Sociology
UNIVERSITY OF TORONTO

Working Papers

Regression-based adjustment for time-varying onfounders

Geoffrey T. Wodtke
University of Toronto

UT Sociology Working Paper No. 2017-01
February, 2017

REGRESSION-BASED ADJUSTMENT FOR TIME-VARYING CONFOUNDERS

Geoffrey T. Wodtke

University of Toronto

December 14, 2016

Corresponding Author: Geoffrey T. Wodtke, Department of Sociology, University of Toronto, 725 Spadina Avenue, Toronto, ON M5S SJ4, Canada. Email: geoffrey.wodtke@utoronto.ca

Acknowledgements: The author thanks Markus Schafer, Andrew Miles, and participants in the Department of Statistical Sciences seminar series at the University of Toronto for helpful comments and discussions on previous iterations of this study. This research was supported by a Connaught New Researcher Award from the University of Toronto.

Keywords: marginal structural models, structural nested mean models, inverse probability of treatment weighting, regression with residuals, time-varying confounders

ABSTRACT

Social scientists are often interested in estimating the marginal effects of a time-varying treatment on an end-of-study continuous outcome. With observational data, estimating these effects is complicated by the presence of time-varying confounders affected by prior treatments, which may lead to bias in conventional regression and matching estimators. In this situation, the inverse-probability-of-treatment-weighted (IPTW) estimator remains unbiased if treatment assignment is sequentially ignorable and the conditional probability of treatment is correctly modeled, but this method is not without limitations. In particular, it is difficult to use with continuous treatments, and it is relatively inefficient. This article proposes an alternative regression-based estimator – two-stage regression-with-residuals (RWR) – that may overcome some of these limitations in practice. It is unbiased for the marginal effects of a time-varying treatment if treatment assignment is sequentially ignorable, the treatment effects of interest are invariant across levels of the confounders, and a model for the conditional mean of the outcome is correctly specified. The performance of the RWR estimator relative to the IPTW estimator is evaluated with a series of simulation experiments and with an empirical example based on longitudinal data from the Panel Study of Income Dynamics. Results indicate that it may outperform the IPTW estimator, at least in certain situations.

1. INTRODUCTION

In the social sciences, researchers are often interested in estimating the marginal, or population average, effects of a time-varying treatment, intervention, or exposure on an end-of-study continuous outcome. Many exposures of interest to social scientists vary over time, and theories of human development and the life course suggest that their effects are likely time-dependent (Bronfenbrenner and Morris 1992; Elder 1998). For example, families move and neighborhoods change, and it is commonly hypothesized that the effects of neighborhood context on academic achievement may depend on the duration or timing of exposure throughout the early life course (Sharkey and Faber 2014; Wodtke et al. 2011, 2016). Similarly, children move between schools, classrooms, and teachers as they grow up, and the effects of these contextual factors on educational outcomes are also widely thought to be a function of exposure duration or timing (Cunha and Heckman 2007; Halpern-Manners et al. 2009). Finally, the structure of the family with which an individual resides is ever evolving, and the effects of family composition on academic achievement could be lagged, cumulative, or heterogeneous during different stages of development (Heard 2007; Krein and Beller 1988). Analyses of how neighborhoods, schools, or families affect human capital formation and status attainment should therefore involve estimating the effects of time-varying contextual trajectories rather than just the effects of point-in-time exposures.

With observational data, estimating the marginal effects of different treatment trajectories is complicated by the presence of time-varying confounders affected by prior treatments. A confounder is a variable that affects both treatment assignment and the outcome, and if left uncontrolled, it will lead to bias in estimates of marginal treatment effects. In the time-varying setting, the values of a confounder may change over time as a function of prior treatments—that

is, selection into different treatments at each point in time may be based in part on time-varying determinants of the outcome that are themselves affected by treatments received at an earlier point in time. In this situation, conventional regression and matching estimators that naively condition on time-varying confounders in an attempt to mitigate confounding bias may induce other forms of bias that are just as problematic. Specifically, naively conditioning on a time-varying confounder affected by prior treatments can lead to over-control and endogenous selection biases (Elwert and Winship 2014; Robins 1999; Robins et al. 2000). Because of these complications, alternative methods are generally required for estimating the marginal effects of a time-varying treatment with observational data.

The most widely used alternative method for estimating the marginal effects of a time-varying treatment is inverse-probability-of-treatment-weighted (IPTW) estimation of a marginal structural model (MSM; Robins 1999; Robins et al. 1994, 2000). A MSM is a model for the marginal mean of the potential outcomes associated with each possible treatment trajectory, and IPTW estimation involves fitting a regression of the outcome on prior treatments to a weighted pseudo-sample in which the time-varying confounders have been appropriately balanced across treatment levels at each time point. Covariate balance is achieved by weighting each observation by the inverse probability that the subject is exposed to her observed level of treatment at each time point conditional on her observed covariate history at each time point. Under the assumptions that treatment assignment is sequentially ignorable (i.e., that treatment is not confounded by unmeasured variables) and that the conditional probability of treatment is correctly modeled, this method is unbiased for the marginal effects of a time-varying treatment even when observed time-varying confounders are affected by prior treatments.

Although IPTW estimation of a MSM avoids the biases associated with conventional regression and matching estimators in the time-varying setting, it is not without limitations. In particular, it is difficult to use with continuous treatments (Naimi et al. 2014), and it is relatively inefficient (Lunceford and Davidian 2004; Robins et al. 1994). In this article, I propose an alternative regression-based estimator for the marginal effects of a time-varying treatment that also avoids the biases associated with conventional regression and matching estimators in the time-varying setting and that may overcome several of the limitations associated with the IPTW estimator. Specifically, I propose regression-with-residuals (RWR) estimation of a constrained structural nested mean model (SNMM) in which the treatment effects of interest are assumed to be invariant across levels of the time-varying confounders. A SNMM is a particular decomposition of the conditional mean of the potential outcomes given the confounders into a set of moderated, or subpopulation average, treatment effects and then a set of so-called "nuisance" associations between the confounders and potential outcomes (Robins 1994; Robins et al. 2007; Author Forthcoming). Under the assumption that there is no effect moderation across levels of the confounders, the treatment coefficients in an appropriately constrained SNMM are equivalent to marginal effects, and thus RWR can be used to estimate the marginal effects of a time-varying treatment in this situation.

RWR estimation differs only slightly from conventional least squares regression. The method proceeds in two stages: first, the confounders at each time point are regressed on all prior variables and then residualized, and second, the outcome is regressed on prior treatments and the residualized confounders to estimate the treatment effects of interest (Almirall et al. 2010, 2013; Author Forthcoming). By residualizing the confounders in the first stage, which purges them of their association with prior treatments, RWR avoids the biases that arise with conventional

regression and matching estimators in the time-varying setting. Furthermore, because it involves a type of regression-based adjustment for observed confounding that does not require modeling the conditional probability of treatment, RWR estimation may be better suited for use with continuous treatments and may also be more efficient than IPTW estimation. Nevertheless, RWR is based on a number of strong modeling assumptions, and if they are violated in practice, this may negate the method's purported advantages.

In the sections that follow, I first define the marginal effects of interest using the potential outcomes framework, explain the conditions needed for their identification from observed data, and outline the problems that arise when estimating these effects using conventional regression or matching. Next, I review IPTW estimation of a MSM and then introduce the SNMM and its RWR estimator, with a focus on the situation where treatment effects are invariant across levels of the confounders. Finally, I investigate the performance of RWR relative to IPTW using a series of simulation experiments, and then I illustrate an application of RWR to estimate the marginal effects of exposure to a disadvantaged neighborhood during childhood and adolescence on achievement test scores using data from the Panel Study of Income Dynamics (PSID). I conclude with a discussion of the limitations associated with regression-based adjustment for time-varying confounders.

2. NOTATION, ESTIMANDS, AND IDENTIFICATION

This section formally defines marginal effects in the longitudinal setting using the potential outcomes framework (Holland 1986; Rubin 1974). For expositional simplicity, I focus on an example with a binary treatment measured at two points in time, a single binary confounder

measured at two points in time, and a continuous end-of-study outcome measured after the treatments and confounders.

Let a_t denote exposure to a binary treatment at time $t \in \{1,2\}$, where $a_t = 1$ indicates exposure to treatment and $a_t = 0$ indicates the absence of treatment at this time point. Next, let $Y_i(a_1, a_2)$ be the end-of-study continuous outcome for subject i had she previously received the treatment sequence $\{a_1, a_2\}$. In this framework, each subject is conceived to have a set of potential outcomes corresponding to all possible treatment sequences, and contrasts between different potential outcomes define causal effects. For example, $Y_i(0,0)$ is the potential outcome for subject i had she never received treatment, $Y_i(1,1)$ is her outcome had she always received treatment, and the difference between these potential outcomes, $Y_i(1,1) - Y_i(0,0)$, is the individual causal effect of being always versus never treated for subject i. The set $\{Y_i(0,0), Y_i(1,0), Y_i(0,1), Y_i(1,1)\}$ represents all possible end-of-study potential outcomes. In practice, only the single potential outcome corresponding to the treatment sequence actually received is ever observed, and the other potential outcomes are counterfactuals that cannot be observed. Causal effects are therefore defined in terms of contrasts that are fundamentally unobservable, and as a result, their estimation always requires strong untestable assumptions, which I discuss in detail below. Finally, let C_{i1} denote the binary confounder for subject i at time $t = 1$, which is measured just prior to treatment at time $t = 1$. In addition, let $C_{i2}(a_1)$ denote the binary confounder for subject i at time $t = 2$, which is measured after treatment at time $t = 1$ but before treatment at time $t = 2$. Because $C_{i2}(a_1)$ is measured after treatment at time $t = 1$, it is indexed by a_1 as a potential outcome to reflect that it may be affected by prior treatment. The set $\{C_{i2}(0), C_{i2}(1)\}$ represents all possible potential outcomes for this variable.

Based on this notation, the average marginal effects of treatment on the end-of-study outcome can be defined using two different causal functions, one for each time point. The first causal function is

$$\lambda_1(a_1) = E\big(Y_i(a_1, 0) - Y_i(0,0)\big),$$

which gives the average marginal effect of being treated only at time $t = 1$ rather than never being treated. The second causal function is

$$\lambda_2(a_1, a_2) = E\big(Y_i(a_1, a_2) - Y_i(a_1, 0)\big),$$

which gives the average marginal effect of being treated at time $t = 2$ rather than not being treated at time $t = 2$ had subjects initially been exposed to treatment a_1. When $a_1 = 0$, this function gives the average marginal effect of being treated only at time $t = 2$ rather than never being treated, and when $a_1 = 1$, it gives the average marginal effect of being always treated versus being treated only at time $t = 1$. At each time point, these functions isolate the average marginal effect of one additional wave of treatment versus no additional waves of treatment, and thus their sum,

$$\begin{aligned}\lambda_1(a_1) + \lambda_2(a_1, a_2) &= E\big(Y_i(a_1, 0) - Y_i(0,0)\big) + E\big(Y_i(a_1, a_2) - Y_i(a_1, 0)\big) \\ &= E\big(Y_i(a_1, a_2) - Y_i(0,0)\big),\end{aligned}$$

gives the average marginal effect of being always treated versus never treated. Taken together, these expressions capture the distal, proximal, and cumulative marginal effects of a time-varying treatment on an end-of-study outcome.

Now, let the set $\{C_{i1}, A_{i1}, C_{i2}, A_{i2}, Y_i\}$ represent the observed data in temporal order. The average marginal effects defined previously can be identified from these data under the assumptions of consistency, positivity, and sequential ignorability of treatment assignment (Robins 1999; Robins et al. 1994, 2000). The consistency assumption states that the observed

outcome is consistent with one of the potential outcomes – that is, $Y_i = A_{i1}A_{i2}Y_i(1,1) + A_{i1}(1-A_{i2})Y_i(1,0) + (1-A_{i1})A_{i2}Y_i(0,1) + (1-A_{i1})(1-A_{i2})Y_i(0,0)$. The positivity assumption states that it is possible for there to be both treated and untreated individuals in each subset of the target population defined by the confounders – that is, it must be possible for each population member to be exposed to all levels of treatment regardless of their covariate history. The sequential ignorability assumption is formally expressed in two parts:

$$Y_i(a_1, a_2) \perp A_{i1} | C_{i1} \; \forall \; (a_1, a_2) \text{ and } Y_i(a_1, a_2) \perp A_{i2} | C_{i1}, A_{i1}, C_{i2} \; \forall \; (a_1, a_2).$$

where $\perp$ denotes statistical independence. In words, this assumptions states that the potential outcomes are conditionally independent of treatment at each time point given the observed past. Substantively, it states that there are not any variables other than the observed confounders that directly affect selection into treatment and the outcome.[1]

Figure 1 displays a directed acyclic graph (Elwert 2013; Pearl 2009) that describes a time-dependent causal system in which this assumption is satisfied. In this graph, nodes represent variables, and arrows between nodes represent direct causal effects. The graph shows that the confounder at each time point directly affects both future treatment and the outcome, that treatment at each time point directly affects the outcome, and that treatment at time 1 also indirectly affects the outcome through the confounder at time 2. The sequential ignorability assumption is satisfied in this figure because the only unobserved variable, denoted by U_i, does not directly affect treatment; rather, it directly affects only the observed confounder and the outcome. Thus, with observational data, the key to estimating the marginal effects of a time-

[1] These effects can also be identified if the potential outcomes are marginally independent of the observed treatment at each time point – that is, if there is not any confounding of treatment whatsoever, whether observed or not. This assumption is met by design in sequentially randomized experimental studies, but it is rarely satisfied with data from an observational research design.

varying treatment is to measure and properly adjust for all variables that directly affect both treatment and the outcome.

3. OVER-CONTROL AND ENDOGENOUS SELECTION BIAS

This section explains the problems associated with conventional regression and matching estimators for the marginal effects of a time-varying treatment in the presence of a time-varying confounder affected by prior treatments. These problems occur even when the identification assumptions outlined previously are satisfied. I use the term "conventional regression and matching" to refer to the broad class of methods that adjust for observed confounders by conditioning on these variables either directly via their inclusion in a linear regression model for the outcome or indirectly via matching on the propensity score.

Conditioning naively on a time-varying confounder affected by past treatment engenders two forms of bias. First, it leads to bias from over-control of intermediate pathways, which is depicted graphically in Figure 2. Conditioning on a variable that lies on the causal pathway from treatment to the outcome partials out the indirect effect of treatment operating through this variable (VanderWeele 2015). More specifically, Figure 2 shows that conditioning on the confounder at time $t = 2$ would partial out the indirect effect of treatment at time $t = 1$ that operates through future levels of the confounder. Thus, conventional regression and matching estimators for the average marginal effect of treatment at time $t = 1$, and by extension, for the cumulative marginal effect of treatment at both time periods, would be biased in this situation.

Second, conditioning naively on a time-varying confounder affected by past treatment may lead to endogenous selection bias, which is depicted graphically in Figure 3.[2] Conditioning on an outcome of two or more variables induces a non-causal association between them (Elwert and Winship 2014; Pearl 2009). More specifically, Figure 3 shows that conditioning on the confounder at time $t = 2$ would induce a non-causal association between prior treatment and the unobserved variable, U_i. Furthermore, because U_i affects Y_i, conditioning on the confounder at time $t = 2$ would induce a non-causal association between prior treatment and the outcome. This induced non-causal association between prior treatment and the outcome would also lead to bias in conventional regression and matching estimators for the marginal effect of treatment at time $t = 1$, and by extension, and for the cumulative marginal effect of treatment at both time periods.

In sum, a failure to measure and properly adjust for all variables that directly affect both treatment and the outcome would lead to confounding bias in estimates of the marginal effects defined previously. But even when the sequential ignorability assumption is satisfied and there is no unobserved confounding of treatment, conventional regression and matching estimators suffer from over-control and endogenous selection biases because they improperly adjust for time-varying confounders affected by prior treatments. Estimating marginal effects in the longitudinal setting therefore demands the use of alternative methods that can adjust for observed confounding while avoiding these other forms of bias.

[2] Endogenous selection bias is also commonly referred to as "collider-stratification bias" in this literature (e.g., Wodtke et al. 2016; Author Forthcoming).

4. THE MSM AND ITS IPTW ESTIMATOR

This section explains the MSM and its IPTW estimator, which provide an alternative method for estimating the marginal effects of a time-varying treatment that avoids the limitations associated with conventional regression and matching. A MSM formally relates the marginal effects of interest, $\lambda_1(a_1)$ and $\lambda_2(a_1, a_2)$, to the marginal mean of the potential outcomes as follows:

$$E\big(Y_i(a_1, a_2)\big) = E\big(Y_i(0,0)\big) + E\big(Y_i(a_1, 0) - Y_{i3}(0,0)\big) + E\big(Y_i(a_1, a_2) - Y_i(a_1, 0)\big) =$$

$$E\big(Y_i(0,0)\big) + \lambda_1(a_1) + \lambda_2(a_1, a_2) = \alpha_{00} + \beta_{10}a_1 + a_2(\beta_{20} + \beta_{21}a_1),$$

where the intercept term, $E\big(Y_i(0,0)\big) = \alpha_{00}$, is the mean under no treatment and the causal functions for each time t, $\lambda_1(a_1) = E\big(Y_i(a_1, 0) - Y_i(0,0)\big) = \beta_{10}a_1$ and $\lambda_2(a_1, a_2) = E\big(Y_i(a_1, a_2) - Y_i(a_1, 0)\big) = a_2(\beta_{20} + \beta_{21}a_1)$, have been parameterized under the constraint that they must equal zero when treatment at time t is equal to zero. This is a saturated MSM with main effects for treatment at each time point and a cross-time interaction between treatments at different time points. In this model, β_{10} is the average marginal effect of being treated only at time $t = 1$ versus never being treated; β_{20} is the average marginal effect of being treated only at time $t = 2$ versus never being treated; $\beta_{20} + \beta_{21}$ is the average marginal effect of being always treated versus being treated only at time $t = 1$; and $\beta_{10} + \beta_{20} + \beta_{21}$ is the average marginal effect of being always treated versus never treated.

A MSM is estimated from the observed data by the method of IPTW. This involves regressing the outcome on prior treatments using a weighted pseudo-sample in which the time-varying confounders have been appropriately balanced across treatment levels at each time point. More specifically, IPTW estimates are computed by fitting the following regression model,

$$E(Y_i|A_{i1}, A_{i2}) = \alpha^*_{00} + \beta^*_{10}A_{i1} + A_{i2}(\beta^*_{20} + \beta^*_{21}A_{i1}),$$

to the observed data using the method of weighted least squares (WLS) with weights for each subject given by

$$w_i = \frac{1}{P(A_{i1} = a_{i1}|C_{i1})} \times \frac{1}{P(A_{i2} = a_{i2}|C_{i1}, A_{i1}, C_{i2})}.$$

The asterisks on the parameters in this model are used to denote the distinction between the counterfactual mean differences parameterized in an MSM and the conditional mean differences parameterized in an observed data regression. The denominator of w_i is the conditional probability that a subject is exposed to her actual treatment at each time t given prior treatments and confounders. At each time point, weighting by w_i balances prior confounders across levels of treatment by giving more weight to subjects with confounder histories that are underrepresented in a treatment group and less weight to subjects with confounder histories that are overrepresented in a treatment group.

Figure 4 presents a stylized graph that illustrates the effect of weighting the observed data by w_i. In the weighted pseudo-sample, treatment at each time point is independent of prior measures of the time-varying confounder. Thus, conditioning on these measures directly in a regression model for the outcome is unnecessary, and a simplified regression that conditions only on prior treatments can be fit to the weighted data in order to estimate the marginal effects of interest. By using weights to adjust for the time-varying confounders rather than conditioning naively on these variables, IPTW estimation of an MSM avoids bias due to over-control and endogenous selection.

Because effect estimates based on w_i can be imprecise and may have non-normal sampling distributions, IPTW estimation is typically based on a stabilized version of the weight, such as

$$sw_i = \frac{P(A_{i1}=a_{i1})}{P(A_{i1} = a_{i1}|C_{i1})} \times \frac{P(A_{i2} = a_{i2}|A_{i1})}{P(A_{i2} = a_{i2}|C_{i1}, A_{i1}, C_{i2})}.$$

This weight also balances prior confounders across levels of treatment at each time point, but it is less variable than w_i and yields effect estimates with an approximately normal sampling distribution (Robins 1999; Robins et al. 1994, 2000). It differs from w_i only in that the numerator is replaced with the probability that a subject is exposed to her actual treatment at each time t conditional on prior treatments. Several different versions of the stabilized weight are commonly used in practice. For example, when also adjusting for a set of time-invariant confounders, denoted by V_i, another common version of the stabilized weight is

$$sw_i = \frac{P(A_{i1}=a_{i1}|V_i)}{P(A_{i1}=a_{i1}|V_i,C_{i1})} \times \frac{P(A_{i2}=a_{i2}|V_i,A_{i1})}{P(A_{i2}=a_{i2}|V_i,C_{i1},A_{i1},C_{i2})},$$

where the probabilities in both the numerator and denominator are also made conditional on these factors. With this version of the weight, the regression model for the outcome must now adjust directly for V_i, but this is not problematic in general because these variables are unaffected by treatment.

The conditional probabilities that compose the weights are unknown quantities that must be estimated from the observed data. This is typically accomplished by estimating a series of generalized linear models (GLM) for the conditional probability of treatment at each time point, although alternative methods are also available (e.g., McCaffrey et al. 2004). With a binary treatment, the conditional probabilities are usually estimated using a GLM with the logit or probit link function. With more complex treatments, more complex models may be required. Note, however, that the functional form of these models is nontrivial. In addition to the consistency, positivity, and sequential ignorability assumptions outlined previously, IPTW estimation also requires correctly specified models for the denominator probabilities in the weight. In other words, IPTW estimation is only unbiased under the additional modeling

assumption that the relationship between measured variables and treatment assignment at each time point has been correctly specified.

IPTW estimation of a MSM avoids bias due to over-control and endogenous selection and is thus preferred over conventional regression and matching in the time-varying setting, but it is not without limitations. First, IPTW estimation can be difficult to implement with continuous treatments, as opposed to binary or polytomous treatments, because in this situation the method additionally requires identifying a correct parametric form for the treatment density function at each time point as well as properly accounting for potential heteroscedasticity in the conditional treatment variance (Naimi et al. 2014). In many social science applications, satisfying these additional requirements is very difficult, and as a result, researchers often resort to arbitrary dichotomization or quantile binning of continuous treatments in order to simplify estimation of the weights (e.g., Sampson et al. 2008; Sharkey and Elwert 2011; Wodtke 2013; Wodtke et al. 2011). Dichotomizing or binning continuous treatments, however, is associated with a loss of information, reduced statistical power, and residual confounding (Royston et al. 2006). The practice should therefore be avoided unless absolutely necessary.

Second, IPTW estimation is relatively inefficient (Lunceford and Davidian 2004; Robins et al. 1994). For example, when its motivating assumptions are met in the point-in-time setting, the conventional regression estimator for the marginal effect of treatment has a smaller sampling variance than the IPTW estimator, as is implied by standard large sample theory (Lunceford and Davidian 2004). Similarly, in the point-in-time setting, augmenting the IPTW estimator by additionally adjusting for confounders directly in the outcome regression also improves its efficiency (Robins et al. 1994). These results suggest that regression-based adjustment for observed confounding is in general more efficient than adjustment via weighting, at least when

the assumptions that motivate regression-based adjustment are satisfied. In addition, regression-based adjustment for observed confounding also does not require a correct model for the conditional probability of treatment at each time point, which makes it easier to use with continuous treatments. Thus, if regression-based methods can be adapted to overcome the biases associated with over-control and endogenous selection in the time-varying setting, they may outperform IPTW estimation of an MSM, at least in certain situations.

5. THE SNMM AND ITS RWR ESTIMATOR

This section explains the SNMM and its RWR estimator. A SNMM is a model for the moderated, rather than marginal, effects of a time-varying treatment (Almirall et al. 2010, 2013; Robins 1994; Robins et al. 2007). Moderated, or subpopulation average, effects refer to the potentially different effects of treatment among all the different subsets of the population defined by prior levels of the confounders. Although a SNMM is in general a model for moderated effects, there are certain situations in which these moderated effects are equivalent to marginal effects. In particular, when the effects of treatment are invariant across levels of the prior confounders, the moderated effects in a SNMM are equivalent to the marginal effects of interest, and thus a simplified SNMM can be fit by the method of RWR—an adaptation of conventional least squares regression that avoids biases from over-control and endogenous selection—to estimate them.

A SNMM is based on the following decomposition of the conditional mean of $Y_i(a_1, a_2)$ given C_{i1} and $C_{i2}(a_1)$ into causal and non-causal components:

$$E\big(Y_i(a_1,a_2)|C_{i1},C_{i2}(a_1)\big) = E\big(Y_i(0,0)\big) + \big[E(Y_i(0,0)|C_{i1}) - E\big(Y_i(0,0)\big)\big] + E(Y_i(a_1,0) -$$

$$Y_i(0,0)|C_{i1}) + \big[E\big(Y_i(a_1,0)\big|C_{i1},C_{i2}(a_1)\big) - E(Y_i(a_1,0)|C_{i1})\big] + E\big(Y_i(a_1,a_2) -$$

$$Y_i(a_1, 0)|C_{i1}, C_{i2}(a_1)\big) = \alpha_{00} + \varepsilon_1(C_{i1}) + \mu_1(C_{i1}, a_1) + \varepsilon_2\big(C_{i1}, a_1, C_{i2}(a_1)\big) +$$

$$\mu_2(C_{i1}, a_1, C_{i2}(a_1), a_2).$$

In this decomposition, $\alpha_{00} = E\big(Y_i(0,0)\big)$ is the mean of the potential outcomes under no treatment; $\mu_1(C_{i1}, a_1) = E(Y_i(a_1, 0) - Y_i(0,0)|C_{i1})$ and $\mu_2(C_{i1}, a_1, C_{i2}(a_1), a_2) = E\big(Y_i(a_1, a_2) - Y_i(a_1, 0)|C_{i1}, C_{i2}(a_1)\big)$ are causal functions that capture the moderated effects of treatment at each time t given prior measures of the confounders; and $\varepsilon_1(C_{i1}) = \big[E(Y_i(0,0)|C_{i1}) - E\big(Y_i(0,0)\big)\big]$ and $\varepsilon_2\big(C_{i1}, a_1, C_{i2}(a_1)\big) = \big[E\big(Y_i(a_1, 0)\big|C_{i1}, C_{i2}(a_1)\big) - E(Y_i(a_1, 0)|C_{i1})\big]$ are so-called nuisance functions that capture both causal and possibly non-causal associations between the confounders and outcome. They are called "nuisance" functions because they do not contain any information about the causal effects of treatment.

The moderated effects in a SNMM represent the average causal effect of one additional wave of treatment at each time point among specific subpopulations defined in terms of the prior confounders. More specifically, $\mu_1(C_{i1}, a_1) = E(Y_i(a_1, 0) - Y_i(0,0)|C_{i1})$ is the average causal effect of being exposed to treatment only at time $t = 1$ among the subpopulations defined by levels of C_{i1}. Similarly, $\mu_2(C_{i1}, a_1, C_{i2}(a_1), a_2) = E\big(Y_i(a_1, a_2) - Y_i(a_1, 0)|C_{i1}, C_{i2}(a_1)\big)$ is the average causal effect of being exposed to treatment at time $t = 2$ among the subpopulations defined by levels of both $C_{i1}$and $C_{i2}(a_1)$ had subjects initially been exposed to treatment a_1.

Any parameterization of $\mu_1(C_{i1}, a_1)$ must satisfy the constraint that it is equal to zero when treatment at time $t = 1$ is equal to zero. With a binary treatment and confounder, a saturated parameterization for the first causal function is

$$\mu_1(C_{i1}, a_1) = E(Y_i(a_1, 0) - Y_i(0,0)|C_{i1}) = a_1(\theta_{10} + \theta_{11}C_{i1}),$$

which includes a main effect for treatment at time $t = 1$ and then an interaction between this treatment and the confounder measured at time $t = 1$. In this expression, θ_{10} is the average

causal effect of treatment at time $t = 1$ among subjects with $C_{i1} = 0$, and $\theta_{10} + \theta_{11}$ is the average causal effect of treatment at time $t = 1$ among subjects with $C_{i1} = 1$. A saturated parameterization for the second causal function is

$$\mu_2(C_{i1}, a_1, C_{i2}(a_1), a_2) = E\big(Y_i(a_1, a_2) - Y_i(a_1, 0)|C_{i1}, C_{i2}(a_1)\big) = a_2\big(\theta_{20} + \theta_{21}C_{i1} + \theta_{22}C_{i2}(a_1) + \theta_{23}C_{i1}C_{i2}(a_1) + \theta_{24}a_1 + \theta_{25}C_{i1}a_1 + \theta_{26}a_1C_{i2}(a_1) + \theta_{27}C_{i1}a_1C_{i2}(a_1)\big),$$

which includes a main effect for treatment at time $t = 2$ and then all possible interactions between this treatment, prior treatment, and the prior confounders. Different combinations of the theta parameters in this expression give the average causal effect of treatment at time $t = 2$ among all possible subpopulations defined by the prior confounders had subjects initially been exposed to treatment a_1.

The nuisance functions in a SNMM capture both causal and possibly non-causal associations between the confounders at each time point and the outcome. More specifically, $\varepsilon_1(C_{i1}) = \big[E(Y_i(0,0)|C_{i1}) - E\big(Y_i(0,0)\big)\big]$ represents the association between C_{i1} and the outcome had subjects never been exposed to treatment, and $\varepsilon_2\big(C_{i1}, a_1, C_{i2}(a_1)\big) = \big[E\big(Y_i(a_1, 0)\big|C_{i1}, C_{i2}(a_1)\big) - E(Y_i(a_1, 0)|C_{i1})\big]$ represents the association between $C_{i2}(a_1)$ and the outcome among the subpopulations defined by C_{i1} had they been exposed to the treatment sequence $\{a_1, 0\}$. An important property of the nuisance functions is that they have mean zero given the past—that is, the unconditional mean of $\varepsilon_1(C_{i1})$ is equal to zero and the conditional mean of $\varepsilon_2\big(C_{i1}, a_1, C_{i2}(a_1)\big)$ given C_{i1} is also equal to zero.[3]

[3] This is because $E\big(\varepsilon_1(C_{i1})\big) = E\big[E(Y_i(0,0)|C_{i1}) - E\big(Y_i(0,0)\big)\big] = E\big(E(Y_i(0,0)|C_{i1})\big) - E\big(Y_i(0,0)\big) = E\big(Y_i(0,0)\big) - E\big(Y_i(0,0)\big) = 0$ and $E\big(\varepsilon_2\big(C_{i1}, a_1, C_{i2}(a_1)\big)|C_{i1}\big) = E\big[E\big(Y_i(a_1, 0)\big|C_{i1}, C_{i2}(a_1)\big) - E(Y_i(a_1, 0)|C_{i1})|C_{i1}\big] = E\big[E\big(Y_i(a_1, 0)\big|C_{i1}, C_{i2}(a_1)\big)|C_{i1}\big] - E(Y_i(a_1, 0)|C_{i1}) = E(Y_i(a_1, 0)|C_{i1}) - E(Y_i(a_1, 0)|C_{i1}) = 0$.

Any parameterization of the nuisance functions must satisfy these zero mean constraints. A saturated parameterization for the first nuisance function is

$$\varepsilon_1(C_{i1}) = \left[E(Y_i(0,0)|C_{i1}) - E\big(Y_i(0,0)\big)\right] = \big(C_{i1} - E(C_{i1})\big)\gamma_{10},$$

where γ_{10} gives the associational effect of the confounder at time $t = 1$ on the outcome had subjects never been exposed to treatment and $\big(C_{i1} - E(C_{i1})\big)$ is a residual transformation of the confounder that centers it around its population mean. It satisfies the zero mean constraint because

$$E\left(\big(C_{i1} - E(C_{i1})\big)\gamma_{10}\right) = \big(E(C_{i1}) - E(C_{i1})\big)\gamma_{10} = 0.$$

A saturated parameterization of the second nuisance function is

$$\varepsilon_2\big(C_{i1}, a_1, C_{i2}(a_1)\big) = \left[E\big(Y_i(a_1,0)\big|C_{i1}, C_{i2}(a_1)\big) - E(Y_i(a_1,0)|C_{i1})\right] = \big(C_{i2}(a_1) - E(C_{i2}(a_1)|C_{i1})\big)(\gamma_{20} + \gamma_{21}C_{i1} + \gamma_{22}a_1 + \gamma_{23}a_1C_{i1}),$$

where different combinations of the gamma parameters give the associational effects of the confounder at time $t = 2$ on the outcome and $\big(C_{i2}(a_1) - E(C_{i2}(a_1)|C_{i1})\big)$ is a residual transformation of the confounder that centers it around its conditional mean. It satisfies the zero mean constraint because

$$E\left(\big(C_{i2}(a_1) - E(C_{i2}(a_1)|C_{i1})\big)(\gamma_{20} + \gamma_{21}C_{i1} + \gamma_{22}a_1 + \gamma_{23}a_1C_{i1})|C_{i1}\right) = \big(E(C_{i2}(a_1)|C_{i1}) - E(C_{i2}(a_1)|C_{i1})\big)(\gamma_{20} + \gamma_{21}C_{i1} + \gamma_{22}a_1 + \gamma_{23}a_1C_{i1}) = 0.$$

Combining the expressions for the causal and nuisance functions outlined previously yields the following saturated SNMM:

$$E\big(Y_i(a_1,a_2)|C_{i1}, C_{i2}(a_1)\big) = \alpha_{00} + \big(C_{i1} - E(C_{i1})\big)\gamma_{10} + a_1(\theta_{10} + \theta_{11}C_{i1}) + \big(C_{i2}(a_1) - E(C_{i2}(a_1)|C_{i1})\big)(\gamma_{20} + \gamma_{21}C_{i1} + \gamma_{22}a_1 + \gamma_{23}a_1C_{i1}) + a_2\big(\theta_{20} + \theta_{21}C_{i1} + \theta_{22}C_{i2}(a_1) + \theta_{23}C_{i1}C_{i2}(a_1) + \theta_{24}a_1 + \theta_{25}C_{i1}a_1 + \theta_{26}a_1C_{i2}(a_1) + \theta_{27}C_{i1}a_1C_{i2}(a_1)\big),$$

This SNMM is nearly identical to a conventional saturated regression model with all possible main effects and interaction terms except that, in several places, it includes terms that resemble residual transformations of the confounders rather than the untransformed values themselves.

A SNMM can be estimated by the method of RWR (Almirall et al. 2010, 2013; Author Forthcoming). This method involves appropriately residualizing the confounders at each time point and then regressing the outcome on prior treatments and the residualized confounders. More specifically, RWR estimates are computed in two stages. First, estimates of $E(C_{i1})$ and $E(C_{i2}|C_{i1}, A_{i1})$ are obtained from the observed data, and then residual terms equal to $\left(C_{i1} - \hat{E}(C_{i1})\right)$ and $\left(C_{i2} - \hat{E}(C_{i2}|C_{i1}, A_{i1})\right)$ are computed using these estimates. With a binary confounder, $E(C_{i1}) = P(C_{i1} = 1)$ and $E(C_{i2}|C_{i1}, A_{i1}) = P(C_{i2} = 1|C_{i1}, A_{i1})$ are typically estimated using GLMs with the logit or probit link function, but alternative methods are also available. With other types of confounders, other types of models can be used.

Second, after obtaining the appropriate residual terms in first stage, the following regression model for the outcome is estimated via least squares:

$$E(Y_i|C_{i1}, A_{i1}, C_{i2}, A_{i2}) = \alpha^*_{00} + \left(C_{i1} - \hat{E}(C_{i1})\right)\gamma_{10} + A_{i1}(\theta^*_{10} + \theta^*_{11}C_{i1}) + \Big(C_{i2} - \hat{E}(C_{i2}|C_{i1}, A_{i1})\Big)(\gamma_{20} + \gamma_{21}C_{i1} + \gamma_{22}A_{i1} + \gamma_{23}A_{i1}C_{i1}) + A_{i2}(\theta^*_{20} + \theta^*_{21}C_{i1} + \theta^*_{22}C_{i2} + \theta^*_{23}C_{i1}C_{i2} + \theta^*_{24}A_{i1} + \theta^*_{25}C_{i1}A_{i1} + \theta^*_{26}A_{i1}C_{i2} + \theta^*_{27}C_{i1}A_{i1}C_{i2}).$$

This model is just the observed-data analogue to the saturated SNMM described previously, and as before, the asterisks are used to denote the distinction between the counterfactual mean differences parameterized in an SNMM and the conditional mean differences parameterized in an observed-data regression. It differs from a conventional saturated regression model only in that estimated residual terms from the first stage have been substituted into the nuisance

functions. Least squares estimation of this model is unbiased for the moderated effects of treatment under the identification assumptions outlined previously.

Figure 5 displays a stylized graph that illustrates the effect of residualizing the confounders at each time point. In particular, it shows that residualizing C_{i2} based on the observed past purges this variable of its association with prior treatment, thereby neutralizing the problematic arrow from A_{i1} to C_{i2} that leads to over-control and endogenous selection bias in conventional regression and matching estimators. In other words, by including the residual transformation of C_{i2} in a regression model for the outcome, rather than just including the untransformed variable, RWR estimation of moderated treatment effects in a SNMM also avoids bias due to over-control and endogenous selection.

Thus far, I have focused on estimating moderated effects in a saturated SNMM, but under the additional assumption that the effect of treatment at each time point is invariant across levels of the prior confounders, these effects are equivalent to marginal effects, and in this situation, the SNMM can be simplified (i.e., appropriately constrained) by omitting all of the treatment-by-confounder interaction terms in the causal functions. This constrained SNMM can then be fit using RWR to estimate the marginal effects of interest. More specifically, if the effect of treatment at time $t = 1$ is the same among the different subpopulations defined by levels of C_{i1}, then $\mu_1(C_{i1}, a_1)$ can be re-expressed as follows:

$$\mu_1(C_{i1}, a_1) = E(Y_i(a_1, 0) - Y_i(0,0)|C_{i1}) = E\big(Y_i(a_1, 0) - Y_i(0,0)\big) = \lambda_1(a_1) = \beta_{10}a_1,$$

which is just the average marginal effect of treatment at time $t = 1$. Similarly, if the effect of treatment at time $t = 2$ is the same among the different subpopulations defined by levels of both C_{i1} and $C_{i2}(a_1)$, then $\mu_2(C_{i1}, a_1 C_{i2}(a_1), a_2)$ can also be re-expressed as follows:

$$\mu_2(C_{i1}, a_1, C_{i2}(a_1), a_2) = E\big(Y_i(a_1, a_2) - Y_i(a_1, 0)|C_{i1}, C_{i2}(a_1)\big) =$$

$$E\big(Y_i(a_1, a_2) - Y_i(a_1, 0)\big) = \lambda_2(a_1, a_2) = a_2(\beta_{20} + \beta_{21}a_1),$$

which is just the average marginal effect of treatment at time $t = 2$. Thus, provided that the assumption of no treatment effect moderation is satisfied, RWR estimation of the following constrained SNMM,

$$E\big(Y_i(a_1, a_2)|C_{i1}, C_{i2}(a_1)\big) = \alpha_{00} + \big(C_{i1} - E(C_{i1})\big)\gamma_{10} + \beta_{10}a_1 + \big(C_{i2}(a_1) -$$

$$E(C_{i2}(a_1)|C_{i1})\big)(\gamma_{20} + \gamma_{21}C_{i1} + \gamma_{22}a_1 + \gamma_{23}a_1C_{i1}) + a_2(\beta_{20} + \beta_{21}a_1),$$

would also be unbiased for the marginal effects of interest.

Moreover, as is the case with any model, a SNMM can be further simplified, or constrained, by invoking additional functional form assumptions. For example, suppose that the causal effects of treatment at each time point and the associational effects of the confounder at each time point are both additive. Translating these additional functional form assumptions into parametric constraints yields the following SNMM:

$$E\big(Y_i(a_1, a_2)|C_{i1}, C_{i2}(a_1)\big) = \alpha_{00} + \big(C_{i1} - E(C_{i1})\big)\gamma_{10} + \beta_{10}a_1 + \big(C_{i2}(a_1) -$$

$$E(C_{i2}(a_1)|C_{i1})\big)\gamma_{20} + \beta_{20}a_2,$$

which closely resembles the conventional linear and additive regression model commonly employed in the social sciences except that it includes main effects for residual transformations of the confounders. If, in addition to the identification assumptions outlined previously, all of these functional form assumptions are satisfied, RWR estimation of this constrained SNMM would also be unbiased for the marginal effects of interest.

With the RWR estimator, however, these types of simplifying functional form assumptions are nontrivial. Even when all of the identification assumptions outlined previously are satisfied, this method is only unbiased if both the causal and nuisance functions in the SNMM are correctly specified. When attempting to estimate marginal effects in a constrained SNMM, the

assumption of correctly specified causal functions implies that the RWR estimator will be biased for these effects if there is any treatment effect moderation across levels of the confounders. In addition, the assumption of correctly specified nuisance functions implies that the RWR estimator will also be biased if either the first-stage models for the confounders or the associational effects of the confounders in the second-stage model for the outcome are incorrectly specified. Thus, RWR estimation of marginal effects in a constrained SNMM is premised on a number of strong modeling assumptions. These modeling assumptions are very similar to those required for estimating marginal effects in the point-in-time setting with a conventional linear and additive regression model fit by ordinary least squares (OLS), and their violation would lead to similar types of misspecification bias (e.g., Morgan and Winship 2015; Winship and Elwert 2010).

Notwithstanding these strong modeling assumptions, RWR estimation of marginal effects in a constrained SNMM may overcome some of the limitations associated with IPTW estimation of an MSM. First, because RWR estimation does not require models for the conditional probability of treatment at each time point, it may be easier to use with continuous treatments. Second, RWR estimation may be relatively more efficient than IPTW estimation because it involves only a slight modification of OLS, which is the minimum-variance estimator when its motivating assumptions are satisfied. This gain in efficiency, however, may be offset in practice by bias resulting from different types of model misspecification, given the strong modeling assumptions that motivate RWR estimation of marginal effects in a constrained SNMM. In the next section, I attempt to shed light on this potential bias-variance tradeoff by investigating the performance of RWR relative to IPTW for estimating marginal effects under several different types of model misspecification.

6. SIMULATION EXPERIMENTS

All simulation experiments are based on 10,000 simulated datasets with two time points. In each simulation, I generate a continuous time-varying confounder $\{C_{i1}, C_{i2}\}$, a binary treatment $\{A_{i1}, A_{i2}\}$, and a continuous end-of-study outcome, Y_i. Specifically, $[C_{i1}] \sim N(\mu_1 = 0, \sigma^2 = 1)$ and $[C_{i2}|C_{i1}, A_{i1}] \sim N(\mu_2 = 0.4C_{i1} + 0.2A_{i1}, \sigma^2 = 1)$; $[A_{i1}|C_{i1}] \sim Bernoulli\big(p_1 = \Phi(0.4C_{i1})\big)$ and $[A_{i2}|C_{i1}, A_{i1}, C_{i2}] \sim Bernoulli\big(p_2 = \Phi(0.2A_{i1} + 0.4C_{i2})\big)$, where Φ is the standard normal cumulative distribution function; and $[Y_i|C_{i1}, A_{i1}, C_{i2}, A_{i2}] \sim N(\mu_3 = 0.4(C_{i1} - \mu_1) + A_{i1}(0.1 + \theta_{11}C_{i1}) + (C_{i2} - \mu_2)(0.4 + \gamma_{21}C_{i1} + \gamma_{22}A_{i1}) + A_{i2}(0.2 + \theta_{21}C_{i1} + 0.1A_{i1}), \sigma^2 = 1)$, where $\{\gamma_{21}, \gamma_{22}\}$ are parameters used to modify the associational effects of the confounder and $\{\theta_{11}, \theta_{21}\}$ are parameters used to modify the magnitude of treatment effect moderation in different simulations. In all simulations, the average marginal effects of interest are identified, and their true values are set to $\lambda_1(1) = 0.1$, $\lambda_2(0,1) = 0.2$, $\lambda_2(1,1) = 0.3$.

The simulation experiments are designed to compare the performance of the RWR estimator and the IPTW estimator for these marginal effects. With the IPTW estimator, correctly specified models for the conditional probability of treatment are always used to estimate stabilized weights. In other words, the modeling assumptions on which the IPTW estimator is based are satisfied in every simulation. With the RWR estimator, the associational effects of the confounders are always assumed to be linear and additive, and the causal effects of treatment are always assumed to be invariant across levels of the confounders. These modeling assumptions are satisfied in some simulations but not in others either because the true associational effects of the confounders are made to be non-additive (i.e., $\gamma_{21} \neq 0$ and $\gamma_{22} \neq 0$) or because the true causal effects of treatment are made to vary across levels of the confounders (i.e., $\theta_{11} \neq 0$ and

$\theta_{21} \neq 0$). The performance of the two estimators is assessed in terms of their bias, standard deviation, and root-mean squared error (RMSE). Stata code for replicating these simulations is provided in the Online Supplement.

Table 1 presents results from a set of simulation experiments that compare the performance of the RWR estimator for marginal effects in a constrained SNMM with correctly specified causal and nuisance functions (i.e., $\gamma_{21} = \gamma_{22} = \theta_{11} = \theta_{21} = 0$) to that of the IPTW estimator with correctly specified models for the conditional probability of treatment. In these experiments, the sample size is varied from "small" ($n = 500$) to "large" ($n = 2000$), and performance measures are compared across estimators. As expected, results indicate that both estimators are unbiased for the marginal effects of interest across all scenarios. And as conjectured, results also indicate that the RWR estimator is less variable than the IPTW estimator. As a result, it consistently outperforms the IPTW estimator in terms of RMSE.

Table 2 presents results from a second set of simulation experiments that evaluate the large sample performance of the RWR estimator for marginal effects in a constrained SNMM when its causal functions are misspecified – that is, when the causal effects of treatment are assumed to be invariant across levels of the confounders but in fact they are not. In these experiments, the magnitude of causal function misspecification is varied, and performance measures are compared across estimators. Specifically, a "minor" degree of causal function misspecification corresponds with setting θ_{11} and θ_{21} equal to 0.05, while "moderate" and "severe" levels of misspecification correspond with two and four times this value, respectively. As expected, results from these experiments show that the RWR estimator is biased for the marginal effects of interest across all scenarios, while the IPTW estimator remains unbiased because it is not premised on any assumptions about the absence of treatment effect moderation. Nevertheless, the RWR estimator

still outperforms the IPTW estimator in terms of RMSE under "minor" and "moderate" levels of causal function misspecification owing to its substantially smaller variance.

Table 3 presents results from a third set of simulation experiments that evaluate the large sample performance of the RWR estimator for marginal effects in a constrained SNMM when its nuisance functions are misspecified – that is, when the associational effects of the confounders are assumed to be additive but in fact they are not. In these experiments, the magnitude of nuisance function misspecification is varied, and performance measures are compared across estimators. Specifically, a "minor" degree of nuisance function misspecification corresponds with setting γ_{21} and γ_{22} equal to 0.05, while "moderate" and "severe" levels of misspecification correspond with two and four times this value, respectively. As expected, results from these experiments show that the RWR estimator is biased for the marginal effects of interest across all scenarios, while the IPTW estimator remains unbiased because it is not premised on any assumptions about correctly modeling the associational effects of the confounders on the outcome. Nevertheless, the RWR estimator still outperforms the IPTW estimator in terms of RMSE under "minor" and "moderate" levels of nuisance function misspecification owing to its substantially smaller variance.

In sum, these simulations indicate that RWR estimation of marginal effects in a constrained SNMM outperforms IPTW estimation when the modeling assumptions on which RWR is based are satisfied because this estimator is considerably less variable. Moreover, the RWR estimator also appears to outperform the IPTW estimator in terms of RMSE even when these modeling assumptions are violated to a minor or moderate degree because in this situation its lower variance outweighs the bias it suffers due to model misspecification. When the modeling assumptions required for RWR estimation of marginal effects are badly violated,

however, its lower variance is outweighed by large biases due to misspecification, and thus the IPTW estimator performs better.

7. AN EMPIRICAL EXAMPLE

In this section, I provide an empirical illustration of how RWR might be used to estimate the marginal effects of a time-varying treatment under the assumptions outlined previously. The example is designed to highlight the problems associated with conventional regression and matching estimators in the time-varying setting as well as the potential advantages of RWR estimation over IPTW estimation for marginal effects. To this end, I focus on an example with a continuous treatment, and I adopt an identification strategy that almost certainly leads to severe over-control and endogenous selection biases when coupled with conventional methods.

More specifically, I focus on estimating the distal, proximal, and cumulative marginal effects of exposure to a disadvantaged neighborhood on adolescent math achievement using data from the PSID main panel and child development supplement (CDS; Michigan Survey Research Center 2014).[4] The PSID main panel began in 1968 with a probability sample of about 4,800 households. These households were re-interviewed annually until 1997 and then biannually thereafter. The CDS is a subcomponent of the PSID designed to provide detailed information about human capital formation during the early life course. It began in 1997 with a sample of 3,563 young children from main panel households. These children were then re-interviewed in

[4] Some of the data used in this analysis are based on "sensitive data files" from the PSID, which were obtained under special contractual arrangements designed to protect the anonymity of respondents. These data are not available from the author. Persons interested in obtaining sensitive data files from the PSID should contact psidhelp@isr.umich.edu. A full set of replication files for this analysis, sans any sensitive data, are available from the author upon request.

2002-2003 and in 2007. The analytic sample for this analysis includes the 1,135 children who were interviewed at the 1997 wave of the CDS when they were between age 3 and 7. These sample members are matched to census tracts using the PSID restricted-use geocode file, and data on the composition of census tracts come from the Geolytics Neighborhood Change Database (GeoLytics 2013).

Treatment in this analysis, denoted by A_{it}, is the socioeconomic composition of a sample member's neighborhood (i.e., census tract of residence). I use principal components analysis to generate a composite measure of neighborhood composition based on a number of interrelated tract characteristics: the poverty rate, the unemployment rate, median household income, the proportion of households that are female-headed, aggregate levels of education, and the proportion of adults in professional or managerial occupations. This is a continuous measure for which higher values represent more disadvantaged neighborhoods and lower values represent more advantaged neighborhoods. It is standardized to have mean zero and unit variance. The outcome in this analysis, denoted by Y_{it}, is math achievement. Math achievement is measured using age-normalized scores on the applied problems subtest from the Woodcock-Johnson Psycho-educational Battery–Revised (Woodcock and Johnson 1989). This variable is also standardized to have mean zero and unit variance. In addition, I measure and adjust for the race, gender, and birth cohort of the sample member as well as for the completed education of the sample member's primary caregiver. Race and gender are both dummy coded—1 for black and 0 for nonblack, and 1 for female and 0 for male. Birth cohort is measured in years, as is the education level of the primary caregiver. These variables are all time-invariant and collectively denoted by V_i.

My identification strategy in this analysis is, first, to construct sequential measurements of neighborhood context and math achievement throughout the early life course, and second, to then adjust for lagged levels of achievement, in addition to the time-invariant factors outlined previously, in an effort to control for confounding of neighborhood effects on an end-of-study measure of math achievement taken during adolescence. Although not without limitations (e.g., Morgan and Winship 2015), analyses that adjust for lagged levels of the outcome in order to estimate the effects of future levels of the treatment on future levels of the outcome arguably provide some of the stronger grounds for causal inference with data from an observational research design (Pearl 2009; VanderWeele 2015). Because this analysis controls for lagged measures of achievement, it is based on a type of "value-added" identification strategy commonly employed in research on student learning (Chetty et al. 2014; Rowan et al. 2002).

More specifically, the time index, t, in this analysis is used to distinguish between measures taken at "baseline" ($t = 0$), "childhood" ($t = 1$), and "adolescence" ($t = 2$). The baseline measure of achievement, Y_{i0}, is taken at the 1997 wave of the CDS, when sample members are age 3 to 7, and it is used only as a control variable. Several years later, the childhood measure of treatment, A_{i1}, is taken at the 1999 wave of the PSID main panel, when sample members are age 5 to 9. Another several years later, the childhood measure of achievement, Y_{i1}, is taken at the 2002-2003 wave of the CDS, when samples members are age 8 to 12. This measure is also treated only as a control variable. Then, at the 2005 wave of the PSID main panel, the adolescent measure of treatment, A_{i2}, is taken when sample members are age 11 to 15. And finally, the adolescent measure of the outcome, Y_{i2}, is taken at the 2007 wave of the CDS, when sample members are age 13 to 17. Thus, to summarize, these data have the following temporal structure: $\{V_i, Y_{i0}, A_{i1}, Y_{i1}, A_{i2}, Y_{i2}\}$, and the goal is to estimate the average marginal

effects of A_{i1} and A_{i2} on Y_{i2} under the key identification assumption that treatment assignment is sequentially ignorable conditional on the time-invariant controls and lagged measures of achievement.[5]

To this end, I first estimate by OLS—for comparative purposes only—a conventional linear and additive regression model for $E(Y_{i2}|V_i, Y_{i0}, A_{i1}, Y_{i1}, A_{i2})$, the conditional mean of adolescent achievement given prior treatments, lagged measures of achievement, and time-invariant factors. OLS estimation of the treatment coefficients in this model would be unbiased for the marginal effects of interest under the identification assumptions of consistency, positivity, and sequential ignorability; under the modeling assumptions that both the associational effects of the control variables and the causal effects of treatment are linear and additive, which implies that there must not be any treatment effect moderation; and finally, under the additional and unconventional assumption that childhood treatment, A_{i1}, does not affect childhood achievement, Y_{i1}. This last assumption is needed to rule out the possibility of over-control and endogenous selection biases, but because childhood exposure to a disadvantaged neighborhood likely affects childhood achievement, it is almost certainly violated in this analyses, and thus OLS estimation will be biased, as outlined previously.

Second, I estimate by RWR a linear and additive SNMM. This involves fitting the following regression model,

$$E(Y_{i2}|V_i, Y_{i0}, A_{i1}, Y_{i1}, A_{i2}) = \alpha_{00}^* + \big(V_i - E(V_i)\big)\gamma_{01}^* + \big(Y_{i0} - E(Y_{i0})\big)\gamma_{02}^* + \beta_{10}^* A_{i1} + \big(Y_{i1} - E(Y_{i1}|V_i, Y_{i0}, A_{i1})\big)\gamma_{10}^* + \beta_{20}^* A_{i2},$$

[5] Missing values are simulated for all variables using multiple imputation with 100 replications, and estimates are combined across datasets using the methods proposed by Rubin (1987).

to the observed data in two stages. In the first stage, $\{V_i, Y_{i0}\}$ are residualized by centering them around their sample means, and Y_{i1} is residualized by centering it around its estimated conditional mean, which is computed from a linear and additive OLS regression of Y_{i1} on $\{V_i, Y_{i0}, A_{i1}\}$. Then, in the second stage, the SNMM is estimated by fitting a linear and additive OLS regression of Y_{i2} on prior treatments and the residual terms from the first stage. RWR estimation of the treatment coefficients in this model, $\{\beta_{10}^*, \beta_{20}^*\}$, is unbiased for the marginal effects of interest under the identification assumptions outlined previously and under the modeling assumptions that both the causal and nuisance functions of the SNMM are correctly specified. This latter assumption implies that the associational effects of the control variables must be linear and additive; that the models used to estimate the residual terms in the first stage must be linear and additive; and that the effects of treatment must be linear and additive, which implies that they are invariant across levels of all control variables. Although these are strong assumptions, they are in fact weaker than those required of conventional OLS regression, as RWR estimation need not assume that Y_{i1} is unaffected by A_{i1}. This is because childhood achievement, Y_{i1}, is appropriately residualized with respect to the observed past, $\{V_i, Y_{i0}, A_{i1}\}$, before it is included as a control in the model for adolescent achievement, Y_{i2}.

Finally, I estimate by IPTW a linear and additive MSM. This involves fitting the following regression model,

$$E(Y_{i2}|V_i, A_{i1}, A_{i2}) = \alpha_{00}^* + \big(V_i - E(V_i)\big)\gamma_{01}^* + \beta_{10}^* A_{i1} + \beta_{20}^* A_{i2},$$

to the observed data using the method of WLS with weights for each subject equal to

$$sw_i = \frac{f\big(A_{i1}\big|V_i; \mu_1, \sigma_1^2\big)}{f\big(A_{i1}\big|V_i, Y_{i0}; \mu_2, \sigma_2^2\big)} \times \frac{f\big(A_{i2}\big|V_i, A_{i1}; \mu_3, \sigma_3^2\big)}{f\big(A_{i2}\big|V_i, Y_{i0}, A_{i1}, Y_{i1}; \mu_4, \sigma_4^2\big)}.$$

In these expressions, $\left(V_i - E(V_i)\right)$ is estimated as above by centering V_i around its sample mean, and because treatment in this analysis is continuous, $f(\ )$ denotes a probability density function with mean μ and variance σ^2. To model the treatment densities in the denominator of the weight, I use the normal density function, and I estimate the mean and variance parameters using predicted values and residual variances, respectively, from linear and additive OLS regressions of A_{i1} on $\{V_i, Y_{i0}\}$ and A_{i2} on $\{V_i, Y_{i0}, A_{i1}, Y_{i1}\}$. Similar procedures are used to estimate the parameters in the numerator of the weight, except that lagged measures of the outcome are omitted from the OLS regressions for treatment.

IPTW estimation of the treatment coefficients in this model, $\{\beta_{10}^*, \beta_{20}^*\}$, is unbiased for the marginal effects of interest under the identification assumptions outlined previously; under the modeling assumption that both the associational effects of the time-invariant factors and the causal effects of treatment on the outcome are linear and additive, which implies that there must not be any treatment effect moderation across levels of the time-invariant factors; and under the additional assumption that the treatment densities in the denominator of the weight are correctly modeled. This latter assumption implies that treatment at each time point must be normally distributed conditional on the observed past with homoscedastic variance and a conditional mean equal to a linear and additive function of prior treatments and controls. The only difference between IPTW and RWR estimation in this analysis is that the former method uses weights to adjust for lagged measures of achievement, while the latter method adjusts for these measures directly in the outcome model after they have been appropriately residualized. Thus, in contrast to RWR estimation, IPTW estimation does not require that the effects of treatment be invariant across prior levels of achievement.

Table 4 presents point estimates and bootstrap standard errors for the marginal effects of interest.[6] Specifically, it presents estimates for $\lambda_1(1) = E\big(Y_{i2}(1,0) - Y_{i2}(0,0)\big)$, the distal marginal effect of childhood exposure to a "disadvantaged" neighborhood one standard deviation above the national mean of the composite disadvantage index rather than a "middle class" neighborhood at the national mean; for $\lambda_2(a_1, 1) = E\big(Y_{i2}(a_1, 1) - Y_{i2}(a_1, 0)\big)$, the proximal marginal effect of adolescent exposure to a "disadvantaged" neighborhood one standard deviation above the national mean of the composite disadvantage index rather than a "middle class" neighborhood at the national mean, which is assumed to be invariant across childhood levels of contextual disadvantage; and finally, for $\lambda_1(1) + \lambda_2(1,1) = \big(Y_{i2}(1,1) - Y_{i2}(0,0)\big)$, the cumulative marginal effect of sustained exposure throughout childhood and adolescence to a "disadvantaged" neighborhood one standard deviation above the national mean of the composite disadvantage index rather than a "middle class" neighborhood at the national mean.

The first column of Table 4 presents conventional OLS regression estimates. They suggest that the proximal effect of adolescent exposure to a disadvantaged neighborhood on adolescent math achievement is negative and statistically significant, but they also suggest, counterintuitively, that the childhood effect is positive and that the cumulative effect is negligible. Note, however, that OLS estimates of the childhood effect, and consequently, also the cumulative effect, almost surely suffer from bias due to over-control and endogenous selection. This is because these estimates are based on a model that naively conditions on childhood achievement, and any effect of childhood exposure to a disadvantaged neighborhood on

[6] For computational simplicity, I compute standard errors by bootstrapping within multiply imputed datasets, rather than by multiply imputing missing data within each bootstrap sample, even though the former approach is somewhat conservative (Schomaker and Heumann 2016).

adolescent achievement likely operates through its effect on achievement measured earlier during the course of development.

The second and third columns of Table 4 present RWR and IPTW estimates, respectively. They are very similar, except that the RWR estimates are slightly more precise. Both sets of estimates indicate that the distal effect of childhood exposure to a disadvantaged neighborhood on adolescent math achievement is negative, but these estimates are substantively small and not statistically significant. Both sets of estimates also indicate that the adolescent and cumulative effects of exposure to a disadvantaged neighborhood are negative and highly significant. For example, sustained exposure to a disadvantaged neighborhood one standard deviation above the national mean of the composite disadvantage index, rather than sustained exposure to a middle class neighborhood at the national mean, is estimated to reduce adolescent math achievement by about 0.15 standard deviations.

8. DISCUSSION

Estimating the marginal effects of a time-varying treatment from observational data is challenging because conventional regression and matching estimators do not properly adjust for time-varying confounders affected by prior treatments. By contrast, IPTW estimation of an MSM can properly adjust for this type of confounder, but the method is difficult to use with continuous treatments and relatively inefficient. This article proposes an alternative method for estimating marginal effects in the longitudinal setting: RWR estimation of a constrained SNMM in which the effect of treatment at each time point is assumed to be invariant across levels of prior confounders. This method is easier to use with continuous treatments because, unlike IPTW, it does not require a model for the conditional probability of treatment at each time point. In

addition, a series of simulation experiments suggest that RWR estimation is more efficient than IPTW estimation when its motivating assumptions are satisfied. Thus, when estimating marginal effects from observational data in the longitudinal setting, fitting a constrained SNMM by the method of RWR may provide an alternative to IPTW estimation of a MSM that is both easier to implement and relatively more efficient, at least in certain situations.

Despite these advantages, RWR estimation of marginal effects is not without limitations. First, unlike IPTW estimation, RWR estimation requires that the effect of treatment at each time point must not vary across levels of the prior confounders—that is, it requires that there must not be any treatment effect moderation. If the observed confounders do in fact moderate the effect of treatment, then RWR estimation of a constrained SNMM in which the effects of treatment are incorrectly assumed to be invariant would suffer from misspecification bias. This type of bias is essentially the same as that which afflicts linear and additive conventional regression models in the point-in-time setting when there is in fact treatment effect moderation (Morgan and Winship 2015). In this situation, the "main effects" of treatment in a linear and additive SNMM, for example, do not represent average marginal effects because the RWR estimator does not appropriately average the different moderated effects together. Although simulation experiments indicate that RWR estimation of marginal effects in a constrained SNMM is somewhat robust to this type of misspecification bias, IPTW estimation may perform better in many empirical applications given the ubiquity of treatment effect moderation in the social sciences (Morgan and Winship 2015; Xie 2007; Xie et al. 2012).

Second, and also unlike IPTW estimation, RWR estimation of marginal effects in a constrained SNMM requires that its nuisance functions be correctly specified. With a small number of confounders and time periods, correctly specifying the nuisance functions is feasible,

but with a large number of confounders and time periods, this becomes much more difficult to accomplish. For each additional confounder or time period in the analysis, additional first-stage models and residual terms are needed, and the potential complexity of these models increases with the dimension of the data. Similarly, the number of associational effects on the outcome also increases with the number of confounders and time periods, as does the potential complexity of these effects, which may now depend on levels of all the other confounders in the model. Thus, in empirical applications with many confounders and time periods, the assumption of correctly specified nuisance functions may become untenable. Although simulation experiments indicate that RWR estimation of marginal effects is somewhat robust to this type of misspecification bias, IPTW estimation may perform better with high-dimensional data because it requires comparatively fewer modeling assumptions in this context.

Nevertheless, IPTW estimation may suffer from a similar set of limitations in practice. For example, both the number and potential complexity of models for the conditional probability of treatment also increase with the number of time periods and confounders in the analysis. With many confounders and time periods, identifying correct models for the conditional probability of treatment at each time point becomes much more difficult, and thus the modeling assumptions on which IPTW estimation is based may also become untenable in this situation. Furthermore, in most empirical applications (e.g., Sharkey and Elwert 2011; Wodtke et al. 2011; Wodtke 2013), including the analysis of the PSID discussed previously, IPTW estimation is implemented with a version of the stabilized weight that requires adjustment for time-invariant confounders directly in the regression model for the outcome. With this implementation of the method, IPTW estimation also requires an assumption about the absence of effect moderation. Specifically, when regression adjustment rather than weighting is used to control for confounding by time-

invariant factors, the "main effects" of treatment in an outcome model estimated by IPTW only represent average marginal effects if the time-invariant confounders do not moderate the effects of treatment.

A variety of diagnostic tools are available for use with IPTW estimation that may help to mitigate, at least in part, the limitations outlined previously (e.g., Austin and Stuart 2015; Cole and Hernán 2008). For example, hypothesis tests and graphical displays can be used to assess whether the confounders are appropriately balanced across treatment levels, and by extension, whether a given model for the conditional probability of treatment is suitable. Similar methods can be used to evaluate the fit of competing treatment models or to test for the presence of effect moderation by time-invariant confounders, if appropriate. These types of diagnostic tools can also be used to investigate the modeling assumptions on which RWR estimation of marginal effects is based. For example, hypothesis tests can be constructed to evaluate whether there is any effect moderation across levels of the time-varying confounders or to compare competing models with different specifications for the nuisance functions. If these diagnostic tests provide little evidence of effect moderation or other forms of model misspecification, interpreting the treatment coefficients in a constrained SNMM estimated by RWR as average marginal effects would be more defensible.

It is also possible to combine IPTW and RWR estimation in different ways (Almirall et al. 2014; Author Forthcoming). This hybrid approach may mitigate some of the limitations associated with RWR in isolation while also preserving some of the efficiency gains associated with regression-based adjustment for observed confounding. For example, with a hybrid approach, weighting could be used to adjust for a subset of the time-varying confounders (e.g., those that appear to moderate the effects of treatment), while the remaining confounders (e.g.,

those that do not appear to moderate the effects of treatment) are adjusted for directly in an SNMM fit to an appropriately weighted pseudo-sample by the method of RWR. In addition, an SNMM can also be estimated using the G-estimator (Robins 1994; Vansteelandt and Joffe 2014), which requires a weaker set of modeling assumptions compared to RWR. Provided that the effects of treatment do not vary across levels of the confounders, G-estimation of marginal effects in a constrained SNMM would be unbiased if either the nuisance functions are correctly specified or models for the conditional probability of treatment are correctly specified. This double-robustness property of the G-estimator provides a degree of protection against misspecification bias, but it comes at the cost of higher variance relative to the RWR estimator (Almirall et al. 2010) and it is difficult to implement with existing software (Vansteelandt and Joffe 2014).

In addition, it is possible to recover estimates of average marginal effects from RWR estimates of moderated effects in an unconstrained SNMM by appropriately averaging these effects together. This type of post-processing procedure would involve iteratively averaging the moderated causal effects in an unconstrained SNMM over the conditional distribution of each confounder given prior confounders and treatments. Such a procedure would provide another method for accommodating effect moderation in longitudinal analyses of average marginal effects using RWR and an SNMM, although it remains unclear whether this approach would still outperform IPTW estimation in terms of relative efficiency.

Because RWR estimation of an SNMM involves only a slight modification of conventional least squares regression, this method is based on familiar computations that can be easily carried out with off-the-shelf statistical software (see the Online Supplement for details). In brief, it is implemented by first regressing the confounders at each time point on prior treatments and

covariates and then computing residuals, and second, by regressing the outcome on these residuals and prior treatments. Valid standard errors can be computed via bootstrapping (Almirall et al. 2014; Efron and Tibshirani 1993).

In sum, RWR estimation of average marginal effects in a constrained SNMM provides an alternative to IPTW estimation of an MSM that is flexible, efficient, and easy to implement. This approach, however, is premised on a number of strong assumptions, many of which are similar to those associated with conventional least squares regression in the point-in-time setting. Because these assumptions may be difficult to satisfy in certain situations—for example, in applications with a large number of time periods and confounders—researchers may want to employ both methods along with formal diagnostic procedures in an attempt to determine which performs best in practice. Given the difficult challenges associated with analyses of time-varying treatments in the longitudinal setting, it is important for researchers to have at their disposal a variety of different statistical tools capable of estimating marginal effects, each with their own advantages and disadvantages in different situations.

REFERENCES

Almirall, Daniel, Daniel F. McCaffrey, Rajeev Ramchand, and Susan A. Murphy. 2013. "Subgroups Analysis when Treatment and Moderators are Time-varying." Prevention Science 14:169-178.

Almirall, Daniel, Thomas Ten Have, and Susan A. Murphy. 2010. "Structural Nested Mean Models for Assessing Time-Varying Effect Moderation." Biometrics 66:131-139.

Austin, Peter C. and Elizabeth A. Stuart. 2015. "Moving towards Best Practice when using

Inverse Probability of Treatment Weighting (IPTW) to Estimate Causal Treatment Effects in Observational Studies." Statistics in Medicine 34:3661-3679.

Bronfenbrenner, Urie and Pamela A. Morris. 1998. "The Ecology of Developmental Processes." Pp. 993-1028 in Theoretical Models of Human Development (5th Edition), edited by William Damon and Richard Lerner. Hoboken, NJ: John Wiley & Sons.

Chetty, Raj, John N. Friedman and Jonah E. Rockoff. 2014. "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates." American Economic Review 104:2593-632.

Cole, Stephen R. and Miguel A. Hernán. 2008. "Constructing Inverse Probability of Treatment Weights for Marginal Structural Models." American Journal of Epidemiology 168:656-664.

Cunha, Flavio and James J. Heckman. 2007. "The technology of skill formation." American Economic Review 97:31-47.

Efron, Bradley and Robert J. Tibshirani. 1993. An Introduction to the Bootstrap. New York: Chapman and Hall.

Elder, Glen H. 1998. "The Life Course as Developmental Theory." Child Development 69:1-12.

Elwert, Felix. 2013. "Graphical Causal Models." Pp. 245-271 in Handbook of Causal Analysis for Social Research, edited by S. L. Morgan. New York: Springer.

Elwert, Felix and Christopher Winship. 2014. "Endogenous Selection Bias: The Problem of Conditioning on a Collider Variable." Annual Review of Sociology 40:31-53.

GeoLytics, Inc. 2013. Neighborhood Change Database, 1970-2010 Tract Data. New Brunswick, NJ: GeoLytics.

Halpern-Manners, Andrew, John R. Warren, and Jennie E. Brand. 2009. "Dynamic measures of

primary and secondary school characteristics: Implications for school effects research." Social Science Research 38:397-411.

Heard, Holly E. 2007. "Fathers, Mothers, and Family Structure: Family Trajectories, Parent Gender, and Adolescent Schooling." Journal of Marriage and Family 69:435-450.

Holland, Paul W. 1986. "Statistics and Causal Inference." Journal of the American Statistical Association 81:945-960.

Krein, Sheila F. and Andrea H. Beller. 1988. "Educational Attainment of Children from Single-parent Families: Differences by Exposure, Gender, and Race." Demography 25:221-234.

Lunceford, Jared K. and Marie Davidian. 2004. "Stratification and Weighting via the Propensity Score in Estimation of Causal Treatment Effects: A Comparative Study." Statistics in Medicine 23:2937-2960.

McCaffrey, Daniel F., Greg Ridgeway, and Andrew R. Morral. 2004. "Propensity Score Estimation with Boosted Regression for Evaluating Causal Effects in Observational Studies." Psychological Methods 9:403-425.

Michigan Survey Research Center. 2014. "Panel Study of Income Dynamics Public- and Restricted-Use Datasets." Ann Arbor, MI: Institutute for Social Research.

Morgan, Stephen L. and Christopher Winship. 2015. Counterfactuals and Causal Inference: Methods and Principles for Social Research (2nd Ed.). New York: Cambridge University Press.

Naimi, Ashley I., Erica E. M. Moodie, Nathalie Auger, and Jay S. Kaufman. 2014. "Constructing Inverse Probability Weights for Continuous Exposures: A Comparison of Methods." Epidemiology 25:292-299.

Pearl, Judea. 2009. Causality (2nd Edition). New York: Cambridge University Press.

Robins, James M. 1994. “Correcting for Noncompliance in Randomized Trials Using Structural Nested Mean Models.” Communications in Statistics-Theory and Methods 23:2379-2412.

Robins, James M. 1999. “Association, Causation, and Marginal Structural Models.” Synthese 121:151-179.

Robins, James M., Miguel A. Hernan, and Andrea Rotnitzky. 2007. “Effect Modification by Time-Varying Covariates.” American Journal of Epidemiology 166:994-1002.

Robins, James M., Miguel A. Hernan, and Babette Brumback. 2000. “Marginal Structural Models and Causal Inference in Epidemiology.” Epidemiology 11:550-560.

Robins, James M., Andrea Rotnitzky, and Lue Ping Zhao. 1994. “Estimation of Regression Coefficients when Some Regressors are not Always Observed.” Journal of the American Statistical Association 89:846-866.

Rowan Brian, Richard Correnti, and Robert J. Miller. 2002. “What large-scale survey research tells us about teacher effects on student achievement.” Teachers College Record 104:1525-1567.

Royston, Patrick, Douglas G. Altman, and Willi Sauerbrei. 2006. “Dichotomizing Continuous Predictors in Multiple Regression: A Bad Idea.” Statistics in Medicine 25:127-141.

Rubin, Donald B. 1974. “Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies.” Journal of Educational Psychology 66:688-701.

Sampson, Robert J., Patrick Sharkey, and Stephen W. Raudenbush. 2008. “Durable Effects of Concentrated Disadvantage on Verbal Ability among African-American Children.” Proceedings of the National Academy of Sciences 105:845-852.

Schomaker, Michael and Christian and Heumann. 2016. “Bootstrap Inference when using Multiple Imputation.” arXiv Working Paper 1602.07933v1.

Sharkey, Patrick and Felix Elwert. 2011. "The Legacy of Disadvantage: Multigenerational Neighborhood Effects on Cognitive Ability." American Journal of Sociology 116:1934-1981.

Sharkey, Patrick and Jacob W. Faber. 2014. "Where, When, Why, and For Whom Do Residential Contexts Matter? Moving Away from the Dichotomous Understanding of Neighborhood Effects." Annual Review of Sociology 40:559-579.

VanderWeele, Tyler. 2015. Explanation in Causal Inference: Methods for Mediation and Interaction. New York: Oxford University Press.

Vansteelandt, Stijn and Marshall Joffe. 2014. "Structural Nested Models and G-estimation: The Partially Realized Promise." Statistical Science 29:707-731.

Winship, Christopher and Felix Elwert. 2010. "Effect Heterogeneity and Bias in Main-Effects-Only Regression Models." Pp. 327-336 in Heuristics, Probability, and Causality: A Tribute to Judea Pearl edited by Rina Dechter, Hector Geffner, and Joseph Y. Halpern. London: College Publications.

Wodtke, Geoffrey T. 2013. "Duration and Timing of Exposure to Neighborhood Poverty and the Risk of Adolescent Parenthood." Demography 50:1765-1788.

Wodtke, Geoffrey T., David J. Harding, and Felix Elwert. 2011. "Neighborhood Effects in Temporal Perspective: The Impact of Long-term Exposure to Concentrated Disadvantage on High School Graduation." American Sociological Review 76:713–736.

Wodtke, Geoffrey T., Felix Elwert, and David J. Harding. 2016. "Neighborhood effect heterogeneity by family income and developmental period." American Journal of Sociology 121:1168–1222.

Woodcock, Richard W and M. Bonner Johnson. 1989. Tests of Achievement, Standard Battery

(Form B). Chicago, IL: Riverside.

Xie, Yu. 2007. "Otis Dudley Duncan's Legacy: The Demographic Approach to Quantitative Reasoning in Social Science." Research in Social Stratification and Mobility 25:141-156.

Xie, Yu, Jennie E. Brand, and B. Jann. 2012. "Estimating Heterogeneous Treatment Effects with Observational Data." Sociological Methodology 42:314-347.

TABLES

Table 1. Performance of the RWR estimator with correctly specified causal and nuisance functions and of the IPTW estimator with correctly specified models for the conditional probability of treatment

Simulation/parameter	IPTW estimator			RWR estimator			
	Bias	SD	RMSE	Bias	SD	RMSE	REL RMSE
Small sample (n=500)							
$\lambda_1(1)$	0.003	0.185	0.185	0.002	0.143	0.143	0.772
$\lambda_2(0,1)$	0.000	0.169	0.169	-0.001	0.130	0.130	0.771
$\lambda_2(1,1)$	0.003	0.172	0.172	0.000	0.134	0.134	0.778
Medium sample (n=1000)							
$\lambda_1(1)$	0.002	0.131	0.131	-0.001	0.100	0.100	0.768
$\lambda_2(0,1)$	0.001	0.118	0.118	0.000	0.091	0.091	0.775
$\lambda_2(1,1)$	-0.002	0.121	0.121	0.000	0.094	0.094	0.777
Large sample (n=2000)							
$\lambda_1(1)$	0.001	0.093	0.093	0.000	0.071	0.071	0.768
$\lambda_2(0,1)$	0.000	0.084	0.084	0.000	0.065	0.065	0.769
$\lambda_2(1,1)$	-0.001	0.085	0.085	-0.001	0.067	0.067	0.780

Notes: RMSE denotes the root mean squared error, and REL RMSE denotes the root mean squared error of the RWR estimator relative to the IPTW estimator. Results are based on 10,000 simulations. See the Online Supplement for details.

Table 2. Large sample (n=2000) performance of the RWR estimator with misspecified causal functions and of the IPTW estimator with correctly specified models for the conditional probability of treatment

Simulation/parameter	IPTW estimator			RWR estimator			
	Bias	SD	RMSE	Bias	SD	RMSE	REL RMSE
Minor mispecification							
$\lambda_1(1)$	0.001	0.092	0.092	-0.020	0.070	0.073	0.796
$\lambda_2(0,1)$	0.000	0.086	0.086	-0.021	0.065	0.068	0.796
$\lambda_2(1,1)$	0.000	0.088	0.088	0.017	0.067	0.069	0.780
Moderate mispecification							
$\lambda_1(1)$	0.001	0.093	0.093	-0.042	0.071	0.083	0.896
$\lambda_2(0,1)$	0.000	0.085	0.085	-0.041	0.065	0.077	0.908
$\lambda_2(1,1)$	0.001	0.090	0.090	0.036	0.067	0.076	0.843
Severe mispecification							
$\lambda_1(1)$	0.001	0.095	0.095	-0.086	0.072	0.112	1.181
$\lambda_2(0,1)$	0.001	0.087	0.087	-0.082	0.066	0.105	1.202
$\lambda_2(1,1)$	0.000	0.098	0.098	0.073	0.068	0.100	1.017

Notes: RMSE denotes the root mean squared error, and REL RMSE denotes the root mean squared error of the RWR estimator relative to the IPTW estimator. Results are based on 10,000 simulations. See the Online Supplement for details.

Table 3. Large sample (n=2000) performance of the RWR estimator with mispecified nuisance functions and of the IPTW estimator with correctly specified models for the conditional probability of treatment

Simulation/parameter	IPTW estimator			RWR estimator			
	Bias	SD	RMSE	Bias	SD	RMSE	REL RMSE
Minor mispecification							
$\lambda_1(1)$	-0.001	0.095	0.095	-0.028	0.072	0.077	0.812
$\lambda_2(0,1)$	0.000	0.086	0.086	-0.026	0.065	0.070	0.810
$\lambda_2(1,1)$	0.001	0.086	0.086	0.024	0.067	0.071	0.827
Moderate mispecification							
$\lambda_1(1)$	0.001	0.097	0.097	-0.056	0.072	0.091	0.932
$\lambda_2(0,1)$	0.002	0.087	0.087	-0.051	0.066	0.083	0.951
$\lambda_2(1,1)$	0.001	0.087	0.087	0.050	0.067	0.084	0.965
Severe mispecification							
$\lambda_1(1)$	0.000	0.103	0.103	-0.111	0.074	0.133	1.288
$\lambda_2(0,1)$	0.002	0.090	0.090	-0.102	0.067	0.122	1.353
$\lambda_2(1,1)$	0.002	0.090	0.090	0.100	0.069	0.121	1.344

Notes: RMSE denotes the root mean squared error, and REL RMSE denotes the root mean squared error of the RWR estimator relative to the IPTW estimator. Results are based on 10,000 simulations. See the Online Supplement for details.

Table 4. Estimated marginal effects of exposure to disadvantaged neighborhoods on end-of-study math achievement

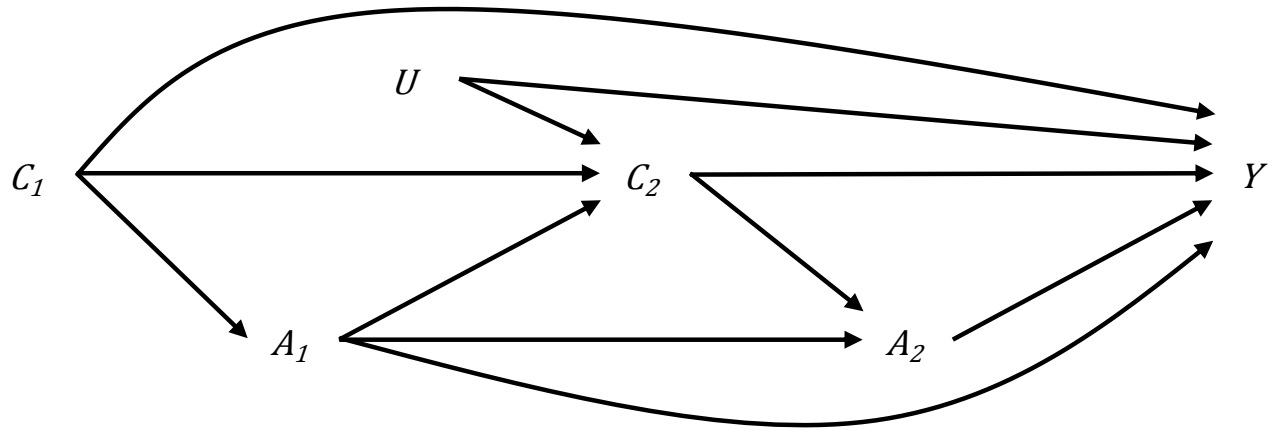
Model/parameters	OLS			RWR			IPTW		
	Est.	SE		Est.	SE		Est.	SE	
Childhood effect									
$\lambda_1(1)$	0.066	(0.041)		-0.044	(0.044)		-0.044	(0.045)	
Adolescence effect									
$\lambda_2(a_1,1)$	-0.102	(0.040)	*	-0.102	(0.039)	**	-0.111	(0.040)	**
Cumulative effect									
$\lambda_1(1)+\lambda_2(1,1)$	-0.035	(0.033)		-0.146	(0.037)	***	-0.154	(0.038)	***

Notes: Sample includes respondents who were interviewed at the 1997 wave of the CDS between age 3 and 7. Results are combined estimates from 100 imputations. Estimates are adjusted for lagged measures of the outcome and for age, race, gender, and parental education at baseline. The outcome is standardized to have zero mean and unit variance.

†$p < 0.10$, *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$ for two-sided tests of no effect.

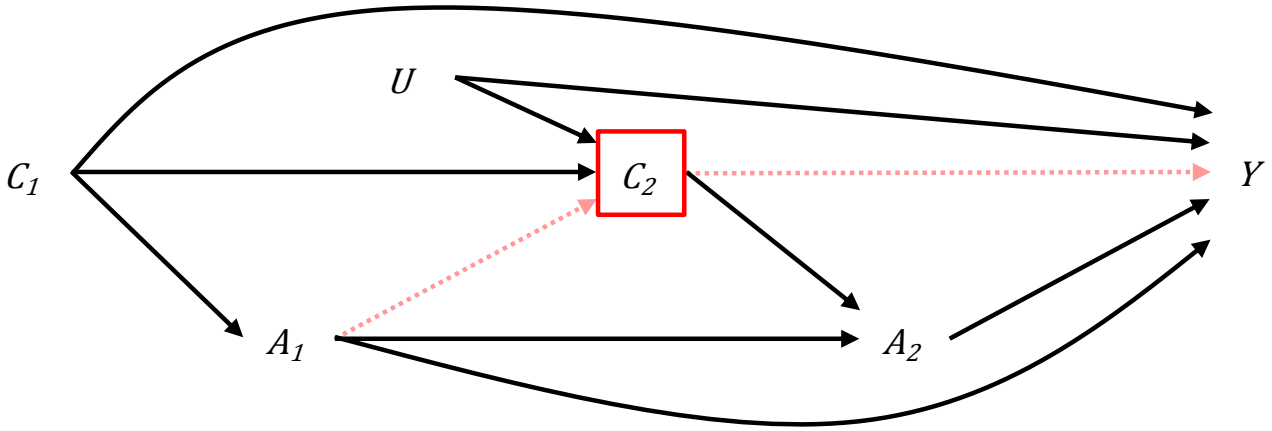
FIGURES

Figure 1. Assumed causal relationships between a time-varying treatment, a time-varying confounder, and the outcome



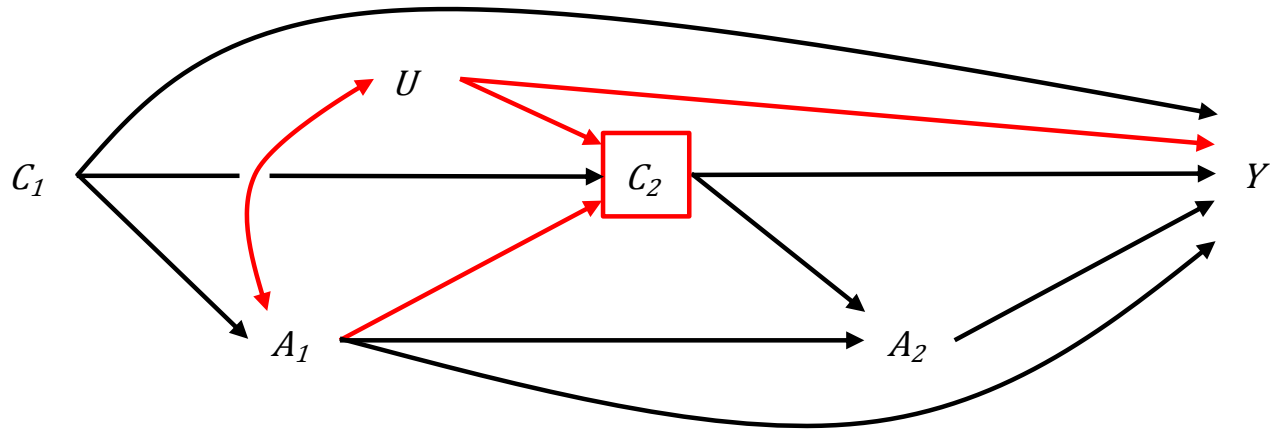
Notes: C_t is an observed time-varying confounder; A_t is a time-varying treatment; Y is the end-of-study outcome; and U is an unobserved covariate.

Figure 2. Graphical depiction of over-control bias induced by conditioning on a time-varying confounder affected by past treatment



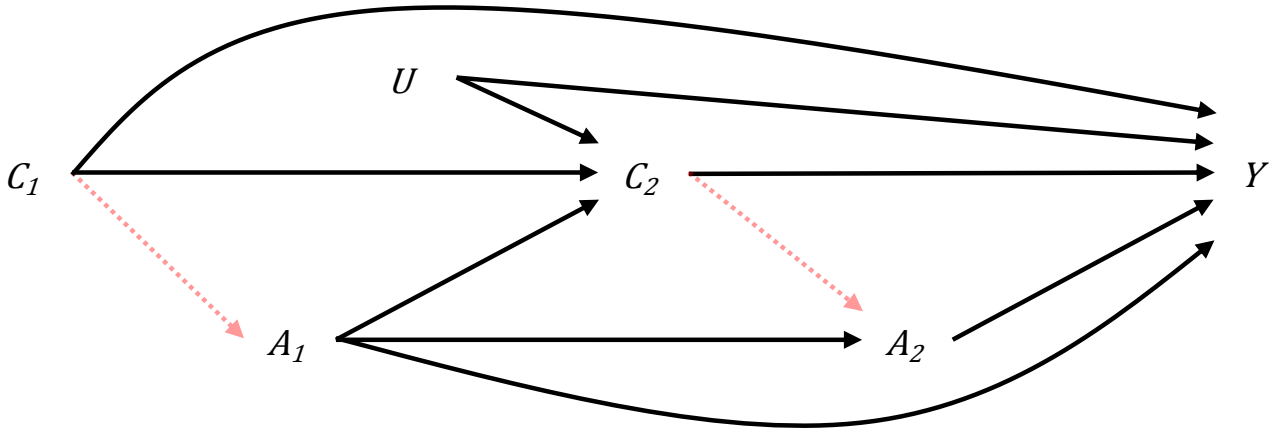
Notes: C_t is an observed time-varying confounder; A_t is a time-varying treatment; Y is the end-of-study outcome; and U is an unobserved covariate.

Figure 3. Graphical depiction of collider-stratification bias induced by conditioning on a time-varying confounder affected by past treatment



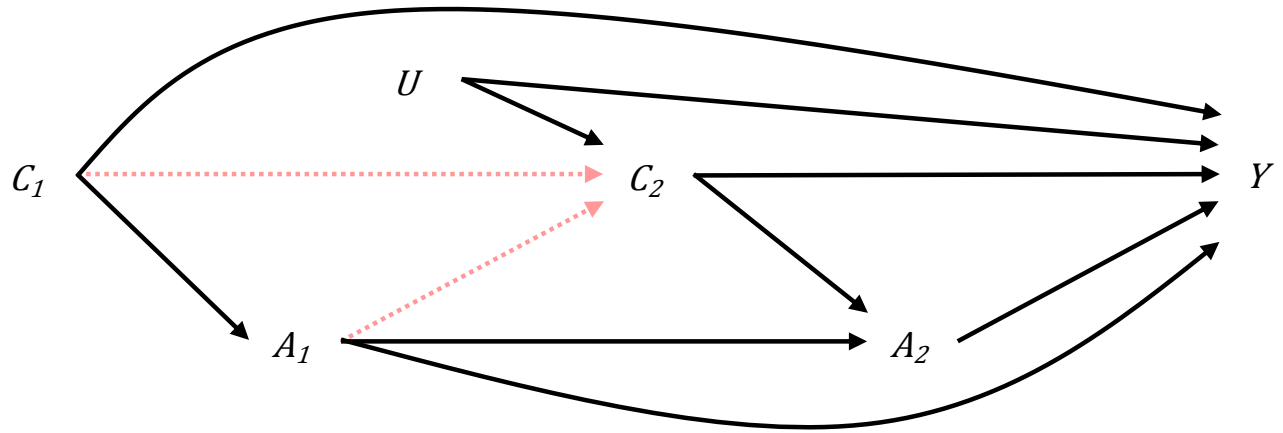
Notes: C_t is an observed time-varying confounder; A_t is a time-varying treatment; Y is the end-of-study outcome; and U is an unobserved covariate.

Figure 4. Graphical depiction of the consequences of weighting by the inverse probability of treatment



Notes: C_t is an observed time-varying confounder; A_t is a time-varying treatment; Y is the end-of-study outcome; and U is an unobserved covariate.

Figure 5. Graphical depiction of the consequences of residualizing the time-varying confounders based on prior treatment and covariates



Notes: C_t is an observed time-varying confounder; A_t is a time-varying treatment; Y is the end-of-study outcome; and U is an unobserved covariate.

ONLINE SUPPLEMENT

Stata Code for Simulation Experiments

```
#delimit ;
capture clear all ;
set more off ;

/************************
DEFINE SIMULATION PROGRAM
************************/
capture program drop sim ;
program define sim, rclass ;
syntax [, obs(integer 1) gamma21(real 0.00) gamma22(real 0.00)
theta11(real 0.00) theta21(real 0.00)] ;
drop _all ;
/*SIMULATE DATA*/
set obs `obs' ;
gen c1 = invnorm(uniform()) ;
gen a1 = 0.4*c1 + invnorm(uniform()) >= 0 ;
gen c2 = 0.4*c1 + 0.2*a1 + invnorm(uniform()) ;
gen a2 = 0.2*a1 + 0.4*c2 + invnorm(uniform()) >= 0 ;
gen y = 0.4*(c1 - 0) + a1*(0.1 + `theta11'*c1) + (c2 - (0.2*c1 +
0.2*a1))*(0.4 + `gamma21'*c1 + `gamma22'*a1) + a2*(0.2 + `theta21'*c1
+ 0.1*a1) + invnorm(uniform()) ;
/*RWR ESTIMATION*/
reg c1 ;
predict c1r, resid ;
reg c2 a1 c1 ;
predict c2r, resid ;
gen a1a2 = a1*a2 ;
reg y c1r a1 c2r a2 a1a2 ;
return scalar lamda1_rwr = _b[a1] - 0.1 ;
return scalar lamda20_rwr = _b[a2] - 0.2 ;
return scalar lamda21_rwr = _b[a2] + _b[a1a2] - 0.3 ;
/*IPTW ESTIMATION*/
probit a1 ;
predict p1 ;
probit a1 c1 ;
predict p2 ;
gen iptw1 = ((p1*a1)+(1-p1)*(1-a1))/((p2*a1)+(1-p2)*(1-a1)) ;
probit a2 a1 ;
predict p3 ;
probit a2 a1 c2 ;
predict p4 ;
```

```
gen iptw2 = ((p3*a2)+(1-p3)*(1-a2))/((p4*a2)+(1-p4)*(1-a2)) ;
gen fnlwt = iptw1*iptw2 ;
reg y a1 a2 a1a2 [pw=fnlwt] ;
return scalar lamda1_iptw = _b[a1] - 0.1 ;
return scalar lamda20_iptw = _b[a2] - 0.2 ;
return scalar lamda21_iptw = _b[a2] + _b[a1a2] - 0.3 ;
end ;

/*********************************
TABLE 1: ALL ASSUMPTIONS SATISFIED
*********************************/
/***SMALL SAMPLE***/
simulate lamda1_rwr=r(lamda1_rwr) lamda20_rwr=r(lamda20_rwr)
lamda21_rwr=r(lamda21_rwr) lamda1_iptw=r(lamda1_iptw)
lamda20_iptw=r(lamda20_iptw) lamda21_iptw=r(lamda21_iptw), reps(10000)
seed(1): sim, obs(500) ;
sum * ;
/***MEDIUM SAMPLE***/
simulate lamda1_rwr=r(lamda1_rwr) lamda20_rwr=r(lamda20_rwr)
lamda21_rwr=r(lamda21_rwr) lamda1_iptw=r(lamda1_iptw)
lamda20_iptw=r(lamda20_iptw) lamda21_iptw=r(lamda21_iptw), reps(10000)
seed(2): sim, obs(1000) ;
sum * ;
/***LARGE SAMPLE***/
simulate lamda1_rwr=r(lamda1_rwr) lamda20_rwr=r(lamda20_rwr)
lamda21_rwr=r(lamda21_rwr) lamda1_iptw=r(lamda1_iptw)
lamda20_iptw=r(lamda20_iptw) lamda21_iptw=r(lamda21_iptw), reps(10000)
seed(3): sim, obs(2000) ;
sum * ;

/***********************************
TABLE 2: SNMM CAUSAL FNS MISSPECIFIED
***********************************/
/***MINOR MISSPECIFICATION***/
simulate lamda1_rwr=r(lamda1_rwr) lamda20_rwr=r(lamda20_rwr)
lamda21_rwr=r(lamda21_rwr) lamda1_iptw=r(lamda1_iptw)
lamda20_iptw=r(lamda20_iptw) lamda21_iptw=r(lamda21_iptw), reps(10000)
seed(4): sim, obs(2000) theta11(0.05) theta21(0.05) ;
sum * ;
/***MODERATE MISSPECIFICATION***/
simulate lamda1_rwr=r(lamda1_rwr) lamda20_rwr=r(lamda20_rwr)
lamda21_rwr=r(lamda21_rwr) lamda1_iptw=r(lamda1_iptw)
lamda20_iptw=r(lamda20_iptw) lamda21_iptw=r(lamda21_iptw), reps(10000)
seed(5): sim, obs(2000) theta11(0.10) theta21(0.10) ;
sum * ;
```

```
/***SEVERE MISSPECIFICATION***/
simulate lamda1_rwr=r(lamda1_rwr) lamda20_rwr=r(lamda20_rwr)
lamda21_rwr=r(lamda21_rwr) lamda1_iptw=r(lamda1_iptw)
lamda20_iptw=r(lamda20_iptw) lamda21_iptw=r(lamda21_iptw), reps(10000)
seed(6): sim, obs(2000) theta11(0.20) theta21(0.20) ;
sum * ;

/************************************
TABLE 3: SNMM NUISANCE FNS MISSPECIFIED
************************************/
/***MINOR MISSPECIFICATION***/
simulate lamda1_rwr=r(lamda1_rwr) lamda20_rwr=r(lamda20_rwr)
lamda21_rwr=r(lamda21_rwr) lamda1_iptw=r(lamda1_iptw)
lamda20_iptw=r(lamda20_iptw) lamda21_iptw=r(lamda21_iptw), reps(10000)
seed(7): sim, obs(2000) gamma21(0.05) gamma22(0.05) ;
sum * ;
/***MODERATE MISSPECIFICATION***/
simulate lamda1_rwr=r(lamda1_rwr) lamda20_rwr=r(lamda20_rwr)
lamda21_rwr=r(lamda21_rwr) lamda1_iptw=r(lamda1_iptw)
lamda20_iptw=r(lamda20_iptw) lamda21_iptw=r(lamda21_iptw), reps(10000)
seed(8): sim, obs(2000) gamma21(0.10) gamma22(0.10) ;
sum * ;
/***SEVERE MISSPECIFICATION***/
simulate lamda1_rwr=r(lamda1_rwr) lamda20_rwr=r(lamda20_rwr)
lamda21_rwr=r(lamda21_rwr) lamda1_iptw=r(lamda1_iptw)
lamda20_iptw=r(lamda20_iptw) lamda21_iptw=r(lamda21_iptw), reps(10000)
seed(9): sim, obs(2000) gamma21(0.20) gamma22(0.20) ;
sum * ;
```