Conditioning on Post-treatment Quantities with Structural Mean Models*

Luke Keele†

May 7, 2013

Abstract

In this paper, I demonstrate how structural mean models (SMMs) can be used to estimate quantities based on post-treatment variables. SMMs are semiparametric models where the parameters of the model or models correspond to meaningful functions of expected potential outcomes for subjects actually exposed to the treatment. SMMs provide a flexible framework for estimating causal effects with noncompliance even with binary outcomes. More widely used methods such as two-stage least square impose stronger distributional assumptions on the data. Next, I show how SMMs can be used to estimate treatment effect modification by a post-treatment variable. Often an intermediate variable between treatment and outcome may alter the causal effect. Effect modification of the type can be easily accommodated in the SMM framework. I demonstrate the utility of SMMs with applications to the get-out-the-vote literature.

1 Introduction

Researchers in political science have increasingly turned to randomized experiments as a means of drawing causal inferences about political processes. Randomized experiments are an attractive research methodology since randomization of a treatment makes treated and control groups equal on average in terms of all observed and unobserved characteristics. Thus the only differences between groups should be receipt of the treatment. Moreover, since allocation of the treatment is controlled by the analyst, experiments can also guarantee that the treatment is temporally prior to the outcome. This avoids the bias that can result from conditioning on post-treatment variables (Rosenbaum 1984).

*This research was supported in part by NIMH grant RC4-MH-092722-01. For helpful comments, I thank Marc Meredith.

†Associate Professor, 211 Pond Lab, Penn State University, University Park, PA 16802 Phone

While randomized experiments can often avoid the complications that arise from conditioning on post-treatment variables, in many cases utilizing post-treatment variables can reveal important information. For example, in experiments where subjects may fail to comply with assigned treatment status, conditioning on post-treatment variables is necessary to learn about the treatment effect for those subjects that comply (Angrist, Imbens and Rubin 1996). In another example, information about causal mechanisms is revealed through the use of post-treatment variables (Imai et al. 2011).

In this article, we explore strategies for post-treatment variables in randomized experiments. Here, we use the framework of structural mean models to both articulate assumptions and develop estimable statistical models. Structural mean models (SMMs) were developed as a framework for the analysis randomized trials with non-compliance (Robins 1989, 1994).[1] SMMs are a class of semiparametric models where the parameters of the model or models correspond to functions of expected potential outcomes for subjects actually exposed to the treatment. The semiparametric structure of these models allow analysts to use statistical techniques that impose weak constraints on the data.

In this article, we highlight two applications of SMMs. Building on the literature in biostatistics, we show how SMMs can be applied to randomized trials with noncompliance with a specific focus on contexts where the outcome is measured with a binary variable. While the instrumental variables (IV) is well-understood as supplying statistical framework for analyzing experiments with noncompliance (Sovey and Green 2011; Gerber and Green 2012), what is less well understood is that use of IV with binary outcomes require additional assumptions for estimation of nonsaturated models. We start with a summary of the large number of strategies that have been proposed for IV with binary outcomes. Here, we carefully delineate the additional assumptions that are required for each method. We, then, provide details about generalized structural mean models for binary data. SMMs for binary data rely

[1]Structural mean models are special cases of structural nested mean models. The basic structure of the two models is the same, but structural nested mean models refer to time-varying treatments while structural mean models refer to ordinary single shot treatments.

on weak assumptions that are often justified by design and avoid the stronger assumptions required for the methods that are most commonly used.

Next, we outline how SMMs can be used to estimate effect modification by a post-treatment variable. The statistical analysis of an experiment entails estimation of treatment effects. Often there is heterogeneity in these treatment effects. Such treatment heterogeneity is often referred to as effect modification. Treatment effect modification occurs when treatment effects differ across strata of pretreatment variables. For example, a treatment effect may be larger among women than men. Effect modification may also occur when an intermediate variable–one measured after treatment–alters the treatment effect on the final outcome. Such post-treatment effect modification, may reveal insights into both the patterns of effects in an experiment, but might also be useful in the design of future interventions. SMMs can be used to semiparametrically estimate such patterns of effect modification.

Structure of the Article. We begin with an introduction to structural mean models. While these models are widely used in biomedical applications, they are rare in the social sciences. In particular, we highlight the semi-parametric structure of these models. Next, we review methods of estimation for experiments with noncompliance and a binary outcome. We compare and contrast the assumptions needed for estimation across several methods including SMMs. We then provide a re-analysis of a set of field experiments from Green, Gerber and Nickerson (2003). Next, we show how SMMs can be used to estimate effect modification by a post-treatment variable. After discussing the assumptions needed for identification, we then apply SMMs to data from field experiments on habitual voting. We demonstrate how SMMs can shed light on whether voting in one election heightens the effect of an intervention design to increase turnout.

2 Structural Mean Models

2.1 Notation

We begin by defining the potential outcome or counterfactual notation to be used throughout the article. We denote Y, R, Z to represent the following observed quantities: Z is a randomization assignment indicator with $Z = 1$ denoting assignment to treatment and $Z = 0$ control. Later, we refer to Z as an instrument. We let R be a second indicator which records whether a unit is actually exposed to a treatment. When subjects fully comply with treatment assignment $R = Z$, but $R \neq Z$ if some subjects do not comply. For the moment, Y is a continuous outcome measure. For each of these observed quantities, we assume there are units i for $(i = 1, \ldots, n)$. Hereafter, we omit subscripts for i and assume these are individual level quantities.

We define potential outcomes in the usual way following the Neyman-Rubin causal model (Holland 1986). Here, $Y(r, z)$ is the potential outcome if Z is set to z and R is set to r. We also define a second potential outcome as $R(z)$, which is the potential exposure status if Z is fixed at z. SMMs are designed to estimate effects of actual exposure to treatment as represented by R as opposed to the effect of assignment to treatment represented by Z. The expected contrast

$$\mathrm{E}(Y - Y(0,0)|R)$$

defines the average causal effect of R. Notice that the contrast here is between the observed outcome Y and $Y(0,0)$, the treatment free potential outcome. In the SMM framework, the causal contrasts are often defined in terms of $Y(0,0)$ the potential outcome that is treatment free. The treatment free potential outcome becomes particularly relevant when subjects select whether $R = 1$ or not.

To identify this causal effect, in the SMM framework, the following set of assumptions

must hold. First, each unit's potential outcomes must be mutually independent. This is often referred to as the stable unit treatment value assumption or SUTVA. Second, the effect of Z on the outcome must come only through R or $Y(r,z) = Y(r)$ (Angrist, Imbens and Rubin 1996). This assumption is commonly referred to as the 'exclusion restriction.' Since we assume that the exclusion restriction holds we denote the potential treatment outcomes as $Y(0) = Y(0,1) = Y(0,0)$. Two additional assumptions are also required for identification. These assumptions, however, unlike SUTVA and the exclusion restriction are justified by the design in randomized trials. First, Z must be independent of both potential outcomes $R(z)$ and $Y(r,z)$. If an analyst uses random mechanism to assign values of Z, this assumption should hold in expectation. Thus the design of the experiment make this assumption true and verifiable. Next, we assume that units in the control group cannot access treatment and only receive the control so that $Pr(R = 0|Z = 0) = 1$. Such a restriction again can be justified by the experiment design and is referred to as a 'no-contamination' restriction (Cuzick et al. 2007).

In experimental designs where the no-contamination restriction is not credible, SMMs require an additional assumption. SMMs are also identified when the no effect modification (NEM) assumption holds. In words, the NEM assumption constrains the causal effects among the treated to be equal for those randomized to treatment and those randomized to control (Hernán and Robins 2006). The exact form of the NEM assumption depends on the SMM. Below, I formally define the NEM for one common SMM. The NEM assumption is analogous but not identical to the more familiar assumption of monotonic selection of exposure to treatment by the units (Imbens and Angrist 1994; Angrist, Imbens and Rubin 1996). For some types of SMMs, identification holds under the monotonicity assumption. Next, I describe the additive SMM, the simplest of SMMs.

2.2 Identification in Additive Structural Mean Models

SMMs are like generalized linear models in that they are formulated for specific link functions. For example, the additive SMM is for situations where the identity link is judged appropriate, and the multiplicative SMM is for situations with log links. All SMMs, however, rely on a basic counterfactual comparison between the observed outcome and the treatment-free potential outcome $Y(0)$. For any SMM, this contrast takes the following form:

$$E(Y|R,Z) - E\{Y(0)|R,Z\} = \eta_s(R)\phi$$

As above, the causal contrast is between the observed outcome and the treatment free potential outcome. Here, $\eta_s(R)\phi$ denotes a structural model with an unspecified functional form. The model is referred to as a structural model since the model refers to counterfactual quantities and causal contrasts instead of to observed associations.[2] Association models describe correlations that can be estimated from observed data, unlike a structural model that describes counterfactual quantities that cannot be observed without a set of assumptions for identification. Later, we will need to introduce an association model for one particular type of SMM.

For the additive SMM, we place the following structure on $\eta_s(R)\phi$

$$E(Y|R,Z) - E\{Y(0)|R,Z\} = (\phi_0 + \phi_1 Z)R \tag{1}$$

The parameters of the additive model corresponds to specific counterfactual comparisons. The parameter $\phi_0 = E\{Y(1) - Y(0)|R = 1, Z = 0\}$, and $\phi_0 + \phi_1 = E\{Y(1) - Y(0)|R = 1, Z = 1\}$. Thus the first parameter is the average causal effect among those who choose

[2]The term "structural model" has a number of highly specific definitions within economics. See Heckman and Vytlacil (2007) for a discussion of the definitions of structural models in economics. The definition of a structural model I use here comes from a strand of research that focuses on counterfactuals and formulates the definition of a structural model in terms of unobservable counterfactuals (Pearl 1995; Robins, Hernan and Brumback 2000). The counterfactual definition of structural models is quite different from some of the definitions typically used in economics.

treatment but are assigned to control, and the sum of the first and second parameter is the average causal effect among those who are assigned to and choose treatment. Under the no-contamination restriction, by design, $\{R = 1, Z = 0\}$ is a measure zero event and $\{R = 1\} = \{R = 1, Z = 1\}$ which implies the following restriction on the structural model:

$$\phi_0 + \phi_1 = \phi = E\{Y(1) - Y(0)|R = 1\}.$$

When Z is randomized, $E\{Y(0)|Z = 0\}$ is alway non-parametrically identified from the data using $E(Y|Z = 0)$. With noncompliance, it is less obvious that $E\{Y(0)|Z = 1\}$ is identified from observable quantities in the data. Under SMMs, we use the following expansion to derive which observable quantities correspond to this counterfactual:

$$\begin{aligned} E\{Y(0)|Z = 1\} &= E\{Y(0)|R = 1, Z = 1\}E\{R|Z = 1\} \\ &+ E\{Y(0)|R = 0, Z = 1\}E\{1 - R|Z = 1\} \end{aligned}$$

If Z is randomized the following is true:

$$E\{Y(0)|Z = 1\} = E\{Y(0)|Z = 0\}$$

This implies that we can write the key counterfactual quantity in terms of observed quantities

$$E\{Y(0)|R = 1, Z = 1\} = \frac{E(Y|Z = 0) - E\{(1 - R)Y|Z = 1\}}{E(R|Z = 1)}$$

This implies that identification holds for the following parameter:

$$\phi = E\{Y(1) - Y(0)|R = 1\}$$

or we can re-express this in terms of the contrast between observed outcomes and the treat-

ment free potential outcome with the following model

$$\phi = E(Y - Y(0)|R, Z = 1)$$

This model is said to be saturated since it imposes no a prior restrictions on the functional form of the estimate. In other words, since it is nonparametric. In this model, ϕ represents the expected change in outcome for subjects who were exposed to $R = 1$ if they had had their exposure set to $R = 0$.

If the no-contamination assumption does not hold, we must assume that NEM holds. Assuming NEM, identification proceeds in a different manner. Again under randomization $E\{Y(0)|Z = 1\} = E\{Y(0)|Z = 0\}$, which can be rewritten in conjunction with (1) as

$$E\{Y - (\phi_0 + \phi_1)R|Z = 1\} = E(Y - \phi_0 R|Z = 0)$$

Under the NEM assumption, this corresponds to constraining $\phi_1 = 0$ which implies

$$\phi_0 = E\{Y(1) - Y(0)|R = 1\}$$

2.3 Estimation in Additive Structural Mean Models

For the additive SMM, estimation proceeds via G-estimation (Robins 1994). Under G-estimation, when the structural model is saturated, the estimator is nonparametric. When the structural model is not saturated, G-estimation is semiparametric which places weak functional form constraints on the model. To weaken functional form assumptions, G-estimation only estimates the parameter for the treatment effect parameter, while all other parameters in the model have a nonparametric form. G-estimation is usually described in the context of estimators for causal effects with inverse probability weighting. The principle of G-estimation, however, is quite general.

Under G-estimation, the consistent estimate for the treatment effect is the value of ϕ

that yields no mean difference between the observed and treatment free potential outcome. For example, take the intention to treat (ITT) estimator, which is the effect of treatment assignment as opposed to compliance. The standard estimator for the ITT is

$$E(Y|Z=1) - E(Y|Z=0)$$

Under G-estimation, we write our estimator for the ITT as

$$Y(0) = Y - \phi Z$$

In a randomized experiment, we observe $Y(0)$, treatment free potential outcomes, when $Z = 0$, if so the consistent estimator for ϕ is the value of ϕ that makes this equality hold.[3] For a saturated additive SMM, the nonparametric G-estimator is

$$\hat{\phi}_0 = \frac{E(Y|Z=1) - E(Y|Z=0)}{E(R|Z=1) - E(R|Z=0)}$$

This estimator, of course, corresponds to the usual Wald estimator (Angrist, Imbens and Rubin 1996). Moreover, this implies that the additive SMM estimates the same complier average causal effect identified under instrumental variables. In sum, additive SMMs are equivalent to instrumental variables (IV) methods

To demonstrate the semiparametric structure of SMMs, we define x as a pretreatment covariate. If we wished to included x, we could write the structural model as:

$$\mathrm{E}(Y - Y(0)|R, Z=1) = \eta_s(R, x)\phi = \phi_0 + \phi_1 R + \phi_2 Rx$$

This model allows for covariate-exposure interactions in the structural model. Here, ϕ_2 defines the change in the average effect of the exposure R for a unit increase in x. Under a fully parametric approach, we would estimate all the parameters in the model using the linear

[3]We have to assume that additive rank preservation holds as well

functional form above. SMMs are semiparametric since both ϕ_0 and ϕ_2 are left unspecified. SMMs are much like the Cox model for survival data. In the Cox model, the baseline hazard is left unspecified. In SMMs all parameters other than the treatment effect are left unspecified. Thus the parameters ϕ_0 and ϕ_2 are left as infinite-dimensional parameters with an unspecified functional form. Thus SMMs make fewer functional form assumptions and are less prone to bias from model misspecification (Robins 1992, 1994). In fact, much of the strength of SMMs stems from the fact that they provide a semiparametric framework for IV estimation. These semiparametric models are asymptotically consistent (Robins 1992) which may be a concern in design with noncompliance but small numbers of units. When samples sizes are small, estimates from a saturated model should accompany estimates from non-saturated models. To summarize, for saturated models, the additive SMM corresponds to the standard nonparametric Wald estimator, but for unsaturated models, SMMs rely on semiparametric estimation. We now turn to the case where $Y \in \{0, 1\}$.

3 IV with Binary Outcomes

Binary outcome measures are common in the social sciences. While the application of IV methods in any setting requires the assumptions outlined above, when the outcome is binary many methods of estimation impose additional assumptions. The exact form of these assumptions vary from estimator to estimator (Clarke and Windmeijer 2012). When the outcome is binary, analysts can use three different approaches (1) bounds (nonparametric but not point identified), (2) fully parametric, (3) semiparametric, and (4) nonparametric but point identified. We do not discuss the bounds approach which uses the weakest assumptions. See Clarke and Windmeijer (2012) for an overview of bounds approaches.

3.1 Current Approaches

We begin with a brief overview of the wide variety of methods available to apply IV to binary outcomes. As is often the case, there is a tension between ease of use and the strength of the assumptions need for estimability. Methods with strong assumptions tend to

be easy to use, while approaches with weaker assumptions are less easy to use. There is one exception to this rule when R and S are binary. When this is true, one can apply the Wald estimator, which is nonparametric. When analysts include pre-treatment covariates or have continuous exposures, the simplest method is to apply the usual two-stage least squares (2SLS) estimator to the binary outcome. Often, 2SLS will provide reasonable answers, but it imposes strong assumptions on probability bounds and constant effects of exposure to the randomized treatment (Imbens 2001). One alternative is to apply Heckman-style models either with a two step estimation process or with maximum likelihood (Heckman 1979). Heckman-style models also impose very strong functional form assumptions and are very sensitive to minor violations of distributional assumptions (Copas 1988; Little 1985). Freedman and Sekhon (2010) also note that issues often arise in the estimation of these models where the numerics are fragile and convergence is uncertain.

A less restrictive approach is through a variety of what we denote as "plug-in" estimators, where the logic of 2SLS is adapted to nonlinear models. Here, developments for probit and logit models have proceeded separately. Rivers and Vuong (1988) develop two different plug-in methods for probit models. The first we call the plug-in prediction method, and the second we call the plug-in residual method. Both start with a first stage model for exposure as a function of treatment status. In each, different quantities from the first stage model are used in the second stage model of the outcome.

We describe the plug-in prediction approach first. Here, in the first stage one fits a linear selection model: $R = \hat{\alpha}_0 + \hat{\alpha}_1 Z$. Then in a second stage model, we fit $Y = \Phi(\hat{\beta}_0 + \hat{\beta}_1 \hat{R})$. In the second stage probit model, $\hat{\beta}_1$ is an estimate of the CACE on the latent probability scale. The plug-in residual method known as a "control-function" estimator works on a similar principle. As before, one fits a linear selection model for R, but instead of using the fitted values for R, we calculate the fitted residual, $\hat{V} = R - \hat{\alpha}_0 + \hat{\alpha}_1 Z$. In stage two, we fit $Y = \Phi(\beta_0 + \beta_1 R + \beta_3 \hat{V})$. Now the estimated coefficient on R is a consistent estimator for the CACE. Nagelkerke et al. (2000) constructed plug-in IV estimators for binary outcomes using logistic link functions in

an similar way. As an alternative they propose fitting logistic regression models for the first stage model. Blundell and Powell (2004) propose a semiparametic control function approach for when R is a linear function.

While plug-in approaches are easy to implement, they rely on strong assumptions. When the selection model is nonlinear, neither estimator is consistent (Clarke and Windmeijer 2012). In fact analytic expressions can be derived to show the bias of these estimators (Cai, Small and Have 2011). Other studies have used simulation to show that both plug-in approaches have higher levels of bias compared to methods with more flexible functional forms (Cai et al. 2012; Vansteelandt et al. 2011). Moreover, both the Heckman-style models and the plug-in methods impose additional assumptions even when the model is saturated. As we note below, SMMs are completely nonparametric when the model is saturated.

To avoid the functional form assumptions needed for parametric approaches, two different semiparametric estimators (besides SMMs) have been developed. Abadie and Gardeazabal (2003) constructs a weighted estimator based on the 'local average response function' (LARF). Tan (2006) proposes a doubly robust approach based on a weighted two-stage estimator. While both approaches are robust, neither have seen wide use in the applied literatures. Structural mean models provide one additional semiparametric approach to IV with a binary outcome.

3.2 Generalized Structural Mean Models

Extendings SMMs to the binary case is not without its challenges. Robins (1999) showed that standard methods of G-estimation cannot be applied to binary outcomes when compliance status is nonignorable. (Vansteelandt and Goetghebeur 2003) found a solution using what has come to be known as the double-logistic SMM. The double-logistic SMM models differences in the probability that observed Y is 1, $P(Y = 1)$ and the probability that the treatment free potential outcome is 1, $P(Y(0) = 1)$, through the following model

$$\text{logit}\{E[Y|R, Z = 1]\} - \text{logit}\{E[Y(0)|R, Z = 1]\} = \eta_s'(R)\psi_0 \tag{2}$$

where $\text{logit}(a) = \log\{a/(1-a)\}$. In this model, the probability of success for subjects who chose to be exposed to the treatment ($R_i = 1$) versus not exposed ($R_i = 0$) is determined by ψ_0, depending on the functional form of $\eta_s'(R)\psi_0$. Here ψ_0 corresponds to the probability of success for participating versus not participating in the intervention. Estimation requires that the analyst specify what is known as an association model which models the outcome as function of treatment and compliance status:

$$\{P(Y = 1)|R, Z = 1)\} = \eta_a(Z)\beta$$

We denote this model as an association model with the subscript 'a' because it describes an observed correlation in the data and is not mean to represent a causal comparison. Specifically, the association model describes the correlation between outcome and compliance given treatment status. As such, it is not a structural model. It is needed since $Y(0)$ is never observed with R except when $R = 0$. This implies that the conditional mean of the potential treatment free outcome is not directly observed. When the model is saturated, the association model has the following parameterization

$$\Lambda(\beta_0 + \beta_1 Z + \beta_2 R + \beta_4 ZR) = \eta_a(Z)\beta$$

where Λ is the cumulative distribution function for the logistic distribution function, and the parameters in the association model are fit with a logistic regression. The double-logistic SMM estimator is the solution to the following moment condition

$$E[\text{expit}\{\beta_0 + \beta_1 + (\beta_2 + \beta_4 - \psi_0)R\}Z = 1] = E[\text{expit}\{\beta_0(\beta_2 - \psi_0)R\}Z = 0] \tag{3}$$

where $\text{expit}(\cdot) = \exp(\cdot)/\{1 + \exp(\cdot)\}$. The estimate for ψ_0 is the value such that the av-

erage difference between the observed outcomes in the control condition are equal to the predicted treatment free outcomes in the treated condition. Given that both the structural and associational models are logistic regressions, this method is often called a double-logistic SMM.[4]

When the models are saturated and the no-contamination assumption holds, a nonparametric estimator is available. Let $p_{y=1|r,z}$ be the observed proportion of successes on Y given R and Z. The nonparametric estimator for ψ_0 is

$$\psi_0 = \frac{p_{y=1|r=z=1}}{1 - p_{y=1|r=z=1}} \left[\frac{p_{r=1|z=1}}{p_{y=1|r=z=0} - p_{y=1,r=0|z=1}} - 1 \right] \tag{4}$$

Again one advantage of the SMM framework is that with saturated models, we need not impose additional assumptions on the data. We should note that estimates of ψ_0 from (4) are on the log-odds scale such that $\exp(\psi_0)$ is an odds-ratio. One could also use the logistic CDF to convert this quantity to the probability scale.

For unsaturated models, Vansteelandt and Goetghebeur (2003) develop a weighted Newton-Raphson algorithm for semiparametric estimation. They prove that this method of estimation is consistent when the association model is correctly specified, which will hold by design in a randomized trial. Alternatively, Clarke, Palmer and Windmeijer (2011) demonstrate how the double-logistic SMM can be estimated with generalized method of moments. The model can also be parameterized with probit link functions. Note the model is semiparametric in that an estimate is produced for ψ_0, all other parameters in the model have unspecified functional forms.

The double-logistic approach is attractive since it places weak constraints on the data (Clarke and Windmeijer 2012). A number of simulation studies have found that the double-logistic SMM performs better than plug-in methods in terms of bias and robustness to misspecification (Cai, Small and Have 2011; Cai et al. 2012; Vansteelandt et al. 2011). The

[4]One drawback to the double-logistic SMM is that the association model may be uncongenial with the structural model which will affect convergence of the estimating algorithm (Vansteelandt and Goetghebeur 2003). This appears to be rare in practice.

double-logistic SMM also has a local robustness property such that if either the association model or the structural model is misspecified, the estimate for ψ_0 is always consistent under the null hypothesis $\psi_0 = 0$. This robustness property guarantees that when the models are misspecified estimates of ψ_0 will have little bias when the true effect is close to zero. We should note that concerns about model specification do not arise for saturated models which occur in randomized designs where the no-contamination restriction holds. We now turn to an empirical application to compare the double-logistic SMM to a plug in approach.

3.3 Application

One literature in political science studies methods for increasing voter turnout through the use of randomized field experiments. This research both focuses on the effectiveness of various get-out-the-vote methods and tests social psychological theories about voters (Green, McGrath and Aronow 2013). One entry in this literature focused on the effectiveness of door-to-door canvassing (Green, Gerber and Nickerson 2003). In this study, the researchers conducted six separate field experiments in the following cities: Bridgeport, Columbus, Detroit, Minneapolis, Raleigh, and St. Paul in November 2001. In each city, voters were randomized to either receive face-to-face contact from local staffers encouraging them to vote or were not contacted.

The elections in the field experiment were all local elections that ranged from school board to city council elections. As we might expect, many of the households randomized to the treatment were not available for the face-to-face message encouraging them to vote. While the intention-to-treat (ITT) effects are easily estimable, in this context, complier effects are of greater interest. Importantly, the design ensures that the no-contamination restriction holds. That is, it is impossible for control households to access the canvassing treatment. This is important since this implies that the double-logistic SMM is identified under the weakest set of assumptions.

In the original analysis, the analysts estimated complier effects using both 2SLS and a

bivariate pro bit with partial observability. In all six field experiments, both turnout rates (the outcome) and compliance with the treatment were quite low. In some cities, turnout and compliance were lower than 15%. Skew in the binary distributions of R and Y is more likely to lead to biased estimates under fully parametric approaches (Cai, Small and Have 2011; Cai et al. 2012). Moreover, compliance was measured with a binary covariate while the approach of Rivers and Vuong (1988) assumes compliance status is continuous. As such, the plug-in approach is not congenial to the data from these experiments. Here, we conduct a basic comparison of the double-logistic SMMs and the plug-in approach of Nagelkerke et al. (2000). We estimate the complier average causal effect for each of the six field experiments. To ensure the comparability of the estimates, we estimate the compiler average causal effect on the odds scale. In sum, we compare estimates of $\exp(\psi_0)$ from a plug-in prediction estimator and the double-logistic SMM. In the biomedical statistics literature, this quantity is often referred to as the causal odds-ratio. For the plug-in method, we use predictor substitution instead of residual substitution.

Table 1 contains estimates of the causal odds ratio and associated 95% confidence intervals for each of the six cities. We estimate saturated models, which implies that we do not include any pretreatment covariates. The specification for each approach is based on the outcome, compliance status, and treatment assignment status.[5] We also include in the table the percentage of subjects in the experiment that that voted irrespective of treatment status, and the percentage of compliers. In general, we find that the plug-in method over estimates the causal odds ratio. In three cases, however, the discrepancy is minimal. In two cities, Bridgeport and Columbus the differences between the plug-in and double logistic estimates is rather large. For the Bridgeport experiment, the estimate of the causal odds ratio is 4.35 with the plug-in method, while it is 3.17 for the SMM. In the Columbus experiment, the plug-in method estimate is 3.56 while it is 2.62 for the SMM. We should note that in both cases the confidence intervals do overlap. The discrepancy appears to be driven by

[5]The original analysis also included dummy variables for each set of staffers that went door-to-door.

the distribution of the outcome variable: when the turnout proportion is small, the bias in the estimate appears to increase. In general, SMMs provide a robust method for estimating compliance effects when outcomes are binary. Next, we explore a second extension of SMMs.

Table 1: Causal Odds Ratio for the in Complier Effect in Six GOTV Field Experiments

			Plug-in Method			
	Bridgeport	Columbus	Detroit	Raleigh	St. Paul	Minneapolis
Odds Ratio	4.35	3.56	1.49	0.91	1.78	1.72
95% C.I.	[1.59, 12.1]	[0.48, 26.09]	[0.97, 2.01]	[0.67, 1.21]	[1.06, 3.01]	[0.69, 4.27]
			Double Logistic SMM			
	Bridgeport	Columbus	Detroit	Raleigh	St. Paul	Minneapolis
Odds Ratio	3.17	2.62	1.41	0.91	1.75	1.67
95% C.I.	[1.08, 9.29]	[0.31, 22.17]	[0.96, 2.05]	[0.70, 1.19]	[1.03, 2.96]	[0.64, 4.29]
Percent Voting[a]	12.7	9.0	44.6	29.1	39.5	25.9
Percent Compliant	14.2	6.5	15.3	16.4	16.5	9.2

Note: The plug-in method we use is that of predictor substitution. The prediction from a first stage model of compliance is substituted into a second stage logistic regression model of the outcome. [a]This refers to the percentage of people that voted in the entire sample irrespective of treatment status.

4 Effect Modification by Post-Treatment Variables

Thus far the chief utility of SMMs has been to provide an alternative approach to IV estimation when outcomes are binary. In the application, the double-logistic SMMs appeared to be particularly useful when the distribution of a binary outcome was heavily skewed. In this section, we explore an extension of SMMs. Often analysts seek to understand whether treatment effect heterogeneity occurs. Here, treatment effects are thought to differ across the level of a pretreatment covariate. In a medical trial, for example, the effects of a randomly administered drug may differ in its effectiveness depending on the sex of the subject. Effect modification may also occur across levels of a posttreament covariate. Such effect modification arises in designs with longitudinal data collection. In these designs, after the randomized intervention, there are multiple waves of data collection. We might suspect that

the effect of treatment might vary across levels of a covariate collected at an intermediate wave. Such longitudinal effect modification can be accommodated in the SMM framework while retaining a more flexible functional form through semiparametric estimation (Stephens, Keele and Joffe 2013). While additional assumptions are necessary for identification, SMMs of this type can provide either a sensitivity analysis for some experimental designs or be used to develop hypotheses for future interventions.

4.1 The Causal Model

First, we redefine the causal model in Equation 2 in the following way. For the moment, we assume that $R = Z$ such that there is full compliance with assigned treatment status. Relaxing this assumption does little to alter the nature of the causal model. First, I define $\mathbf{S}$ as potential post-treatment effect modifiers which are a set of intermediate covariates observed after treatment but prior to the outcome, also possibly multivariate. I also define $\mathbf{x}$ as a set of covariates measured before treatment. The causal contrast is now

$$\begin{aligned}\text{logit}(\text{E}[Y|\mathbf{S}, Z = 1, \mathbf{x}]) - \text{logit}(\text{E}[Y(0)|\mathbf{S}, Z = 1, \mathbf{x}]) \\ = \eta_s'(\mathbf{S}, R, \mathbf{x})\psi = f_1(Z)\psi_1 + f_2(Z, \mathbf{S})\psi_2 \quad (5)\end{aligned}$$

where $f_1(\cdot)$ and $f_2(\cdot)$ are arbitary known functions up to an unknown p-dimensional parameter $\psi_0 = (\psi_{01}, \psi_{02})$, where $p = dim(\psi_1) + dim(\psi_2)$. One model for a binary treatment Z and a single intermediate variable S might assume the following forms for $f_1(\cdot)$ and $f_2(\cdot)$

$$\text{logit}(P(Y = 1|S, Z = 1, \mathbf{x})) - \text{logit}(P(Y(0) = 1|S, Z = 1, \mathbf{x})) = \psi_1 Z + \psi_2 ZS \quad (6)$$

Interpretation of the quantities from this causal model is straightforward. The second parameter ψ_2 quantifies the differential in the effect of intervention Z at varying levels of S. When $\psi_2 = 0$, homogeneous causal effects are estimated across values of S.

We contrast this causal model with a more familiar one. First we alter Equation 2

$$\text{logit}\{\text{E}[Y|\mathbf{x}, Z=1]\} - \text{logit}\{\text{E}[Y(0)|\mathbf{x}, Z=1]\} = \eta_s'(Z, \mathbf{x}, Z\mathbf{x})\psi. \tag{7}$$

Under this parameterization $\eta_s'(Z, \mathbf{x}, Z\mathbf{x})\psi = (\alpha_0 + \alpha_1\mathbf{x} + \psi_0 Z + \psi_1 Z\mathbf{x})$ which allows for the possibility that the causal odds ratio varies over levels of $\mathbf{x}$. Such a parameterization allows for effect modification by pretreatment covariates as opposed to the model in (6) which allows for effect modification by post-treatment covariates. The SMM in (6) can also easily accommodate when $R \neq Z$. When this is true, the model requires the usual IV assumptions, but the model in (6) now allows for the possibility that the causal odds ratio for compliers varies over levels of S. See Stephens, Keele and Joffe (2013) for details on the estimation of SMMs with post-treatment effect modification. Estimation is often aided by a rich set of pretreatment covariates. The chief difference between models (7) and (6) is that treatment might affect S which then alters the treatment on later outcomes. Given that S is post-treatment assumptions for identification are altered from the usual SMM.

4.2 Identifiability

Identification under causal model (6) does require additional assumptions; in fact these models cannot be nonparametrically identified. Here, I discuss the conditions for identifiability. See for Vansteelandt and Goetghebeur (2004) for a related theorem and proof. For the moment, we ignore the binary nature of Y_i and the logistic link. Consider the following three structural mean models

$$\begin{aligned}
\text{E}(Y_i - Y_{i,0}|S, Z=1, \mathbf{x}) &= \psi_1 Z + \psi_2 ZS \\
\text{E}(Y_i - Y_{i,0}|S, Z=1, \mathbf{x}) &= \psi_1 Z + \psi_2 Z\mathbf{x} \\
\text{E}(Y_i - Y_{i,0}|S, Z=1, \mathbf{x}) &= \psi_1 Z + \psi_2 ZS + \psi_3 Z\mathbf{x}
\end{aligned} \tag{8}$$

This system of equation is not identified since there is not enough information to disentangle

the effects of $\mathbf{x}$ and S as there are only two identifying restrictions in (5). For identification, we must restrict one set of the multiplicative terms to be zero. To identify effect modification by a post-treatment covariate, we must restrict the $\mathbf{x}$ multiplicative terms to be zero. Under this set of modeling restrictions, identification holds. The restriction on pre-treatment effect modification leads to approximately averaged effects. No-interaction assumptions are often necessary for identification when conditioning on post-treatment quantities (Ten Have et al. 2007; Vansteelandt and Goetghebeur 2004; Robins and Greenland 1992; Hernán and Robins 2006).

Given that nonparametric identification is not possible, how should we interpret the results of SMMs with post-treatment effect modification? We argue that these models are useful if carefully applied. First, the results should not be given a strong causal interpretation. It is probably unwise to for this model to be the sole form of analysis, but the model can be useful as a sensitivity analysis in conjunction with other models. In general, SMMs of this type might also aid the design of future interventions or be used to generate additional hypotheses. As we demonstrate in the application, these models serve as a useful secondary test for what is know as downstream experimental analysis. Moreover these SMMs retain their local robustness property (Stephens, Keele and Joffe 2013), so an analyst is unlikely to falsely conclude that a treatment is without effect.

4.3 Application

Green and Gerber (2002) introduced the concept of "downstream" experimental analysis. In a downstream experimental analysis, researchers take advantage of the exogenous variation generated by randomized interventions to examine indirect effects from experiments. They propose an approach based on instrumental variables for estimating the indirect effects studied in a downstream analysis. We propose that SMMs with post-treatment effect modification can serve as useful method for the study of downstream effects.

The study of whether voting is habitual has been one fruitful application of downstream

experimental analysis (Gerber, Green and Shachar 2003). It is thought that voting may be habitual: once a citizen starts voting a habit forms and repeated voting is more likely. To demonstrate the application of SMMs to post-treatment effect modification, we use data on habitual voting. Our application is a downstream analysis of the well-known social pressure field experiment on voter turnout (Gerber, Green and Larimer 2008). In that study, voters were sent mail messages that told voters that voting status is a matter of public record that could be disclosed to family, friends or neighbors. Voters that received social pressure messages were considerably more likely to vote than those in the control condition. Here, we use an augmented version of the original data from this experiment. Specifically, we use data from Davenport et al. (2010) who added the turnout status from a series of subsequent elections for those who participated in the experiment. The social pressure experiment was fielded in August of 2006 before an August primary election in Michigan. In the original study, Gerber, Green and Larimer (2008) estimated the effect of the social pressure treatments on turnout in that August primary. The augmented data contains the voting status for four later elections starting with the general election of November 2006 to the general election of November 2008. In the analysis that follows, we restrict the outcome to voting status in the November 2006 general election.

The simplest method for a downstream analysis of habitual voting is to calculate an ITT estimate for outcomes other than those in the study. In the habitual voting example, we can calculate ITT estimates for the effect of the social pressure intervention on the November election instead of the August primary which was the original outcome of interest. It is this strategy that is used in (Davenport et al. 2010). One alternative to the ITT method for downstream analysis uses IV methods (Green and Gerber 2002). We introduce some notation to outline the IV approach to downstream analysis. Z is the randomized assignment indicator with $Z = 1$ denoting assignment to the social pressure treatment and $Z = 0$ denotes assignment to control. We let R be a second indicator which records whether someone voted in the August 2006 election following the field experiment. Now, Y is a binary outcome

measure, which records whether someone voted in a downstream election, here, the November 2006 general election. In the IV approach to downstream analysis proposed by Green and Gerber (2002), Z serves as an instrument while R becomes an intermediate outcome which transmits the effect of the intervention to the outcome Y. The estimand is the effect of the treatment as mediated by voting in the intermediate election.

Identification of the estimand in the IV approach to downstream analysis requires the usual IV assumptions. First, treatment status must be ignorable and SUTVA must hold. In this application instrument status is ignorable under the design, and there is little reason to suspect a SUTVA violation. Moreover, the social pressure intervention had strong effects on turnout in August 2006, so the instrument is not weak. Given that the control group cannot access the social pressure treatment, the no-contamination assumption is justified by the design. The key assumption in the downstream analysis is the exclusion restriction: it must be true that the social pressure intervention in August cannot have a direct effect on turnout in November. If we use the November 2006 election as the outcome, we have good reason to doubt the exclusion restriction. For the exclusion restriction to be valid, we must assume that no one remembers the social pressure intervention when deciding to vote in the general election three months later. In one study of habitual voting, the length of time between the treatment and downstream outcome was spaced out to a year to bolster the exclusion restriction (Gerber, Green and Shachar 2003). That is, once a year has passed between the initial treatment and the downstream election, it is more likely that there is no indirect effect. Gerber and Green (2012) discuss other methods for evaluating the exclusion restriction in this context. Estimation of the downstream effect under IV is straightforward as the usual 2SLS estimator can be applied. Of course, given that the intermediate and outcome variables are binary one might opt for the double-logistic SMM to avoid stronger functional form assumptions.

To apply an SMM with post-treatment effect modification, we first redefine R, voting in the August election, as S a postreatment effect modifier. The logic for effect modification is

straightforward. If voting is habitual, the effect of the social pressure treatment should only persist for those who voted in August. We specify the following conditional SMM

$$\text{logit}(P(Y = 1|S, Z = 1, \mathbf{x})) - \text{logit}(P(Y(0) = 1|S, Z = 1, \mathbf{x})) = \psi_1 Z + \psi_2 ZS$$

The parameter ψ_2 quantifies the differential in the effect of the social pressure intervention Z at varying levels of S voting in August. The SMM allows the effect of the intervention in August to be directly transmitted to the November outcome, but we expect the effect of the intervention on voting in November to be larger among those who voted in August.

For the empirical analysis, we study two of the social pressure interventions. The one intervention we examine disclosed voting records of everyone on the block. We term this the "Neighbors" treatment. A second intervention disclosed voting records of everyone in the household. We term this the "Self" treatment. In the statistical models, we compare each of these treatments to a control group that did not receive a mailing of any kind. For each treatment, we report three different estimates. The first is the ITT estimate of the social pressure intervention on turnout in the November election. Second, we use IV for a downstream analysis. We implement IV via the double-logistic SMM. Finally, we use an SMM with post-treatment effect modification, where voting in August is allowed to modify the effect of the treatment.[6]

Table 2 contains the results for both treatments across the three different identification strategies. We report all results as odds ratios. We first review the results for the Neighbors treatment. First, for those who received the neighbors social pressure treatment, the odds of voting are 8% higher in the November election, and the treatment effect is statistically significant as the 95% confidence interval is bound away from 1. If we assume that the treatment in August has no direct on voting in November, the odds of voting are 68% percent higher in the treated group, which is also statistically significant. Next under the SMM specification with treatment effect modification, we stratify the treatment effect estimate

[6]In this model, we condition on age, sex, and voting history for the last six elections.

Table 2: Estimating the Downstream Effects for 2006 Social Pressure Field Experiment

	Self		Neighbors	
Downstream ITT	1.03		1.08	
	[0.99, 1.06]		[1.04, 1.12]	
Downstream IV	1.23		1.68	
	[0.98, 1.53]		[1.41, 1.99]	
Effect Modification				
	Didn't Vote in August	Voted in August	Didn't Vote in August	Voted in August
	0.97	1.18	1.02	1.25
	[0.89, 1.06]	[1.01, 1.39]	[0.94, 1.12]	[1.08, 1.47]

Note: All point estimates are odds ratios with associated 95% confidence interval. Downstream IV estimate via double-logistic SMM. For Downstream IV estimate, voting in August primary is the intermediate post-treatment outcome. Voting in the August primary serves as the effect modifier for estimates in last row of the table. Treatment was in effect before August 2006 primary. Downstream outcome is voting in November 2006 general election. Each contrast is between one social pressure condition and the control condition that did not receive any mail in the field experiment.

by voting status in August. For those who failed to vote in the August primary, the social pressure intervention increased the odds of voting by a mere 2% and the 95% confidence interval now brackets 1. For voters that did vote in August, we observe that treatment raises the odds of voting in November by 25% and the estimate is statistically significant. One advantage of the SMM effect modification approach in this context is that it provides us with an estimand that exactly matches the theory of habitual voting. That is, we estimate a treatment effect among those who voted at least once and may have picked up the voting habit. We observe the same pattern with the self treatment but with smaller effects. Here, in the ITT analysis, the treatment raises the odds of voting by 3%, but in the IV analysis, the odds of voting among the treated are 23% higher. Both estimates, however, are not statistically significant. When we allow treatment effect modification, we do find that the odds of voting in November are 18% higher for those treated in August, and the estimate is just statistically significant.

The congruent findings across the IV approach and effect modification approach suggest a clear empirical pattern about habitual voting. While the IV approach relies on a suspect exclusion restriction and effect modification requires restricting the form of treatment interactions, taken together consistent results provide better evidence for the expected empirical pattern. We think the strategy, here, is successful in that for both IV and SMM with effect modification identification is not entirely justified by the experimental design. Each approach requires very different assumptions for identification. Agreement across the methods then forms a sensitivity analysis. A sensitivity analysis based on finding congruent results across different identification assumptions.

4.4 Conclusion

In many experiments with simple designs, statistical analysis often need not venture beyond a difference of means. But in other cases more sophisticated forms of analysis are required. In particular complications always arise when analyst condition on quantities

that are post-treatment. In these situations, identification of causal effects always relies on strong untestable assumptions. One cautious approach would be to only and always report intention-to-treat analyses where identification is rarely questionable. While such caution may be admirable, it is often worth extracting additional information from experimental interventions. So long as such analyses are reported with ITT estimates, readers can clearly distinguish which estimates rely on which assumptions. In general, the enemy is not assumptions but the use of assumptions without transparency.

Here, we have introduced structural mean models and focused on two instances where they can more usefully extract information from randomization than more commonly used tools. SMMs provide a general framework for estimating treatment effects with noncompliance. The SMM approach provides a nonparametric estimator when models are saturated, and SMMs allow for semiparametric estimation of causal quantities when models are not saturated. While simple experiments can often rely on fully nonparametric forms of analysis that is less true in analyses that involve conditioning on post-treatment quantities. The semiparametric form of SMMs imposes fewer functional form assumptions on the data than more popular, conventional statistical models such regression with two stage least squares. The advantages of SMMs are more apparent when the outcome is a binary measure. While a wide variety of estimation methods have been proposed for IV methods with binary outcomes, most require fairly strong distribution assumptions, which is not true of SMMs. In the empirical application, one standard binary IV method appeared to be biased when the distributions of the compliance and outcome measures were highly skewed. Moreover, SMMs retain a local robustness property for null effects.

We also highlighted a second useful extension of SMMs: estimating treatment effects conditional on post-treatment covariates. Here, the treatment effect is allowed to vary along levels of a post-treatment measure. While SMMs do require strong assumptions for identification under this extension, we argue this technique serves several useful purposes. In a downstream experimental analysis, SMMs with post-treatment effect modification are com-

plimentary to the IV approach proposed by Green and Gerber (2002). Both methods rely differing sets of untestable assumptions, and when the results are consistent across both approaches conclusions are significant strengthened versus either method in isolation. Second, such effect modification may provide useful hypothesis generation for future interventions.

References

Abadie, Alberto and Javier Gardeazabal. 2003. "The Economic Costs of Conflict: A Case Study of the Basque Country." *American Economic Review* 93(1):112–132.

Angrist, Joshua D., Guido W. Imbens and Donald B. Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91(434):444–455.

Blundell, Richard W. and James L. Powell. 2004. "Endogeneity in semiparametric binary response models." *Review of Economic Studies* 71(3):655–679.

Cai, Bing, Dylan S. Small and Thomas R. Ten Have. 2011. "Two-stage instrumental variable methods for estimating the causal odds ratio: Analysis of bias." *Statistics in Medicine* 30(15):1809–1824.

Cai, Bing, Sean Hennessy, James H. Flory, Daohang Sha, Thomas R. Ten Have and Dylan S. Small. 2012. "Simulation Study of Instrumental Variable Approaches With An Application to a Study of the Antidiabetic Effect of Bezafibrate." *Pharmacoepidemiology and Drug Safety* 21(2):114–120.

Clarke, Paul S. and Frank Windmeijer. 2012. "Instrumental Variable Estimators for Binary Outcomes." *Journal of the American Statistical Association* 107(500):1638–1652.

Clarke, Paul S., Tom Palmer and Frank Windmeijer. 2011. "Estimating Structural Mean Models with Multiple Instrumental Variables Using Generalized Method of Moments."

Copas, J.B. 1988. "Binary Regression Models for Contaminated Data." *Journal of The Royal Statistical Society B* 50:225–265.

Cuzick, Jack, Peter Sasieni, Jonathan Myles and Jonathan Tyrer. 2007. "Estimating the effect of treatment in a proportional hazards model in the presence of non-compliance and contamination." *Journal of The Royal Statistical Society, Series B* 69(4):565–588.

Davenport, Tiffany C., Alan S. Gerber, Donald P. Green, Christopher W. Larimer, Christopher D. Mann and Costas Panagopoulos. 2010. "The Enduring Effects of Social Pressure: Tracking Campaign Experiments Over a Series of Elections." *Political Behavior* 32(3):423–430.

Freedman, David A. and Jasjeet S. Sekhon. 2010. "Endogeneity in Probit Response Models." *Political Analysis* 18(2):138–150.

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 Christopher W. Larimer. 2008. "Social Pressure and Voter Turnout: Evidence From a Large-Scale Field Experiment." *American Political Science Review* 102(1):33–48.

Gerber, Alan S., Donald P. Green and Ron Shachar. 2003. "Voting May Be Habit-Forming: Evidence from a Randomized Field Experiment." *American Journal of Political Science* 47(3):540–550.

Green, Donald P. and Alan S. Gerber. 2002. "The Downstream Benefits of Experimentation." *Political Analysis* 10(4):394–402.

Green, Donald P., Alan S. Gerber and David W. Nickerson. 2003. "Getting Out the Vote in Local Elections: Results from Six Door-to-Door Canvassing Experiments." *Journal of Politics* 65(4):1083–1096.

Green, Donald P., Mary C. McGrath and Peter M. Aronow. 2013. "Field Experiments and the Study of Voter Turnout." *Journal of Elections, Public Opinion, and Parties* 23(1):27–48.

Heckman, James A. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47:153–161.

Heckman, James J. and Edward J. Vytlacil. 2007. Economic Evaluation of Social Programs, Part 1: Causal Models, Structural Models and Econometric Policy Evaluation. In *Handbook of Econometrics*. Vol. 6 New York, NY: Elsevier chapter 70, pp. 4779–4874.

Hernán, Miguel A. and James M. Robins. 2006. "Instruments for Causal Inference: An Epidemiologists Dream." *Epidemiology* 17(4):360–372.

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.

Imbens, Guido W. 2001. "Comment on: Estimation of limited dependent variable models with dummy endogenous regressors: simple strategies for empirical practice." *Journal of Business & Economic Statistics* 19(1):17–20.

Imbens, Guido W. and Joshua D. Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62(2):467–476.

Little, Roderick J. 1985. "A Note About Models for Selectivity Bias." *Econometrica* 53(6):1469–1474.

Nagelkerke, N., V. Fidler, R. Bersen and M. Borgdorff. 2000. "Estimating treatment effects in randomized clinical trials in the presence of non-compliance." *Statistics in Medicine* 19(4):1849–1864.

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

Rivers, Douglas and Quang H. Vuong. 1988. "Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models." *Journal of Econometrics* 39(3):347–366.

Robins, James M. 1989. The Analysis of Randomized and Non-Randomized AIDS Treatment Trials Using A New Approach To Causal Inference In Longitudinal Studies. In *Healther Service Research Methodology: A Focus on AIDS*, ed. L. Sechrest, H. Freeman and A. Mulley. Washington D.C.: US Public Health Service, National Center for Health Services Research pp. 113–159.

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

Robins, James M.and Andrea Rotnitzky. 1992. Recovery of Information and Adjustment for Dependent Censoring Using Surrogate Markers. In *AIDS Epidemiology-Methodological Issues*, ed. N Jewell, K. Dietz and V. Farewell. Boston: Birkhauser pp. 297–331.

Robins, J.M. and S. Greenland. 1992. "Identifiability and Exchangeability For Direct and Indirect Effects." *Epidemiology* 3(2):143–155.

Rosenbaum, Paul R. 1984. "The Consequences of Adjusting For a Concomitant Variable That Has Been Affected By The Treatment." *Journal of The Royal Statistical Society Series A* 79(1):41–48.

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

Stephens, Alisa, Luke J. Keele and Marshall Joffe. 2013. "Estimating Post-Treatment Effect Modification With Generalized Structural Mean Models."

Tan, Zhiqiang. 2006. "Regression and Weighting Methods for Causal Inference Using Instrumental Variables." *Journal of the American Statistical Association* 101(476):1607–1618.

Ten Have, Thomas R., Marshall Joffe, Kevin G. Lynch, Gregory K. Brown, Stephen A. Maisto and Aaron T. Beck. 2007. "Causal Mediation Analyses with Rank Preserving Models." *Biometrics* 63(3):926–934.

Vansteelandt, Stijn and Els Goetghebeur. 2003. "Causal Inference with Generalized Structural Mean Models." *Journal of the Royal Statistical Society, Series B* 65(4):817–835.

Vansteelandt, Stijn and Els Goetghebeur. 2004. "Using Potential Outcomes as Predictors of Treatment Activity Via Strong Structural Mean Models." *Statistica Sinica* 14(3):907–925.

Vansteelandt, Stijn, Jack Bowden, Manoochehr Babanezhad and Els Goetghebeur. 2011. "On Instrumental Variables Estimaton of Causal Odds Ratios." *Statistical Science* 26(3):403–422.