The Econometrics Journal

*Econometrics Journal* (2022), volume **25**, pp. 628–648.
https://doi.org/10.1093/ectj/utac018



# Evaluating (weighted) dynamic treatment effects by double machine learning

HUGO BODORY[†], MARTIN HUBER[‡] AND LUKÁŠ LAFFÉRS[§]

[†]*Vice-President's Board (Research & Faculty), University of St. Gallen, Varnbüelstrasse 14, 9000 St. Gallen, Switzerland.*
Email: hugo.bodory@unisg.ch

[‡]*Dept. of Economics, University of Fribourg, Bd. de Pérolles 90, 1700 Fribourg, Switzerland.*
Email: martin.huber@unifr.ch

[§]*Dept. of Mathematics, Matej Bel University, Tajovskeho 40, 97401 Banská Bystrica, Slovakia.*
Email: lukas.laffers@gmail.com

First version received: 27 August 2021; final version accepted: 20 December 2021.

**Summary:** We consider evaluating the causal effects of dynamic treatments, i.e., of multiple treatment sequences in various periods, based on double machine learning to control for observed, time-varying covariates in a data-driven way under a selection-on-observables assumption. To this end, we make use of so-called Neyman-orthogonal score functions, which imply the robustness of treatment effect estimation to moderate (local) misspecifications of the dynamic outcome and treatment models. This robustness property permits approximating outcome and treatment models by double machine learning even under high-dimensional covariates. In addition to effect estimation for the total population, we consider weighted estimation that permits assessing dynamic treatment effects in specific subgroups, e.g., among those treated in the first treatment period. We demonstrate that the estimators are asymptotically normal and $\sqrt{n}$-consistent under specific regularity conditions and investigate their finite sample properties in a simulation study. Finally, we apply the methods to the Job Corps study.

**Keywords:** *Dynamic treatment effects*, *double machine learning*, *efficient score*.

**JEL codes:** *C21*.

## 1. INTRODUCTION

In many empirical problems, policy makers and researchers are interested in the causal effects of sequences of interventions or treatments, i.e., dynamic treatment effects. Examples include the impact of sequences of training programmes (for instance, a job application training followed by a language courses) on the employment probabilities of job seekers or the effect of sequential medical interventions (for instance, a surgery combined with rehabilitation training) on health. As treatment assignment is typically nonrandom, causal inference about distinct sequences of treatments requires controlling for confounders jointly affecting the various treatments and the outcome of interest. An assumption commonly imposed in the literature is sequential conditional independence, which implies that the treatment in each period is unconfounded conditional on past treatment assignments, past outcomes, and the history of observed covariates up to the respective treatment assignment. Due to increasing data availability, the number of observed covariates that may potentially serve as control variables to justify the sequential conditional independence

assumption has been growing in many empirical contexts, which poses the question of how to optimally control for such a wealth of information in the estimation process.

This paper combines the semiparametrically efficient estimation of dynamic treatment effects under sequential conditional independence, with the double machine learning (DML) framework outlined in Chernozhukov et al. (2018) to control for observed covariates in a data-driven way. More specifically, treatment effect estimation is based on the efficient score function belonging to the class of doubly robust estimation as discussed in Robins et al. (1994) and Robins and Rotnitzky (1995), and relies on plug-in estimates of the dynamic treatment propensity scores (the conditional treatment probabilities given histories of covariates and past treatments) and conditional mean outcomes (given histories of treatments, covariates, and past outcomes). We obtain these plug-in estimates by machine learning, which permits algorithmically controlling for covariates with the highest predictive power for the treatments and outcomes.

To safeguard against overfitting bias due to correlations between the estimation steps, the plug-in models and the treatment effects are estimated in different parts of the data, whose role is subsequently swapped to prevent not using parts of the data for effect estimation (and thereby increasing the variance). We show that our estimator satisfies the so-called Neyman (1959) orthogonality discussed in Chernozhukov et al. (2018), and is thus asymptotically normal and $\sqrt{n}$-consistent under specific regularity conditions despite the data-driven estimation of the plug-ins. One restriction is that the convergence of the plug-in estimates to the true models as a function of the covariates is not too slow, which is satisfied if each of the estimators converges at a rate faster than $n^{-1/4}$. When using lasso as machine learner, this implies a form of approximate sparsity, meaning that the number of important covariates for obtaining a decent approximation of the plug-ins is small relative to the sample size. However, the set of these important confounders need not be known a priori, which is particularly useful in high-dimensional data with a vast number of covariates that could potentially serve as control variables.

As a further contribution, we discuss the DML-based estimation of weighted dynamic treatment effects where the weight is defined as a function of the baseline covariates. This permits, for instance, assessing treatment sequences among those treated or not treated in the first period, and therefore provides a rather general framework for the definition of interesting subpopulations. Also for this estimator, based on a weighted version of the efficient score function, we show Neyman (1959) orthogonality and $\sqrt{n}$-consistency under specific restrictions on the convergence rates of the plug-in estimators, which now also include the estimated weighting function.

Furthermore, we investigate the method's finite sample behaviour in a simulation study and find the point estimators to perform rather decently in the simulation designs considered. As an empirical contribution, we assess the effects of various treatment sequences in the US Job Corps study on an educational intervention for disadvantaged youth. We find that attending vocational training in the two initial years of the programme likely increases the employment probability four years after the start of Job Corps when compared to no instruction. In contrast, the relative performance of sequences of vocational vs. academic classroom training is less clear.

The literature on dynamic treatment effects goes back to Robins (1986), who proposes a dynamic causal framework along with an estimation approach known as g-computation for recursively modelling outcomes at some point in time, as functions of the (histories of) observed covariates and treatments under the sequential conditional independence assumption. G-computation was originally implemented by parametric maximum likelihood estimation of nested structural models for the outcomes in all periods, requiring the (in general tedious) estimation of the conditional densities of all time-varying covariates. Robins (1998) suggested an alternative, less complex modelling approach based on so-called marginal structural models representing

outcomes in specific treatment states as functions of time-constant covariates only. In order to also control for time-varying confounding, such marginal models need (in the spirit of Horvitz and Thompson, 1952) to be combined with weighting by the inverse of the dynamic treatment propensity scores, see for instance Robins et al. (1999) and Robins et al. (2000). The propensity scores in each period are typically estimated by sequential logit regressions, but see Imai and Ratkovic (2015) for an alternative, empirically likelihood-based approach that aims at finding propensity score specifications that maximise covariate balance. Lechner (2009) considers inverse probability weighting (IPW) by the dynamic treatment propensity scores alone (i.e., without the use of marginal outcome models), while Lechner and Miquel (2010) apply propensity score matching and Blackwell and Strezhnev (2022) direct matching on the covariates.

Doubly robust estimators of dynamic treatment effects comprise methods that are consistent if either the sequential treatment propensity scores or nested outcome models are correctly specified. This includes estimation based on the sample analogue of the efficient influence function (underlying the semiparametric efficiency bounds) provided in Robins (2000), which is a function of both the nested treatment and outcome models.[1] In contrast, Bang and Robins (2005) propose a doubly robust estimator that is based on estimating potential outcomes by nested models of conditional mean outcomes (given the covariate histories as well as past and current treatment assignments) in all periods, a form of g-computation that does not require tedious likelihood estimations of conditional densities as initially proposed in Robins (1986). Here, doubly robustness comes from the fact that a weight based on the nested treatment propensity scores is included as additional covariate in conditional mean estimation.

van der Laan and Gruber (2012) demonstrate that this approach fits the framework of Targeted Maximum Likelihood Estimation (TMLE) of van der Laan and Rubin (2006), which obtains doubly robustness through updating initial conditional outcome estimates by regressing them on a function of the nested propensity scores in each period, and offers a refined estimator. Specifically, they suggest estimating nuisance parameters by the super learner of van der Laan et al. (2007), an ensemble method for machine learning. In contrast, the approach suggested in this paper does not rely on the likelihood estimation of marginal structural models, nor of nested structural models requiring the estimation of conditional covariate densities. Similar to TMLE, our approach is based on combining nested conditional mean outcomes with propensity score estimation. Different from TMLE, however, we base estimation on the efficient influence function, which does not iteratively update the nested outcomes. In addition, we also consider weighted treatment effect estimation as a function of baseline covariates. As we estimate the plug-in parameters by machine learning as recently also considered in Tran et al. (2019), we formally show that our approach fits the DML framework of Chernozhukov et al. (2018), and discuss regularity conditions under which $\sqrt{n}$-consistency is attained. Farbmacher et al. (2022) used similar techniques to analyse DML in the context of mediation analysis.

Lewis and Syrgkanis (2021) propose an alternative DML estimator of dynamic treatment effects. It is based on residualising or debiasing the outcome and the treatment by purging the effects of observed confounders, using machine learning and regressing the debiased outcome on the debiased treatment in a specific period. This approach may also be applied to continuous (rather than discrete) treatments, but, in contrast to our method, assumes partial linearity in the outcome model. Finally, Viviano and Bradic (2021) suggest a further doubly robust method that can be combined with machine learning, but replaces weighting by the inverse of the propensity

[1] Yu and van der Laan (2006) discuss an alternative doubly robust approach based on combining propensity scores with the estimation of marginal structural models.

scores (as applied in our paper) by a dynamic version of covariate balancing as discussed in Zubizarreta (2015) and Athey et al. (2018).

This paper proceeds as follows. Section 2 introduces the concepts of dynamic treatment effects in the potential outcome framework, presents the identifying assumptions, and discusses identification. Section 3 proposes an estimation procedure based on DML, and shows $\sqrt{n}$-consistency and asymptotic normality under specific conditions. Section 4 extends the procedure to the evaluation of weighted dynamic treatment effects. Section 5 provides a simulation study. Section 6 presents an empirical application to data from Job Corps, an educational programme for disadvantaged youth. Section 7 concludes.

## 2. DEFINITION OF DYNAMIC TREATMENT EFFECTS AND IDENTIFICATION

We are interested in the causal effect of a sequence of discretely distributed treatments and will for the sake of simplicity focus on the case of two sequential treatments in the subsequent discussion. To this end, denote by $D_t$ and $Y_t$ the treatment (e.g., a training programme) and the outcome (e.g., employment) in period $T = t$. Therefore, $D_1$ and $D_2$ are the treatments in the first and second periods, respectively, and may take values $d_1, d_2 \in \{0, 1, \ldots, Q\}$, with 0 indicating nontreatment and $1, \ldots, Q$ the different treatment choices (where $Q$ denotes the number of nonzero treatments). Let $Y_2$ denote the outcome of interest measured in the second period after the realisation of treatment sequence $D_1$ and $D_2$.[2] To define the dynamic treatment effects of interest, we make use of the potential outcome framework, see for instance Rubin (1974). Denoting by $\underline{d}_2$ a specific treatment sequence $(d_1, d_2)$ with $d_1, d_2 \in \{0, 1, \ldots, Q\}$, then $\underline{D}_2 \equiv (D_1, D_2)$, and $Y_2(\underline{d}_2)$ denotes the potential outcome hypothetically realised when the treatments are set to that sequence $\underline{d}_2$. We also define $\{0, 1, \ldots, Q\}^2 = \{0, 1, \ldots, Q\} \times \{0, 1, \ldots, Q\}$.

We aim at evaluating the average treatment effect (ATE) of two distinct treatment sequences in the population,

$$\Delta(\underline{d}_2, \underline{d}_2^*) = E[Y_2(\underline{d}_2) - Y_2(\underline{d}_2^*)],$$

with $\underline{d}_2 \neq \underline{d}_2^*$ such that the sequences differ either in $d_1$ or in both $\underline{d}_2$.[3] Examples are the evaluation of a sequence of two binary treatments vs. no treatment, e.g., $\underline{d}_2 = (1, 1)$ and $\underline{d}_2^* = (0, 0)$, or the effect of the first treatment when holding the second treatment constant, $\underline{d}_2 = (1, d_2)$ and $\underline{d}_2^* = (0, d_2)$, with $d_2 \in \{0, 1\}$. The latter parameter is known as the controlled direct effect in causal mediation analysis, see for instance Pearl (2001), assessing the net effect of the first treatment when setting the second treatment to be $D_2 = d_2$ for everyone.[4] Throughout the paper we assume that the stable unit treatment value assumption (SUTVA, Rubin, 1980) holds such that $\Pr(\underline{D}_2 = \underline{d}_2 \Rightarrow Y_2 = Y_2(\underline{d}_2)) = 1$. This rules out interaction effects and general equilibrium effects, and implicitly assumes that treatments are uniquely defined.

[2] We do not consider the evaluation of treatment effects on outcomes in the first period, as this corresponds to the conventional static treatment framework as, for instance, considered in Chernozhukov et al. (2018).

[3] In the case of $\underline{d}_2, \underline{d}_2^*$ sharing the same $d_1$, but differing in terms $d_2$, the identification problem collapses to the standard case with one treatment period (namely $T = 2$) under the condition that $D_1 = d_1$. The case of a single treatment period also prevails when considering the effects on $Y_1$, i.e., the outcome in period $T = 1$, which only permits assessing the effect of $D_1$. In either case, the standard DML framework for single treatment periods can be applied as, e.g., outlined in Belloni et al. (2017) such that we do not consider these scenarios in this paper.

[4] From the perspective of causal mediation analysis, our paper complements the study of Farbmacher et al. (2022), who apply DML to the estimation of so-called natural direct and indirect effects. In the latter case, $D_2$ is not prescribed to have the same value $d_2$ for everyone, but corresponds to the potential value $D_2(d_1)$, i.e., the hypothetical treatment state of $D_2$ that would be 'naturally chosen' (i.e., without prescription) as a consequence of $D_1 = d_1$.

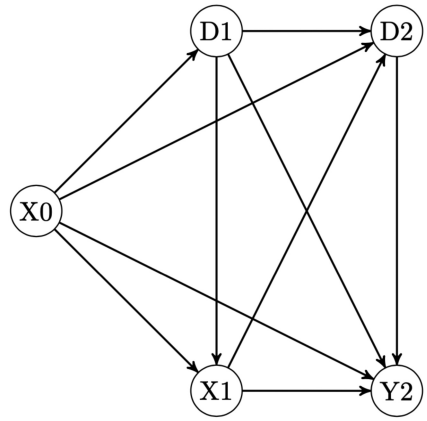


**Figure 1.** Causal paths under sequential conditional independence.

Identification relies on a sequential conditional independence assumption, requiring that the treatment in each period is conditionally independent of the potential outcomes, conditional on previous treatments and (histories of) observed covariates measured prior to treatment, which might include past outcomes, too. Let, to this end, $X_t$ denote the observed characteristics in period $T = t$. $X_0$ consists of pre-treatment characteristics measured prior to the first treatment $D_1$, while $X_1$ (which may contain $Y_1$) is measured prior to $D_2$, but may be influenced by $D_1$ as well as $X_0$. Covariates in a particular period may therefore be affected by previous covariates and treatments, implying that confounding may be dynamic in the sense that identification relies on time-varying observables rather than on baseline covariates alone. Figure 1 provides a graphical illustration using a directed acyclic graph, with arrows representing causal effects. Each of $D_1$, $D_2$, and $Y_2$ might be causally affected by distinct and statistically independent sets of unobservables not displayed in Figure 1, but none of these unobservables may jointly affect $D_1$ and $Y_2$ given $X_0$, or $D_2$ and $Y_2$ given $D_1$, $X_0$, and $X_1$.

Formally, the first assumption invokes conditional independence of the treatment in the first period $D_1$ and the potential outcomes $Y_2(\underline{d}_2)$ given $X_0$ as commonly invoked in the treatment evaluation literature, see, e.g., Imbens (2004). It rules out unobserved confounders jointly affecting $D_1$ and $Y_2(\underline{d}_2)$ conditional on $X_0$.

ASSUMPTION 2.1 (CONDITIONAL INDEPENDENCE OF THE FIRST TREATMENT): $Y_2(\underline{d}_2) \perp D_1 | X_0$, *for* $\underline{d}_2 \in \{0, 1, \ldots, Q\}^2$, *where '*$\perp$*' denotes statistical independence.*

*The second assumption invokes conditional independence of the second treatment* $D_2$ *given the first treatment* $D_1$*, and the (history of) covariates* $X_0$ *and* $X_1$*, which we denote by* $\underline{X}_1 = (X_0, X_1)$ *to ease notation. It rules out unobserved confounders jointly affecting* $D_2$ *and* $Y_2(\underline{d}_2)$ *conditional on* $D_1$ *and* $\underline{X}_1$.

ASSUMPTION 2.2 (CONDITIONAL INDEPENDENCE OF THE SECOND TREATMENT): $Y_2(\underline{d}_2) \perp D_2 | D_1, X_0, X_1$, *for* $\underline{d}_2 \in \{0, 1, \ldots, Q\}^2$.

*The third assumption imposes common support, meaning that the treatment in each period is not a deterministic function of the respective observables in the conditioning set, which rules out conditional treatment probabilities (or propensity scores) of* 0 *or* 1. *This implies that conditional*

*on each value of the observables occurring in the population, subjects with distinct treatment assignments* $\{0, 1, \ldots, Q\}$ *exist.*

ASSUMPTION 2.3 (COMMON SUPPORT): $\Pr(D_1 = d_1|X_0) > 0$, $\Pr(D_2 = d_2|D_1, \underline{X}_1) > 0$ *for* $d_1, d_2 \in \{0, 1..., Q\}$.

*To ease notation, we henceforth denote the propensity scores by* $p^{d_1}(X_0) = \Pr(D_1 = d_1|X_0)$ *and* $p^{d_2}(D_1, \underline{X}_1) = \Pr(D_2 = d_2|D_1, \underline{X}_1)$. *Furthermore, we denote the conditional mean outcome in the second period by* $\mu^{Y_2}(\underline{D}_2, \underline{X}_1) = E[Y_2|\underline{D}_2, X_0, X_1]$ *and the nested conditional mean outcome in the first period by*

$$\nu^{Y_2}(\underline{D}_2, X_0) = \int E[Y_2|\underline{D}_2, X_0, X_1 = x_1]dF_{X_1=x_1|D_1,X_0},$$

*where* $F_{X_1=x_1|D_1,X_0}$ *denotes the conditional distribution function of* $X_1$ *given* $(D_1, X_0)$ *at value* $x_1$. *For a fixed vector of treatments* $\underline{D}_2 = \underline{d}_2$, *the quantity* $\nu^{Y_2}(\underline{d}_2, X_0)$ *is equal to* $E[E[Y_2|\underline{D}_2 = \underline{d}_2, X_0, X_1]|D_1 = d_1, X_0]$, *and this suggests that it can be obtained by a sequential estimation of nested conditional means. This is the approach followed in this paper, as it avoids the estimation of conditional covariate distributions, which might be cumbersome if covariates are high dimensional.*

As, for instance, discussed in Tran et al. (2019), Assumptions 2.1–2.3 permit identifying the mean potential outcome $E[Y(\underline{d}_2)]$ based on the following expression:

$$E[Y(\underline{d}_2)] = E[\psi^{\underline{d}_2}], \text{ where}$$

$$\psi^{\underline{d}_2} = \frac{I\{D_1 = d_1\} \cdot I\{D_2 = d_2\} \cdot [Y_2 - \mu^{Y_2}(\underline{d}_2, \underline{X}_1)]}{p^{d_1}(X_0) \cdot p^{d_2}(d_1, \underline{X}_1)} + \frac{I\{D_1 = d_1\} \cdot [\mu^{Y_2}(\underline{d}_2, \underline{X}_1) - \nu^{Y_2}(\underline{d}_2, X_0)]}{p^{d_1}(X_0)} + \nu^{Y_2}(\underline{d}_2, X_0).$$

This follows from the fact that $\psi^{\underline{d}_2} = E[Y(\underline{d}_2)]$, which corresponds to the efficient score function of dynamic treatment effects, as discussed in Robins (2000), has a zero mean property: $E[\psi^{\underline{d}_2} - E[Y(\underline{d}_2)]] = 0$.

## 3. ESTIMATION OF THE COUNTERFACTUAL WITH K-FOLD CROSS-FITTING

We subsequently propose an estimation strategy for the counterfactual $E[Y(\underline{d}_2)]$ with $\underline{d}_2 \in \{0, 1, \ldots, Q\}^2$ and show its $\sqrt{n}$-consistency under specific regularity conditions. Define

$$\psi^{\underline{d}_2}(W, \eta, \Psi_0^{\underline{d}_2}) = \frac{I\{D_1 = d_1\} \cdot I\{D_2 = d_2\} \cdot [Y_2 - \mu^{Y_2}(\underline{d}_2, \underline{X}_1)]}{p^{d_1}(X_0) \cdot p^{d_2}(d_1, \underline{X}_1)} + \frac{I\{D_1 = d_1\} \cdot [\mu^{Y_2}(\underline{d}_2, \underline{X}_1) - \nu^{Y_2}(\underline{d}_2, X_0)]}{p^{d_1}(X_0)} + \nu^{Y_2}(\underline{d}_2, X_0) - \Psi_0^{\underline{d}_2},$$

where $\mathcal{W} = \{W_i|1 \leq i \leq N\}$ with $W_i = (Y_{2i}, D_{1i}, D_{2i}, X_{0i}, X_{1i})$ for all $i$ denotes the set of observations and $I\{\cdot\}$ denotes the indicator function. The true nuisance parameters are denoted by $\eta_0 = (p_0^{d_1}(X_0), p_0^{d_2}(D_1, \underline{X}_1), \mu_0^{Y_2}(\underline{D}_2, \underline{X}_1), \nu_0^{Y_2}(\underline{D}_2, X_0))$, their estimates by

$\hat{\eta} = (\hat{p}^{d_1}(X_0), \hat{p}^{d_2}(D_1, \underline{X}_1), \hat{\mu}^{Y_2}(\underline{D}_2, \underline{X}_1), \hat{\nu}^{Y_2}(\underline{D}_2, X_0))$. Let $\Psi_0^{\underline{d}_2} = E[Y(\underline{d}_2)]$ denote the true counterfactual.[5]

We suggest estimating the $\Psi_0^{\underline{d}_2}$ using the following algorithm that combines orthogonal score estimation and sample splitting. Further below we will outline the conditions under which this estimation strategy leads to $\sqrt{n}$-consistent estimates for the counterfactual.

ALGORITHM 3.1: Estimation of $E[Y(\underline{d}_2)]$

(1) Split $\mathcal{W}$ in $K$ subsamples. For each subsample $k$, let $n_k$ denote its size, $\mathcal{W}_k$ the set of observations in the sample, and $\mathcal{W}_k^C$ the complement set of all observations not in $k$.
(2) For each $k$, use $\mathcal{W}_k^C$ to estimate the model parameters of $p^{d_1}(X_0)$ and $p^{d_2}(d_1, \underline{X}_1)$. Split $\mathcal{W}_k^C$ into two nonoverlapping subsamples and estimate the model parameters of the conditional mean $\mu^{Y_2}(\underline{d}_2, \underline{X}_1)$, and the nested conditional mean $\nu^{Y_2}(\underline{d}_2, X_0)$ in the distinct subsamples. Predict the models among $\mathcal{W}_k$, where the predictions are denoted by $\hat{p}_k^{d_1}(X_0)$, $\hat{p}_k^{d_2}(d_1, \underline{X}_1)$, $\hat{\mu}_k^{Y_2}(\underline{d}_2, \underline{X}_1)$, $\hat{\nu}_k^{Y_2}(\underline{d}_2, X_0)$.
(3) For each $k$, obtain an estimate of the moment condition for each observation $i$ in $\mathcal{W}_k$, denoted by $\hat{\psi}_{i,k}^{\underline{d}_2}$ :

$$\hat{\psi}_{i,k}^{\underline{d}_2} = \frac{I\{D_{1i} = d_1\} \cdot I\{D_{2i} = d_2\} \cdot [Y_{2i} - \hat{\mu}_k^{Y_2}(\underline{d}_2, \underline{X}_{1i})]}{\hat{p}_k^{d_1}(X_{0i}) \cdot \hat{p}_k^{d_2}(d_1, \underline{X}_{1i})} + \frac{I\{D_{1i} = d_1\} \cdot [\hat{\mu}_k^{Y_2}(\underline{d}_2, \underline{X}_{1i}) - \hat{\nu}_k^{Y_2}(\underline{d}_2, X_{0i})]}{\hat{p}_k^{d_1}(X_{0i})} + \hat{\nu}_k^{Y_2}(\underline{d}_2, X_{0i}).$$

(4) Average the estimated scores $\hat{\psi}_{i,k}^{\underline{d}_2}$ over all observations across all $K$ subsamples to obtain an estimate of $\Psi^{\underline{d}_2}$ in the total sample, denoted by $\hat{\Psi}^{\underline{d}_2} = 1/n \sum_{k=1}^{K} \sum_{i=1}^{n_k} \hat{\psi}_{i,k}^{\underline{d}_2}$.

[5] We note that the two-periods framework considered in this paper easily extends to more treatment periods; see the general formula for multiple treatment periods provided in Tran et al. (2019) (equation (9) in their section 4.1). For instance, in the case of three treatment periods and using analogous notation, the efficient score function takes the following form:

$$\begin{aligned}\psi^{\underline{d}_3}(W, \eta, \Psi_0^{\underline{d}_3}) = {} & \frac{I\{D_1 = d_1\} \cdot I\{D_2 = d_2\} \cdot I\{D_3 = d_3\} \cdot [Y_3 - \xi^{Y_3}(\underline{d}_3, \underline{X}_2)]}{p^{d_1}(X_0) \cdot p^{d_2}(d_1, \underline{X}_1) \cdot p^{d_3}(\underline{d}_2, \underline{X}_2)} \\ & + \frac{I\{D_1 = d_1\} \cdot I\{D_2 = d_2\} \cdot [\xi^{Y_3}(\underline{d}_3, \underline{X}_2) - \mu^{Y_3}(\underline{d}_3, \underline{X}_1)]}{p^{d_1}(X_0) \cdot p^{d_2}(d_1, \underline{X}_1)} \\ & + \frac{I\{D_1 = d_1\} \cdot [\mu^{Y_3}(\underline{d}_3, \underline{X}_1) - \nu^{Y_3}(\underline{d}_3, X_0)]}{p^{d_1}(X_0)} \\ & + \nu^{Y_3}(\underline{d}_3, X_0) - \Psi_0^{\underline{d}_3},\end{aligned}$$

where

$$\begin{aligned}\xi^{Y_3}(\underline{D}_3, \underline{X}_2) &= E[Y_3 | \underline{D}_3, \underline{X}_2], \\ \mu^{Y_3}(\underline{D}_3, \underline{X}_1) &= \int \xi^{Y_3}(\underline{D}_3, \underline{X}_1, X_2 = x_2) dF_{X_2 = x_2 | \underline{D}_2, \underline{X}_1}, \\ \nu^{Y_3}(\underline{D}_3, X_0) &= \int \mu^{Y_3}(\underline{D}_3, X_0, X_1 = x_1) dF_{X_1 = x_1 | D_1, X_0}.\end{aligned}$$

As a remark concerning step 2 of the algorithm, it may appear nonstandard to estimate $\mu^{Y_2}(\underline{d}_2, \underline{X}_1)$ and $\nu^{Y_2}(\underline{d}_2, X_0)$ in distinct subsamples. This approach aims at avoiding correlations between both estimation steps and, thus, overfitting bias, because the estimate of $\mu^{Y_2}(\underline{d}_2, \underline{X}_1)$ is used as a plug-in parameter for estimating $\nu^{Y_2}(\underline{d}_2, X_0)$.

In order to achieve $\sqrt{n}$-consistency for counterfactual estimation, we make the following assumption on the prediction quality of the machine learners when estimating the nuisance parameters. Closely following Chernozhukov et al. (2018), we introduce some further notation. Let $(\delta_n)_{n=1}^{\infty}$ and $(\Delta_n)_{n=1}^{\infty}$ denote sequences of positive constants with $\lim_{n\to\infty} \delta_n = 0$ and $\lim_{n\to\infty} \Delta_n = 0$, while $\delta_n \geq n^{-1/2}$. Furthermore, let $c, \epsilon, C$ and $q$ be positive constants such that $q > 4$, and let $K \geq 2$ be a fixed integer. Also, for any random vector $Z = (Z_1, \ldots, Z_l)$, let $\|Z\|_q = \max_{1\leq j\leq l} \|Z_j\|_q$, where $\|Z_l\|_q = \left(E\left[|Z_l|^q\right]\right)^{\frac{1}{q}}$. In order to ease notation, we assume that $n/K$ is an integer. For the sake of brevity we omit the dependence of probability $\Pr_P$, expectation $E_P(\cdot)$, and norm $\|\cdot\|_{P,q}$ on the probability measure $P$.

ASSUMPTION 3.1 (REGULARITY CONDITIONS AND QUALITY OF PLUG-IN PARAMETER ESTIMATES): *For all probability laws $P \in \mathcal{P}$ the following conditions hold for the random vector $(Y_2, D_1, D_2, X_0, X_1)$ for all $d_1, d_2 \in \{0, 1, \ldots, Q\}$:*

*(a)* $\|Y_2\|_q \leq C$,
$\left\|E[Y_2^2|D_1 = d_1, D_2 = d_2, \underline{X}_1]\right\|_{\infty} \leq C^2$,

*(b)* $\Pr(\epsilon \leq p_0^{d_1}(X_0) \leq 1 - \epsilon) = 1$,
$\Pr(\epsilon \leq p_0^{d_2}(d_1, \underline{X}_1) \leq 1 - \epsilon) = 1$,

*(c)* $\left\|Y_2 - \mu_0^{Y_2}(\underline{d}_2, \underline{X}_1)\right\|_2 = E\left[\left(Y_2 - \mu_0^{Y_2}(\underline{d}_2, \underline{X}_1)\right)^2\right]^{\frac{1}{2}} \geq c$.

*(d) Given a random subset $I$ of $[n]$ of size $n_k = n/K$, the nuisance parameter estimator $\hat{\eta} = \hat{\eta}((W_i)_{i\in I^c})$ satisfies the following conditions. With $P$-probability no less than $1 - \Delta_n$:*

$$\begin{aligned}
\|\hat{\eta} - \eta_0\|_q &\leq C,\\
\|\hat{\eta} - \eta_0\|_2 &\leq \delta_n,\\
\left\|\hat{p}^{d_1}(X_0) - 1/2\right\|_{\infty} &\leq 1/2 - \epsilon,\\
\left\|\hat{p}^{d_2}(D_1, \underline{X}_1) - 1/2\right\|_{\infty} &\leq 1/2 - \epsilon,\\
\left\|\hat{\mu}^{Y_2}(\underline{D}_2, \underline{X}_1) - \mu_0^{Y_2}(\underline{D}_2, \underline{X}_1)\right\|_2 \times \left\|\hat{p}^{d_1}(X_0) - p_0^{d_1}(X_0)\right\|_2 &\leq \delta_n n^{-1/2},\\
\left\|\hat{\mu}^{Y_2}(\underline{D}_2, \underline{X}_1) - \mu_0^{Y_2}(\underline{D}_2, \underline{X}_1)\right\|_2 \times \left\|\hat{p}^{d_2}(D_1, \underline{X}_1) - p_0^{d_2}(D_1, \underline{X}_1)\right\|_2 &\leq \delta_n n^{-1/2},\\
\left\|\hat{\nu}^{Y_2}(\underline{D}_2, X_0) - \nu_0^{Y_2}(\underline{D}_2, X_0)\right\|_2 \times \left\|\hat{p}^{d_1}(X_0) - p_0^{d_1}(X_0)\right\|_2 &\leq \delta_n n^{-1/2}.
\end{aligned}$$

The only nonprimitive condition is the condition (d). It puts restrictions on the quality of the nuisance parameter estimators. Condition (a) states that the distribution of the outcome does not have unbounded moments. (b) refines the common support condition such that the propensity scores are bounded away from 0 and 1. Finally, (c) states that the covariates $\underline{X}_1$ do not perfectly predict the conditional mean outcome.

For demonstrating the $\sqrt{n}$-consistency of our estimator of the mean potential outcome, we show that it satisfies the requirements of the DML framework in Chernozhukov et al. (2018) by

first verifying linearity and Neyman orthogonality of the score (see Appendix S1.1). Then, as $\psi^{\underline{d}_2}(W, \eta, \Psi_0^{\underline{d}_2})$ is smooth in $(\eta, \Psi_0^{\underline{d}_2})$, it is sufficient that the plug-in estimators converge with a rate faster than $n^{-1/4}$ for achieving $n^{-1/2}$-convergence for the estimation of $\hat{\Psi}^{\underline{d}_2}$ as postulated in Theorem 3.1. This convergence rate of $n^{-1/4}$ has been shown to be achieved by many commonly used machine learners under specific conditions, such as lasso, random forests, boosting and neural nets, see for instance Belloni et al. (2014), Luo and Spindler (2016), Wager and Athey (2018), Farrell et al. (2021), Syrgkanis and Zampetakis (2020), and Singh (2021).

THEOREM 3.1. *Under Assumptions 2.1–3.1, it holds for estimating* $E[Y(\underline{d}_2)]$ *based on Algorithm 3.1:*

$\sqrt{n}\left(\hat{\Psi}^{\underline{d}_2} - \Psi_0^{\underline{d}_2}\right) \to N(0, \sigma_{\psi^{\underline{d}_2}})$, *where* $\sigma_{\psi^{\underline{d}_2}} = E[(\psi^{\underline{d}_2} - \Psi_0^{\underline{d}_2})^2]$. *Moreover, the asymptotic variance* $\sigma_{\psi^{\underline{d}_2}}$ *may be consistently estimated by:* $\hat{\sigma}_{\psi^{\underline{d}_2}} = 1/n \sum_{k=1}^{K} \sum_{i=1}^{n_k} (\hat{\psi}_{i,k}^{\underline{d}_2} - \hat{\Psi}^{\underline{d}_2})^2$.

The proof of Theorem 3.1 is provided in Appendix S1.1.

From Theorem 3.1, which demonstrates the $\sqrt{n}$-consistent estimation of the mean potential outcomes, it follows that the ATE can be $\sqrt{n}$-consistently estimated, too, as it is defined as a linear combination (and more specifically, as the difference) of the potential outcomes: $\hat{\Delta}(\underline{d}_2, \underline{d}_2^*) = \hat{\Psi}^{\underline{d}_2} - \hat{\Psi}^{\underline{d}_2^*}$. Accordingly, the estimator of the asymptotic variance of the ATE corresponds to $\hat{\sigma}_{\Delta(\underline{d}_2, \underline{d}_2^*)} = 1/n \sum_{k=1}^{K} \sum_{i=1}^{n_k} \left(\hat{\psi}_{i,k}^{\underline{d}_2^*} - \hat{\psi}_{i,k}^{\underline{d}_2} - (\hat{\Psi}^{\underline{d}_2} - \hat{\Psi}^{\underline{d}_2^*})\right)^2$.

## 4. EVALUATION OF WEIGHTED DYNAMIC TREATMENT EFFECTS

Lechner and Miquel (2010) show that, under our assumptions, one may identify treatment effects for specific subgroups that are defined as a function of the distribution of the baseline covariates $X_0$. To this end, let $S$ denote a binary indicator for belonging to the subgroup of interest that satisfies $S \perp Y_2(\underline{d}_2)|X_0$, as $S$ may be selective in $X_0$, but not with respect to the post-treatment covariates $X_1$ after controlling for $X_0$. Furthermore, denote by $g(X_0) = \Pr(S = 1|X_0)$ the probability of being in that group conditional on $X_0$. Interesting examples for such subgroup are the treated or nontreated populations in the first period, obtained by defining $S = I\{D_1 = d_1\}$ with $d_1 \in \{0, 1, \ldots, Q\}$, in order to assess whether the treatment effect varies across treatment states. We note that $S$ might even be a deterministic function of $X_0$, such that $g(X_0)$ equals one for specific values of the baseline covariates and zero otherwise. This permits assessing the conditional average treatment effect (CATE) in a subgroup defined upon values of $X_0$ as well as effect heterogeneity across subgroups. For instance, we may set $S = 1$ for females and $S = 0$ otherwise in order to evaluate the CATE conditional on gender.

We can identify the mean potential outcomes conditional on $S = 1$ based on reweighting by $g(X_0)$, see, e.g., Hirano et al. (2003), who use this approach for weighted ATE evaluation based on IPW. That is,

$$E[Y_2(\underline{d}_2)|S = 1] = E\left[\frac{S \cdot Y_2(\underline{d}_2)}{\Pr(S = 1)}\right] = E\left[\frac{g(X_0)}{\Pr(S = 1)} \cdot E[Y_2(\underline{d}_2)|X_0]\right]$$
$$= E\left[\frac{S}{\Pr(S = 1)} \cdot E[Y_2(\underline{d}_2)|X_0]\right], \tag{4.1}$$

where the first equality follows from basic probability theory, and the remaining ones from the fact that $S \perp Y_2(\underline{d}_2)|X_0$ and the law of iterated expectations. This suggests the following identification

approach:

$$
\begin{aligned}
E[Y_2(\underline{d}_2)|S=1] &= E[\psi^{\underline{d}_2,S=1}], \text{ where} \\
\psi^{\underline{d}_2,S=1} &= \frac{g(X_0)}{\Pr(S=1)} \cdot \frac{I\{D_1=d_1\} \cdot I\{D_2=d_2\} \cdot [Y_2 - \mu^{Y_2}(\underline{d}_2, \underline{X}_1)]}{p^{d_1}(X_0) \cdot p^{d_2}(d_1, \underline{X}_1)} \\
&\quad + \frac{g(X_0)}{\Pr(S=1)} \cdot \frac{I\{D_1=d_1\} \cdot [\mu^{Y_2}(\underline{d}_2, \underline{X}_1) - \nu^{Y_2}(\underline{d}_2, X_0)]}{p^{d_1}(X_0)} \\
&\quad + \frac{S}{\Pr(S=1)} \cdot \nu^{Y_2}(\underline{d}_2, X_0). \qquad (4.2)
\end{aligned}
$$

Note that the term $\frac{S}{\Pr(S=1)} \cdot \nu^{Y_2}(\underline{d}_2, X_0)$ in (4.2) corresponds to $\frac{S}{\Pr(S=1)} \cdot E[Y_2(\underline{d}_2)|X_0]$ in (4.1). Appendix S1.2 shows that the moment condition $E[\psi^{\underline{d}_2,S=1} - E[Y_2(\underline{d}_2)|S=1]] = 0$ holds, such that $E[\psi^{\underline{d}_2,S=1}]$ identifies the weighted mean potential outcome, and proves Neyman orthogonality. It demonstrates that DML is $\sqrt{n}$-consistent and asymptotically normal under Assumption 4.1 below. The latter formalises the rate restrictions on the plug-in estimates, which now also contain an estimate of $g(X_0)$ denoted by $\hat{g}(X_0)$. To this end, Algorithm 3.1 outlined in Section 3 is applied to estimate $E[Y_2(\underline{d}_2)|S=1]$ by using modified moment conditions in steps 3 and 4.

More specifically, the previously used $\hat{\psi}_{i,k}^{\underline{d}_2}$ computed in some subsample $k$ is replaced by

$$
\begin{aligned}
\hat{\psi}_{i,k}^{\underline{d}_2,S=1} &= \hat{g}_k(X_{0i}) \cdot \frac{I\{D_{1i}=d_1\} \cdot I\{D_{2i}=d_2\} \cdot [Y_{2i} - \hat{\mu}_k^{Y_2}(\underline{d}_2, \underline{X}_{1i})]}{\hat{p}_k^{d_1}(X_{0i}) \cdot \hat{p}_k^{d_2}(d_1, \underline{X}_{1i})} \\
&\quad + \hat{g}_k(X_{0i}) \cdot \frac{I\{D_{1i}=d_1\} \cdot [\hat{\mu}_k^{Y_2}(\underline{d}_2, \underline{X}_{1i}) - \hat{\nu}_k^{Y_2}(\underline{d}_2, X_{0i})]}{\hat{p}_k^{d_1}(X_{0i})} + S_i \cdot \hat{\nu}_k^{Y_2}(\underline{d}_2, X_{0i}). \qquad (4.3)
\end{aligned}
$$

In step 4, the estimated scores $\hat{\psi}_{i,k}^{\underline{d}_2,S=1}$ are averaged over all observations across all $K$ subsamples and divided by an estimate of $\Pr(S=1)$ to obtain an estimate of $\Psi_0^{\underline{d}_2,S=1} = E[Y_2(\underline{d}_2)|S=1]$ based on $\hat{\Psi}^{\underline{d}_2,S=1} = \left[\sum_{k=1}^{K}\sum_{i=1}^{n_k} \hat{\psi}_{i,k}^{\underline{d}_2}\right] \Big/ \left[\sum_{k=1}^{K}\sum_{i=1}^{n_k} S_i\right]$.

The following assumption refines the conditions of Assumption 3.1 such that asymptotic normality holds for the DML estimator based on (4.2).

We redefine the vector of nuisance parameters such that it includes the $g$ function. To this end, we denote by $\chi_0 = (g_0(X_0), \eta_0)$ the true nuisance parameters and by $\hat{\chi} = (\hat{g}_0(X_0), \hat{\eta})$ their estimates.

ASSUMPTION 4.1 (REGULARITY CONDITIONS AND QUALITY OF PLUG-IN PARAMETER ESTIMATES): *For all probability laws $P \in \mathcal{P}$ the following conditions hold for the random vector $(Y_2, D_1, D_2, X_0, X_1, S)$ for all $d_1, d_2 \in \{0, 1, \ldots, Q\}$:*

*(a)* $\|Y_2\|_q \leq C$,
$\left\|E[Y_2^2|D_1=d_1, D_2=d_2, \underline{X}_1]\right\|_\infty \leq C^2$,

*(b)* $\Pr(\epsilon \leq p_0^{d_1}(X_0) \leq 1-\epsilon) = 1$
$\Pr(\epsilon \leq p_0^{d_2}(d_1, \underline{X}_1) \leq 1-\epsilon) = 1$,

(c) $\left\| Y_2 - \mu_0^{Y_2}(\underline{d}_2, \underline{X}_1) \right\|_2 = E\Big[ \Big( Y_2 - \mu_0^{Y_2}(\underline{d}_2, \underline{X}_1) \Big)^2 \Big]^{\frac{1}{2}} \geq c.$

(d) *Given a random subset I of* $[n]$ *of size* $n_k = N/K$, *the nuisance parameter estimator* $\hat{\chi} = \hat{\chi}((W_i)_{i \in I^C})$ *satisfies the following conditions. With P-probability no less than* $1 - \Delta_n$:

$$
\begin{aligned}
\|\hat{\chi} - \chi_0\|_q &\leq C, \\
\|\hat{\chi} - \chi_0\|_2 &\leq \delta_n, \\
\left\| \hat{p}^{d_1}(X_0) - 1/2 \right\|_\infty &\leq 1/2 - \epsilon, \\
\left\| \hat{p}^{d_2}(D_1, \underline{X}_1) - 1/2 \right\|_\infty &\leq 1/2 - \epsilon, \\
\left\| \hat{\mu}^{Y_2}(\underline{D}_2, \underline{X}_1) - \mu_0^{Y_2}(\underline{D}_2, \underline{X}_1) \right\|_2 \times \left\| \hat{p}^{d_1}(X_0) - p_0^{d_1}(X_0) \right\|_2 &\leq \delta_n n^{-1/2}, \\
\left\| \hat{\mu}^{Y_2}(\underline{D}_2, \underline{X}_1) - \mu_0^{Y_2}(\underline{D}_2, \underline{X}_1) \right\|_2 \times \left\| \hat{p}^{d_2}(D_1, \underline{X}_1) - p_0^{d_2}(D_1, \underline{X}_1) \right\|_2 &\leq \delta_n n^{-1/2}, \\
\left\| \hat{\nu}^{Y_2}(\underline{D}_2, X_0) - \nu_0^{Y_2}(\underline{D}_2, X_0) \right\|_2 \times \left\| \hat{p}^{d_1}(X_0) - p_0^{d_1}(X_0) \right\|_2 &\leq \delta_n n^{-1/2}. \\
\left\| \hat{\mu}^{Y_2}(\underline{D}_2, \underline{X}_1) - \mu_0^{Y_2}(\underline{D}_2, \underline{X}_1) \right\|_2 \times \|\hat{g}(X_0) - g_0(X_0)\|_2 &\leq \delta_n n^{-1/2}, \\
\left\| \hat{\nu}^{Y_2}(\underline{D}_2, X_0) - \nu_0^{Y_2}(\underline{D}_2, X_0) \right\|_2 \times \|\hat{g}(X_0) - g_0(X_0)\|_2 &\leq \delta_n n^{-1/2}.
\end{aligned}
$$

Assumption 4.1 can be satisfied if the plug-in estimator $\hat{g}(X_0)$ converges to its true value $g_0(X_0)$ with rate faster than $n^{-1/4}$ just like the estimators of the other nuisance terms. Then, the ATE in the subgroup, denoted by

$$\Delta(\underline{d}_2, \underline{d}_2^*, S = 1) = E[Y_2(\underline{d}_2) - Y_2(\underline{d}_2^*) | S = 1],$$

is $\sqrt{n}$-consistently estimated, as postulated in Theorem 4.1.

THEOREM 4.1. *Under Assumptions 2.1–2.3 and 4.1, it holds for estimating* $E[Y_2(\underline{d}_2)|S = 1]$ *based on Algorithm 3.1:*

$\sqrt{n}\Big(\hat{\Psi}^{\underline{d}_2, S=1} - \Psi_0^{\underline{d}_2, S=1}\Big) \to N(0, \sigma_{\psi^{\underline{d}_2, S=1}})$, *where* $\sigma_{\psi^{\underline{d}_2, S=1}} = E[(\psi^{\underline{d}_2, S=1} - \Psi_0^{\underline{d}_2, S=1})^2]$. *Moreover, the asymptotic variance* $\sigma_{\psi^{\underline{d}_2, S=1}}$ *may be consistently estimated by:*

$\hat{\sigma}_{\psi^{\underline{d}_2, S=1}} = 1/n \sum_{k=1}^K \sum_{i=1}^{n_k} (\hat{\psi}_{i,k}^{\underline{d}_2, S=1} - \hat{\Psi}^{\underline{d}_2, S=1})^2$.

The proof of Theorem 4.1 is provided in Appendix S1.2.

In analogy to the discussion after Theorem 3.1, it follows from Theorem 4.1 that, not only the mean potential outcomes, but also the ATE (i.e., the difference in mean potential outcomes) conditional on $S = 1$ can be estimated $\sqrt{n}$-consistently: $\hat{\Delta}(\underline{d}_2, \underline{d}_2^*, S = 1) = \hat{\Psi}^{\underline{d}_2, S=1} - \hat{\Psi}^{\underline{d}_2^*, S=1}$. The corresponding estimator of the asymptotic variance is given by $\hat{\sigma}_{\Delta(\underline{d}_2, \underline{d}_2^*, S=1)} = 1/n \sum_{k=1}^K \sum_{i=1}^{n_k} \left( \hat{\psi}_{i,k}^{\underline{d}_2, S=1} - \hat{\psi}_{i,k}^{\underline{d}_2^*, S=1} - (\hat{\Psi}^{\underline{d}_2, S=1} - \hat{\Psi}^{\underline{d}_2^*, S=1}) \right)^2$.

**Table 1.** Confounding based on $\beta_{X_0} = \beta_{X_1} = 0.4/i^4$.

| Number of covariates | Sample size | Pseudo-$R^2$ (%) $\hat{p}^{d_1}(X_0)$ | pseudo-$R^2$ (%) $\hat{p}^{d_2}(d_1, \underline{X}_1)$ | $R^2$ (%) $\hat{\mu}^{Y_2}(\underline{X}_1)$ |
|---|---|---|---|---|
| 50 | 2,500 | 15 | 29 | 38 |
| 50 | 10,000 | 13 | 26 | 36 |
| 100 | 2,500 | 17 | 33 | 41 |
| 100 | 10,000 | 14 | 27 | 37 |
| 500 | 2,500 | 37 | 71 | 62 |
| 500 | 10,000 | 19 | 36 | 42 |

## 5. SIMULATION STUDY

This section provides a simulation study to investigate the finite sample behaviour of our DML method for dynamic treatment effects based on the following data generating process:

$$
\begin{aligned}
Y_2 &= D_1 + D_2 + X_0'\beta_{X_0} + X_1'\beta_{X_1} + U,\\
D_1 &= I\{X_0'\beta_{X_0} + V > 0\},\\
D_2 &= I\{0.3D_1 + X_0'\beta_{X_0} + X_1'\beta_{X_1} + W > 0\},\\
X_0 &\sim N(0, \Sigma_0), \quad X_1 \sim N(0, \Sigma_1),\\
U, V, W &\sim N(0, 1), \text{ independently of each other.}
\end{aligned}
$$

Outcome $Y_2$ is a function of the observed variables $D_1$, $D_2$, $X_0$, $X_1$, and the unobserved scalar $U$. The treatment effects of both $D_1$ and $D_2$ are equal to 1. $D_1$ is a function of $X_0$ and the unobserved scalar $V$. $D_2$ is a function of both pre- and post-treatment covariates $X_0$ and $X_1$, the first treatment $D_1$, and the unobservable scalar $W$. Both $X_0$ and $X_1$ are vectors of covariates of dimension $p$, drawn from a multivariate normal distribution with zero mean and covariance matrices $\Sigma_0$ and $\Sigma_1$, respectively. $U$, $V$, $W$ are random and standard normally distributed. We consider two sample sizes of $n = 2{,}500$ and 10,000, running 1,000 simulations for the smaller and 250 simulations for the larger sample sizes.

In our simulations, we set $p$, the number of covariates in $X_1$ and $X_0$, respectively, to 50, 100, or 500. $\Sigma_0$ and $\Sigma_1$ are defined based on setting the covariance of the $i$th and $j$th covariate in $X_0$ or $X_1$ to $0.5^{|i-j|}$. The coefficients $\beta_{X_0}$ and $\beta_{X_1}$ gauge the impacts of the covariates on $Y_2$, $D_2$, and $D_1$, respectively, and thus the magnitude of confounding. The $i$th element in the coefficient vectors $\beta_{X_0}$ and $\beta_{X_1}$ is set to $0.4/i^4$ for $i = 1, \ldots, p$, implying a quadratic decay of covariate importance in terms of confounding. As reported in Table 1, this specification implies that the $R^2$ statistic based on linearly predicting $Y_2$ by $\underline{X}_1$ ranges from 36 to 62%, depending on the number of covariates and the sample size. Furthermore, the Nagelkerke (1991) pseudo-$R^2$ when predicting $D_1$ by $X_0$ and $D_2$ by $D_1, \underline{X}_1$ based on probit models ranges from 13 to 37% and 26 to 71%, respectively. These figures point to a substantial level of confounding as it may be reasonably encountered in empirical applications.

We investigate the performance of ATE estimation when comparing the sequences of obtaining both treatments ($\underline{d}_2 = (d_1 = 1, d_2 = 1)$) vs. no treatment ($\underline{d}_2^* = (d_1 = 0, d_2 = 0)$) in the total population based on Theorem 3.1 and in the treated in the first period based on Theorem 4.1. The nuisance parameters, i.e., the linear and probit specifications of the outcome and treatment

**Table 2.** Simulation results based on $\beta_{X_0} = \beta_{X_1} = 0.4/i^4$.

| Covariates | Sample size | True effect | Absolute bias | Standard deviation | Average SE | RMSE | Coverage in % |
|---|---|---|---|---|---|---|---|
| | | | | ATE: $\hat{\Delta}(\underline{d}_2, \underline{d}_2^*)$ (all) | | | |
| 50 | 2,500 | 2 | 0.027 | 0.07 | 0.069 | 0.075 | 91.6 |
| 50 | 10,000 | 2 | 0.007 | 0.035 | 0.034 | 0.036 | 94.4 |
| 100 | 2,500 | 2 | 0.04 | 0.072 | 0.069 | 0.083 | 88.7 |
| 100 | 10,000 | 2 | 0.011 | 0.035 | 0.034 | 0.037 | 94.4 |
| 500 | 2,500 | 2 | 0.063 | 0.07 | 0.068 | 0.094 | 83.4 |
| 500 | 10,000 | 2 | 0.02 | 0.035 | 0.034 | 0.04 | 89.6 |
| | | | | ATE on selected: $\hat{\Delta}(\underline{d}_2, \underline{d}_2^*, S = 1)$ | | | |
| 50 | 2,500 | 2 | 0.027 | 0.076 | 0.087 | 0.081 | 96.5 |
| 50 | 10,000 | 2 | 0.006 | 0.037 | 0.043 | 0.038 | 95.6 |
| 100 | 2,500 | 2 | 0.042 | 0.079 | 0.087 | 0.089 | 94.0 |
| 100 | 10,000 | 2 | 0.011 | 0.037 | 0.043 | 0.039 | 96.4 |
| 500 | 2,500 | 2 | 0.064 | 0.075 | 0.088 | 0.099 | 91.5 |
| 500 | 10,000 | 2 | 0.019 | 0.038 | 0.043 | 0.043 | 95.2 |

*Notes:* SE and RMSE denote the standard error and the root mean squared error, respectively. Coverage is based on 95% confidence intervals.

equations, are estimated by lasso regressions using the default options of the *SuperLearner* package provided by van der Laan et al. (2007) for the statistical software R.Three-fold cross-fitting is used for the estimation of the treatment effects. We drop observations whose products of estimated treatment propensity scores in the first and second period, $\hat{p}^{d_1}(X_0) \cdot \hat{p}^{d_2}(D_1, \underline{X}_1)$, are close to zero, namely smaller than a trimming threshold of 0.01 (or 1%). This avoids an explosion of the propensity score-based weights, and thus of the variance when estimating the mean potential outcomes by the sample analogue of identification result (4.3), where the product of the propensity scores enters the denominator for reweighing the outcome. Our estimation procedure is available in the *causalweight* package for R by Bodory and Huber (2018).

Table 2 presents the main findings when estimating the ATE in the total population, $\hat{\Delta}(\underline{d}_2, \underline{d}_2^*)$, and among the subgroup of treated in the first period, $\hat{\Delta}(\underline{d}_2, \underline{d}_2^*, S = 1)$. Irrespective of the number of covariates, the absolute biases go to zero as the sample size increases. Furthermore, the standard deviations and root mean squared errors (RMSE) of the ATE estimators are roughly cut by half when quadrupling the sample size, as implied by $\sqrt{n}$-consistency. The levels of the standard deviations and RMSEs are somewhat higher for $\hat{\Delta}(\underline{d}_2, \underline{d}_2^*, S = 1)$ than for $\hat{\Delta}(\underline{d}_2, \underline{d}_2^*)$, which comes from the additional weighting step due to targeting the treated subpopulation with $S = 1$. We also observe that the average standard errors (average SE) based on the asymptotic variance approximations appear to converge at $\sqrt{n}$-rate; however, among the subgroup of treated, they slightly overestimate the true standard deviations. In general, the coverage rates based on 95% confidence intervals approach the nominal value with increasing sample size.

## 6. EMPIRICAL APPLICATION

We apply our DML approach to evaluate the effects of training sequences provided by the Job Corps programme on employment. Job Corps is the largest US programme offering vocational training and academic classroom instruction for disadvantaged individuals aged 16 to 24. It is financed by the US Department of Labor and currently has about 50,000 participants every year. Besides vocational credentials, students may obtain a high school diploma or equivalent qualifications. Individuals meeting specific low-income requirements can participate in Jobs Corps without any costs.

A range of studies analyses the impact of Job Corps based on an experimental study with randomised access to the programme between November 1994 and February 1996. In particular, Schochet et al. (2001) and Schochet et al. (2008) discuss in detail the study design and report the average effects of random programme assignment on a broad range of outcomes. Their findings suggest that Job Corps increases educational attainment, reduces criminal activity, and increases employment and earnings, at least for some years after the programme. Flores et al. (2012) assess the impact of a continuously defined treatment, namely the length of exposure to academic and vocational instruction on earnings, and find positive effects. As the length of the treatment is (in contrast to programme assignment) not random, they impose a conditional independence assumption and control for baseline characteristics at Job Corps assignment. Colangelo and Lee (2021) suggest DML-based estimation of continuous treatment effects and apply it to assess the employment effects of Job Corps. In contrast to these contributions on continuous treatment doses of Job Corps, we consider discrete sequences of multiple treatments and also control for post-treatment confounders rather than baseline covariates only.

Several contributions assess specific causal mechanisms of the programme. Flores and Flores-Lagunes (2009) find a positive direct effect of programme assignment on earnings when controlling for work experience, which they assume to be conditionally independent given observed covariates. Also Huber (2014) imposes a conditional independence assumption and estimates a positive direct health effect when controlling for the mediator employment. Using a partial identification approach permitting mediator endogeneity, Flores and Flores-Lagunes (2010) compute bounds on the direct and indirect effects of Job Corps assignment on employment and earnings mediated by obtaining General Educational Development (GED) certificate, high school degree, or vocational degree. Under their strongest set of assumptions, the results point to a positive direct effect net of obtaining a degree. Frölich and Huber (2017) use an instrumental variable strategy based on two instruments to disentangle the earnings effect of being enrolled in Job Corps into an indirect effect, via hours worked, and a direct effect, likely related to a change in human capital. Their results point to the existence of an indirect rather than a direct mechanism. Even though our framework of analysing sequences of treatments is in terms of statistical issues somewhat related to the evaluation of causal mechanisms, it relies on distinct identifying assumptions more than the previously mentioned studies, which, e.g., do not consider controlling for post-treatment confounders.

For the empirical analysis, we use a sample based on the data source of Schochet et al. (2019). The newly constructed dataset is provided by Huber (2022). Our sample consists of 11,313 individuals with completed follow-up interviews four years after randomisation, out of which 6,828 and 4,485 were randomised in and out of Job Corps, respectively. We exploit the sequential structure of academic education and vocational training in the programme to define dynamic treatment states. Since most of the education and training activities were taken in the first two years, we focus on the latter when generating a sequence of binary treatments for each observation.

**Table 3.** Sequences of treatments.

| Dynamic treatments | | | in Job Corps | Observations |
|---|---|---|---|---|
| Code | Year 1 | Year 2 | | |
| 00 | no educ/train | no educ/train | No | 4,485 |
| 11 | no educ/train | no educ/train | Yes | 320 |
| 12 | no educ/train | acad educ | Yes | 43 |
| 13 | no educ/train | voc train | Yes | 42 |
| 21 | acad educ | no educ/train | Yes | 1,328 |
| 22 | acad educ | acad educ | Yes | 341 |
| 23 | acad educ | voc train | Yes | 183 |
| 31 | voc train | no educ/train | Yes | 1,279 |
| 32 | voc train | acad educ | Yes | 109 |
| 33 | voc train | voc train | Yes | 573 |
| Missings | | | | 2,610 |

*Notes: no educ/train* means not participating in any Job Corps programme related to education or training measures. *acad educ* and *voc train* stand for academic education and vocational training, respectively, offered by Job Corps.

The treatment states in our application can take four different values: $d_1, d_2, d_1^*, d_2^* \in \{0, 1, 2, 3\}$. State 0 refers to no instruction offered due to being randomised out of Job Corps (control group), state 1 to no instruction despite being randomised in (never takers in the denomination of Angrist et al., 1996), 2 to academic education among programme participants, and 3 to vocational training among programme participants. If individuals participate in both academic education and vocational training in a specific year, we assign the code of the treatment that was attended to a larger extent in terms of completed hours.

Table 3 reports various sequences of treatments in the data along with the corresponding number of observations. For instance, the treatment sequence 00 refers to those 4,485 control group members that were randomised out, and did not participate in any education activities offered by Job Corps. Furthermore, 320 individuals assigned to Job Corps do not participate in any form of education either, as indicated by the sequence 11. We also note that for 2,610 out of the 11,313 individuals, information on the treatment sequences is missing. The literature explains the missing values by a random skip logic error, due to which asking questions about treatment participation was randomly omitted for a subset of survey participants, see page J.5 in Schochet et al. (2003). In our analysis we drop the control group with treatment sequence 00, but make use of it in a placebo test outlined further below. Furthermore, for several potential comparisons of treatment sequences, small sample issues and/or problems of a lack of common support in propensity scores (and thus, covariates) arise. For this reason, we confine our evaluation to comparing treatment sequence 33 (vocational training in both years) to either 22 (academic education in both years), 21 (academic education in the first year), or 11 (no participation in either year).

Our outcome variable is a binary employment indicator measured four years after randomisation. Table 4 reports the mean outcome across various treatment sequences, which ranges from 77 to 89%. It also provides the sequence-specific numbers of cases with missing outcomes that are dropped from the analysis, which appear quite low. We aim at estimating the ATE of treatment sequences $\underline{d}_2$ vs. $\underline{d}_2^*$ among individuals whose treatment in the first year corresponds to the first-year-treatment of either $\underline{d}_2$ or $\underline{d}_2^*$. An alternative would be to assess the ATE in the total sample randomised into Job Corps (which would thus also include individuals with different first-year

**Table 4.** Mean outcome conditional on treatment sequence.

| Treatment code | Employment | |
|---|---|---|
| | Mean | Missings |
| 00 | 0.78 | 46 |
| 11 | 0.78 | 7 |
| 21 | 0.78 | 17 |
| 22 | 0.82 | 1 |
| 33 | 0.89 | 3 |
| Missings | 0.77 | 37 |

*Notes:* The first column provides the codes of the treatment sequences, see Table 3; the second column gives the average employment per sequence; the third one gives the number of missing observations.

**Table 5.** Number of covariates.

| Raw variables | $X_0$ | $X_1$ |
|---|---|---|
| Dummy | 295 | 575 |
| Categorical | 53 | 13 |
| Numeric | 26 | 226 |
| Total | 374 | 814 |
| Processed variables | $X_0$ | $X_1$ |
| Dummy | 883 | 1,201 |
| Numeric | 26 | 226 |
| Total | 909 | 1,427 |

*Notes:* $X_0$ and $X_1$ denote regressors measured prior to the first and second periods, respectively.

treatments from the ones evaluated), but this proved to be problematic due to lacking common support in terms of treatment propensity scores.

We make use of a large set of potential control variables that also include covariates, which have been identified as important confounders in several articles assessing the sensitivity of programme evaluations to the inclusion and omission of such confounders in observational labour market studies. Biewen et al. (2014), for instance, conclude that imposing conditional independence assumptions requires the availability of rich data on employment and benefit histories, and socioeconomic characteristics. Lechner and Wunsch (2013) point to the importance of factors such as health, caseworker assessments, regional information, timing of unemployment and programme start, pre-treatment outcomes, job search behaviour, and labour market histories. In line with these findings, our covariates comprise information about socioeconomic characteristics, pre-treatment labour market histories, education and training, job search activities, welfare receipt, health, crime, and how one learnt about the existence of Job Corps. Table S1 in Appendix S2 reports more details on these features, including variable descriptions and distributions across treatment sequences.

We condition on observed characteristics $X_t$ in periods $t \in \{0, 1\}$. $X_0$ denotes control variables measured at baseline prior to the first treatment $D_1$, whereas $X_1$ is observed one year after randomisation, but prior to the second treatment $D_2$. Table 5 provides the number and types of variables assigned to $X_0$ and $X_1$. Our raw data include 1,188 characteristics. After some data manipulations based on generating dummies for values of categorical variables and missing items in

**Table 6.** Effect estimates with a trimming threshold of 0.01.

| $\underline{d}_2$ | $\underline{d}_2^*$ | $\hat{E}[Y_2(\underline{d}_2^*)\|S=1]$ | $\hat{\Delta}(\underline{d}_2, \underline{d}_2^*, S=1)$ | SE | $p$-value | Observations | Trimmed |
|---|---|---|---|---|---|---|---|
| 33 | 22 | 0.76 | 0.1 | 0.06 | 0.11 | 3,783 | 507 |
| 33 | 21 | 0.82 | 0.05 | 0.03 | 0.07 | 3,783 | 43 |
| 33 | 11 | 0.81 | 0.08 | 0.03 | 0.02 | 2,346 | 22 |

*Notes:* $\underline{d}_2$ and $\underline{d}_2^*$ indicate the treatment sequences under treatment and nontreatment, respectively. $\hat{E}[Y_2(\underline{d}_2^*)|S=1]$ denotes the mean potential outcome under nontreatment conditional on $S=1$, where $S$ is an indicator for the first treatment corresponding to either the first treatment in $\underline{d}_2$ or $\underline{d}_2^*$. $\hat{\Delta}(\underline{d}_2,\underline{d}_2^*, S=1)$ provides the ATE estimate, SE the standard error. The last column gives the number of observations dropped according to the trimming rule $p^{d_1}(X_0) \cdot p^{d_2}(D_1,\underline{X}_1) < 0.01$.

dummy or categorical variables, we end up with, all in all, 2,336 regressors. Missing observations in numerical variables were replaced by the mean values of the nonmissing items. Furthermore, we standardised numerical covariates to have a zero mean and a standard deviation of 0.5.

We estimate $\Delta(\underline{d}_2, \underline{d}_2^*, S=1)$, with $S=1$ if the first treatment corresponds to either the first treatment in $\underline{d}_2$ or $\underline{d}_2^*$, based on 3-fold cross-fitting, and the random forest (see Breiman, 2001) as machine learner of the nuisance parameters. To this end, we use the *SuperLearner* package with default options provided by van der Laan et al. (2007) for the statistical software R. Our motivation for choosing the random forest is that it is (in the spirit of kernel regression) a nonparametric estimator that does not impose functional form assumptions (like linearity) on the conditional outcome or treatment models. As in our simulation study, we drop observations whose products of propensity scores in the first and second period are smaller than 0.01 to impose common support in our sample, and avoid an explosion in the propensity score-based weights. For a visual assessment of the common support, Appendix S3 provides plots with the propensity score distributions across all treatment sequences considered in this application. In general, common support is rather decent for the first period propensity scores $\hat{p}^{d_1}(X_0)$, while the overlap is weaker for the scores in the second period $\hat{p}^{d_2}(D_1, \underline{X}_1)$, especially at the boundaries of the distributions.

Table 6 presents the results for our three different comparisons of treatment sequences. As displayed in the first row, we find no statistically significant increase in employment when attending two years of vocational training, rather than two years of academic classroom training. Even though the point estimate $\hat{\Delta}(\underline{d}_2, \underline{d}_2^*, S=1)$ suggests an increase of 10 percentage points in the employment probability (starting from a counterfactual probability of 76%), the $p$-value is beyond any conventional level of statistical significance. For the comparison of vocational training to academic education in the first year or no training in either year presented in the second and third rows, however, the effects of 5 and 8 percentage points are statistically significant at the 10% and 5% levels, respectively. We therefore conclude that vocational training appears to increase the employment probability four years after randomisation into Job Corps, while it is less clear whether it performs relatively better than academic classroom training. The results are qualitatively similar when increasing the trimming threshold for the products of the propensity scores to 0.03, see Table 7. However, the $p$-value of the effect of vocational vs. no training is now somewhat higher (4%).

To partially assess the validity of the conditional independence assumptions imposed in this application, we conduct a placebo test based on comparing the outcomes of two control groups, as for instance discussed in Athey and Imbens (2017). The first control group are the never takers, i.e., those randomised into Job Corps who never attended any form of instruction with treatment sequence 11. The second control group are those randomised out, and thus without access to Job Corps instruction with treatment sequence 00. We estimate the pseudo-treatment effect of

**Table 7.** Effect estimates with a trimming threshold of 0.03.

| $\underline{d}_2$ | $\underline{d}_2^*$ | $\hat{E}[Y_2(\underline{d}_2^*)\|S=1]$ | $\hat{\Delta}(\underline{d}_2, \underline{d}_2^*, S=1)$ | SE | $p$-value | Observations | Trimmed |
|---|---|---|---|---|---|---|---|
| 33 | 22 | 0.78 | 0.07 | 0.05 | 0.23 | 3,783 | 1,940 |
| 33 | 21 | 0.82 | 0.05 | 0.03 | 0.05 | 3,783 | 587 |
| 33 | 11 | 0.81 | 0.07 | 0.03 | 0.04 | 2,346 | 170 |

*Notes:* $\underline{d}_2$ and $\underline{d}_2^*$ indicate the treatment sequences under treatment and nontreatment, respectively. $\hat{E}[Y_2(\underline{d}_2^*)|S=1]$ denotes the mean potential outcome under nontreatment conditional on $S=1$, where $S$ is an indicator for the first treatment corresponding to either the first treatment in $\underline{d}_2$ or $\underline{d}_2^*$. $\hat{\Delta}(\underline{d}_2,\underline{d}_2^*, S=1)$ provides the ATE estimate, SE the standard error. The last column gives the number of observations dropped according to the trimming rule $\hat{p}^{d_1}(X_0) \cdot \hat{p}^{d_2}(D_1,\underline{X}_1) < 0.03$.

**Table 8.** Placebo test with a trimming threshold of 0.01.

| ATE estimate | SE | $p$-value | Observations | Trimmed |
|---|---|---|---|---|
| 0 | 0.02 | 0.92 | 4,752 | 196 |

*Notes:* The ATE estimate provides the pseudo-treatment effect when comparing the employment outcomes of never takers (treatment sequence 11) and those randomised out (treatment sequence 00) conditional on baseline covariates $X_0$. The last column states the number of observations dropped according to the trimming rule: $p(X_0) < 0.01$.

Job Corps on the employment outcome using the DML approach for assessing static (rather than dynamic) treatments, as for instance discussed in Chernozhukov et al. (2018). To this end, we consider sequence 11 as pseudo-treatment and sequence 00 as nontreatment and control for the baseline covariates $X_0$ based on the random forest as machine learner of the nuisance parameters. As neither group attended any training, the true ATE is equal to zero. As shown in Table 8, the estimated ATE is indeed approximately zero with a $p$-value of 92%. This provides some statistical support for the satisfaction of the conditional independence assumption, at least with respect to the baseline covariates $X_0$.

## 7. CONCLUSION

In this paper, we combined dynamic treatment evaluation with DML under sequential selection-on-observables assumptions, which avoids ad hoc pre-selection of control variables. This approach appears particularly fruitful in high-dimensional data with many potential control variables. We suggested estimators for the (weighted) average effects of sequences of treatments (with the so-called controlled direct effect being a special case) based on Neyman-orthogonal score functions, sample splitting, and machine learning-based plug-in estimates of conditional mean outcomes and treatment propensity scores. We demonstrated the $\sqrt{n}$-consistency and asymptotic normality of the treatment effect estimators under specific regularity conditions, and analysed their finite sample behaviour in a Monte Carlo simulation. Finally, we applied our method to the Job Corps data to analyse the effects of distinct sequences of educational programmes, and found positive employment effects for vocational training when compared to no programme participation.

## ACKNOWLEDGEMENTS

We have benefitted from comments by Jelena Bradic, Saraswata Chaudhuri, Yingying Dong, Arturas Juodis, Frank Kleibergen, Jonathan Roth, Vasilis Syrgkanis, Davide Viviano, and

seminar participants at McGill University, the University of Amsterdam, the University of Duisburg-Essen, the University of Bolzano/Bozen, and the University of California Irvine (all online). Lafférs acknowledges support provided by the Slovak Research and Development Agency under contract no. APVV-17-0329 and VEGA-1/0692/20.

## REFERENCES

Angrist, J., G. Imbens and D. Rubin (1996). Identification of causal effects using instrumental variables. *Journal of American Statistical Association 91*, 444–72 (with discussion).

Athey, S. and G. W. Imbens (2017). The state of applied econometrics: Causality and policy evaluation. *Journal of Economic Perspectives 31*(2), 3–32.

Athey, S., G. W. Imbens and S. Wager (2018). Approximate residual balancing: Debiased inference of average treatment effects in high dimensions. *Journal of the Royal Statistical Society Series B 80*, 597–623.

Bang, H. and J. Robins (2005). Doubly robust estimation in missing data and causal inference models. *Biometrics 61*, 962–72.

Belloni, A., V. Chernozhukov and C. Hansen (2014). Inference on treatment effects after selection among high-dimensional controls. *Review of Economic Studies 81*, 608–50.

Belloni, A., V. Chernozhukov, I. Fernández-Val and C. Hansen (2017). Program evaluation and causal inference with high-dimensional data. *Econometrica 85*(1), 233–98.

Biewen, M., B. Fitzenberger, A. Osikominu and M. Paul (2014). The effectiveness of public-sponsored training revisited: The importance of data and methodological choices. *Journal of Labor Economics 32*(4), 837–97.

Blackwell, M. and A. Strezhnev (2022). Telescope matching for reducing model dependence in the estimation of the effects of time-varying treatments: An application to negative advertising. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* (to appear).

Bodory, H. and M. Huber (2018). The causalweight package for causal inference in R. SES Working Paper 493, University of Fribourg.

Breiman, L. (2001). Random forests. *Machine Learning 45*, 5–32.

Chernozhukov, V., D. Chetverikov, M. Demirer, E. Duflo, C. Hansen, W. Newey and J. Robins (2018). Double/debiased machine learning for treatment and structural parameters. *Econometrics Journal 21*, C1–68.

Colangelo, K. and Y.-Y. Lee (2021). Double debiased machine learning nonparametric inference with continuous treatments. arXiv preprint arXiv:2004.03036v6.

Farbmacher, H., M. Huber, L. Laffers, H. Langen and M. Spindler (2022). Causal mediation analysis with double machine learning. *Econometrics Journal 25*(2), 277–300.

Farrell, M. H., T. Liang and S. Misra (2021). Deep neural networks for estimation and inference. *Econometrica 89*(1), 181–213.

Flores, C. A. and A. Flores-Lagunes (2009). Identification and estimation of causal mechanisms and net effects of a treatment under unconfoundedness. IZA DP No. 4237.

Flores, C. A. and A. Flores-Lagunes (2010). Nonparametric partial identification of causal net and mechanism average treatment effects. mimeo, University of Florida.

Flores, C. A., A. Flores-Lagunes, A. Gonzalez and T. C. Neumann (2012). Estimating the effects of length of exposure to instruction in a training program: The case of job corps. *The Review of Economics and Statistics 94*, 153–71.

Frölich, M. and M. Huber (2017). Direct and indirect treatment effects: Causal chains and mediation analysis with instrumental variables. *Journal of the Royal Statistical Society: Series B 79*(5), 1645–66.

Hirano, K., G. W. Imbens and G. Ridder (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica 71*, 1161–89.

Horvitz, D. G. and D. J. Thompson (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association 47*, 663–85.

Huber, M. (2014). Identifying causal mechanisms (primarily) based on inverse probability weighting. *Journal of Applied Econometrics 29*, 920–43.

Huber, M. (2022). Replication data for: Evaluating (weighted) dynamic treatment effects by double machine learning. Harvard Dataverse, V2. 10.7910/DVN/FS0KBA. https://doi.org/10.7910/DVN/FS0KBA.

Imai, K. and M. Ratkovic (2015). Robust estimation of inverse probability weights for marginal structural models. *Journal of the American Statistical Association 110*, 1013–23.

Imbens, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review. *Review of Economics and Statistics 86*, 4–29.

Lechner, M. (2009). Sequential causal models for the evaluation of labor market programs. *Journal of Business and Economic Statistics 27*, 71–83.

Lechner, M. and R. Miquel (2010). Identification of the effects of dynamic treatments by sequential conditional independence assumptions. *Empirical Economics 39*, 111–37.

Lechner, M. and C. Wunsch (2013). Sensitivity of matching-based program evaluations to the availability of control variables. *Labour Economics 21*, 111–21.

Lewis, G. and V. Syrgkanis (2021). Double/debiased machine learning for dynamic treatment effects. *Advances in Neural Information Processing Systems 34*, 22695–707.

Luo, Y. and M. Spindler (2016). High-dimensional $l_2$ boosting: Rate of convergence. arXiv preprint 1602.08927.

Nagelkerke, N. J. D. (1991). A note on a general definition of the coefficient of determination. *Biometrika 78*, 691–2.

Neyman, J. (1959). Optimal asymptotic tests of composite statistical hypotheses. In U. Grenander (Ed.), *Probability and Statistics*, 416–44. New York, NY: Wiley.

Pearl, J. (2001). Direct and indirect effects. In J. Breese and D. Koller (Eds.), *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence*, 411–20. San Francisco, CA, USA: Morgan Kaufman.

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 Modelling 7*, 1393–512.

Robins, J. M. (1998). Marginal structural models. *1997 Proceedings of the American Statistical Association, Section on Bayesian Statistical Science*, 1–10. Alexandria, VA: American Statistical Association. https://cdn1.sph.harvard.edu/wp-content/uploads/sites/343/2013/03/msm-web.pdf.

Robins, J. M. and A. Rotnitzky (1995). Semiparametric efficiency in multivariate regression models with missing data. *Journal of the American Statistical Association 90*, 122–29.

Robins, J. M. (2000). Marginal structural models versus structural nested models as tools for causal inference. In M. E. Halloran and D. Berry (Eds.), *Statistical Models in Epidemiology, the Environment, and Clinical Trials*, 95–133. New York: Springer.

Robins, J. M., S. Greenland and F.-C. Hu (1999). Estimation of the causal effect of a time-varying exposure on the marginal mean of a repeated binary outcome. *Journal of the American Statistical Association 94*, 687–700.

Robins, J. M., M. A. Hernan and B. Brumback (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology 11*, 550–60.

Robins, J. M., A. Rotnitzky and L. Zhao (1994). Estimation of regression coefficients when some regressors are not always observed. *Journal of the American Statistical Association 90*, 846–66.

Rubin, D. (1980). Comment on 'Randomisation analysis of experimental data: The Fisher randomisation test' by D. Basu. *Journal of American Statistical Association 75*, 591–3.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomised and nonrandomised studies. *Journal of Educational Psychology 66*, 688–701.

Schochet, P. Z., J. Bellotti, R.-J. Cao, S. Glazerman, A. Grady, M. Gritz, S. McConnell, T. Johnson and J. Burghardt (2003). *National Job Corps Study: Data Documentation and Public Use Files: Volume iv*. Princeton, NJ: Mathematica Polic Rearch, Inc.

Schochet, P. Z., J. Burghardt and S. Glazerman (2001). *National Job Corps Study: The Impacts of Job Corps on Participants Employment and Related Outcomes*. Washington, DC: Mathematica Policy Research, Inc.

Schochet, P. Z., J. Burghardt and S. McConnell (2008). Does Job Corps work? Impact findings from the national Job Corps study. *American Economic Review 98*, 1864–86.

Schochet, P. Z., J. Burghardt and S. McConnell (2019). Replication data for: Does Job Corps Work? Impact Findings from the National Job Corps Study. 10.3886/E113269V1, Nashville, TN, USA. https://doi.org/10.3886/E113269V1.

Singh, R. (2021). Debiased kernel methods. arXiv preprint arXiv:2102.11076V2. doi: 10.48550/arXiv.2102.11076.

Syrgkanis, V. and M. Zampetakis (2020). Estimation and inference with trees and forests in high dimensions. arXiv preprint arXiv:2007.03210.

Tran, L., C. Yiannoutsos, K. Wools-Kaloustian, A. Siika, M. van der Laan and M. Petersen (2019). Double robust efficient estimators of longitudinal treatment effects: Comparative performance in simulations and a case study. *International Journal of Biostatistics 15*(2), 1–27.

van der Laan, M. J. and S. Gruber (2012). Targeted minimum loss based estimation of causal effects of multiple time point interventions. *International Journal of Biostatistics 8*(1), 1–37.

van der Laan, M. J., E. C. Polley and A. E. Hubbard (2007). Super learner. *Statistical Applications in Genetics and Molecular Biology 6*(1), 1–21.

van der Laan, M. J. and D. Rubin (2006). Targeted maximum likelihood learning. *International Journal of Biostatistics 2*, 1–38.

Viviano, D. and J. Bradic (2021). Dynamic covariate balancing: Estimating treatment effects over time. arXiv preprint 2103.01280. 2103.01280.

Wager, S. and S. Athey (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association 113*, 1228–42.

Yu, Z. and M. J. van der Laan (2006). Double robust estimation in longitudinal marginal structural models. *Journal of Statistical Planning and Inference 136*(3), 1061–89.

Zubizarreta, J. R. (2015). Stable weights that balance covariates for estimation with incomplete outcome data. *Journal of the American Statistical Association 110*, 910–22.

## SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Online Appendix
Replication Package

*Co-editor Victor Chernozhukov handled this manuscript.*