The International Journal of Biostatistics

Volume 8, *Issue* 1 2012 *Article* 14

Targeted Maximum Likelihood Estimation for Dynamic Treatment Regimes in Sequentially Randomized Controlled Trials

Paul H. Chaffee, *University of California, Berkeley*
Mark J. van der Laan, *University of California, Berkeley*

Recommended Citation:
Chaffee, Paul H. and van der Laan, Mark J. (2012) "Targeted Maximum Likelihood Estimation for Dynamic Treatment Regimes in Sequentially Randomized Controlled Trials," *The International Journal of Biostatistics*: Vol. 8: Iss. 1, Article 14.
DOI: 10.1515/1557-4679.1406

Targeted Maximum Likelihood Estimation for Dynamic Treatment Regimes in Sequentially Randomized Controlled Trials

Paul H. Chaffee and Mark J. van der Laan

Abstract

Sequential Randomized Controlled Trials (SRCTs) are rapidly becoming essential tools in the search for optimized treatment regimes in ongoing treatment settings. Analyzing data for multiple time-point treatments with a view toward optimal treatment regimes is of interest in many types of afflictions: HIV infection, Attention Deficit Hyperactivity Disorder in children, leukemia, prostate cancer, renal failure, and many others. Methods for analyzing data from SRCTs exist but they are either inefficient or suffer from the drawbacks of estimating equation methodology. We describe an estimation procedure, targeted maximum likelihood estimation (TMLE), which has been fully developed and implemented in point treatment settings, including time to event outcomes, binary outcomes and continuous outcomes. Here we develop and implement TMLE in the SRCT setting. As in the former settings, the TMLE procedure is targeted toward a pre-specified parameter of the distribution of the observed data, and thereby achieves important bias reduction in estimation of that parameter. As with the so-called Augmented Inverse Probability of Censoring Weight (A-IPCW) estimator, TMLE is double-robust and locally efficient. We report simulation results corresponding to two data-generating distributions from a longitudinal data structure.

KEYWORDS: semi-parametric efficient estimation, targeted maximum likelihood estimation, estimation methods, sequential randomized controlled trials, dynamic treatment regimes

Author Notes: This work was supported by NIH grant 5R01AI74345-5

1 Introduction

1.1 Background

The treatment of many types of afflictions involves ongoing therapy—that is, application of therapy at more than one point in time. Therapy in this context often involves treatment of patients with drugs, but need not be limited to drugs. For example, the use of pill organization devices ("pillboxes") has been studied as a means to improve drug adherence (Petersen et al., 2007), and others (Moodie et al., 2009) have studied the optimum time at which infants should stop breastfeeding.

A common setting for ongoing treatment therapy involves randomization to initial treatment (or randomization to initial treatment within subgroups of the population of interest), followed by later treatments which may also be randomized, or randomized to a certain subset of possible treatments given that certain intermediate outcomes were observed after the initial treatment. Examples from the literature include treatment by antipsychotic medications for reduction in severity of schizophrenia symptoms (Tunis et al., 2006), treatment of prostate cancer by a sequence of drugs determined by success or failure of first-line treatment (Bembom and van der Laan, 2007), when HIV patients should switch treatments (Orellana et al. 2010, van der Laan and Petersen 2007) and many others.

Suppose, for example, that every subject in a prostate cancer study is randomized to an initial pair of treatments (A or B, say), and if a subject's tumor size increases or does not decrease, the subject is again randomized to A or B at the second treatment point. On the other hand, if the subject does well on the first treatment (tumor size decreases, say), then he or she is assigned the same treatment at the second time point as the first. The general term for multiple time point treatments in which treatments after the first-line are assigned in response to intermediate outcomes is *dynamic treatment regimes* or *dynamic treatment rules* (Murphy et al., 2001). If the intermediate outcome in such SRCTs is affected by initial treatment, and in turn affects decisions at the second time-point treatment as well as the final outcome, then it is a so-called "time-dependent confounder."

1.2 Existing Procedures

A number of methods have been proposed to estimate parameters associated with such a study. This article describes implementation of targeted maximum likelihood estimation for two time-point longitudinal data structures, and is based on the framework developed for general longitudinal data structures presented in van der Laan (2010a,b).

Lunceford et al. (2002) develop inverse probability of treatment weighted (IPTW) estimators and an estimating equation estimator suitable for analysis of survival times from a leukemia clinical trial. Wahed and Tsiatis (2004) propose an estimating equation-based estimator which uses the efficient influence curve for estimating treatment policy-specific parameters in two-stage clinical trials. They later extended those methods to account for right-censoring in such trials (Wahed and Tsiatis, 2006). Guo and Tsiatis (2005) develop what they call a "Weighted Risk Set Estimator" for use in two-stage trials where the outcome is a time-to-event (such as death). Tunis et al. (2006) use IPTW methods, Marginal Structural Models and the so-called "g-estimation" method for analyzing the causal effect of a "continuous" treatment regime of atypical antipsychotic medications on severity of schizophrenia symptoms. This study/analysis involved no time-dependent confounders, however. Laber et al. (2009) use Q-learning to estimate optimal dynamic treatment regimes in Attention Deficit Hyperactivity Disorder in children. Miyahara and Wahed (2010) used weighted Kaplan-Meier estimators for estimating treatment-specific survival rates. Orellana et al. (2010) use structural marginal mean models, IPTW and the so-called augmented inverse probability of censoring weighted (A-IPCW) estimators with a view toward estimating optimal treatment regimes for switching to HAART therapy among HIV-positive patients. Bembom and van der Laan (2007) apply simple g-computation and IPTW estimation procedures in analyzing the optimum response of prostate cancer patients to randomized first-line treatment followed by second-line treatment which was either 1) the same as the first line treatment if that had been deemed successful, or 2) randomized to three remaining treatments if the first line had failed. The data used for the latter analysis has recently been re-analyzed using stabilized IPTW estimation by Wang et al. (2012). The latter article was the subject of discussion articles, among them a general presentation of the methods described here (Chaffee and van der Laan, 2012). This type of trial and data closely resembles what we simulate and analyze in the present study, though we add baseline covariates and more than two levels of success in the intermediate biomarker covariate in order to generalize the data structure to more types of scenarios.

We present a new estimator for this longitudinal data structure: the targeted maximum likelihood estimator (TMLE). TMLE has application in a wide range of data structures and sampling designs (van der Laan and Rose, 2011). Though this estimator can be applied to a broad range of data structures of longitudinal type, we focus here on the estimation of treatment-rule-specific mean outcomes. This also covers static treatment regimes for the given data structures.

In the next section we describe the data structure and define the likelihood for the scenarios we intend to analyze. Once we have specified a counterfactual target parameter of interest and equated it with a well-defined mapping from condi-

tional distributions of the data to a real number, we describe TMLE in broad outline, and in particular, the implementation of two different estimators grounded in the general TMLE approach. Specifically we present the so-called efficient influence curve for certain parameters of interest and show the relationship between elements of this object and elements of the targeted maximum likelihood estimators. Following these general descriptions we present simulation results, including details of specific treatment rules, data generation and results in terms of bias, variance and relative mean squared error. A short discussion of the results follows.

2 Data Structure and Likelihood

In the settings of interest here, a randomly sampled subject has data structure $O = (L(0), A(0), L(1), A(1), Y = L(2)) \sim P_0$, where $L(0)$ indicates a vector of baseline covariates, $A(0)$ is initial randomized treatment, $L(1)$ is, say, an intermediate biomarker (which we first consider as binary), $A(1)$ is the second time point treatment (which we also take as binary), $Y = L(2)$ is the clinical outcome of interest and P_0 is the joint distribution of O. We take the data to be *n i.i.d.* copies of O. We also assume $A(1)$ can be set in response to $L(1)$. The patient's full treatment is therefore $(A(0), A(1))$, and specific realizations of $(A(0), A(1))$ may or may not constitute realizations of a specific dynamic treatment rule. Such "rules" are dynamic in the sense that the regimen can be assigned according to a patient's response to treatment over time. However, even if $A(0)$ and $A(1)$ are both unconditionally randomized, parameters of the distribution of the above data can nevertheless be identified which correspond with dynamic treatment regimens.

The likelihood of the data described above can be factorized as

$$p(O) = \prod_{j=0}^{2} P[L(j) \mid \bar{L}(j-1), \bar{A}(j-1)] \prod_{j=0}^{1} P[A(j) \mid \bar{L}(j), \bar{A}(j-1)], \qquad (1)$$

where $\bar{A}(j) = (A(0), A(1), ..., A(j))$ and $\bar{L}(j)$ is similarly defined. Factorizing the likelihood in this way is suggested by the time–ordering of the variables in O.

For simplicity, we introduce the notation $Q_{L(j)},\ j = 0, 1, 2$ to denote the factors of (1) under the first product and $g_{A(j)},\ j = 0, 1$ for those under the second; the latter we refer to as the *treatment and/or censoring mechanism*. Thus in the simpler notation we have

$$p(O) = \prod_{j=0}^{2} Q_{L(j)} \prod_{j=0}^{1} g_{A(j)} = Qg.$$

Analogously to point treatment contexts, we define a treatment-specific mean for the multiple time-point data structure where here a particular treatment means a specific treatment course over time. Instead of a static treatment regime, we define a *treatment rule*, $d = (d_0, d_1)$ for the treatment points $(A(0), A(1))$, which is the set of mappings $d_0 : \mathscr{D}_0 \to \mathscr{A}_0, d_1 : \mathscr{D}_1 \to \mathscr{A}_1$, where $\mathscr{A}_j, j = 0, 1$ is the set of possible values for $A(j)$, $\mathscr{D}_0$ is the support of $L(0)$ and $\mathscr{D}_1$ is the support of $(L(0), A(0), L(1))$. We can express the overall rule as $d(\bar{L}(1)) = (d_0(L(0)), d_1(\bar{L}(1)))$. Under this definition we can easily express either static or dynamic treatment rules, or a combination of the two (see examples in section 4.1).

The G-formula is defined as the product across all nodes, excluding intervention nodes, of the conditional distribution of each node given its parent nodes in the model, and with the values of the intervention nodes fixed according to the static or dynamic intervention of interest. This formula thus expresses the distribution of $\bar{L}$ under the dynamic intervention $\bar{A} \equiv (A(0), A(1)) = d(\bar{L})$:

$$P^d(\bar{L}) = \prod_{j=0}^{2} Q^d_{L(j)}(\bar{L}(j)), \tag{2}$$

where

$$Q^d_{L(j)}(\bar{L}(j)) \equiv P(L(j) \mid \bar{L}(j-1), \bar{A}(j-1) = d(\bar{L}(j-1))).$$

The superscript d here indicates that the conditional distribution of each node given its parent L nodes is also conditional on treatment being set according to the specified treatment rule. We reserve subscript d to refer to counterfactually-defined variables.

2.1 Causal and Statistical Models

We assume a structural causal model (SCM, see Pearl, 2000) and associated causal model $\mathscr{M}^F$, which includes all possible distributions compatible with the assumed SCM. Elements of the *observed* data model, $\mathscr{M}$, can be thought of as being indexed by the elements of $\mathscr{M}^F$, i.e., $\mathscr{M} = \left\{P_{P_{U,X}} : P_{U,X} \in \mathscr{M}^F\right\}$.

Suppose now that we are interested in the outcomes of individuals had their treatment regimen been assigned according to some rule, d. Given a particular SCM such as the one defined above, we can write Y_d, the so-called counterfactual outcome under rule d, as the value Y would have taken on under the intervention where A is set to the value dictated by $d(\bar{L})$ as specified by the SCM.

With the counterfactual outcome Y_d now defined in terms of the solution to a system of structural equations given by the SCM, we define a corresponding counterfactual parameter $\Psi^F(P_{U,X}) = EY_d$. Using (2),

$$\Psi^F(P_{U,X}) = EY_d = \sum_{l(0),l(1)} E\left(Y_d \mid L(0) = l(0), L_d(1) = l(1)\right) \prod_{j=0}^{1} Q_{L_d(j)}(\bar{l}(j)), \quad (3)$$

where $Q_{L_d(j)} \equiv P(L_d(j) \mid \bar{L}_d(j-1))$ and we omit the subscript d on $L(0)$ since it is prior to any treatment. In words, this parameter is the mean outcome under $P_{U,X}$ when treatment is set according to $\bar{A} = d(\bar{L})$.

For the parameter of interest here, EY_d, the sequential randomization assumption (SRA), $Y_d \perp A(j) \mid Pa(A(j))$ for $j = 0, 1$, is sufficient for identification of the causal parameter $\Psi^F(P_{U,X})$ and a particular parameter of the observed data distribution $\Psi(P_0)$ for some Ψ (Robins, 1986). In particular, the SRA implies

$$\begin{aligned}
\Psi^F(P_{U,X}) &\equiv EY_d \\
&\overset{SRA}{=} \sum_{l(0),l(1)} E\left(Y \mid L(0) = l(0), L(1) = l(1), \bar{A} = d(\bar{L})\right) \times \\
&\qquad P\left(L(1) = l(1) \mid L(0) = l(0), A(0) = d_0\right) \times \\
&\qquad P(L(0) = l(0)) \\
&= \Psi(P_0). \qquad (4)
\end{aligned}$$

which is the so-called *identifiability result*.

Note that this parameter depends only on the Q part of the likelihood and we therefore also write $\Psi(P_0) = \Psi(Q_0)$. Note also that the first two factors in the summand are undefined if either $P\left(\bar{A} = d(\bar{L}) \mid L(0) = l(0), L(1) = l(1)\right) = 0$ or $P\left(A(0) = d_0 \mid L(0) = l(0)\right) = 0$ for any $(l(0), l(1))$, and so we require these two conditional probabilities to be positive (the so-called *positivity assumption*).

In this article we present a method for semi-parametric efficient estimation of causal effects. This is achieved through estimation of the parameters of the G-computation formula given above. The method is based on n independent and identically distributed observations of O, and our statistical model $\mathscr{M}$, corresponding to the causal model $\mathscr{M}^F$, makes no assumptions about the conditional distributions $Q_{L(j)},\ j = 0, 1, 2$.

The parameter EY_d can be approximated by generating a large number of observations from the intervened distribution P_d and taking the mean of the final outcome, in this case $L(2)$. The joint distribution P_d can itself be approximated

by simulating sequentially from the conditional distributions $Q_{L_d(j)}, j = 0,1,2$ to generate the observed values $L(j)$.

EY_d can also be computed analytically:

$$\begin{aligned}\Psi(Q_0) \equiv EY_d &= \sum_y y \sum_{l(0),l(1)} P_d[l(0),l(1),y] \\ &\overset{SRA}{=} \sum_y y \sum_{l(0),l(1)} P[Y = y \mid \bar{A} = d(\bar{L}), L(0) = l(0), L(1) = l(1)] \times \\ &\qquad P[L(1) = l(1) \mid L(0) = l(0), A(0) = d_0(L(0))] \times P[L(0) = l(0)] \\ &= \sum_y y \sum_{l(0),l(1)} Q^d_{L(2)}(l(0),l(1),y)\, Q^d_{L(1)}(l(0),l(1))\, Q^d_{L(0)}(l(0)),\end{aligned}$$

The last expression is equivalent to the sum given in (4) if Y is binary. If $L(0)$ is continuous, the sum over $l(0)$ is replaced by an integral. The integral is replaced in turn by the empirical distribution if the expression above is approximated from a large number of observations. In that case the last line reduces to

$$\Psi(Q_0) = \frac{1}{n} \sum_{i=1}^{n} \sum_y y \sum_{l(1)} Q^d_{L(2)}(L(0)_i, l(1), y)\, Q^d_{L(1)}(L(0)_i, l(1)). \tag{5}$$

The latter expression represents a well-defined mapping from the conditional distributions $Q_{L(j)}$ to the real line. Given an estimator $Q_n \equiv \prod_{j=0}^2 Q_{L(j)_n}$ of $Q_0 \equiv \prod_{j=0}^2 Q_{L(j)}$ we arrive at the substitution estimator $\Psi(Q_n)$ of $\Psi(Q_0)$.

Next we describe the targeted maximum likelihood estimator (TMLE) of the relevant parameters of the G-computation formula. The TMLE is a double-robust and locally efficient substitution estimator. The methods described here extend naturally to data structures with more time points, and/or more than one time-dependent confounder per time point (van der Laan, 2010a).

3 Targeted Maximum Likelihood Estimator

With the above parameter now established to be a well-defined mapping from the distribution of the data to the real line, we turn to the estimation of the conditional distributions, $Q_{L(j)}$ which are the domains of the function defining the parameter of interest, $\Psi(Q_0)$.

3.1 Basic Description

In targeted maximum likelihood estimation we begin by obtaining an initial estimator of Q_0; we then update this estimator with a fluctuation function that is tailored specifically to remove bias in estimating the particular parameter of interest. Naturally, this means that the fluctuation function is a function of the parameter of interest. There are, of course, various methods for obtaining an initial estimator: one can propose a parametric model for each factor $Q_{L(j)}$ and estimate the coefficients using maximum likelihood, or one can employ machine learning algorithms which use the data itself to build a model. The former method involves using standard software if the factors $L(j)$ are binary. Each of these general methods in turn has many variants. We favor machine learning, and in particular the Super Learner approach (van der Laan et al., 2007). Utilization of machine learning for the initial estimator $Q^{(0)}$ of Q can improve efficiency of the TMLEs (as well as A-IPCW) in a randomized trial. In observational settings, in which g is unknown, the utilization of machine learning can improve both bias and variance; in either context we feel it is important to employ machine learning to achieve the best performance from the double-robust estimators.

Upon obtaining an initial estimate $Q^{(0)}$ of Q_0, the next step in TMLE is to apply a fluctuation function to this initial estimator that is the least favorable parametric submodel through the initial estimate, $Q^{(0)}$ (van der Laan and Rubin, 2006). This parametric submodel through $Q^{(0)}$ is chosen so that estimation of $\Psi(Q_0)$ is "hardest in the sense that the parametric Cramer-Rao Lower Bound for the variance of an unbiased estimator is maximal among all parametric submodels," (van der Laan, 2010a). Since the Cramer-Rao lower bound corresponds with a standardized L_2 norm of $d\Psi(Q_n(\varepsilon))/d\varepsilon$ evaluated at $\varepsilon = 0$, this is equivalent to selecting the parametric submodel for which this derivative is maximal w.r.t. this L_2 norm.

We also seek an (asymptotically) efficient estimator. This too is achieved with the above described fluctuated update $Q_n(\varepsilon)$ because the score of our parametric submodel at zero fluctuation equals the efficient influence curve of the pathwise derivative of the target parameter, Ψ (also evaluated at $\varepsilon = 0$).

TMLE thus essentially consists in 1) selecting a submodel $Q_g(\varepsilon)$ possibly indexed by nuisance parameter g, and 2) a valid loss function $\mathscr{L}(Q,O) : (Q,O) \rightarrow \mathscr{L}(Q,O) \in \mathbb{R}$. Given these two elements, TMLE solves

$$P_n \left\{ \frac{d}{d\varepsilon} [\mathscr{L}(Q_n^*(\varepsilon))]_{\varepsilon=0} \right\} = 0,$$

so if this "score" is equal to the efficient influence curve, $D^*(Q_n^*, g_n)$, then we have that Q_n^* solves $P_n D^*(Q_n^*, g_n) = 0$. Here we used the notation $P_n f \equiv \frac{1}{n}\sum_{i=1}^n f(O_i)$. Now a result from semi-parametric theory is that solving this efficient score for

the target parameter yields, under regularity conditions (including the requirement that Q_n and g_n consistently estimate Q_0 and g_0, respectively), an asymptotically linear estimator with influence curve equal to $D^*(Q_0, g_0)$. The TMLE of the target parameter is therefore efficient. Moreover, the TMLE is double-robust in that it is a consistent estimator of $\Psi(Q_0)$ if either Q_n or g_n is consistent.

TMLE acquires these properties by choosing the fluctuation function, $Q_n(\varepsilon)$, such that it includes a term derived from the efficient influence curve of $\Psi(Q_0)$.

The following theorem presents the efficient influence curve for a parameter like the ones described above. The content of the theorem will make it immediately apparent why the fluctuation function described subsequently takes the form it does; i.e., it will be seen how the terms in the efficient influence curve lead directly to the form of the fluctuation function, $Q_{L(j)n}(\varepsilon)$.

3.2 Efficient Influence Curve

We repeat here Theorem 1 from van der Laan (2010a).

Theorem 1 *The efficient influence curve for $\Psi(Q_0) = E_0 Y_d$ at the true distribution P_0 of O can be represented as*

$$D^* = \Pi(D_{IPCW} \mid T_Q),$$

where

$$D_{IPCW}(O) = \frac{I(\bar{A} = d(\bar{L}))}{g(\bar{A} = d(\bar{L}) \mid X)} Y - \psi.$$

T_Q is the tangent space of Q in the nonparametric model, X is the full data (in the present context the full data X would be defined as $(L(0), L(1)_d, L(2)_d)$ and Π denotes the projection operator onto T_Q in the Hilbert space $L_0^2(P_0)$ of square P_0-integrable functions of O, endowed with inner product $\langle h_1, h_2 \rangle = E_{P_0} h_1 h_2(O)$.

This subspace,

$$T_Q = \sum_{j=0}^{2} T_{Q_{L(j)}}$$

is the orthogonal sum of the tangent spaces $T_{Q_{L(j)}}$ of the $Q_{L(j)}$-factors, which consists of functions of $L(j), Pa(L(j))$ with conditional mean zero, given the parents $Pa(L(j))$ of $L(j), j = 0, 1, 2$. Recall also that we denote $L(2)$ by 'Y.'

Let

$$D_j^*(Q,g) = \Pi(D_j \mid T_{Q_{L(j)}}).$$

Then

$$\begin{aligned}
D_0^* =& E(Y_d \mid L(0)) - \psi, \\
D_1^* =& \frac{I[A(0) = d_0(L(0))]}{g[A(0) = d_0(L(0)) \mid X]} \left\{C_{L(1)}(Q_0)(1) - C_{L(1)}(Q_0)(0)\right\} \times \\
& \qquad \{L(1) - E[L(1) \mid L(0), A(0)]\}, \\
D_2^* =& \frac{I[\bar{A} = d(\bar{L})]}{g[\bar{A} = d(\bar{L}) \mid X]} \left\{L(2) - E[L(2) \mid \bar{L}(1), \bar{A}(2)]\right\},
\end{aligned}$$

where, for $\delta = \{0,1\}$ *we used the notation*

$$C_{L(1)}(Q_0)(\delta) \equiv E(Y_d \mid L(0), A(0) = d(L(0)), L(1) = \delta).$$

We note that

$$E[Y_d \mid L(0), A(0) = d_0(L(0)), L(1)] = E[Y \mid \bar{L}(1), \bar{A} = d(\bar{L})].$$

We omit the rest of the theorem as presented in van der Laan (2010a) as it pertains to data structures with up to T time points, $T \in \mathbb{N}$.

As mentioned above, TMLE solves the efficient influence curve equation, $P_n D^*(Q_n^*, g_n)$. This is accomplished by adding a covariate to an initial estimator $Q_{L(j)}^{(0)}$ as follows. (Here $L(j)$ is taken as binary.)

$$logit[Q_{L(j)n}(\varepsilon)] = logit[Q_{L(j)n}^{(0)}] + \varepsilon C_{L(j)}(Q_n, g_n), \tag{6}$$

where, for example,

$$C_{L(1)}(Q,g) \equiv \frac{I[A(0) = d_0(L(0))]}{g[A(0) = d_0(L(0)) \mid X]} \left\{C_{L(1)}(Q_0)(1) - C_{L(1)}(Q_0)(0)\right\},$$

with $C_{L(1)}(Q_0)(\delta)$ as defined in Theorem 1, and

$$C_{L(2)}(Q,g) \equiv \frac{I(\bar{A} = d(\bar{L})))}{g(\bar{A} = d(\bar{L})) \mid X)}.$$

It immediately follows that this choice of $Q_{L(j)}(\varepsilon)$ yields a score that is equal to the efficient influence curve at $\varepsilon = 0$ as claimed.

3.3 Implementation of the TMLEs

Below we briefly describe two different procedures for the fitting of ε, which we call the *one-step* and *iterative* approaches, and which result in two distinct targeted maximum likelihood estimators. The iterative approach estimates a common ε for all factors for which a fluctuation function is applied, and the one-step estimator fits each factor separately. In the latter case 'ε' in equation (6) should be replaced with 'ε_j.'

In a forthcoming article we present yet another method of fitting ε (see the working paper, Chaffee and van der Laan, 2011). The method there invovles solving the efficient influence curve equation directly, rather than indirectly by solving the score equation.

It's worth noting that the number of different TMLEs is not limited to the number of methods for fitting the fluctuation function. Targeted maximum likelihood estimators can also be indexed by different initial estimators, $Q^{(0)}$. The class of TMLEs is defined by the fact that they all apply a specific fluctuation function to the initial estimator $Q^{(0)}$ (which is explicitly designed so that the derivative of the loss function at zero fluctuation is equal to the efficient influence curve), independent of the choice of $Q^{(0)}$, and a loss function for the purposes of estimating ε.

Of course, some choices for $Q^{(0)}$ are better than others in that they will be better approximations of Q_0. Doing a good job on the initial estimator has important performance consequences.

One-Step TMLE

The one-step TMLE exploits the fact that estimates of the conditional distributions of Y and Y_d are not required in order to compute the clever covariate term of $Q_{L(2)}(\varepsilon)$, the latter being the final Q_0 term in the time-ordering of the factors (for a two-stage sequential randomized trial). This allows one to update $Q^{(0)}_{L_d(2)} \equiv P(Y_d = 1 \mid L_d(1), L(0)) = E_{Q^{(0)}}[Y_d \mid L_d(1), L(0)]$ with its fluctuation $\varepsilon_2 C_{L(2)}(Q,g)$ first, then use this updated (i.e., fluctuated) estimate $Q^*_{L(2)}$ in the updating step of the $Q_{L(1)}$ term. We remind the reader that the efficient influence curve—and hence $C_{L(j)}(Q,g)$—is parameter-specific, and therefore different parameters (which in our context amounts to different EY_d indexed by d) will have different realizations of the clever covariates.

As with the maximum likelihood estimator (discussed in section 4), both estimators (one-step and iterative) require an initial estimate $Q^{(0)}_{L(j)}$ of $Q_{L(j)}$ for

$j = 0,1,2$, where $Q^{(0)}_{L(0)} \equiv P_{Q^{(0)}}(L(0))$ will just be estimated by the empirical distribution of $L(0)$. Thus the estimates $Q^{(0)}_{L(j)}$, $j = 1,2$ would just be, e.g., the ML estimates if that is how one obtains an initial estimate of Q_0. Upon obtaining these initial estimates of Q_0, one then computes an "updated" estimate $Q^*_{L(2)}$ by fitting the coefficient ε_2 using (in this case of binary factors), logistic regression. The estimate of ε_2 is thus an MLE. This means computing a column of values of $C_{L(2)}$ (one value per observation) and then regressing the outcome $L(2)$ on this variable using the logit of the initial prediction (based on $Q^{(0)}_{L(2)}$) as offset. That is, for each observation a predicted value of $L(2)$ on the logit scale is generated based on the previously obtained $Q^{(0)}_{L(2)}$. Then $\varepsilon_{2,n}$ is found by regressing $L(2)$ on the computed column $C_{L(2)}$ with $logit\left(Q^{(0)}_{L(2)}\right)$ as offset. (This is achieved in R with the offset argument in the *glm* function.)

Note that this clever covariate, $C_{L(2)}$, requires an estimate of $g(\bar{A} \mid X) = g(\bar{A} \mid L(0), L(1))$ (the latter equality valid under the sequential randomization assumption). With $A(0)$ random and $A(1)$ a function of $L(1)$ only, and if $L(1)$ is binary or discrete, this estimate is easily obtained non-parametrically. If $L(1)$ is continuous, some modeling will be required.

Having obtained an estimate $Q^*_{L(2)}$ (which is parameter-dependent, and hence *targeted* at the parameter of interest), one then proceeds to update the estimate of $Q_{L(1)}$ by fitting the coefficient $\varepsilon_{1,n}$—again using logistic regression if $L(1)$ is binary. Note that $C_{L(1)}(Q,g)$ involves an estimate of $Q_{L(2)}$. Naturally, we use our best (parameter-targeted) estimate for this, $Q^*_{L(2)}$, which was obtained in the previous step. $Q^* = (Q^*_{L(1)}, Q^*_{L(2)})$ now solves the efficient influence curve equation, and iterating the above procedure will not result in an updated estimate of Q^*—i.e., the k^{th} iteration estimate $\varepsilon^{(k)}$ will be zero for all $k > 1$ if the procedure is repeated using the Q^* obtained in the previous round as initial estimator. Armed now with the updated estimate $Q^* \equiv (Q^*_{L(1)}, Q^*_{L(2)})$, we obtain the one-step TMLE, $\Psi(Q^*)$, from the G-computation formula (5) for our parameter of interest with Q^* in place of Q_0.

When $L(1)$ is multilevel—say, four levels—one can model $Q_{L(1)}$ as follows. Code each of the categories for $L(1) \in \{0,1,2,3\}$ as a binary indicator variable,

$L(1,m),\ m=0,1,2,3$:

$$P[L(1)=m \mid Pa(L(1))] = P[L(1)=m \mid L(1)\geq m, Pa(L(1))] * P[L(1)\geq m \mid Pa(L(1))]$$
$$= P[L(1,m)=1 \mid L(1)\geq m, Pa(L(1))] \prod_{m'=0}^{m-1} \left\{1 - P[L(1,m')=1 \mid L(1)\geq m', Pa(L(1))]\right\}$$
$$= Q_{L(1,m)}(1, \bar{L}(1,m-1)=0, Pa(L(1))) \prod_{m'=0}^{m-1} Q_{L(1,m')}(0, \bar{L}(1,m'-1)=0, Pa(L(1))),$$

where $\bar{L}(1,s) = (L(1,s), L(1,s-1), ..., L(1,0))$. In this way, the conditional density of each binary factor of $L(1)$, $Q_{L(1,m)}$, can be estimated using logistic regression. We now denote $Q_{L(1)} = \prod_{m=0}^{3} Q_{L(1,m)}$.

To estimate these binary conditional densities, one creates a new data set analogous to a repeated measures data set, in which the number of rows corresponding to each observation is determined by the value of m for which $L(1,m)=1$. For example, suppose that for individual i, $L(1)_i = 2$ and therefore $L(1,2)_i = 1$. Then i will contribute three rows of data where the values in the cells for each row are identical except for two columns: a column that denotes an indicator and an adjacent column corresponding to the increasing values of m from 0 to 2. The rows for the indicator column for this individual are 0 up until $m=2$ (at which the indicator is 1), and the next row is the first row for the next individual in the dataset. One now performs a logistic regression of the column corresponding to the indicator on the parents of $L(1)$, including the column for m.

Now with conditional densities for these binary indicator variables in hand, one can proceed with the targeting step. Each $Q_{L(1,m)},\ m=0,1,2,3$ is updated by adding a clever covariate term. The terms are again derived from the corresponding part of the efficient influence curve associated with the likelihood of the data, as factorized according to this new data structure with binary indicator variables (see Appendix I). One can see from these terms that the updating proceeds as above for the binary $L(1)$ case, i.e., one computes $C_{L(2)}$ first, then the terms $C_{L(1,m)},\ m=0,1,2,3$ in sequence backwards in time, starting with $C_{L(1,3)}$, and performs logistic regression to obtain the estimates of ε. Again, this process of computing the clever covariates and estimating the corresponding $\varepsilon's$ converges in one round.

Iterative TMLE

The procedure here corresponds to estimating ε with the MLE,

$$\varepsilon_n = \underset{\varepsilon}{argmax} \prod_{j=1}^{2} \prod_{i=1}^{n} Q_{L(j),n}(\varepsilon)(O_i).$$

In contrast to the one-step approach, here we estimate a single/common ε for all factors $Q_{L(j)}$, $j = 1,2$.

This iterative approach requires treating the observations as repeated measures. Thus, (assuming $L(1)$ binary for the moment), each observation contributes two rows of data, and instead of a separate column for $L(1)$ and $L(2)$, the values from these columns are alternated in a single column one might call "outcome." Thus the first two rows in the data set correspond to the first observation. Both rows are the same for this first observation except for three columns: those for outcome, offset and clever covariate. There are no longer separate columns for $L(1)$ and $L(2)$, nor for the offsets, and there is likewise a single column for $C_{L(j)}$. The rows for all three columns alternate values corresponding to $j = 1$ and $j = 2$ (as described for $L(j)$).

If $L(1)$ is multi-level, the repeated measures for each observation consists of the rows described in the previous section, plus one row for $L(2)$.

Maximum likelihood estimation of ε is then carried out by running logistic regression on the outcome with $C_{L(j)}$ as the sole covariate (this column contains $C_{L(1,m)}$ for all non-degenerate m if $L(1)$ is discrete), and with $logit\left(Q^{(0)}_{L(j)}\right)$ as offset. This value of ε_n is used as coefficient for the clever covariates in the $Q_{L(j)}(\varepsilon)$ terms for the next iteration. Note that $C_{L(1)} : (Q_n, g_n) \rightarrow C_{L(1)}(Q_n, g_n)$. Thus for the k^{th} iteration ($k = 1,2,...$), $C^{(k)}_{L(1)} = C^{(k)}_{L(1)}(Q^{(k-1)}_n, g_n)$, and g_n is not updated. The process can be iterated till convergence. Convergence is hardly required, however, if the difference $|\psi^{(k-1)}_n - \psi^{(k)}_n|$ is much smaller than $var(\psi^{(k-1)}_n)$. Here $\psi^{(k)}_n \equiv \Psi(Q^{(k)}(\varepsilon))$ is the k^{th} iteration TMLE of the parameter, and the estimated variance, $var_n(\psi^{(k-1)}_n)$ can be used in place of the true variance. Our simulations suggest that the iterated values of $\psi^{(k)}_n$ are approximately monotonic, and in any case, the value of $|\varepsilon_n|$ for successive iterations typically diminishes more than an order of magnitude. The latter fact implies that successive iterations always produce increasingly smaller values of the absolute difference $|\psi^{(k-1)}_n - \psi^{(k)}_n|$, which means that once this difference meets the above stated criterion, the process is complete for all practical purposes.

4 Simulations

We simulated data corresponding to the data structure described in section 2 under varying conditions. The specifics of the data generation process are given in the appendix. The conditions chosen illustrate the double-robustness property of TMLE and EE, and behavior at various sample sizes. We report on simulations in which

$A(0)$ was assigned randomly but $A(1)$ was assigned in response to an individual's $L(1)$; the latter corresponding to an individual's intermediate response to treatment $A(0)$. The specification of these dynamic regimes are given in the following section.

Simulations were divided into two main cases: binary $L(1)$, and discrete $L(1)$ with four levels. For each simulated data set, we computed the estimate of the target parameter $\Psi(P_0) \equiv EY_d$ for three specific rules using the following estimators: 1) One-step TMLE; 2) Iterative TMLE; 3) Inverse Probability of Treatment Weighting (IPTW); 4) Efficient Influence Curve Estimating Equation Methodology (EE); 5) Maximum Likelihood Estimation using the G-computation formula. In the *Results* subsection we give bias, variance and relative MSE estimates.

Here is a brief description of each of the estimators examined.

- *Maximum Likelihood*
  The (parametric) MLE requires a parametric specification of $Q_{L(j)}$ for computation of the parameter estimate, $\Psi(Q_0)$. The form used (e.g., $Q_{L(j),n} = expit[m(\bar{L}(j-1),\bar{A}(j-1) \mid \beta_n)]$ for some function $m(\cdot \mid \cdot)$) was either that of the correct $Q_{L(j)}$ or a purposely misspecified form, and in either case the MLE of the coefficients β were obtained with common software (namely, the *glm* function in the R language). The estimate of EY_d was then computed using the G-computation formula (5), which, e.g., with binary Y and binary $L(1)$, and using the empirical distribution of $L(0)$ is

$$\begin{aligned}\Psi(Q_0) =& \frac{1}{n}\sum_{i=1}^{n}\sum_{y}y\sum_{l(1)}Q^d_{L(1)}(L(0)_i,l(1))Q^d_{L(2)}(L(0)_i,l(1),y)\\ =& \frac{1}{n}\sum_{i=1}^{n}\Big\{Q^d_{L(1)}(L(0)_i,l(1)=1)Q^d_{L(2)}(L(0)_i,l(1)=1,y=1)\\ &+Q^d_{L(1)}(L(0)_i,l(1)=0)Q^d_{L(2)}(L(0)_i,l(1)=0,y=1)\Big\}.\end{aligned}$$

  The maximum likelihood estimator, which is a substitution estimator, can thus be expressed as $\Psi^{MLE}_n = \Psi\left(Q^{(0)}\right)$, where for each factor $Q^d_{L(j)}$ in the G-computation formula, the corresponding MLE, $Q^{(0)d}_{L(j)}$ is substituted, and where $Q^{(0),d} \equiv Q^{MLE,d}$.
  The estimator thus requires estimations of $Q_{L(j)} \equiv P(L(j) \mid Pa(L(j)))$, which as mentioned above, were correctly specified for one set of simulations and incorrectly specified for another.
- *One-Step TMLE*
  See Implementation section above. The initial estimator of Q_0 is the MLE estimator given above.

- *Iterative TMLE*
  See Implementation section above. Here also the initial estimator of Q_0 is the MLE estimator.
- *IPTW*
  The IPTW estimator is defined to be

$$\psi_n^{IPTW} = \frac{1}{n}\sum_{i=1}^{n} Y_i \frac{I(\bar{A}_i = d(\bar{L})}{g[\bar{A}_i = d(\bar{L}) \mid X_i]}.$$

  As with TMLE, this estimator requires estimation of $g[\bar{A} = d(\bar{L}) \mid X]$, which for binary factors and binary treatment is a straightforward non-parametric computation. The IPTW estimator is known to become unstable when there are ETA violations, or practical ETA violations. Adjustments to the estimator that compensate for these issues have been proposed (Bembom and van der Laan, 2008). In the simulations at hand, $g[\bar{A} = d(\bar{L}) \mid \bar{L}]$ was bounded well away from 0 and 1 but was nevertheless not estimated at all (the true distribution of $A \mid X$ was used). However, van der Laan and Robins (2003) show that there is some efficiency gain in estimating $g(\bar{A} \mid \bar{L})$ over using the known true g.
- *Estimating Equation Method*
  This method solves the efficient influence curve estimating equation in ψ. That is,

$$\psi_n^{EE} = P_n E_{Q_n}(Y_d \mid L(0)) + \frac{1}{n}\sum_i \left\{ D_{1,n}^*(O_i) + D_{2,n}^*(O_i) \right\},$$

  with $D_{1,n}^*, D_{2,n}^*$ as given in Theorem 1 except that the true conditional expectations of Y and of Y_d in the expressions for D_1^* and D_2^* are replaced with their respective sample estimates. The only difference between this estimator and the so-called augmented inverse probability of censoring weights (A-IPCW) estimator is in the way the expression for the efficient influence curve is derived. The results for the A-IPCW estimator should be identical to those for the one we describe here.
  Just as with the TMLE, this estimator requires model specifications of $Q_{L(j)}$, $j = 1,2$ for estimation of $E(Y_d \mid L(0))$ and for the elements of D_1^*, D_2^* that involve conditional expectations of Y_d and of Y. Here again we used the ML estimates of $Q_{L(j)}$, under both correct and incorrect model specification scenarios, i.e., we used $Q_n = Q^{(0)}$ for the factors involving estimates of Q_0 in the estimating equation above. (See description of the Maximum Likelihood Estimator above.)

4.1 Some Specific Treatment Rules

We considered several treatment rules, one set for binary $L(1)$ (three different rules), and a necessarily different set (also three separate rules) for the discrete $L(1)$ case. This permits easy computation of the natural parameters of interest $EY_{d_i} - EY_{d_j}$, for $i \neq j$, where in our case, $i, j = 1,2,3$. Indeed such parameters are arguably the ultimate parameters of interest to researchers utilizing longitudinal data of the type described here, since they implicitly give the optimum treatment rule among those considered. As the number of discrete levels of $L(1)$ increases, one can begin considering indexing treatment rules by threshold levels θ of $L(1)$ such that, e.g., assuming binary $A(0)$ and $A(1)$, one could set A(1) according to $A(1) = [1 - A(0)]I(l(1) < \theta) + [A(0)]I(l(1) \geq \theta)$.

Binary L(1)

In the binary $L(1)$ case, we considered the following three treatment rules

- *Rule 1.* $A(0) = 1$, $A(1) = A(0) * I(L(1) = 1) + (1 - A(0)) * I(L(1) = 0)$. In words, set treatment at time 0 to treatment 1, and if the patient does well on that treatment as defined by $L(1) = 1$, continue with same treatment at time 1. Otherwise, switch at $A(1)$ to treatment 0.
- *Rule 2.* $A(0)$ either 0 or 1, and $A(1) = A(0)$. That is, $A(0)$ can be either 0 or 1, but whatever it is, stay on the same treatment at time 1, independent of patient's response to treatment $A(0)$.
- *Rule 3.* $A(0) = 0$, $A(1) = A(0) * I(L(1) = 1) + (1 - A(0)) * I(L(1) = 0)$. In words, set treatment at time 0 to 0 and if the patient does well, stay on treatment 0 at time 1, otherwise switch to treatment 1 at $A(1)$. This is identical to Rule 1 except that patients start on treatment 0 instead of treatment 1.

Note that estimation of, or evaluation of, a rule-specific parameter does not require that patients were actually assigned treatment in that manner, i.e., according to the rule. If patients were assigned treatment randomly, then one simply needs to know which individuals in fact followed the rule in order to estimate the rule-specific mean outcome. In this case, and with $P(A(0) = 1) = P(A(1) = 1) = 0.5$, one could also construct the simple, consistent estimator $(1/n_d)\sum_i Y_i I(\bar{A}_i = d(\bar{L}_i))$, where $n_d = \sum_i (\bar{A}_i = d(\bar{L}_i))$, but this estimator is inefficient relative to the double-robust estimators.

On the other hand, if treatment was indeed assigned according to, e.g., one of the above treatment rules, then $L(1)$ is a time-dependent confounder. These are really the cases of interest. If one's estimator does not adjust for confounding in

these cases it will be biased. All the estimators we compared attempt to adjust for confounding in one way or another.

Discrete L(1) with Four Values

With discrete-valued $L(1)$ ($L(1) \in \{0,1,2,3\}$), the treatment rules were necessarily modified slightly to accommodate the additional values. The analog of rule 1 above, for example, is of the form

- $A(0) = 1, A(1) = A(0) * I(L(1) > l(1)) + (1 - A(0)) * I(L(1) \leq l(1))$ for some $l(1) \in \{0,1,2,3\}$.

4.2 Simulation Results

Notes on the tables

Estimates of bias, variance and relative mean squared error (Rel MSE) are presented for the TMLEs and several comparison estimators. We define relative MSE for each estimator as the ratio of its MSE to that of an efficient, unbiased estimator. The efficiency bound here is the variance of the efficient influence curve. Thus for each estimator ψ_n of ψ_0,

$$\text{Rel MSE} \equiv \frac{(E(\psi_n) - \psi_0)^2 + var(\psi_n)}{var(D^*(Q,g))/n},$$

where D^* is the efficient influence curve for the relevant parameter, Ψ^F. In fact, the value used in these computations for $var(D^*(Q,g))$ is itself an estimate computed from taking the variance of $D^*(Q_0,g_0)(O)$ from a large number of observations generated from P_0.

The bias values shown are not accurate to much less than 10^{-3}. This is because the true parameter values were also obtained by simulation from the true P_d for each rule d with a large number of observations. Thus bias estimates that appear to be smaller than this should be viewed as simply being $< 10^{-3}$. We indicate these estimates with an asterisk.

Qm, gc denotes results where g (the treatment mechanism) was correctly specified, but $Q_{L(2)}$ was purposely misspecified. Qc, gc are simulations for which both Q and g are correctly specified. In an SRCT, we expect g to be known and thus did not perform analyses with a misspecified g. For each trial scenario we present results for both Qc, gc and Qm, gc. Note that the IPTW estimator is not affected by whether or not Q_n is correctly specified, since it does not estimate Q_0.

Varying numbers of simulations were done under the different scenarios. The number of simulations under each configuration (i.e., a given scenario and either Qc, gc or Qm, gc) ranged from 1990 to 5000 depending on computation time.

Confidence Intervals and Coverage Estimates

Table 3 gives influence curve-based estimates of the true coverage for computed 95% confidence intervals for the two TMLEs. The latter were computed for each simulated data set by estimating the variance of the efficient influence curve using that data set.

Scenario I: Binary L(1); A(1) Assigned in Response to L(1)

For brevity we only include the performance of the estimators for a single parameter, EY_1. The results for the other treatment-rule-specific parameters are similar.

Scenario II: Discrete L(1); A(1) Assigned in Response to L(1)

With discrete $L(1)$ we modeled the binary factors $Q_{L(1,m)}$ similarly to the way these factors were generated, i.e., using a hazard approach (see Appendix II). Thus each binary factor is modeled with logistic regression: as with the binary case, an initial estimate $Q^{(0)}_{L(1,m)}$ is obtained by logistic regression (where this estimator could be correctly or incorrectly specified) and a corresponding fluctuation function applied.

Small Sample Results

We also simulated data under scenario II above for a sample size of 30. We anticipated efficiency differences (if any) between the iterative and one-step TMLEs would show up at this very small sample size (see Discussion section). We saw no significant difference in the variance of these two estimators, however. The performance of the TMLEs at this sample size is remarkable, particularly under model misspecification, and we felt these results warranted a separate table (see Table 4).

4.3 Discussion

Relative efficiency for the ML estimator is almost always ≤ 1. The semi-parametric efficiency bound does not apply in general to that of an estimator based on a parametric model. Even so, when Q is correctly specified, the variance of the ML

Qc, gc

		TMLE (1-step)	TMLE (Iter)	IPTW	MLE	EE
$n = 100$	Bias	3.0e-3	2.8e-3	-1.5e.3	1.2e-3	1.8e-3
	Var	3.9e-3	3.9e-3	1.1e-2	3.9e-3	3.8e-3
	Rel MSE	1.3	1.3	3.9	1.3	1.3
$n = 250$	Bias	*	*	-2.4e-3	1.0e-3	*
	Var	1.3e-3	1.3e-3	4.6e-3	1.3e-3	1.3e-3
	Rel MSE	1.1	1.1	3.9	1.1	1.1
$n = 500$	Bias	*	*	-1.7e-3	*	*
	Var	6.3e-4	6.3e-4	2.3e-3	6.3e-4	6.3e-4
	Rel MSE	1.1	1.1	4.0	1.1	1.1

Qm, gc

		TMLE (1-step)	TMLE (Iter)	IPTW	MLE	EE
$n = 100$	Bias	3.9e-3	3.5e-3	1.5e-3	-1.2e-1	-1.2e-3
	Var	4.5e-3	4.5e-3	1.1e-2	2.8e-3	4.1e-3
	Rel MSE	1.6	1.5	3.9	6.3	1.4
$n = 250$	Bias	1.4e-3	1.1e-3	-2.4e-3	-1.3e-1	-1.3e-3
	Var	1.7e-3	1.7e-3	4.6e-3	1.1e-3	1.6e-3
	Rel MSE	1.4	1.4	3.9	14.6	1.4
$n = 500$	Bias	*	*	-1.7e-3	-1.3e-1	*
	Var	8.7e-4	8.6e-4	2.3e-3	5.7e-4	8.3e-4
	Rel MSE	1.5	1.5	4.0	28.5	1.4

Table 1: Scenario I Results: Performance of the various estimators in estimating EY_1 at various sample sizes. 'Qc, gc': Q correctly specified, g correctly specified; 'Qm, gc': Q misspecified, g correctly specified. Iterative TMLE estimates in this table were for the 5th iteration. Asterisks indicate bias $<$ 10e-3.

Qc, gc

		TMLE (1-step)	TMLE (Iter)	IPTW	MLE	EE
$n = 100$	Bias	-3.1e-3	-3.0e-3	-3.2e-3	-2.6e-3	-3.3e-3
	Var	5.5e-3	5.5e-3	2.0e-2	4.9e-3	5.4e-3
	Rel MSE	1.1	1.1	4.0	1.0	1.1
$n = 200$	Bias	-1.5e-3	-1.4e-3	3.8e-3	1.2e-3	-1.5e-3
	Var	2.6e-3	2.6e-3	1.2e-2	2.3e-3	2.6e-3
	Rel MSE	1.0	1.0	4.1	0.9	1.0
$n = 500$	Bias	*	*	1.3e-3	*	*
	Var	1.0e-3	1.0e-3	4.3e-3	9.0e-4	1.0e-3
	Rel MSE	1.0	1.0	4.2	0.9	1.0

Qm, gc

		TMLE (1-step)	TMLE (Iter)	IPTW	MLE	EE
$n = 100$	Bias	-1.7e-3	-1.7e-3	-3.2e-3	-7.0e-2	-3.2e-3
	Var	5.2e-3	5.2e-3	2.0e-2	2.9e-3	5.1e-3
	Rel MSE	1.0	1.0	4.0	1.5	1.0
$n = 200$	Bias	-1.9e-3	-1.9e-3	3.8e-3	-7.0e-2	-2.2e-3
	Var	2.6e-3	2.6e-3	1.2e-2	1.5e-3	2.6e-3
	Rel MSE	1.0	1.0	4.1	2.5	1.0
$n = 500$	Bias	*	*	1.3e-3	-7.0e-2	*
	Var	1.1e-3	1.1e-3	4.3e-3	6.4e-4	1.1e-3
	Rel MSE	1.1	1.1	4.2	5.5	1.0

Table 2: Scenario II Results: Performance of the various estimators in estimating a single parameter, EY_1, for various sample sizes. 'Qc, gc' means Q correctly specified, g correctly specified, while 'Qm' means Q misspecified. Iterative TMLE estimates in this table were for the 3rd iteration. Asterisks indicate bias $<$ 10e-3.

Scenario I

	n = 100	250	500
Qc,gc			
TMLE (1-step)	0.85	0.92	0.93
TMLE (iter)	0.85	0.92	0.94
Qm,gc			
TMLE (1-step)	0.91	0.93	0.94
TMLE (iter)	0.91	0.93	0.94

Scenario II

	n = 100	200	500
Qc,gc			
TMLE (1-step)	0.88	0.91	0.94
TMLE (iter)	0.89	0.91	0.94
Qm,gc			
TMLE (1-step)	0.91	0.92	0.95
TMLE (iter)	0.91	0.92	0.95

Table 3: Coverage for nominal 95% confidence intervals under both data generation scenarios for the two TMLEs at various sample sizes.

Qc,gc

	Bias	Var	Rel MSE
TMLE (1-step)	-0.016	0.023	1.4
TMLE (iter)	-0.021	0.022	1.4
IPTW	3.7e-3	0.069	4.1
MLE	-0.035	0.021	1.3
EE	-0.027	0.021	1.3

Qm,gc

	Bias	Var	Rel MSE
TMLE (1-step)	-6.5e-3	0.019	1.2
TMLE (iter)	-7.0e-3	0.019	1.1
IPTW	3.7e-3	0.069	4.1
MLE	-3.0e-1	0.070	9.4
EE	-9.8e-3	0.027	1.6

Table 4: Scenario II Data, at $n = 30$: Performance of the various estimators in estimating EY_1. 'Qc, gc' means Q correctly specified, g correctly specified, while 'Qm' means Q misspecified. Iterative TMLE estimates in this table were for the 4th iteration.

estimator appears to be very close to the semi-parametric efficiency bound when $n \geq 200$.

Of particular note is that the TMLE, EE and MLE estimators are already very close to the efficiency bound at $n = 250$ under Qc in the binary $L(1)$ case. Further, the reduction in bias in going to $n = 500$ is small in absolute terms.

Even more noteworthy is the performance of the TMLEs at the small sample size of 30 for the scenario II simulations (discrete $L(1)$). Bias and variance of both estimators are *better* when $Q^{(0)}$ is misspecified. Misspecification in this case consisted in setting $Logit(Q_{L(2)}) = 3 * L(1)$ (compare with the true data generating function given in Appendix II), but using correct specification for $Q_{L(1)}$. With $Q^{(0)}$ misspecified, the bias of both TMLEs is quite small and the variance is very close to the efficiency bound. EE also shows lower bias under incorrect Q, but not lower variance. The better performance under misspecification can be understood by noting that under correct model specification, many more parameters of the model must be fit. We expect that asymptotically, there is a gain in efficiency of the TMLEs and EE if $Q^{(0)}$ is consistently estimated, but these simulations show that a parsimonious model as initial estimator, even if misspecified, can have distinct advantages in TMLE at small sample sizes.

The effect is still noticeable at sample size 100 in the discrete $L(1)$ case. There we also see lower bias of the TMLEs under incorrect model specification than under correct model specification. This phenomenon is not present in the scenario I simulations however.

The advantage of the TMLEs' being substitution estimators also becomes apparent in these small sample results: at $n = 30$, many times the EE and IPTW estimators gave estimates outside the bounds of the model ($EY_d \in [0,1]$). Indeed, under Qm, the EE estimator gave estimates of $EY_1 > 1$ more than 13% of the time. For more extensive performance comparisons between TMLE and other double robust estimators (including the A-IPCW estimator) under various conditions, including sparsity/positivity violation conditions, see, e.g., Porter et al. (2011), Stitelman et al. (2011), Gruber and van der Laan (2010), Stitelman and van der Laan (2010) and van der Laan and Rose (2011).

In general, under incorrect specification of Q we do not expect any of the estimators that estimate Q_0 to be *asymptotically* efficient except for the MLE, which used a much simpler model than the true model and therefore could easily achieve a lower variance bound. Misspecification of Q in all cases meant misspecifying $Q^{(0)}_{L(2)}$ but correctly specifying $Q_{L(1)}$. Thus under Qm, gc the MLE will be biased but the TMLE and EE estimators are double robust and therefore still asymptotically unbiased under correct specification of g. Under the scenarios simulated here g is expected to be known and we therefore omitted simulations in which g is mis-

specified; the latter will of course result in bias of the IPTW estimator. Scenarios in which g is not known, or not completely known are also quite plausible, however; e.g., one can easily imagine settings in which assignment of $A(0)$ and/or $A(1)$ was not done in complete accordance with a defined treatment rule. Nevertheless, even in these cases, with $A(0)$ randomized and $L(1)$ discrete or binary, non-parametric estimation of g would not be difficult. If $A(0)$ is a function of $L(0)$ then some smoothing will be required for the estimate of $g(A(0) \mid L(0))$ and model misspecification is likely to arise.

The two versions of TMLE we've implemented (one-step and iterative) typically agree in their estimate of the parameter to within 1%, and in many cases to within quite a bit less than this. For the two time-point data structure we've simulated, the one-step estimator is conceptually easier to implement than the iterative approach, and slightly faster computationally. As the number of estimated factors increases (either from having multiple time points, multiple covariates in $L(j)$, $1 < j < K$, or both), the iterative method may become the more practical programming choice.

Also noteworthy is that the one-step TMLE requires estimation of two ε's in the binary $L(1)$ case and four ε's in the discrete $L(1)$ case. For the general data structure $(L(0),A(0),...L(K),A(K),L(K+1))$ where intermediate factor $L(j)$ has t_j levels, the number of ε's the one-step estimator must fit is $\sum_{j=1}^{K+1}(t_j-1)$. In contrast, the iterative TMLE performs a fitting of ε that is independent of K and t_j. (Though a new round of fitting occurs for each iteration, the bulk of the fitting occurs in the first iteration.) We thus expected at least a small efficiency advantage for the iterative method, though we have not observed it in the simulations presented here, even in sample sizes as low as 30.

Comparison of the TMLE and Estimating Equation Methods

The fundamental differences between targeted maximum likelihood estimation and estimating equation-based estimation have been detailed in the seminal targeted maximum likelihood paper (van der Laan and Rubin, 2006) and elsewhere (see, e.g., van der Laan and Rose, 2011). The differences bear repeating, however, and we give a synopsis of them here.

The most essential difference is summed up in the fact that a TMLE is defined as a (particular) substitution estimator—i.e, an estimator that can be represented as $\Psi(P_n^*)$ for an estimator P_n^* in the statistical model $\mathscr{M}$—and an EE estimator is not. This difference has important ramifications.

The EE algorithm is defined by writing the efficient influence curve, $D(P)$, as an estimating function $D(\psi,\eta)$ in terms of parameter ψ and nuisance parameter

η, and solving for ψ (van der Laan and Robins, 2003). In general, being able to express $D(P)$ in such a form is not a reasonable requirement for parameters and models. In contrast the TMLE algorithm (described in section 3.1) does not rely on the efficient influence curve's being an estimating function.

The TMLE definition also does not rely on an estimating equation's having a unique solution, while EE is only well defined if the estimating equation has a unique solution in ψ. The existence of multiple solutions of estimating equations is a common phenomenon, just as a log-likelihood can have multiple local maxima (and thus multiple solutions for the associated score equation) even though it has a unique maximum. The TMLE P_n^* of P is not defined as a solution of the equation $0 = \sum_i D^*(P)(O_i)$ in P either (it is not even sensible to state that $P_n D(P) = 0$ has a unique solution in P, since there is a whole class of P's that solve it)—it just happens to solve the efficient score equation $0 = \sum_i D^*(P_n^*)(O_i)$ as a by-product of iteratively maximizing the likelihood (or other loss) along a least-favorable submodel.

Instead of having to deal with multiple solutions of an equation, one might well be faced with an estimating equation with no solution at all (in its parameter space); this can occur, for example, under practical violations of the positivity assumption. In the estimation problem addressed in this article, with positive probability the relevant estimating equation is only solved by a negative number, or number larger than 1. We noted this behavior of the EE estimator—i.e., giving an estimate that's not even a probability—in the Results section above.

As mentioned above, dramatic differences in finite sample performance between EE (of which A-IPCW is an example) and TMLE under practical violations of the positivity assumption have been established in many settings. The erratic behavior of EE in such cases is mainly due to its not respecting the global constraints imposed by the target parameter mapping defined on the statistical model. Such differences in behavior are not expected in a sequentially randomized trial in which the treatment mechanism is known and nicely bounded away from zero, and sample size is reasonably large, but the differences can be quite apparent in observational settings or for very small sample sizes in the sequentially randomized trial setting (as seen here).

Appendix I: Efficient Influence Curve for Discrete $L(1)$

In the following, $D^*_{1,m}$ indicates the efficient influence curve for the m^{th} binary indicator of $L(1), m = 0,1,2,3$, and $Pa(L(1)) = (L(0), A(0))$. We have

$$D^*_{1,0}(O) = \frac{I(A(0) = d_0(L(0)))}{g(d_0(L(0)) \mid X)} \{E(Y_d \mid L(1) = 0, Pa(L(1))) - \sum_{m>0} E\left[Y_d \mid L(1) = m, Pa(L(1))\right] P(L(1) = m \mid L(1) > 0, Pa(L(1)))\} \times \{I(L(1) = 0) - I(L(1) \geq 0) E[I(L(1) = 0) \mid Pa(L(1))]\},$$

where, e.g.,

$$\begin{aligned} &P(L(1) = 2 \mid L(1) > 0, Pa(L(1))) \\ &= \frac{P(L(1) = 2, L(1) > 0 \mid Pa(L(1)))}{P(L(1) > 0 \mid Pa(L(1)))} \\ &= \frac{P(L(1) = 2) \mid Pa(L(1))}{1 - P(L(1) = 0 \mid Pa(L(1)))} \\ &= \frac{P(L(1) = 2 \mid L(1) \geq 2, Pa(L(1))) \prod_{s<2} [1 - P(L(1) = s \mid L(1) \geq s, Pa(L(1)))]}{1 - P(L(1) = 1 \mid Pa(L(1)))} \\ &= P(L(1) = 2 \mid L(1) \geq 2, Pa(L(1))) [1 - P(L(1) = 1 \mid L(1) \geq 1, Pa(L(1)))], \end{aligned}$$

and

$$\begin{aligned} &P(L(1) = 3 \mid L(1) > 0, Pa(L(1))) \\ &= P(L(1) = 3 \mid L(1) \geq 3, Pa(L(1))) \prod_{s=1}^{2} [1 - P(L(1) = s \mid L(1) \geq s, Pa(L(1)))] \\ &= 1 * \prod_{s=1}^{2} [1 - P(L(1) = s \mid L(1) \geq s, Pa(L(1)))]. \end{aligned}$$

Similarly,

$$D^*_{1,1}(O) = \frac{I(A(0) = d_0(L(0)))}{g(d_0(L(0)) \mid X)} \{E(Y_d \mid L(1) = 1, Pa(L(1))) - \sum_{m>1} E\left[Y_d \mid L(1) = m, Pa(L(1))\right] P(L(1) = m \mid L(1) > 1, Pa(L(1)))\} \times \{I(L(1) = 1) - I(L(1) \geq 1) E[I(L(1) = 1) \mid L(1) \geq 1, Pa(L(1))]\},$$

and

$$E[I(L(1)=m) \mid L(1) \geq m, Pa(L(1))] \equiv P(L(1)=m \mid L(1) \geq m, Pa(L(1))).$$

$D^*_{1,2}(O)$ is similar, but $D^*_{1,3}(O) = 0$ since

$$\begin{aligned} I(L(1)=3) - & I(L(1) \geq 3) E[I(L(1)=3) \mid L(1) \geq 3, Pa(L(1))] \\ &= I(L(1)=3) - I(L(1)=3) * E[I(L(1)=3) \mid L(1) \geq 3, Pa(L(1))] \\ &= I(L(1)=3) - I(L(1)=3) * P[L(1)=3 \mid L(1) \geq 3, Pa(L(1))] \\ &= I(L(1)=3) - I(L(1)=3) * 1 = 0. \end{aligned}$$

Thus the efficient influence curve for EY_d is

$$D^*(O) = D^*_0(O) + \sum_{m=0}^{3} D^*_{1,m}(O) + D^*_2(O),$$

with $D^*_0(O)$ and $D^*_2(O)$ exactly as given in Theorem 1.

The expression for clever covariate $C_{L(1,m)}$ follows immediately from $D^*_{1,m}$ as simply the IPCW term times the first bracketed term. So, for example, $C_{L(1,2)}$ would be

$$\begin{aligned} C_{L(1,2)} = & \frac{I(A(0)=d_0(L(0)))}{g(d_0(L(0)) \mid X)} \{E(Y_d \mid L(1)=2, Pa(L(1))) - \\ & \sum_{m>2} E[Y_d \mid L(1)=m, Pa(L(1))] P(L(1)=m \mid L(1)>2, Pa(L(1)))\}. \end{aligned}$$

Appendix II: Data Generation

In this appendix we describe the data generation process for each of the variables in the causal model. There are notable differences in the two major sets of simulations (i.e., the binary $L(1)$ case vs. the discrete $L(1)$ case).

- *L(0)*
  For both binary and discrete $L(1)$ cases, $L(0)$ consisted of four baseline covariates, $L(0) = (W_1, ..., W_4)^T$, three of which were distributed Normally, i.e.,

  $$(W_1, W_2, W_3)^T \sim N(\mu, \Sigma),$$

  with $\mu = (0, -0.35, 0)^T$ and with all off-diagonal terms of Σ set to 0. The fourth baseline covariate W_4 was distributed as a truncated normal, also independent of the other baseline variables. Specifically, let random variable $W' \sim N(5, 1.5^2)$. Then

  $$W_4 = \begin{cases} W' & \text{if } 2 < W' < 8 \\ 0 & \text{otherwise} \end{cases}$$

- *A(0)*
  $A(0)$ was assigned randomly for all simulations, $A(0) \sim Ber(0.5)$
- *L(1)*
  - (1) *Binary* In the binary $L(1)$ case,
    $L(1) \sim Ber([1 + exp(-(Logit[Q_{L(1)}]))]^{-1})$, where

    $Logit[Q_{L(1)}] = \frac{1}{2.5}(2 - W_1 - W_4 - 2W_2^2 + 1.8W_3^2 - 3W_4W_3 + 3A(0) + 2(1 - A(0)))$.

    and with $W_1, ...W_4$ as defined above.
  - (2) *Discrete* The conditional probabilities for each factor $L(1,m)$, $m = 0, 1, 2$, were generated as follows.
    $logit[Q_{L(1,0)}] = \frac{1}{6.5}[-15 - W_1 - W_4 - 2W_2^2 + 1.8W_3^2 - 3W_4W_3 + 3A(0) + 2(1 - A(0))]$,
    $logit[Q_{L(1,1)}] = logit[Q_{L(1,0)}] + 2.8$,

    $logit[Q_{L(1,2)}] = logit[Q_{L(1,1)}] + 4.2$.

- *A(1)*
  - *(1) Binary L(1)* $A(1)$ was set according to

    $$A(1) = \begin{cases} A(0) & \text{if } L(1) = 1 \\ A(0) & \text{with probability 0.5 otherwise} \end{cases}$$

  - *(2) Discrete L(1)* $A(1)$ in the discrete case was set according to

$$A(1) = \begin{cases} A(0) & \text{if } L(1) > 1 \\ A(0) & \text{with probability 0.5 otherwise} \end{cases}$$

- *L(2)*
  - *(1) Binary L(1)* For the binary $L(1)$ simulations, $L(2) \sim Ber([1+exp(-(Logit[Q_{L(2)}]))]^{-1})$, where

$$Logit[Q_{L(2)}] = \tfrac{1}{2.5}(2 - W_1 - W_4 - 2W_2^2 + 1.8W_3^2 - 3W_4W_3 + 3A(0) + 2(1-A(0)) + 2L(1) - 1.5(1-L(1)) + 6 * I(d(\bar{L}) = 1) - 6.5 * I(d(\bar{L}) = 2) - W_1(1-A(0)) + W_4A(1))).$$

  - *(2) Discrete L(1)* For the simulations with discrete L(1), $L(2) \sim Ber([1+exp(-(Logit[Q_{L(2)}]))]^{-1})$, where

$$Logit[Q_{L(2)}] = \tfrac{1}{6}(-7 - W_1 - W_4 - 0.7W_2^2 + 0.6W_3^2 - W_4W_3 + 9A(0) + 3(1-A(0))) + 1.4L(1) - W_1(1-A(0)) + W_4A(1) + 6 * I(d(\bar{L}) = 3).$$

In the above expressions $I(d(\bar{L}) = j), j = 1,2,3$ is equal to 1 if rule j was followed at both treatment time points (as described in section 4.1) and 0 otherwise.

References

Bembom, O. and M. van der Laan (2007): "Statistical methods for analyzing sequentially randomized trials," *Journal of the National Cancer Institute*, 99, 1577–1582.

Bembom, O. and M. van der Laan (2008): "Data-adaptive selection of the truncation level for inverse-probability-of-treatment-weighted estimators," *Berkeley Division of Biostatistics Working Paper Series. Working Paper 230.*

Chaffee, P. and M. J. van der Laan (2011): "Targeted minimum loss based estimation based on directly solving the efficient influence curve equation," *U.C. Berkeley Division of Biostatistics Working Paper Series*, Working Paper 287.

Chaffee, P. and M. J. van der Laan (2012): "Discussion of evaluation of viable dynamic treatment regimes in a sequentially randomized trial of advanced prostate cancer, by Wang et al. 2012," *Journal of the American Statistical Association*.

Gruber, S. and M. J. van der Laan (2010): "A targeted maximum likelihood estimator of a causal effect on a bounded continuous outcome," *The International Journal of Biostatistics*, 6.

Guo, X. and A. Tsiatis (2005): "A weighted risk set estimator for survival distributions in two-stage randomization designs with censored survival data," *International Journal of Biostatistics*, 1.

Laber, E., M. Qian, D. Lizotte, and S. Murphy (2009): "Statistical inference in dynamic treatment regimes," *Revision of Univ. of Michigan, Statistics Department Technical Report 506*.

Lunceford, J., M. Davidian, and A. Tsiatis (2002): "Estimation of survival distributions of treatment policies in two-stage randomization designs in clinical trials," *Biometrics*, 58, 48–57.

Miyahara, S. and A. Wahed (2010): "Weighted Kaplan-Meier estimators for two-stage treatment regimes," *Statistics in Medicine*, 29, 2581–2591.

Moodie, E., R. Platt, and M. Kramer (2009): "Estimating response-maximized decision rules with applications to breastfeeding," *Journal of the American Statistical Association*, 104, 155–165.

Murphy, S., M. van der Laan, and J. Robins (2001): "Marginal mean models for dynamic regimes," *Journal of the American Statistical Association*, 6, 1410–1423.

Orellana, L., A. Rotnitzky, and J. Robins (2010): "Dynamic regime marginal structural mean models for estimation of optimal dynamic treatment regimes," *International Journal of Biostatistics*, 6.

Pearl, J. (2000): *Causality: Models, Reasoning and Inference*, Cambridge: Cambridge University Press.

Petersen, M., Y. Wang, M. van der Laan, D. Guzman, E. Riley, and D. Bangsberg (2007): "Pillbox organizers are associated with improved adherence to HIV antiretroviral therapy and viral suppression: a marginal structural model analysis," *Clinical Infectious Diseases*, 45, 908–15.

Porter, K. E., S. Gruber, M. J. van der Laan, and J. S. Sekhon (2011): "The relative performance of targeted maximum likelihood estimators," *International Journal of Biostatistics*, 7.

Robins, J. (1986): "A new approach to causal inference in mortality studies with sustained exposure periods—application to control of the healthy worker survivor effect," *Mathematical Modeling*, 7, 1393–1512.

Stitelman, O. M., V. D. Gruttola, and M. J. van der Laan (2011): "A general implementation of TMLE for longitudinal data applied to causal inference in survival analysis," *U.C. Berkeley Division of Biostatistics Working Paper Series*, Working Paper 281.

Stitelman, O. M. and M. J. van der Laan (2010): "Collaborative targeted maximum likelihood for time to event data," *The International Journal of Biostatistics*, 6.

Tunis, S., D. Faries, and e. a. Nyhuis, A.W. (2006): "Cost-effectiveness of olanzapine as first-line treatment for schizophrenia: results from a randomize, open-label, 1-year trial," *Value Health*, 9, 77–89.

van der Laan, M. and M. Petersen (2007): "Causal effect models for realistic individualized treatment and intention to treat rules," *The International Journal of Biostatistics*, 3.

van der Laan, M. and J. Robins (2003): *Unified methods for Censored Longitudinal Data and Causality*, New York: Springer Verlag.

van der Laan, M. and S. Rose (2011): *Targeted Learning: Causal Inference for Observational and Experimental Data*, Berlin Heidelberg New York: Springer.

van der Laan, M. and D. Rubin (2006): "Targeted maximum likelihood learning," *The International Journal of Biostatistics*, 2.

van der Laan, M. J. (2010a): "Targeted maximum likelihood based causal inference: Part I," *The International Journal of Biostatistics*, 6.

van der Laan, M. J. (2010b): "Targeted maximum likelihood based causal inference: Part II," *The International Journal of Biostatistics*, 6.

van der Laan, M. J., E. C. Polley, and A. E. Hubbard (2007): "Super learner," *U.C. Berkeley Division of Biostatistics Working Paper Series*, Working paper 222, http://www.bepress.com/ucbbiostat/paper222.

Wahed, A. and A. Tsiatis (2004): "Optimal estimator for the survival distribution and related quantities for treatment policies in two-stage randomization designs in clinical trials," *Biometrics*, 60, 124–133.

Wahed, A. and A. Tsiatis (2006): "Semi-parametric efficient estimation of the survival distribution for treatment policies in two-stage randomization designs in clinical trials with censored data," *Biometrika*, 60, 147–161.

Wang, L., A. Rotnitzky, X. Lin, R. E. Millikan, and P. F. Thall (2012): "Evaluation of viable dynamic treatment regimes in a sequentially randomized trial of advanced prostate cancer," *Journal of the American Statistical Association*, forthcoming.