matching algorithm that matches on X_t. The key parameter, δ_1, will still fail to equal the causal parameter of interest, β_1, when X_t is a time-varying confounder, since X_t is posttreatment for $\underline{A}_{t-1}$. Thus, X_t is in the difficult position of being both an omitted variable *and* a posttreatment variable for the action history. These traditional methods of estimating β_1 fail in the face of time-varying confounders, whether or not we adjust for them, since either approach leads to bias.

One might think that the two traditional estimation procedures would at least provide bounds on the true causal effect, with β_1 falling between γ_1 and δ_1. When the omitted variable bias and the posttreatment bias have the same sign, however, this bounding will fail to hold. This can occur, for instance, when strategic actors attempt to compensate poor performance with beneficial actions. Suppose that there is a strong, positive effect of negative advertising and that trailing campaigns use it to bolster their positions. The omission of polling in a model would lead to an understatement of the negativity effect, since candidates tend to be trailing when they go negative. Positive campaigns would appear stronger than negative campaigns, even though negativity boosts performance. The inclusion of polling in a model would also lead to an understatement of the effect, since it washes away the increase in polls from past negativity. Thus, the true effect of negativity would be higher than either of the traditional methods would predict. Robins (1997) gives a numerical example that has these features.

The Action Model and Inverse Probability of Treatment Weighting

As shown above, the usual single-shot approaches break down when the actions can vary over time. Fortunately, inverse probability of treatment weighting (IPTW) can recover unbiased estimates of causal effects, even in dynamic settings. To see how IPTW works, note that, due to the omitted variables, the distribution of the potential outcomes differs from the distribution of the observed outcomes ($E[Y(\underline{a})] \neq E[Y|\underline{A} = \underline{a}]$). Regression and matching attempt to avoid this problem by finding subsets of the data where those distributions are the same and making comparisons within these subsets. This conditioning removes the omitted variable bias, but it can induce posttreatment bias. Methods that rely on weighting, such as IPTW, avoid these by never explicitly conditioning on the confounders in the outcome model.[10]

Robins, Hernán, and Brumback (2000) show that under the above assumptions, a reweighted version of the observed outcomes will have the same distribution as the potential outcomes. In the campaigns context, the reweighted outcomes for always-positive campaigns will

---

[10]This approach is similar in spirit to Heckman selection models (Achen 1986; Heckman 1976) in the sense that they separate out the selection model (who goes negative) and the outcome model (how negativity affects vote shares). These types of methods, however, rely on instrumental variables. Unfortunately, instrumental-variable methods require effects to be constant in the population in order to estimate the average causal effect (Angrist, Imbens, and Rubin 1996). The approach described here makes no such assumptions.

look like the outcomes if we forced all Democrats to remain positive. The weights in a given week are defined as

$$W_{it} = \frac{1}{\Pr(A_{it}|\underline{A}_{it-1}, \underline{X}_{it})}. \tag{5}$$

In words, the denominator of W_{it} is the probability of observing the action that unit i actually took in that week, conditional on the past. To generate an overall weight for each race, we simply take the product of the weekly weights over time:

$$W_i = \prod_{t=1}^{T} W_{it}. \tag{6}$$

A simple example helps to explain the construction of the weights. Suppose that there were only two weeks in a campaign, with a poll update in between the weeks. A candidate decides to go negative or stay positive in the first week, sees the outcome of the poll, decides to go negative in the second week, and then observes the election results. A candidate who stays positive in week one, trails in the polls, and then goes negative in week two would have the following weight:

$$W_i = \frac{1}{\Pr(\text{pos}_1)} \cdot \frac{1}{\Pr(\text{neg}_2|\text{trail}, \text{pos}_1)}. \tag{7}$$

The first term in the denominator is simply the probability of being positive in the first week. The second term is the probability she would have gone negative in the second week, conditional on trailing and having been positive in the first week. The resulting denominator is the probability of observing the campaign (pos_1, neg_2), conditional on the time-varying covariate, polls.

Why Weighting Works

The weights in IPTW remove any confounding by ensuring that the distribution of action sequences is the same in each level of the confounder. In the reweighted data, the action decisions are unrelated to the measured confounders, and, thus, they cannot account for any remaining differences between action sequences. It is instructive to see how this works in the single-shot case. Let $\Pr_W(\text{neg}|\text{trail})$ be the reweighted probability of observing a negative candidate, conditional on trailing in the polls. We can find this probability by multiplying the original probability by the weight for this type of observation and divide by a normalizing constant:

$\Pr_W(\text{neg}|\text{trail})$

$$= \frac{\frac{1}{\Pr(\text{neg}|\text{trail})} \cdot \Pr(\text{neg}|\text{trail})}{\frac{1}{\Pr(\text{neg}|\text{trail})} \cdot \Pr(\text{neg}|\text{trail}) + \frac{1}{\Pr(\text{pos}|\text{trail})} \cdot \Pr(\text{pos}|\text{trail})} \tag{8}$$

$$= \frac{1}{2}.$$

The denominator simply ensures that the reweighted probabilities will sum to one. Using the same logic, it is clear to see that $\Pr_W(\text{pos}|\text{trail}) = 1/2$ as well. Thus, in the reweighted data, a race is equally likely to go negative as stay positive when they are trailing.

Intuitively, this weighting breaks the links between the action decision and the factors that affect the action decision. Candidates who are pursuing common strategies, where $\Pr(A_{it}|\underline{A}_{it-1}, X_{it})$ is closer to 1, will have lower weights than those candidates with less common strategies. This weighting corrects the deviations from the ideal experiment we would have liked to run. The sequential ignorability assumption is crucial here because we cannot correct for deviations we do not observe. Since there is no connection between the action sequence and the confounders in the reweighted data, we can simply run whatever model we wanted to run in the first place, without conditioning on time-varying confounders. And because we never condition on these variables, we never introduce posttreatment bias as with single-shot approaches.

Estimating the Weights

Of course, without randomization, the probability of going negative will be unknown, leaving Equation (5) to be estimated. To do so, we must model the decision to go negative in each week, conditional on the past. Since the decision is dichotomous, a common approach is to estimate the probability of going negative with a logit model:

$$\Pr(A_{it} = 1|\underline{A}_{it-1}, \underline{X}_{it}; \alpha)$$
$$= (1 + \exp\{-h(\underline{A}_{it-1}, \underline{X}_{it}; \alpha)\})^{-1}, \tag{9}$$

where h is a linear, additive function of the action history, covariate history, and parameters α. For instance, we might have

$$h(\underline{A}_{it-1}, \underline{X}_{it}; \alpha) = \alpha_0 + \alpha_1 A_{it-1} + \alpha_2 X_{it} + \alpha_3 t, \tag{10}$$

which models the action decision as a function of negativity in the last week (A_{t-1}), the most recent poll results (X_{it}), and the week of the campaign (t).

An estimate of the weights requires an estimate of the parameter vector α from this model. We can obtain these estimates, $\hat{\alpha}$, from a pooled logistic regression, treating

each campaign-week as a separate unit. These estimates form the basis for the estimated weights,

$$\widehat{W_i} = \prod_{t=1}^{T} \frac{1}{\Pr(A_{it}|\underline{A}_{it-1}, \underline{X}_{it}; \hat{\alpha})}, \qquad (11)$$

where the denominator is now the predicted action probability (or fitted value) for unit i at each time period.[11] Note that it is not necessary to estimate the same model for all units. For example, incumbents and nonincumbents might require different models because their approaches to the negativity decision are so distinct.

Each observation i is then weighted by $\widehat{W_i}$ in a weighted generalized linear model for the outcome, with form $g(\underline{a}; \beta)$ from Equation (2).[12] Robins (2000) shows this estimation procedure is consistent for the causal parameters, β, under sequential ignorability, positivity, and the correct model for the weights. The most straightforward way to estimate standard errors and confidence intervals is to bootstrap the entire estimation procedure, including the weights (Robins, Hernán, and Brumback 2000). For negative campaigning, this means resampling the set of campaigns (not the set of campaign-weeks), reestimating the weights, and running the weighted outcome model on the resampled data.

Stabilized Weights

If campaigns have vastly different likelihoods of going negative, then the estimated weights from Equation (11) can have extreme variability, which results in low efficiency. We can use a slightly different version of the weights, called the *stabilized weights*, to decrease this variability and increase efficiency. The stabilized weights take advantage of an interesting fact: the numerator of the weights does not change the consistency of the estimation procedure.[13] While it was natural to use the value 1 as the numerator, we can replace it with other functions of the action history that increase efficiency. The usual choice used in the literature (Robins, Hernán, and

Brumback 2000) is

$$SW_i = \prod_{t=1}^{T} \frac{\Pr(A_{it}|\underline{A}_{it-1}; \delta)}{\Pr(A_{it}|\underline{A}_{it-1}, \underline{X}_{it}; \alpha)}, \qquad (12)$$

where the numerator is a model for the marginal probability of action, conditional on past action. When actions are randomized, these stabilized weights will be equal to one since the action probability would be unaffected by the covariates in the denominator.

Of course, the numerator of SW_i is unknown, leaving us with the task of estimating δ. All this requires is an additional logit model for the numerator to estimate the probability of going negative without conditioning on the time-varying covariates. If the outcome model will include interactions with baseline covariates, then both the numerator and the denominator should include those variables. To construct these weights, one simply needs to obtain predicted probabilities from each model for every unit-period. Then, for each unit, take the product of those probabilities across time periods and divide to obtain the estimates $\widehat{SW_i}$.

Estimating the Effect of Going Negative

Pundits and theorists often bemoan the growth in negative campaign advertising in recent decades. Less often do they discuss its effectiveness. An implicit assumption in the air of political discourse is "Of course it works, politicians do it." The prospect of dirtying the waters with such cheap and tawdry tactics is bad enough, and being useless would only add insult to injury. A contingent of political scientists has investigated just how useful negativity is for candidates, without reaching a consensus. In a comprehensive review, Lau, Sigelman, and Rovner describe the state of the literature: "All told, the research literature does not bear out the idea that negative campaigning is an effective means of winning votes" (2007, 1176).

The usual approach to estimating the effectiveness of negative advertising (see Lau and Pomper [2002, 2004] for examples) is to regress election outcomes on a summary measure of the degree of negativity in a campaign along with controls for various static attributes of the race. A crucial problem for these investigations is that campaign tone is a dynamic process, changing from week to week. Furthermore, there are strong time-varying confounders. For instance, poll numbers affect the decision to go negative, but going negative also affects poll numbers. Thus, polling is both pre- and

---

[11] These values are easily found using a combination of the `glm` and `predict` functions in R (R Development Core Team 2011).

[12] The `survey` package in R can implement this weighting for a large class of outcome models (Lumley 2004).

[13] As shown in Robins (2000), the numerator only alters the marginal distribution of the action $\underline{A}$, which does not affect the marginal distribution of the potential outcomes, $Y(\underline{A})$. This is because the marginal distribution of $\underline{A}$ does not affect the distribution of Y conditional on $\underline{A}$, which is the crucial ingredient for the distribution of the potential outcomes.

postaction: a classic time-varying confounder. As shown above, ignoring the polls and conditioning on the polls will both result in biased estimates. We can estimate the effect of time-varying actions, though, using marginal structural models and inverse probability of treatment weighting.

The goal of this application is to estimate the effect of going negative for Democratic candidates in state wide elections. I use data on campaigns for Senate and gubernatorial seats in the cycles of 2000, 2002, 2004, and 2006. For each campaign, I code the advertising tone using data from the University of Wisconsin Advertising Project (Goldstein and Rivlin 2007). To ensure consistency across years, I use a simple measure of negative or contrast ads: does the ad mention the opposing candidate?[14] I use this coding to construct a measure of whether a candidate has "gone negative" in a given week of the campaign based on what percentage of ads is negative.[15] The WiscAds data also provide a proxy for weekly campaign spending: the total number of ads aired in a week. In addition to advertising data, I also collected weekly polling data from various sources,[16] along with baseline covariates, such as predicted competitiveness of the race (as measured by the *Congressional Quarterly* score), incumbency status, number of ads run by each candidate in their primaries, the length in weeks of the campaign, measures of challenger quality and incumbent weakness, and the number of congressional districts in the state. Much of these data come from Lau and Pomper (2004) with additional data collection. In this example, baseline is the day after the final primary.

A Model for Going Negative

In order to estimate the causal parameters from an MSM, we must construct the weights from the previous section. In order to satisfy the assumption of sequential ignorability, we must gather as many covariates as possible that might influence the decision to go negative in a given week and are correlated with the election outcome. This is, of course, a difficult task, but we can often leverage substantive knowledge to guide our models. The dynamic reasons for a candidate to go negative might be numerous, but it is likely that the state of the race, as summarized by the polls in a given week, is at worst a proxy for these factors and at best the most important factor. Indeed, we might think that the candidates doing the worst in the polls are the most likely to go negative. This is why it is crucial to include both precampaign measures of competitiveness and dynamic measures of campaign performance in the form of polling data. Without these data, it would be impossible to differentiate between the effect of negative advertising on the one hand and negative advertising simply indicating weak candidates on the other. Previous literature on the effectiveness of negative advertising has not had access to the kind of polling data available in more recent elections, hampering its ability to address these dynamic-selection issues.

To address these concerns, I include the following time-varying covariates in the weighting model: the Democratic share of the polls in the last week, the share of undecided voters in the last week, past negativity of both Democrats and Republicans, the amount of advertising by both candidates in the last week, and the week of the campaign. It may be the case that these variables do not encompass or proxy all of the factors that influence candidates, which is why it is crucial to assess any inferences using sensitivity analysis, as I do in the next section. With these covariates, I ran two separate pooled-logistic models for the decision to go negative: a separate numerator and denominator model.[17]

These models largely fit with the intuition and theory of campaigns, with high advertising and already negative races being more likely to be negative. Figure 3 shows that there is a strong relationship between polling and the decision to go negative: nonincumbent Democrats in safe seats rarely go negative, but those who are trailing often do. To construct the weights, I combine predicted probabilities from these models according to Equation (12).[18] Due to empirical violations of positivity, I restricted the analysis to common support on baseline covariates,

---

[14]The WiscAds project failed to collect data in 2006, so I acquired and computer-coded the data directly from CMAG, the consultant group which provides the data to WiscAds.

[15]For the analysis below, I used a cutoff of 10%. The results appear unaffected by this choice, as weeks tend to be dominated by only a few ads. If there is one ad that is negative, it pushes the percent negativity quite high.

[16]Polling data come from the *Hotline* daily political briefing for 2000 and 2002 and from http://www.pollster.com for 2004 and 2006.

[17]In order to stabilize the weights further, I include all baseline covariates in both the numerator and denominator models. This means that IPTW will only balance the time-varying covariates, leaving any remaining baseline imbalance. Since this imbalance is time-constant, we can remove it through traditional modeling approaches, and, thus, I include those covariates in the outcome model below.

[18]The weight models are pooled logistic generalized additive models (GAMs), which is what allows for the flexible modeling of the polling. I used the `mgcv` package to fit this model (Wood 2011).

FIGURE 3 The Marginal Relationship between Lagged Polling Numbers and Going Negative for Democratic Nonincumbent Candidates
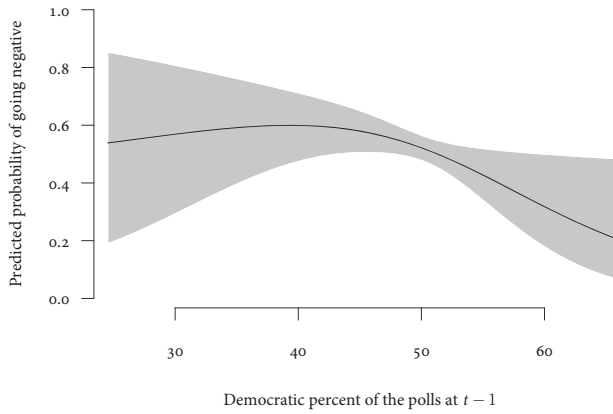
Note: All other variables from the model are held at their mean, or median, depending on the type of variable. The shaded region is a 95% confidence band. Intuitively, trailing Democrats are more likely to go negative than leading Democrats.

which mostly involved removing extremely uncompetitive race.[19]

The Time-Varying Effects of Negativity

The effect of negative advertising is unlikely to be constant across time. Ads closer to Election Day should have a stronger impact than those earlier in the campaign, and marginal structural models allow us to estimate these time-varying effects. I break up the effect into an early campaign effect (the primary through September) and a late campaign effect (October and November). Vote shares are continuous, so a linear MSM is appropriate for the potential outcomes:

$$
\begin{aligned}
E[Y_i(\underline{a})] = {} & \beta_0 + \beta_1 \left(\sum_{T-5}^{T} A_{it} \right) + \beta_2 Z_i \\
& + \beta_3 Z_i \left(\sum_{T-5}^{T} A_{it} \right) + \beta_4 \underline{A}_{iT-6} \\
& + \beta_5 Z_i \, \underline{A}_{iT-6} + \beta_6 X_i,
\end{aligned}
\tag{13}
$$

where Z_i is an indicator for being a Democratic incumbent, X_i is a vector of baseline covariates, and T is the

[19] Note that weeks in which a candidate runs *no* ads are weeks when a candidate *cannot* go negative. These weeks receive a stabilized weight of one, meaning they do not contribute to the weight of their overall campaign. A more thorough analysis would treat the number of ads and the tone of those ads as a joint treatment.

TABLE 1 Estimated Effects of an Additional Week of Negative Advertising in the Last Five Weeks of the Campaign on the Democratic Percent of the Two-Party Vote

Estimator	Democratic Incumbent	Democratic Nonincumbent
Naïve	−0.92 (−1.88, −0.17)	0.43 (−0.21, 1.03)
Control	−0.55 (−1.37, 0.11)	0.50 (−0.11, 1.11)
IPTW	−0.57 (−1.42, 0.24)	0.64 (0.09, 1.12)

Note: Bootstrapped 95% confidence intervals are in parentheses, with those crossing zero set in gray. Inverse probability weighting estimates a strong, positive effect for nonincumbents and a strong, negative effect for incumbents. Note that the competing models fail to bound the IPTW-estimated effect.

week of the election. The summation terms calculate how many of the last five weeks of the campaign the Democrat went negative. This covers October and early November, which is the homestretch of the campaign. The model separately estimates the direct effect of earlier negativity and allows for incumbent status to modify both early and late effects. Following the IPTW approach, I weight each campaign using weights constructed from the above going-negative model.

It is instructive to compare estimates from this model with two competing approaches. First, the *naïve estimator* simply ignores all time-varying covariates and fits Equation (13) to the observed data without weights. Second, the *control estimator* attempts to control for the covariates by including them as additional regressors in Equation (13).[20] These represent the two single-shot methods used by applied researchers: the naïve estimator to guard against posttreatment bias and the control estimator to guard against omitted variable bias.

Table 1 shows the estimated effects of late campaign negativity from all three models broken out by incumbent status.[21] The MSM finds that Democratic incumbents are hurt by going negative, while nonincumbents are helped. Nonincumbents see a 0.64 percentage point increase in the Democratic percent of the two-party vote for every additional week of negative advertising in the last five weeks. Incumbents, on the other hand, drop 0.57 percentage points for the same change. As Figure 4 shows, there is no evidence of a direct effect of earlier negativity

[20] Polls are included as the mean Democratic poll percentage, total number of ads as the average ads per week, and Republican negativity as the overall duration of Republican negativity.

[21] Estimates here are produced using the `svyglm` function in the `survey` package, version 3.22-4 (Lumley 2010). Standard errors and confidence intervals come from bootstrapping this model.

on the final vote in either group. Note that these results control for polls taken at the beginning of the campaign. It is surprising, then, to see effects that are even this large since these baseline polls are highly predictive of the outcome in Senate and gubernatorial elections.

When we compare the three estimators, we find that using a dynamic causal inference approach leads to substantively different conclusions about negative advertising. For nonincumbents, for instance, the two single-shot methods recover what the previous literature has found: no significant effect of negativity (Lau and Pomper 2004; Lau, Sigelman, and Rovner 2007). With the MSM, however, we find that negativity does have a large, statistically significant, and positive effect. Interestingly, the MSM-estimated effect is well outside of the bounds set by the naïve and control estimators. For these nonincumbents, the IPTW estimate is over 18% larger in magnitude than either of the other methods. Thus, "trying it both ways" would be an unsuccessful strategy in this case. For incumbents, we find additional divergence from the literature. Similar to the effects of negativity for incumbents estimated by Lau and Pomper (2004), the naïve estimator finds a large and harmful effect of negativity for incumbents. But this is likely driven by midcampaign confounding, and once I account for this using an MSM, I find no significant effect of negativity for incumbents. Overall, a dynamic causal inference approach leads to different conclusions about the effectiveness of negative advertising for U.S. statewide election than the previous literature had found.

Assessing Model Assumptions

With single-shot causal inference methods such as matching, balance checks are crucial diagnostics (Ho et al. 2006). These checks ensure that the treated and control groups are similar on their background covariates. Usually this takes the form of simple comparisons of covariate means in the treated and control group, though more sophisticated techniques exist. Unfortunately, this simple approach is ill-suited to the dynamic setting since it is unclear what groups to compare. At a given week of the campaign, negative and positive campaigns might differ on a time-varying confounder, but these differences might be due to past negativity.

Under the above assumptions of the IPTW estimator, the decision to go negative is unconfounded in the weighted data, conditional on past negativity. We should expect, then, that the observed actions will be independent of time-varying covariates once we weight by SW_i.

FIGURE 4 **Inverse Probability of Treatment Weighting Estimates of the Time-Varying Effects of Negative Campaigning with Bootstrapped 95% Confidence Intervals**

Effect of a week of negative advertising on Democratic voteshare
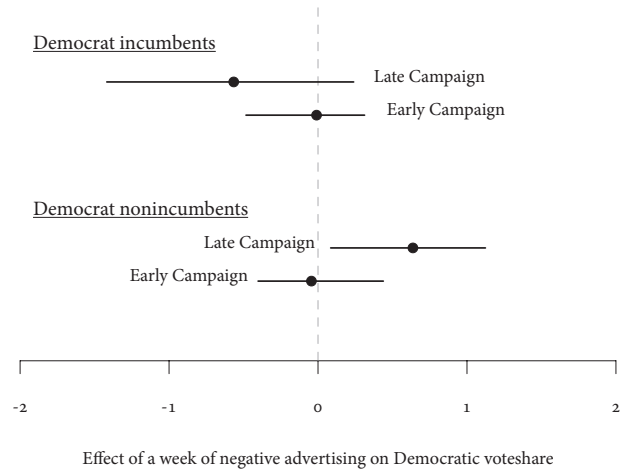
Note: Negative ads are more potent later in the campaign (October and November) than earlier in the campaign, but the direction of the effect is negative for incumbents and positive for nonincumbents.

This independence is, however, conditional on a unit's action history. For instance, suppose we had two campaigns that had remained positive until week t. Then the decision to go negative in week $t + 1$ for these two campaigns should not depend on time-varying covariates, such as polling, in the weighted data. We can assess balance in the weighted data, then, by checking for associations between the action decision and the time-varying covariates that affect that decision, conditional on the action history. If, after reweighting the data and conditioning on past negativity, the decision to go negative is still predictive of past polling, then there is likely residual confounding of the relationship between the outcome and negativity.

Figure 5 shows how the weights reduce this *history-adjusted imbalance* in the campaign-advertising example. It shows the change in the standardized history-adjusted imbalance from the unweighted to the weighted data.[22] In the unweighted data, for instance, Democrats were much more likely to go negative after an attack by Republicans ($R \ Neg_{t-1}$). Once we apply the weights, however, the differences move much closer to zero because IPTW gives

[22]These differences come from an unweighted and weighted pooled regression of the time-varying covariate at t on (a) the baseline covariates, (b) Democratic negativity before week t, and (c) Democratic negativity in week t. The coefficient on (c), divided by its standard error, is the standardized difference.

FIGURE 5 **The Change in History-Adjusted Balance between the Weighted and Unweighted Data as Measured by Standardized Differences between Those Campaign-Weeks That Went Negative versus Those That Remained Positive, Conditional on Baseline Covariates**
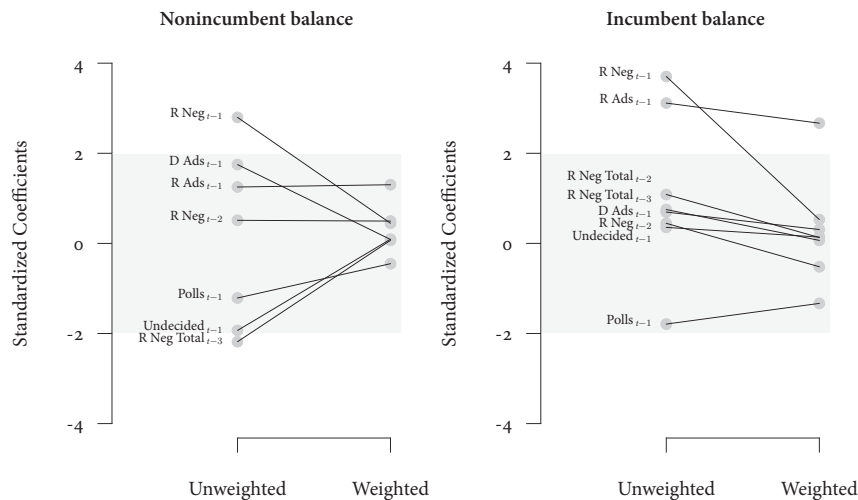
Note: The differences are, all told, closer to zero in the weighted model. "R Neg" is whether the Republican went negative, "Ads" are the number of ads run by the candidates, "R Neg Total" is the total number of Republican negative weeks in the campaign, and "Polls" is the averaged polling numbers for Democrats.

relatively more weight to races that went negative without Republican negativity in the last week.

One stark observation from the diagnostic plot is that confounding exists for incumbents even after weighting for the time-varying confounders. This makes sense for two reasons. First, data-quality issues plague incumbents since their safer races attract less polling. Second, incumbents have stronger positivity problems with extremely safe seats rarely going negative. Furthermore, incumbent campaign-weeks with high total ad volumes almost always feature negativity. These issues prevent the weights from fully eliminating the confounding in the data and should give us pause when interpreting the estimates for incumbents.

Cole and Hernán (2008) propose a series of model checks based on the distribution of the weights, SW_i. They note that the confounding of time-varying covariates is what pushes weights away from 1. A mean weight far below 1 indicates that there are relatively few surprise actions—those that are unlikely given the covariate history. This lack of surprises indicates that the probability of going negative is close to 0 or 1 in some parts of the covariate space, which is a violation of positivity. In some sense, there are very few comparable cases for these units

with high weights. In the matching literature, this is called a lack of common support. A good check for these issues in the weight model is to check the distribution of stabilized weights in each period to ensure that (1) the means at each point in time are close to 1, and (2) the minimum and maximum values are reasonably close to 1. The final distributions of the weights by week are in Figure 6 . Their means are all very close to 1, and the upper bounds are fairly low, indicating well-behaved weights.

Sensitivity Analyses

Causal estimates from an MSM have excellent properties when the assumptions of positivity and sequential ignorability hold. Of these, sequential ignorability is the trickiest, as it requires that, conditional on the covariate and action histories, the action decision is unrelated to the potential outcome. Any residual differences in potential outcomes between treated and control groups we call *unmeasured confounding* or omitted variable bias. Unless we conduct an experiment and randomize the action, this assumption must be justified through substantive knowledge. Since it is impossible to test this assumption, it is

FIGURE 6 Stabilized Weights Over the Course of the Campaigns
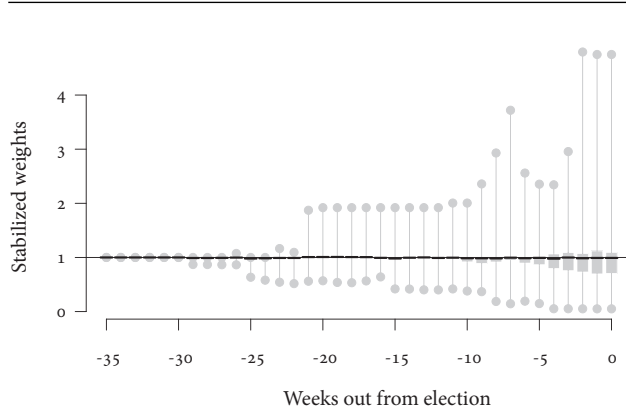
Note: The black lines are the weekly means, the gray rectangles are the weekly inter-quartile ranges, and the thin gray lines denote the range of the weights. Note that campaigns begin at various times so that there are very few campaigns at 30 weeks out, but very many at five weeks out. These weights appear well behaved as their means are close to 1.

vital to include as much information as possible and to conduct a sensitivity analysis of any estimated results.

Robins (1999) proposes a method to investigate the sensitivity of estimates to the presence of unmeasured confounding. Robins quantifies the amount of confounding as

$$q(\underline{x}_t, \underline{a}_{t-1}, a_t^*) = E[Y(\underline{a})|\underline{x}_t, \underline{a}_{t-1}, a_t] \\ - E[Y(\underline{a})|\underline{x}_t, \underline{a}_{t-1}, a_t^*]. \quad (14)$$

This function measures the difference in potential outcomes for some group that shares an action and covariate history. For two campaigns that are observationally equivalent up to week t, q measures the structural advantages of those campaigns that go negative in week t compared to those that remain positive in week t. For instance, the confounding may take the form

$$q(\underline{x}_t, \underline{a}_{t-1}, a_t^*; \alpha) = \alpha[a_t - a_t^*]. \quad (15)$$

This is a simple and symmetric form of omitted variable bias. When α equals zero, then there is no difference between those campaigns that go negative in week t and those that stay positive, given campaign histories. If α is positive, then negative campaigns are instrinsically stronger than those that remain positive. That is, when α is positive, then $Y(\underline{a})$ is higher for $a_t = 1$ (negative campaign weeks) than $a_t = 0$ (positive campaigns weeks). If α is negative, then those candidates who are going negative are worse off. Note that these selection biases are all conditional on the observed covariate history.

The above IPTW estimation procedure assumes that $\alpha = 0$, yet Robins (1999) shows that we can estimate the parameters under any assumption about α. Thus, by setting α to various levels, we can estimate the causal effect under different assumptions about the degree of omitted variable bias. To do so, we have to replace the outcome with a *bias-adjusted* outcome,

$$Y_\alpha \equiv Y - \sum_{k \in (0,1)} \sum_{t=0}^{T} \underbrace{q(\underline{X}_t, \underline{A}_{t-1}, k; \alpha)}_{\text{bias at this history}} \\ \times \underbrace{\Pr(A_t = k | \underline{A}_{t-1}, \underline{X}_t)}_{\text{probability of reaching this history}}, \quad (16)$$

where the first term is simply the observed outcome, Y, and the second term is the overall omitted variable bias, built up from the bias in each time period. We can then reestimate the parameters of the marginal structural model with outcome Y_α instead of Y to get bias-adjusted estimates and bias-adjusted confidence intervals. Note that when $\alpha = 0$, the bias function is zero, so that $Y_0 = Y$ and the usual estimation aligns with the assumption of sequential ignorability. Of course, the probability term is unknown and must be estimated. Fortunately, we have already estimated this function as part of the estimation of the weights, SW_i.

Figure 7 shows how the estimated effect of late-campaign negativity varies across different assumptions about the omitted variable bias, encoded in the parameter α, which runs along the x-axis. The magnitude of α describes how much stronger or weaker the negative campaigns are, on average, in terms of their potential outcomes. This figure also charts the change in the confidence intervals under the various assumptions about bias, with those that cross zero shaded lighter.

When negative campaigns are even 0.5 percentage points stronger than positive campaigns on average, our 95% confidence intervals would overlap with zero for nonincumbents. This might occur if campaigns were attacking their opponents for (unmeasured) behavior such as a scandal or an unpopular vote. Note that these imbalances would have to occur within levels of the covariate history and thus exist after conditioning on polls. If the negative campaigns are instead weaker, perhaps because campaigns go negative when they are in trouble, then results only grow stronger for nonincumbents. The results for incumbents show the reverse pattern. The results are fairly sensitive to the degree of confounding. Note, though, that this confounding would have to be above and beyond any information contained in polls and pre-campaign measures of competitiveness.

FIGURE 7 **Sensitivity of the Results to Deviations from the Sequential Ignorability Assumption**
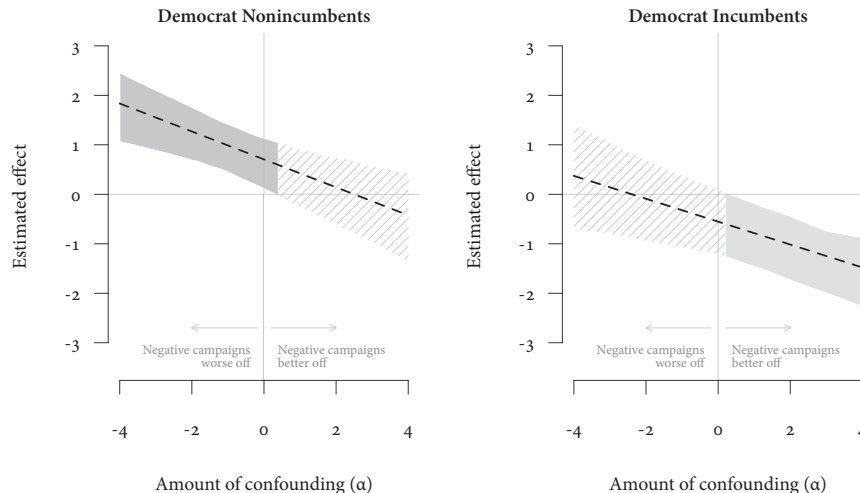
Note: The parameter α indexes assumptions about confounding, where negative values indicate that the observed negative campaigns are inherently weaker than predicted by the observed variables. Positive values assume that those negative campaigns are stronger than predicted.

It is clear that there is some sensitivity to the sequential ignorability assumption, and we could attempt to gather more data (quantitative and qualitative) to justify which direction this confounding is likely to lean. Notably, we could measure a larger set of dynamic campaign features such as scandals and endorsements. It would also help to draw on additional electoral cycles, such as 2008, when incumbents faced strong challengers to overcome the overlap and ignorability issues.

Conclusion

Political actions do not happen all at once. There are sequences of events that unfold over time. As we have seen, this poses strong problems for extant single-shot causal inference methods. This article brings to bear a framework that explicitly models the dynamic sequences and builds methods to test their effects. The original application of MSMs was to epidemiological data. Robins (1997) develops a set of methods called structural nested models with an application to HIV-treatment studies. In that context, the units are patients and doctors change the treatment over time if the patient status worsens. The analogy to politics is suggestive: campaign managers and candidates as doctors, working to save their patient, the election. Of course, candidates face human opposed to vi-

ral opponents, yet this changes only the types of variables needed to satisfy sequential ignorability.

The structural nested model of Robins (1997) provides an alternative approach to dynamic causal inference. These techniques center on modeling the effect of going negative at every possible history, which allows effects that interact with time-varying covariates. The estimation methods resemble backwards induction in game theory. Unfortunately, these structural nested models require models for entire sets of time-varying covariates and complicated computation to estimate while researchers can easily use off-the-shelf software to implement an MSM.

The focus of this article has been the effect of action sequences, yet in many political science situations, actors follow dynamic strategies—updating their actions based on changing conditions. It is likely that the optimal action is actually a strategy, since being able to respond to the current state of affairs is more effective than following a predefined sequence of actions. Hernán et al. (2006) demonstrates that marginal structural models and inverse probability weighting can estimate the effectiveness of strategies with a simple form such as "go negative when polls drop below x%." In addition, structural nested models can estimate the effect of arbitrary strategies. As might be expected, precisely estimating these effects requires larger sample sizes than the effects of simple action sequences.

A crucial path for future research is model development. In this article, I used a fairly simple model to estimate different effects for early and late in the campaign. This is a crude division of the data, and more fine-grained modeling might help to smooth effects over time. Indeed, we would expect that the effect of negativity in week 5 should be quite similar to the effect of negativity in week 6. Better MSMs should be able to handle this type of structure.

Dynamic causal inference is a problem for more than just campaigns. Each subfield of political science analyzes actions that occur over time and have multiple decision points: foreign aid, interest rates, budget allocations, state policies, and even democracy. Indeed, many of the assumptions in this article (or variations thereof) are implicit in time-series cross-sectional models, where the counterfactual framework is rarely discussed in explicit terms. Thus, there is a great opportunity for future work that identifies areas with dynamic causal inference problems and attempts to clarify or improve existing results.

References

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

Achen, Christopher. 1986. *Statistical Analysis of Quasi-Experiments*. Berkeley: University of California Press.

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

Cole, Stephen R., and Constantine E. Frangakis. 2009. "The Consistency Statement in Causal Inference: A Definition or an Assumption?" *Epidemiology* 20(1): 3–5.

Cole, Stephen R., and Miguel A. Hernán. 2008. "Constructing Inverse Probability Weights for Marginal Structural Models." *American Journal of Epidemiology* 168(6): 656–64.

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

Glynn, Adam N. 2012. "The Product and Difference Fallacies for Indirect Effects." *American Journal of Political Science* 56(1): 257–69.

Glynn, Adam N., and Kevin M. Quinn. 2010. "An Introduction to the Augmented Inverse Propensity Weighted Estimator." *Political Analysis* 18(1): 36–56.

Goldstein, Kenneth, and Joel Rivlin. 2007. "Congressional and Gubernatorial Advertising, 2003–2004." Combined File [dataset]. Final release.

Heckman, James. 1976. "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables, and Simple Estimator for Such Models." *Annals of Economic and Social Measurement* 5: 475–92.

Hernán, Miguel A., Emilie Lanoy, Dominique Costagliola, and James M. Robins. 2006. "Comparison of Dynamic Treatment Regimes via Inverse Probability Weighting." *Basic & Clinical Pharmacology Toxicology* 98(3): 237–42.

Ho, Daniel E., Kosuke Imai, Gary King, and Elizabeth A. Stuart. 2006. "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference." *Political Analysis* 15(3): 199–236.

Horvitz, D. G., and D. J. Thompson. 1952. "A Generalization of Sampling without Replacement from a Finite Universe." *Journal of the American Statistical Association* 47(260): 663–85.

Lau, Richard R., and Gerald M. Pomper. 2002. "Effectiveness of Negative Campaigning in U.S. Senate Elections." *American Journal of Political Science* 46(1): 47–66.

Lau, Richard R., and Gerald M. Pomper. 2004. *Negative Campaigning: An Analysis of U.S. Senate Elections Campaigning American Style*. Lanham, MD: Rowman & Littlefield.

Lau, Richard R., Lee Sigelman, and Ivy Brown Rovner. 2007. "The Effects of Negative Political Campaigns: A Meta-analytic Reassessment." *Journal of Politics* 69(4): 1176–1209.

Lumley, Thomas. 2004. "Analysis of Complex Survey Samples." *Journal of Statistical Software* 9(1): 1–19. R package verson 2.2. http://www.jstatsoft.org/v09/i08.

Lumley, Thomas. 2010. "Survey: Analysis of Complex Survey Samples." R package version 3.23-2. http://faculty.washington.edu/tlumley/survey/.

Pearl, Judea. 2010. "On the Consistency Rule in Causal Inference: Axiom, Definition, Assumption, or Theorem?" *Epidemiology* 21(6): 872–75.

Pierson, Paul. 2000. "Not Just What, but When: Timing and Sequence in Political Processes." *Studies in American Political Development* 14(1): 72–92.

R Development Core Team. 2011. *R: A Language and Environment for Statistical Computing*. Vienna, Austria. http://www.R-project.org.

Robins, James M. 1986. "A New Approach to Causal Inference in Mortality Studies with Sustained Exposure Periods—Application to Control of the Healthy Worker Survivor Effect." *Mathematical Modelling* 7(9–12): 1393–1512.

Robins, James M. 1997. "Causal Inference from Complex Longitudinal Data." In *Latent Variable Modeling and Applications to Causality*, ed. M. Berkane. *Lecture Notes in Statistics* 120. New York: Springer-Verlag, 69–117.

Robins, James M. 1999. "Association, Causation, and Marginal Structural Models." *Synthese* 121(1/2): 151–79.

Robins, James M. 2000. "Marginal Structural Models versus Structural Nested Models as Tools for Causal Inference." In *Statistical Models in Epidemiology, the Environment, and Clinical Trials*, ed. M. Elizabeth Halloran and Donald Berry. *The IMA Volumes in Mathematics and Its Applications* 116. New York: Springer-Verlag, 95–134.

Robins, James M., Miguel A. Hernán, and Babette Brumback. 2000. "Marginal Structural Models and Causal Inference in Epidemiology." *Epidemiology* 11(5): 550–60.

Rubin, Donald B. 1978. "Bayesian Inference for Causal Effects: The Role of Randomization." *Annals of Statistics* 6(1): 34–58.

VanderWeele, Tyler J. 2009. "Concerning the Consistency Assumption in Causal Inference." *Epidemiology* 20(6): 880–83.

Wood, Simon N. 2011. "Fast Stable Restricted Maximum Likelihood and Marginal Likelihood Estimation of Semiparametric Generalized Linear Models." *Journal of the Royal Statistical Society (B)* 73(1): 3–36.