Title
Targeted learning of individual effects and individualized treatments using an instrumental variable

Author
Toth, Boriska

Publication Date
2016

Peer reviewed|Thesis/dissertation

Targeted learning of individual effects and individualized treatments using an instrumental variable

by

Boriska Toth

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Biostatistics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Mark van der Laan, Chair
Professor Nicholas Jewell
Professor Jasjeet Sekhon

Fall 2016

Targeted learning of individual effects and individualized treatments using an instrumental variable

Abstract

Targeted learning of individual effects and individualized treatments using an instrumental variable

by

Boriska Toth

Doctor of Philosophy in Biostatistics

University of California, Berkeley

Professor Mark van der Laan, Chair

We consider estimation of causal effects when treatment assignment is potentially subject to unmeasured confounding, but a valid instrumental variable is available. Moreover, our models capture treatment effect heterogeneity, and we allow conditioning on an arbitrary subset of baseline covariates in estimating causal effects. We develop detailed methodology to estimate several types of quantities of interest: 1) the dose-response curve, where our parameter of interest is the projection unto a finite-dimensional working model; 2) the mean outcome under an optimal treatment regime, subject to a cost constraint; and 3) the mean outcome under an optimal intent-to-treat regime, subject to a cost constraint, in which an optimal intervention is done on the instrumental variable. These quantities have a central role for calculating and evaluating individualized treatment regimes. We use semiparametric modeling throughout and make minimal assumptions. Our estimate of the dose-response curve allows treatment to be continuous and makes slightly weaker assumptions than previous research. This work is the first to estimate the effect of an optimal treatment regime in the instrumental variables setting. For each of our parameters of interest, we establish identifiability, derive the efficient influence curve, and develop a new targeted minimum loss-based estimator (TMLE). In accordance with the TMLE methodology, these substitution estimators are asymptotically efficient and double robust. Detailed simulations confirm these desirable properties, and that our estimators can greatly outperform standard approaches. We also apply our estimator to a real dataset to estimate the effect of parents' education on their infant's health.

Contents

List of Tables

Acknowledgments

This thesis is greatly indebted to my advisor Mark van der Laan. I wished to work with a PhD advisor who is not only an authority in his or her field, but is truly intellectually brilliant. I couldn't have asked for more. Mark is an absolute inspiration and has been an invaluable resource to me for his depth of knowledge, insightfulness, and innovativeness in driving statistical research. I was often amazed by his ability to instantly dive into the details of our projects as soon as I stepped into his office, while he juggled a myriad of other collaborations. Mark was equally responsive to my technical questions as he was supportive with administrative and professional matters. In addition, I thank him for proposing the main topic of targeted learning on instrumental variable models, seeing my interest in working at the intersection of biostatistics and econometrics. Mark made this opportunity to do a biostatistics PhD possible in the first place by supporting me in transferring to the program.

I am grateful to Profs. Nicholas Jewell and Jasjeet Sekhon, who served on both my dissertation and qualifying exam committees, and Prof. Michael Anderson, who served on my qualifying exam committee, for all their valuable insights. I'd also like to thank the extraordinarily dedicated staff of the biostatistics and statistics departments. Specifically, Sharon Norris, La Shana Porlaris, and Burke Bundy have each been of great help to me on numerous occasions during this process. I am thankful to have had the opportunity to work with Hank Ibser, Howard D'Abrera, Nusrat Rabbee, Wayne Lee, and Roger Purves in the many gratifying experiences I've had teaching undergraduate statistics.

I fondly acknowledge the many people who made these Berkeley years such an indescribably special time and place, giving me my very closest friends, opening my world through many international friendships, and making me at home here all these years.

I thank my parents for the gift of education.

Most of all, I thank András. He knows why.

Finally, I thank Diána Lolka for the infusion of sweetness, healing, and inspiration.

Chapter 1

Introduction

Utilizing instrumental variables. When estimating a causal effect in an observational study, the problem of unmeasured confounding is a pervasive caveat. It is similarly problematic in inferring a causal effect of a treatment in an experiment where the treatment isn't fully randomized. A classic solution for obtaining a consistent estimate is to use an instrumental variable, assuming one exists. Informally, an instrumental variable, or instrument, is a variable Z that affects the outcome Y only through its effect on the treatment A, and the residual (error) term of the instrument is uncorrelated with the residual term of the outcome (Imbens and Angrist 1994, Angrist et al 1996, Angrist and Krueger 1991). Thus, the instrument produces exogenous variation in the treatment.

Instrumental variables have been used widely in biostatistics and especially econometrics to obtain consistent estimates of a treatment effect. (See (Brookhart et al 2010) for a large collection of references.) They are a basic tool for inferring the causal effect of a clinical treatment or a medication on a health outcome, as large-scale randomization of patients is often not feasible. In these settings, the instrumental variable is usually some attribute that is related to the health care a patient receives, but is not at the level of individual patients. Thus, the instrument is not confounded by factors affecting an individual's response to treatment. For example, (Newhouse and McClellan 1998) exploit regional variation in the availability of catheterization and revascularization procedures as their instrument in estimating the effect of these procedures on reducing mortality in heart attack patients. Another important setting for instrumental variables in health research is when the treatment is randomly assigned, but non-compliance is significant. Then the random treatment assignment serves as an ideal instrument. (van der Laan et al 2007) describe this setting. An example of the use of instruments in social science research is the data analysis we give in chapter 3. In this dataset, the extent to which individuals were affected by a school reform program serves as the instrument for estimating the effect of parents' education on their newborn's health.

Causal effects given arbitrary subgroups of the population. Most commonly, instrumental variables are used to estimate a simple pointwise treatment effect. This work, on the other hand, solves the more complex problem of estimating functions involving an arbitrary subset V of baseline covariates W. In chapter 3, the causal effect of treatment (dose-response curve) as a function of $V \subseteq W$ is estimated. In chapters 4-5, the optimal dynamic treatment rule as a function of covariates V (and under cost constraints) is derived, for the purpose of estimating the mean outcome under this optimal rule. We allow binary or continuous outcomes Y, and make no restrictions on the type of data unless otherwise noted.

Modeling heterogenous treatment effects can improve the precision of a statistical model. Moreover, it is of paramount importance in many applications to make estimates of treatment effect that are conditional on individual characteristics, see for instance (Imai and Strauss 2011). This is especially so in clinical settings. These days there is great interest and computational feasibility in designing individualized treatment regimes based on a patient's characteristics and biomarkers. The paradigm of precision medicine calls for incorporating genetic, environmental and lifestyle variables into treatment decisions. For estimating the dose-response curve in chapter 3, we take the expected causal effect given V, so far an infinite dimensional parameter, and project that function of V unto a user-supplied working model of finite dimension. Our model allows A to be continuous and makes slightly weaker causal assumptions than previous work. Furthermore, it is important when estimating the causal effect as a function of covariates V that V can be a strict subset of all baseline covariates W. As an example, medical data often involves a large space of covariates, and conditioning on many covariates in estimating relevant components of the data-generating distribution can be helpful in: 1) decreasing the variance of estimated conditional means, and 2) ensuring that the instrument induces exogenous variation given the covariates. However, a physician typically has a smaller set of patient variables that are available and that s/he considers reliable predictors. Thus the causal effect as a function of an arbitrary subset of baseline covariates is of great use.

Outcomes under optimal treatment rules. Another natural parameter of interest is the population mean outcome under an optimal treatment rule, pursued in chapters in 4 and 5. Coupling statistical estimation with optimization within the same model is a standard problem across many fields, with a large body of previous work in engineering and statistics. This work is the first to estimate the effect of an optimal dynamic treatment (ODT) regime in the instrumental variables setting. The underlying parameter of interest in an ODT problem is often the actual optimal treatment rule, as a function of covariates. However, that is an infinite dimensional parameter, so instead we target the mean outcome under the optimal rule. We solve the estimation problem both for the setting where the dynamic intervention is on the treatment variable A (chapter 5), and where the intervention is on the instrument Z (chapter 4).

A good example of the usefulness of both these models is for analyzing the get-out-the-vote campaign described in (Arceneaux et al 2006). A large-scale voter mobilization experiment was done in which individuals were randomly assigned (instrument Z) to a treatment of receiving a phone call, and there is unmeasured confounding between an individual actually receiving the phone call (treatment A) and their voting behavior (outcome Y). Many baseline covariates are known. It is of significant interest to both political parties and social scientists to learn the optimal assignment of individuals to the treatment of a phone call when subject to a cost constraint, given their individual characteristics. One parameter of interest is: what would be the mean outcome under the optimal selection of individuals to receive the phone calls? Chapter 5 deals with this scenario of the optimal dynamic treatment regime, when optimally intervening on treatment variable A. Chapter 4, on the other hand, focuses on the optimal dynamic treatment regime when the 'treatment' is actually an intervention done on the instrumental variable Z. In settings where unmeasured confounding is a potential problem, it is often not possible in practice to intervene directly on the treatment variable. It might not be possible to actually control whether individuals pick up the phone. Thus, the mean outcome under optimal intervention on the instrument is also a parameter of interest. We call this the optimal dynamic *intent-to-treat* mechanism, so named because the instrument is often a randomized assignment to treatment or encouragement mechanism. Under our randomization assumption on Z, the optimal dynamic intent-to-treat problem is the same as an optimal dynamic treatment problem when considering Z to be the treatment variable that is unconfounded with Y.

The problem of finding the optimal deterministic treatment rule is NP-hard (Karp 1972). However, when allowing possible non-deterministic treatments, there is a simple closed form solution for the optimal dynamic treatment, or the optimal dynamic intent-to-treat. The optimal rule is to treat all strata with the highest marginal gain per marginal cost, so that the total cost of the policy equals the cost constraint. Chapter 4 gives realistic conditions under which we have a well-behaved estimation problem, with pathwise differentiability. Under these conditions, the optimal solution is a deterministic rule.

Semiparametric methodology. While instrumental variables are widely used to infer causal effects, the majority of studies make use of strong assumptions about the structure of the data and typically rely on parametric models (Terza et al. 2008). In contrast, this work uses semiparametric modelling. Beyond the criteria that there is a valid instrument, we make use of the single structural assumption that the expected value of the outcome is linear in the treatment, conditional on the covariates. This assumption is used in virtually all similar works; however, as we discuss below, even this single assumption can be weakened.

We use targeted minimum loss estimation (TMLE), which is a methodology for semiparametric estimation that has very favorable theoretical properties and can be superior to other estimators in practice (van der Laan and Rubin 2006, van der Laan and Rose 2011). The TMLE procedure targets only those components of the data-generating distribution that are relevant to the statistical parameter of interest. Initial estimates are formed of certain components, by data-adaptively learning on a library of prediction algorithms. The initial estimates are then fluctuated one or more times in a direction that removes bias and optimizes for semiparametric efficiency.

The TMLE method has a robustness guarantee: it produces consistent estimates even when the functional form is not known for all relevant components. The most common such scenario is when the conditional distribution of the outcome cannot be estimated consistently, and one only has information about the form of the distributions generating the instrument and treatment. TMLE also guarantees asymptotic efficiency when all relevant components and nuisance parameters are consistently estimated. Thus, under certain conditions, the TMLE estimator is optimal in having the asymptotically lowest variance for a consistent estimator in a general semiparametric model, thereby achieving the semiparametric Cramer-Rao lower bound (Newey 1990). Another beneficial property is asymptotic linearity. This ensures that TMLE-based estimates are close to normally-distributed for moderate sample sizes, which makes for accurate coverage of confidence intervals.

TMLE has the advantage over other semiparametric efficient estimators that it imposes constraints to ensure that the estimator matches the data well. It is a substitution estimator, meaning that the final estimate is made by evaluating the parameter of interest on the estimates of its relevant components, where these estimates respect bounds observed in the data. These properties have been linked to good performance in sparse data in (Gruber and van der Laan 2010), while we demonstrate performance gains over other estimators in continuous data having sharp boundaries in section 3.3.

Extensive simulation results validate the strong performance of our TMLE-based estimator. In estimating the dose-response curve, TMLE can show enormous bias removal for a moderate gain in variance. It can have both lower bias and lower variance than an incorrectly specified parametric model, due to the vastly better fit resulting from data-adaptive learning. In addition, it can show superior finite-sample performance over other semiparametric efficient estimators for certain types of data, and demonstrates good coverage, with 95% confidence intervals that were typically 1-2% too wide. For the parameters involving optimal dynamic treatments, TMLE demonstrates reduction in finite-sample bias over a consistent initial substitution estimator, and the cross-validated (Zheng and van der Laan 2011) version of TMLE was typically found to have the best performance. The empirical variance of the TMLE-based estimators appears to converge to the semiparametric efficiency bound. Consistency in the case of partial misspecification was confirmed, in the sense of lemma 16. Our simulations also addressed the important question of to what extent improved statistical estimation can lead to better optimization results. We were able to demonstrate significant increases in the value of the mean outcome under the estimated optimal rule, when a larger library of data-adaptive learners achieved a closer fit. The known distribution was used to evaluate the true mean outcome under an estimated optimal rule here.

Chapter 2: Background

We first give background on the basic notions of semiparametric efficiency theory. Next, we describe the targeted minimum loss based-estimation (TMLE) methodology. We conclude with a literature review.

Chapter 3: Estimating the dose-response function using an instrument.

We define a causal parameter of interest for estimating the dose-response curve as a function of covariates $V \subseteq W$. We give a causal model making minimal assumptions and prove identifiability. The canonical gradient is then derived, and three different TMLE-based estimators given for the target parameter. In detailed simulations, we show the performance of each of these estimators, and compare with standard methods, both parametric and semiparametric. We present an application to the (Chou et al 2010) dataset to estimate the effect of parents' education on their infant's health. Finally, we briefly describe a test for unmeasured confounding using the estimators developed.

Chapter 4: Estimating the outcome under optimal dynamic intent-to-treat.

We give assumptions for estimating the mean counterfactual outcome, under optimal assignment of instrumental variable Z and given cost constraints. We then prove a simple closed form solution, and remark on its extension to a continuous or categorical instrument. Next, the canonical gradient is derived, and a TMLE based on a single-step fluctuation given. Desirable theoretical properties for estimation and inference are proved. Finally, detailed simulations are presented to confirm the theoretical results and demonstrate the link between effective optimization and precise statistical estimation.

Chapter 5: Estimating the outcome under optimal dynamic treatment, using an instrument.

We consider the problem of estimating the mean counterfactual outcome under the optimal treatment rule given a cost constraint, when an instrumental variable is needed for identifiability. We state the causal model, assumptions, and identifiability results. The canonical gradient is derived and we check that it is contained in the tangent space of the model. We also derive a remainder term, and describe the theoretical properties of our estimators that follow from it. Two TMLE estimators are presented in this chapter, one that is iterative and based on likelihood maximization, and one that involves solving a numerically challenging equation in a single step.

Chapter 2

Background

2.1 Semiparametric efficiency theory.

(See for instance (van der Vaart 2000) for a reference.) Recall that an *efficient estimator* is one that achieves the optimal asymptotic variance among regular semiparametric estimators. We briefly give a few relevant definitions.

An estimator is *asymptotically linear* if, informally, it is asymptotically equivalent to a sample average. Formally, we have that an estimator Ψ_n^* for estimating true parameter $\Psi(P_0)$ from an iid sample $(O_1, .., O_n)$ is asymptotically linear if
$\sqrt{n}(\Psi_n^* - \Psi(P_0)) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \dot{\Psi}_P(O_i) + o_P(1)$, where $\dot{\Psi}_P$ is a zero mean, finite variance function. $\dot{\Psi}_P$ is called the influence function.

Recall that a parameter Ψ is pathwise differentiable at P_0 relative to a tangent space of a model $\mathcal{P}$ at P_0 if there exists a continuous linear map $\dot{\Psi}_{P_0}$ such that for every score function g in the tangent space and submodel $t \to P_t$ with score function g, we have
$\frac{\Psi(P_t)-\Psi(P_0)}{t} \to \dot{\Psi}_{P_0}\ g$. By the Riesz representation theorem, we have $\dot{\Psi}_{P_0}\ g = \int \tilde{\Psi}_{P_0}\ g dP_0$ where $\tilde{\Psi}_{P_0}$ is an "influence function". The *efficient influence curve* is the unique influence function whose coordinate functions are contained in the closure of the linear span of the tangent space.

An estimator is *efficient* if it is asymptotically linear with the efficient influence curve as its influence function. Thus, we have that for an efficient estimator Ψ_n^* estimating true parameter $\Psi(P_0)$ from an iid sample $(O_1, ..., O_n)$: $\Psi_n^* - \Psi(P_0) = \frac{1}{n} \sum_{i=1}^n D^*(P_0)(O_i) + o_P(\frac{1}{\sqrt{n}})$, where D^* is the efficient influence curve.

The proof that the optimal asymptotic variance is the variance of the efficient influence curve (where optimality is among regular semiparametric estimators) is a short proof. Essentially, one checks the lower bound on variance guaranteed by the Cramer-Rao inequality, in the direction of the hardest submodel (that having greatest variance).

2.2 Targeted maximum likelihood estimation.

The targeted minimum loss-based framework

Targeted minimum loss-based estimation (TMLE) is a method to construct a semi-parametric substitution estimator of a target parameter $\Psi(P_0)$ of a true distribution $P_0 \in \mathcal{M}$, where $\mathcal{M}$ is a semiparametric statistical model (van der Laan and Rubin 2006, van der Laan and Rose 2011). The estimate is based on sampling n i.i.d. data points $(O_1, ..., O_n)$ from P_0. It is consistent and asymptotically efficient under certain conditions.

(1) One first notes that the parameter of interest $\Psi(P_0)$ depends on P_0 only through relevant components Q_0 of the full distribution P_0, in other words, $\Psi(P_0) = \Psi(Q_0)$ [1]. TMLE *targets* these relevant components by only estimating these Q_0 and certain nuisance parameters g_0 [2] that are needed for updating the relevant components. An initial estimate (Q_n^0, g_n) is formed of the relevant components and nuisance parameters. This is typically done using the Super Learner (see below) approach described in (van der Laan et al 2007), in which the best combination of learning algorithms is chosen from a library using cross-validation. (2) Then the relevant components Q_n^0 are fluctuated, possibly in an iterative process, in an optimal direction for removing bias efficiently. (3) Finally, one evaluates the statistical target parameter on the updated relevant components Q_n^*, and arrives at estimate $\psi_n^* = \Psi(Q_n^*)$.

Note that the final estimate of ψ_n^* is formed by evaluating the target parameter on estimates of relevant components that are consistent with a single data-generating distribution, and with the observed bounds of the data. This property of being a *substitution estimator* has been shown to be conducive to good performance in practice (Gruber and van der Laan 2010).

We use notation such as Q_n^0, where the subscript clarifies that an empirical estimate is being made from the sample of size n, while the superscript refers to the estimate being an initial one ("zeroeth" iteration). To fluctuate the initial components Q_n^0 to updated components Q_n^1, one defines a fluctuation

[1] We are abusing notation here for the sake of convenience by using $\Psi(\cdot)$ to denote both the mapping from the full distribution to $\mathbb{R}^d$, and from the relevant components to $\mathbb{R}^d$.

[2] The nuisance parameters are those components g_0 of the efficient influence curve $D^*(Q_0, g_0)$ that $\Psi(Q_0)$ does not depend on.

function $\epsilon \to Q(\epsilon|g_n)$. g_n is an estimate of the nuisance parameters, and the fluctuation of Q_n^0 can depend on g_n, although we sometimes drop the explicit dependency in the notation, and use $Q(\epsilon)$ to denote $Q(\epsilon|g_n)$. One also defines a loss function $L()$, where we set $Q_n^1 = Q_n^0(\epsilon_n^0|g_n)$ by solving for fluctuation $\epsilon_n^0 = \text{argmin}_\epsilon \; L(Q_n^0(\epsilon|g_n), g_n, (O_1, ..., O_n))$. We use the convention that when the fluctuation parameter ϵ is zero, $Q_n^0(\epsilon|g_n) = Q_n^0$. This procedure of updating $Q_n^{k+1} = Q_n^k(\epsilon_n^k|g_n)$ might need to be iterated to convergence. In some versions of TMLE, the nuisance parameters g_n are also updated, using a fluctuation function and loss function similarly. The requirement is to choose the fluctuation and loss functions so that, upon convergence of the components to their final estimate Q_n^* and g_n^*, the efficient influence curve equation is solved:

$$P_n \; D^*(Q_n^*, g_n^*) = 0$$

P_n denotes the empirical distribution $(O_1, ..., O_n)$, and we use the shorthand notation $P_n f = \frac{1}{n}\sum_{i=1}^n f(O_i)$. The equation above is the basis for the guarantees of consistency (under partial misspecification) and asymptotic efficiency (under correct specification of relevant components and nuisance parameters).

To give a few examples, the loss function might be the mean squared error, or the negative log likelihood function. For instance, for the estimator using iterative updating presented in section 3.2, we use fluctuation $\mu_n^1 = \mu_n^0 + \epsilon \cdot C_{Y,n}^0$, with $\mu = E(Y|W,Z)$ and C_Y as defined in section 4.1. The loss function is $L(Q_n^0(\epsilon|g_n), g_n, (O_1, ..., O_n)) = \sum_{i=1}^n \left(Y[i] - \mu_n^0[i] - \epsilon \cdot C_{Y,n}^0[i]\right)^2$.

2.3 Literature review.

Estimating the dose-response curve using an instrument. Let W be a vector of baseline covariates, and $m(W)$ denote the marginal causal effect of treatment given W. Most prior work on estimating the marginal causal effect of a treatment using an instrument deal with either the case where a scalar average effect $E(m(W))$ is estimated, or the entire curve $m(W)$ is estimated. In contrast, our work estimates $E(m(W)|V)$ for V possibly a strict subset of W. (Tan 2010) is another work that lets V be any subset of W and gives estimators for the marginal effect of the treatment on Y, conditional on V and level of treatment. However, their marginal effect is assumed to take a parametric form.

(Ogburn et al) is a recent work that also proposes a semiparametric estimator for the marginal causal effect given a strict subset of the covariates

$V \subseteq W$[3]. They also present an estimator for the best least-squares projection of the true causal effect unto a parametric working model. Their estimators use the method estimating equations, and are efficient and double robust, but are not substitution estimators. In addition, (Ogburn et al) restrict attention to the case of a binary instrument and treatment, and make slightly stronger assumptions about the instrument than we do (for instance, they assume no confounding between the instrument and treatment).

(Abadie 2003) gives an estimator for the treatment effect in compliers as a function of W. However, the instrument propensity score $P(Z|W)$ must be estimated consistently in his approach. Both (van der Laan et al 2007) and (Robins 2004) present semiparametric, consistent, and locally efficient estimators for the effect of treatment on an outcome, as a function of covariates W, as motivated by the setting where Z is the randomized assignment to a binary treatment, and A is the binary compliance with treatment. The counterfactual outcomes are assumed to follow a parametric form $E(Y(A=0)|W,Z,A) = \tilde{m}(W,Z,A)$. The former work gives a solution for binary outcomes using the method of estimating equations, so that their estimator is double robust to misspecification of either $\Pr(Z|W)$ or $E(Y(A=0)|W,Z,A)$.

For the special case of a null V where a scalar average effect is estimated, semiparametric efficient approaches abound (see for instance: Cheng et al 2009; Hong and Nekipelov 2010; Kasy 2009). (Uysal 2011) and (Tan 2006) describe doubly robust estimators, where either the propensity score $Pr(Z|W)$, or the conditional means given the instrument, must be correctly specified.

(Clarke and Windmeijer 2010a and 2010b) discuss a number of approaches for dealing with binary outcomes in the instrumental variables setting.

Optimal dynamic treatments.

(Luedtke and van der Laan 2016a) is a recent work that gives a TMLE estimator for the mean outcome under optimal dynamic treatment given a cost constraint. While we state the ODT problem in chapter 4 in terms of finding the optimal value of an instrumental variable when there is potentially confounding between the treatment and outcome, that problem is very similar to the one solved in (Luedtke and van der Laan 2016a). The main difference is that we allow a cost function that depends on covariates, while only unit costs are considered in the previous work. Hence, there is a simpler closed-form solution to the optimal rule in that work. (Luedtke and van der Laan 2106a) require that the optimal treatment has a unique solution. In (Luedtke and van der Laan 2016b), the authors expand greatly on this issue of possible non-

[3] Ogburn et al's work was accepted for publication around the time this work was completed.

unique solutions. They show that a unique solution for the optimal treatment is needed for pathwise differentiability of the mean outcome, which in turn is necessary for deriving a regular asymptotically linear (RAL) estimator. They also derive a martingale-based estimator that gives root-n confidence intervals, even when pathwise differentiability does not hold. The conditions we require in 4.2 and 5.2 are adopted from these works.

A large body of work focuses on the case of optimal treatment regimes in the unconstrained case, such as (Robins 2004). More recently, various approaches tackle the constrained ODT problem: (Zhang et al 2012) describe a solution that assumes the optimal treatment regime is indexed by a finite dimensional parameter, while (Chakraborty et al 2013) describe a bootstrapping method for learning ODT regimes with confidence intervals that shrink at a slower than root-n rate. (Chakraborty and Moodie 2013) gives a review of recent work on the constrained case.

Chapter 3

Estimating the dose-response function.

3.1 The model and causal parameter of interest

We use the notation that P_0 and E_0 refer to the true probability distribution and expectation, respectively, and P_n and E_n the empirical counterparts. We observe n i.i.d. copies $O_1, \ldots, O_n$ of a random variable $O = (W, Z, A, Y) \sim P_0$, where P_0 is its probability distribution. Here W denotes the measured baseline covariates, and Z denotes the subsequently (in time) realized instrument that is believed to only affect the final outcome Y through the intermediate treatment variable A. The goal of the study is to assess a causal effect of treatment A on outcome Y. We consider the case in which it is believed that A is a function of both the measured W and also unmeasured confounders. As a consequence, methods that rely on the assumption of no unmeasured confounding will likely be biased. Figure 1 shows how the variables in our model are related; the arrows indicate the direction of causation.

Using the structural equation framework of (Pearl 2000), we assume that each variable is a function of other variables that affect it and a random term (also called error term). Let U denote the error terms. Thus, we have
$W = f_W(U_W), Z = f_Z(W, U_Z), A = f_A(W, Z, U_A), Y = f_Y(W, Z, A, U_Y)$
where $U = (U_W, U_Z, U_A, U_Y) \sim P_{U,0}$ is an exogenous random variable, and f_W, f_Z, f_A, f_Y may be unspecified or partially specified (for instance, we might know that the instrument is randomized). U_Y is the term that reflects possible confounding between A and Y. When Y is binary, we assume $U_Y = (\tilde{U}_Y, U'_Y)$, where the $\tilde{U}_Y$ component contains the potentially confounded, residual term,

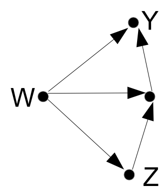

Figure 3.1: Causal diagram

with $\Pr(Y = 1|W, Z, A, \tilde{U}_Y) = \tilde{f}_Y(W, Z, A, \tilde{U}_Y)$[1]

Assumption 1 parts 1)-3) below need to be made to guarantee that Z is a valid instrument for estimating the effect of A on Y. Part 4), in turn, is needed for identifiability of the causal effect.

Assumption 1 ***Assumptions ensuring that Z is a valid instrument:***

1. ***Exclusion restriction.*** *Z only affects outcome Y through its effect on treatment A. Thus, $f_Y(W, Z, A, U_Y) = f_Y(W, A, U_Y)$.*

2. ***Exogeneity of the instrument.*** *$E(U_Y|W, Z) = 0$ for any W, Z.*

[1] The U'_Y term is an exogenous r.v. whose purpose is for sampling binary Y. Let U'_Y be a Unif[-.5,.5] r.v. (we set it to have 0 mean to conform to assumption 2.) Then $Y = 1((U'_Y + .5) < \tilde{f}_Y(W, Z, A, \tilde{U}_Y))$.

3. Z ***induces variation in*** A***.*** $Var_0[E_0(A|Z,W)|W] > 0$ *for all* W*.*

Structural equation for outcome Y***:***

4. $Y = Am_0(W) + \theta_0(W) + U_Y$ *for continuous* Y*, and*
$\Pr(Y = 1|W, A, \tilde{U}_Y) = Am_0(W) + \theta_0(W) + \tilde{U}_Y$ *for binary* Y*,*
where $U_Y = (\tilde{U}_Y, U'_Y)$ *for an exogenous r.v.* U'_Y*.*

In other words, although we don't assume that A is randomized with respect to Y, we do assume that Z is randomized with respect to Y, conditional on W in both cases. The third assumption guarantees that for every value of covariates W, there is variation in the instrument, and that the instrument induces variation in the treatment.

The linearity in A of the structural equation for Y is necessary for identifying the treatment effect using an instrument unless further assumptions are made. In the common case where the treatment A is binary, this assumption always holds, and we have a fully general semi-parametric model that only assumes Z is a valid instrument. It should also be noted that unlike many instrument-based estimators, we don't require the instrument to be randomized with respect to treatment ($U_Z \perp\!\!\!\perp U_A|\ W$ is not necessary).

We use the counterfactual framework of (Pearl 2000) to define the causal parameter of interest. Let counterfactual outcome $Y(a)$ denote the outcome given by the structural equations if the treatment variable were set to $A = a$, and all other variables, including the exogenous terms, were unchanged. We have that $Y(a) = a \cdot m_0(W) + \theta_0(W) + U_Y$ for all possible values $a \in \mathcal{A}$, where $\mathcal{A}$ denotes a support of A. We can now define the marginal causal effect we're interested in as $E_0(Y(a) - Y(0))$ and observe that it equals $a \cdot Em_0(W)$. Similarly, define adjusted causal effects $E_0(Y(a) - Y(0) \mid V)$ conditional on a user supplied covariate $V \subset W$. These causal effects are functions of $m_0(W)$ and the distribution of W.

Causal effect of interest:
The marginal causal effect is $E_0(Y(a) - Y(0)) = a \cdot Em_0(W)$.
The adjusted causal effect is $E_0(Y(a) - Y(0) \mid V) = a \cdot E(m_0(W) \mid V)$, given a user supplied covariate $V \subset W$.

Note that $m_0(W)$ represents the causal effect of one unit of treatment given W.

Notation: Let $\rho_0(Z, W) = \Pr_0(Z|W)$. Let $\Pi_0(Z, W) \equiv E_0(A \mid Z, W)$ be the conditional mean of A given Z, W.

Let $\mu_0(Z,W) \equiv E_0(Y \mid \Pi_0(Z,W),W)$ be the expected value of Y, given W and $\Pi_0(Z,W)$.

The instrumental variable assumption that $E(U_Y|Z,W)=0$ implies

$$E_0(Y \mid \Pi_0(Z,W),W) = \Pi_0(Z,W)m_0(W) + \theta_0(W)$$

Thus, our structural equation model implies a semiparametric regression model for $E_0(Y \mid \Pi_0(Z,W),W)$. Note that for a pair of values z and z_1, we have

$$E_0(Y \mid Z=z,W) - E_0(Y \mid Z=z_1,W) = \{\Pi_0(z,W) - \Pi_0(z_1,W)\}m_0(W)$$

From this equation, we get an identifiability result for m_0, stated below as a formal lemma.

Lemma 1 *Let $\Pi_0(Z,W) \equiv E_0(A \mid Z,W)$. Let $d_{Z,0}$ be the conditional probability distribution of Z, given W. Let $\mathcal{W}$ be a support of the distribution $P_{W,0}$ of W. Let $w \in \mathcal{W}$. By assumption 1 above, $Var(\Pi_0(z,w)|W=w) > 0$, so there exists two values (z,z_1) in a support of $d_{Z,0}(\cdot \mid W=w)$ for which $\Pi_0(z,w) - \Pi_0(z_1,w) \neq 0$. Thus*

$$m_0(w) = \frac{E_0(Y \mid Z=z,W=w) - E_0(Y \mid Z=z_1,W=w)}{\Pi_0(z,w) - \Pi_0(z_1,w)},$$

which demonstrates that $m_0(w)$ is identified as a function of P_0.

Statistical model: The above stated causal model implies the statistical model $\mathcal{M}$ consisting of all probability distributions P of $O=(W,Z,A,Y)$ satisfying the semiparametric regression model $E_P(Y \mid Z,W) = \Pi(P)(Z,W)m(P)(W) + \theta(P)(W)$ for some unspecified functions $m(P),\theta(P)$, and $\Pi(P)(Z,W) = E_P(A \mid Z,W)$. $\Pi(P)(Z,W)$ must satisfy $\mathrm{Var}_P[\Pi(P)(Z,W)|W] > 0$ for all W. Notice that when Z is binary, the semiparametric regression equation is always satisfied for some $m,\ \theta$.

Causal parameter: We define our causal parameter of interest to be the projection of the dose-response curve $E_0(Y(a)-Y(0) \mid V) = aE_0(m_0(W) \mid V)$ on a working model. Let $\{am_\beta(v):\beta\}$ be a working model for $E_0(Y(a)-Y(0) \mid$

V). Specifically, given some weight function $h(A,V)$, let

$$\begin{aligned}
\beta_0 &= \arg\min_{\beta} E_0 \sum_a h(a,V)\{aE(m_0(W) \mid V) - am_\beta(V)\}^2 && (3.1)\\
&= \arg\min_{\beta} E_0 \sum_a h(a,V)a^2\{E(m_0(W) \mid V) - m_\beta(V)\}^2 && (3.2)\\
&= \arg\min_{\beta} E_0 \sum_a h(a,V)a^2\{m_0(W) - m_\beta(V)\}^2 && (3.3)\\
&\equiv \arg\min_{\beta} E_0\ j(V)\{m_0(W) - m_\beta(V)\}^2, && (3.4)
\end{aligned}$$

where we defined $j(V) \equiv \sum_a h(a,V)a^2$.

For example, if V is empty, and $m_\beta(v) = \beta$, then $E_0(Y(a) - Y(0)) = \beta_0 a$. We can also select $V = W$ and $m_\beta(w) = \beta^T w$, in which case $\beta_0^T w$ is the projection of $m_0(w)$ on this linear working model $\{\beta^T W : \beta\}$.

Statistical target parameter: Our target parameter is $\psi_0 = \beta_0$.

Let $\Psi : \mathcal{M} \to \mathbb{R}^d$ be the target parameter mapping so that $\Psi(P_0) = \psi_0 = \beta_0$, which exists under the identifiability assumptions stated in Lemma 1. We note that $\psi_0 = \Psi(P_0) = \Psi(m_0, P_{W,0})$ only depends on P_0 through m_0 and $P_{W,0}$, while m_0, as statistical parameter of P_0, is identified as a function of $\mu_0 = E_0(Y \mid Z, W)$ under the semiparametric regression model $\mu_0 = E_0(Y \mid Z, W) = \pi_0(Z,W)m_0(W) + \theta_0(W)$.

The statistical estimation problem is now defined. We observe n i.i.d. copies of $O = (W, Z, A, Y) \sim P_0 \in \mathcal{M}$, and we want to estimate $\psi_0 = \Psi(P_0)$ defined in terms of the mapping $\Psi : \mathcal{M} \to \mathbb{R}^d$.

Weakening the structural assumption.

We briefly note that the structural assumption $Y = f_Y(W, A, U_Y) = Am(W) + \theta(W) + U_Y$ can be weakened in many cases when Z is a continuous variable. For a general equation $Y = f_Y(W, A, U_Y) = q(W, A) + U_Y$, where $q(W, A)$ is any function, we can write a Taylor approximation for a k-degree polynomial in A as

$$f_Y(W, A, U_Y) = A^k m_k(W) + A^{k-1}m_{k-1}(W) + ... + Am_1(W) + m_0(W) + U_Y$$

Now suppose we have $(k+1)$ values of Z: $(Z_k, Z_{k-1}, ..., Z_0)$. We have that $E(Y|Z_i, W) = E(A^k|Z_i, W)m_k(W) + E(A^{k-1}|Z_i, W)m_{k-1}(W) + ... + m_0(W)$. This means if the equation below is solvable (the matrix shown is not singular),

then we can identify
$(m_k(W), m_{k-1}(W), ..., m_0(W))$.

$$\begin{bmatrix} E(Y|Z_k,W) \\ \vdots \\ E(Y|Z_0,W) \end{bmatrix} = \begin{pmatrix} E(A^k|Z_k,W) & E(A^{k-1}|Z_k,W) & \cdots \\ \vdots & \ddots & \vdots \\ E(A^k|Z_0,W) & E(A^{k-1}|Z_0,W) & \cdots \end{pmatrix} \begin{bmatrix} m_k(W) \\ \vdots \\ m_0(W) \end{bmatrix}$$

3.2 Targeted minimum loss based estimation

Assuming a parametric form $m_\alpha(W)$ for the effect of treatment as a function of covariates

We are interested in both the scenario when the treatment effect function $m(W)$ is unconstrained, and when it has a parametric form $m_0 = m_{\alpha_0}$ for some model $\{m_\alpha : \alpha\}$ and finite-dimensional α. We focus on the first case in this chapter for deriving a detailed TMLE methodology. However, the second case is also widely applicable. In section 3.6 we derive the efficient influence curve for Ψ_0 in the setting of a parametric function $m_\alpha(W)$. TMLE-based estimators can be derived for that model analogously to the three estimators derived in this chapter.

The efficient influence curve of Ψ_0

The efficient influence curve for Ψ is derived in section 3.6. Recall our semi-parametric model, and notation $P_{W,0}, \pi_0, \rho_0(Z,W), m_0(W), \theta_0(W)$, from section 3.1. Also, define $h_1(V) \equiv \sum_a h(a,V)a^2 \frac{d}{d\beta_0} m_{\beta_0}(V)$, which has the same dimension as β_0, where $h(a,V)$ is defined in section 3.1.

Lemma 2 *The efficient influence curve of $\Psi : \mathcal{M} \to \mathbb{R}^d$ is given by*

$$\begin{aligned} &D^*(P_0) = D^*_W(P_0) \\ &+c_0^{-1}\tfrac{h_1(V)}{\sigma^2(W)}(\pi_0(Z,W) - E_0(\pi_0(Z,W) \mid W))(Y - \pi_0(Z,W)m_0(W) - \theta_0(W)) \\ &-c_0^{-1}\tfrac{h_1(V)}{\sigma^2(W)}\left\{(\pi_0(Z,W) - E_0(\pi_0(Z,W) \mid W))m_0(W)\right\}(A - \pi_0(Z,W)) \\ &\equiv D^*_W(P_0) + C_Y(Z,W)(Y - \pi_0(Z,W)m_0(W) - \theta_0(W)) \\ &\qquad\qquad\qquad\qquad\qquad\qquad -C_A(Z,W)(A - \pi_0(Z,W)) \\ &\equiv D^*_W(P_0) + D^*_Y(P_0) - D^*_A(P_0), \end{aligned} \tag{3.5}$$

where

$$c_0 \equiv E_0 \sum_a h(a,V)a^2 \left\{\frac{d}{d\beta_0} m_{\beta_0}(V)\right\}^2,$$

which is a $d \times d$ matrix, and

$$D^*_W(P_0) \equiv c_0^{-1} \sum_a h(a,V) a^2 \frac{d}{d\beta_0} m_{\beta_0}(V)(m_0(W) - m_{\beta_0}(V))$$

$$\begin{aligned}
\sigma^2(W) &= Var_{\rho_0}(\Pi_0(Z,W) \mid W)) \\
h(W) &= c_0^{-1} \frac{h_1(V)}{\sigma^2(W)} \\
C_Y(Z,W) &= h(W)(\pi_0(Z,W) - E_{\rho_0}(\pi_0(Z,W) \mid W)) \\
C_A(Z,W) &= C_Y(Z,W) m_0(W).
\end{aligned}$$

Note that $D^*(P_0)$ will be a vector-valued function in general.

Here is the TMLE estimation procedure for our marginal structural model:

Step 1: Forming initial estimates.

Components of P_0 that need to be estimated: Initial estimates must be formed of *relevant components* $Q_n^0 = (m_n^0(W), P_{W,n})$, and *nuisance parameters* $g_n^0 = (\Pi_n^0(Z,W), E_n^0(\Pi_n^0|W), \mathrm{Var}_n^0(\Pi_n^0|W), \theta_n^0(W))$.

Super Learner. We use the **Super Learner** approach to form initial estimates (van der Laan et al 2007), and software implementation in R (`http://cran.r-project.org/web/packages/SuperLearner/index.html`). Super Learner is a data-adaptive technique to choose the best linear combination of learning algorithms from a library. The objective that is minimized is the cross-validated empirical mean squared error. Each candidate learning algorithm is trained on all the data except for a hold-out test set, and this process is repeated over different hold-out sets so all data points are included in a test set. The linear combination of candidate learners that minimizes MSE over all test sets in chosen. This method has the very desirable guarantees that: 1) if none of the candidate learners converge at a parametric rate, Super Learner asymptotically attains the same risk as the oracle learner, which selects the true optimal combination of learners and 2) if one of the candidate learners uses a parametric model and contains the true data-generating distribution, Super Learner converges at an almost-parametric rate.
See section 3.3 for a list of candidate learning algorithms we use for forming the initial estimates.

Step 2: Fluctuating the relevant components Q_n^0.

We present three versions of TMLE in this paper: one where the relevant components and nuisance parameters are fluctuated iteratively, and two versions of the non-iterative TMLE described below.

Non-iterative TMLE. Suppose we have a fluctuation function $\epsilon \rightarrow Q(\epsilon|g_n)$ so that we can solve for ϵ the equation:

$$P_n D^*(Q_n^0(\epsilon|g_n), g_n) = 0 \tag{3.6}$$

Then the efficient influence curve is satisfied in a single update and there is no need for iteration. This case corresponds to using the loss function $L(Q, g, (O_1, ..., O_n)) = \left|\frac{1}{n}\sum_{i=1}^{n} D^*(Q, g)(O_i)\right|^2$. In a single step, a solution can be found so the loss function takes its lower bound of 0.

It turns out that we can solve equation 3.6 without updating P_W by setting it to its empirical distribution $P_W = P_{W,n}$ of the baseline covariates. Thus, we need to solve

$$P_n D^*(Q_n^* = \{m_n^0(\epsilon), P_{W,n}\}, g_n) = 0 \tag{3.7}$$

where we drop the explicit dependency of $m_n^0(\epsilon)$ on g_n in the notation. Section 3.2 describes versions of this non-iterative estimator that use logistic and linear fluctuations for $m_n^0(\epsilon)$.

Step 3: Obtain final estimate $\beta_n^* = \Psi(m_n^*, P_{W,n})$.

Properties of TMLE.

See section 3.6 for sketches of proofs.

Efficiency

(See van der Laan and Robins 2003, and van der Laan and Rubin 2006.)

Theorem 1 *(Asymptotic efficiency.) Suppose all initial estimates (Q_n^0, g_n^0) are consistent, and that $D^*(Q_n^*, g_n^*)$ belongs to a P_0-Donsker class. Then the final estimate $\Psi(Q_n^*)$ is asymptotically efficient, with*

$$\Psi(Q_n^*) - \Psi(Q_0) = \left[P_n - P_0\right]\ D^*(Q_0, g_0) + o_P(1/\sqrt{n}) \tag{3.8}$$

Consistency under misspecification

Recall our notation $\Pi(Z, W) = E(A|Z, W)$ and $\rho(Z, W) = \Pr(Z|W)$. TMLE yields a consistent estimate for $\Psi^* = \beta_n^*$ under 3 scenarios of partial misspecification of components:

1. Initial estimates Π^0 and ρ^0 are consistent.
2. Initial estimates m^0 and ρ^0 are consistent.
3. Initial estimates m^0 and θ^0 are consistent.

Estimator using a logistic fluctuation for scalar ψ

This estimator has the advantage that it can match the bounds of the observed data in estimating $E(Y|W,\Pi(Z,W))$.

In accordance with the non-iterative TMLE procedure, we want to find ϵ such that $P_n D^*(Q_n^* = \{m_n^0(\epsilon), P_{W,n}\}, g_n) = 0$ according to equation 3.7.

A pre-processing step is done of converting Y-values to the range [0,1] using a linear mapping $Y \to \tilde{Y}$, where $\tilde{Y} = 0$ corresponds to $\min(Y)$ in the dataset and $\tilde{Y} = 1$ to $\max(Y)$. Thus, we can use the mapping $\tilde{Y} = (Y - \min(Y))/(\max(Y) - \min(Y))$. The equation $E(Y \mid \Pi(Z,W),W) = \Pi(Z,W)m(W)+\theta(W)$ can be written as $E(\tilde{Y} \mid \Pi(Z,W),W) = \Pi(Z,W)\tilde{m}(W)+\tilde{\theta}(W)$, where $\tilde{m}(W) = m(W)/(\max(Y) - \min(Y)) \in [-1,1]$ and $\tilde{\theta}(W) = (\theta(W) - \min(Y))/(\max(Y) - \min(Y)) \in [0,1]$. Now initial estimates can be formed of all relevant components and nuisance parameters using the modified data set $(W,Z,A,\tilde{Y})$.

Replacing $m_n^0(\epsilon)$ with $\tilde{m}_n^0(\epsilon)$, we use this fluctuation function in equation 3.7:

$$\tilde{m}_n^0(\epsilon)(W) = 2 \times \text{logistic}(\text{logit}(\frac{\tilde{m}_n^0(W)+1}{2}) + \epsilon^T \cdot h(W)) - 1 \tag{3.9}$$

where logistic() denotes the function $\text{logistic}(x) = \frac{1}{1+e^{-x}}$ and logit() its inverse $\text{logit}(y) = \log\frac{y}{1-y}$. This corresponds to the mapping $f(\epsilon) = \text{logistic}(\text{logit}(f) + \epsilon \cdot h)$ where f is $\tilde{m}_n^0$ scaled to be in $[0,1]$.

Inspecting the efficient influence curve, we have that the first term $P_n\ D_W^*(Q_n^*, g_n) = 0$, because this expression is equivalent to $\beta_n^* = \arg\min_\beta P_{W,n} j(V)\{m_n^*(W) - m_\beta(V)\}^2$, which holds by definition of β_n^*. Also, we have that the $+/-\ h(W)(\pi(Z,W) - E(\pi(Z,W) \mid W))(\pi(Z,W)m(W))$ terms cancel. Thus $D^*(Q,g)$ reduces to $h(W)(\pi(Z,W) - E(\pi(Z,W) \mid W))(Y - A \cdot m(W) - \theta(W))$ so we need to find ϵ such that

$$P_n\ D^*(\tilde{m}_n^0(\epsilon), P_{W,n}, g_n^0) =$$

$$\frac{1}{n}\sum_{i=1}^{n}\ h_n^0(W)(\pi_n^0(Z,W) - E_n^0(\pi_n^0 \mid W))(\tilde{Y} - A \cdot \tilde{m}_n^0(\epsilon)(W) - \tilde{\theta}_n^0(W)) = 0$$

for $\tilde{m}_n^0(\epsilon)(W)$ defined in 3.9.
Since $E_0(\tilde{Y} - A \cdot \tilde{m}_0(W) - \tilde{\theta}_0(W)|Z, W) = 0$, the equation above has a solution ϵ for any reasonable initial estimates ($Q_n^0 = \{\tilde{m}_n^0, P_{W,n}\}, g_n^0$). For $k = \dim(\beta)$, we have a k-dimensional equation in k-dimensional ϵ. When $k = 1$ and we need a scalar ϵ, we can use a bisection method as a computationally simple way to compute ϵ. One first finds left and right boundaries ϵ_1, ϵ_2 such that

$$E_n\ h_n^0(W)(\pi_n^0(Z,W) - E_n^0(\pi_n^0 \mid W))(A \cdot \tilde{m}_n^0(\epsilon_1)(W)) \leq$$
$$E_n\ h_n^0(W)(\pi_n^0(Z,W) - E_n^0(\pi_n^0 \mid W))(\tilde{Y} - \tilde{\theta}_n^0(W)) \leq$$
$$E_n\ h_n^0(W)(\pi_n^0(Z,W) - E_n^0(\pi_n^0 \mid W))(A \cdot \tilde{m}_n^0(\epsilon_2)(W))$$

where E_n denotes the empirical mean. Then one iteratively shrinks the distance between the left and right boundaries ϵ_1 and ϵ_2 until a suitably close approximation to the solution is found.

Once one solves for ϵ and finds $\tilde{m}_n^* = \tilde{m}_n^0(\epsilon)$, one converts back to the original scale for outcome Y, by setting $m_n^* = \tilde{m}_n^* \cdot (\max(Y) - \min(Y))$. Then the parameter of interest is evaluated by finding $\Psi(m_n^*, P_{W,n}) = \beta_n^*$.

When the parameter of interest ψ is vector-valued, solving the efficient influence curve equation using a logistic fluctuation translates to a non-convex multi-dimensional optimization problem with no known analytical solution. Various numerical techniques and software packages are available.

One application of this estimator is to use a tighter bound for $E(Y|\Pi(Z,W),W)$ than the bounds of the data. For instance, when Y is a rare binary outcome, its conditional mean for any value of W might lie in a far smaller interval than $[0,1]$.

Estimator using a linear fluctuation

Once again, we want to find ϵ such that $P_n D^*(Q_n^* = \{m_n^0(\epsilon), P_{W,n}\}, g_n) = 0$ according to 3.7. A TMLE-based estimator that is especially simple to understand and implement involves using a simple linear fluctuation

$$m_n^0(\epsilon)(W) = m_n^0(W) + h(W)^T \cdot \epsilon$$

and solving for ϵ in a single non-iterative step. $h(W) = c_0^{-1} \frac{h_1(V)}{\sigma^2(W)}$ as defined in lemma 2.

As usual, we form initial estimates of all relevant components and nuisance parameters. In solving the efficient influence curve equation 3.7, once again we have that $P_n\ D_W^*(Q_n^*, g_n) = 0$, and we can simplify to get

$$E_n\ h_n^0(W)(\pi_n^0) - E_n^0(\pi_n^0(Z,W) \mid W))(Y - A \cdot m_n^0(W) - \theta_n^0(W))$$

$$= E_n\ h_n^0(W)(\pi_n^0 - E_n^0(\pi_n^0(Z,W) \mid W))(A \cdot h_n^0(W)^T \epsilon)$$

We can solve for (generally vector-valued) ϵ by finding the solution to a simple system of linear equations. As usual, we then set $m_n^* = m_n^0(\epsilon) = m_n^0(W) + h(W)^T \cdot \epsilon$, and evaluate the parameter of interest $\psi_n^* = \Psi(m_n^*, P_{W,n})$ by finding the projection β_n^* of m_n^* unto the working model $\{m_\beta(v) : \beta\}$.

This approach is simple and achieves the same asymptotic guarantees as any of the other formulations of TMLE. However, it has the drawback compared to the version described above using logistic fluctuation that the final estimate $\mu_n^* = \Pi_n^0 \cdot m_n^* + \theta_n^0$ is not constrained to observe the bounds of Y in the data.

Estimator using iterative updating

One estimation method in the TMLE framework we developed involves iteratively updating relevant components and nuisance parameters until convergence to components (Q_n^*, g_n^*) such that the efficient influence curve equation is satisfied: $P_n\ D^*(Q_n^*, g_n^*) = 0$.

As usual, initial estimates are formed of all relevant components Q_n^0 and the nuisance parameters g_n^0. We set P_W to its empirical distribution $P_W = P_{W,n}$ and never update that component. Next, at each iteration until convergence, we fluctuate components as follows:

(i) Let k denote the iteration number. For $\mu = E(Y|W,Z)$, and $C_Y(Z,W) = h(W)(\pi(Z,W) - E(\pi(Z,W) \mid W))$ as defined in section 3.1 and lemma 2, we have:

$$\mu_n^{k+1} = \mu_n^k + \epsilon \cdot C_{Y,n}^k$$

$$\epsilon = \arg\min \sum_{i=1}^{n} (Y[i] - \mu_n^k[i] - \epsilon \cdot C_{Y,n}^k[i])^2$$

Note that by setting $m_n^{k+1} = m_n^k + \epsilon \cdot h_n^k$, and $\theta_n^{k+1} = \theta_n^k + \epsilon \cdot [-h_n^k E(\Pi_n^k|W)]$, where h_n^k refers to $h(W) = c_0^{-1} \frac{h_1(V)}{\sigma^2(W)}$, we have that $\mu_n^{k+1} = m_n^{k+1} \cdot \Pi_n^k + \theta_n^{k+1}$ and thus remains in our marginal structural model.

(ii) Given μ_n^{k+1}, we update $C_{A,n}^{k+1} = C_{Y,n}^k m_n^{k+1}$ and then fluctuate $\Pi_n^k(Z,W) = E_n^k(A|Z,W)$ as follows. If A is continuous, we first replace A with linear transformation $A' \in [0,1]$, where $A' = (A - \min(A))/(\text{range}(A))$, and apply the inverse transformation to get the final $\Pi(Z,W)$.

$$\Pi_n^{k+1} = \Pi_n^k(\epsilon) = \text{logistic}(\text{logit}(\Pi_n^k) + \epsilon \cdot C_{A,n}^{k+1})$$
$$\epsilon = \arg\min \sum_{i=1}^{n} \Big[-A[i] \cdot \log(\Pi_n^k(\epsilon)[i]) - (1 - A[i]) \cdot \log(1 - \Pi_n^k(\epsilon)[i]) \Big]$$

where the logistic function is $\frac{1}{1+e^-x}$ and the logit its inverse. The optimization above is solved using standard logistic regression software, even though the independent variable can be continuous in $[0, 1]$ here. We then update $C_{Y,n}^{k+1} = h_n^k \cdot (\ \Pi_n^{k+1} - E\big[\Pi_n^{k+1}(Z, W|W)\big]\ )$.

(iii) Finally, we update the $E(\Pi(Z, W)|W)$-component to be $E(\Pi_n^{k+1}(Z, W)|W)$, and $\sigma^2(W) = \text{Var}(\Pi(Z, W)|W)$ as $\text{Var}(\Pi_n^{k+1}(Z, W)|W)$, using the initial estimates for the relevant parts of $\Pr(Z|W)$.

This algorithm converges to components (Q_n^*, g_n^*). In each step of updating μ_n^k, Π_n^k, we are solving for ϵ^k to minimize $\sum_{i=1}^n L(P_n^k(\epsilon))(O_i)$, for some loss function L and parametric submodel $P(\epsilon)$. Thus we have $\frac{d}{d\epsilon} \sum_{i=1}^n L(P_n^k(\epsilon))(O_i)|_{\epsilon=\epsilon^k} = 0$. As the algorithm converges, we have that the objective $\sum_{i=1}^n L(P_n^*(\epsilon))(O_i)$ is minimized with $\epsilon = 0$; in other words, the components (Q_n^*, g_n^*) are already optimal for the loss function and do not get fluctuated. Thus, we have $\frac{d}{d\epsilon} \sum_{i=1}^n L(P_n^*(\epsilon))(O_i)|_{\epsilon=0} = 0$.

It is easy to check that for the loss function used to update μ_n^k, we have $\frac{d}{d\epsilon} L(P(\epsilon))|_{\epsilon=0} = C_Y \cdot (Y - \mu) = D_Y^*$, so we have $P_n\ D_Y^* = 0$ upon convergence. Similarly, for the loss function used to update Π_n^k, we have $\frac{d}{d\epsilon} L(P(\epsilon))|_{\epsilon=0} = C_A \cdot (A - \Pi) = D_A^*$, so we have $P_n\ D_A^* = 0$ upon convergence. We have that the first term $P_n\ D_W^* = 0$, because this expression is equivalent to $\beta_n^* = \arg\min_\beta P_{W,n} j(V)\{m_n^*(W) - m_\beta(V)\}^2$, which holds by definition of β_n^*. Thus, $P_n D^*(Q_n^*, g_n^*) = 0$ and we have a valid TMLE procedure.

3.3 Simulation results

We show results from a number of simulations. We compare all three versions of a TMLE-based estimator proposed above to several standard methods: 1) a likewise semiparametric, locally efficient estimator based on the method of estimating equations; 2) two-stage least squares, which is a standard parametric approach; 3) a biased estimate of the causal effect of A on Y ignoring the confounding.

There are two cases we use for the parameter of interest: *Scalar.* We estimate a constant mean causal effect $E(Y(1) - Y(0)) = E(m(W)) = \beta =$

$m_\beta(v)$. *Vector-valued linear.* We use a linear working model $m_\beta(w) = \beta^T \binom{1}{w}$ for $E(Y(1) - Y(0)|W) = m(W)$.

Standard approaches for comparison.

The method of estimating equations

(van der Laan and Robins 2003) presents background on the method of estimating equations. When the efficient influence curve is an explicit function of the parameter of interest Ψ_0, under regularity conditions, one can solve for Ψ using the equation

$$P_n D^*(P) = P_n \; D^*(P_W, \Pi, \mathrm{E}(\Pi|W), \mathrm{Var}(\Pi|W)m, \theta, \Psi) = 0$$

The components of D^* are estimated using Super Learner, just as with TMLE. Estimating equations has the same properties of local efficiency and robustness to misspecification as the TMLE-based estimators: when all relevant components and nuisance parameters are estimated consistently, the estimate is asymptotically efficient, and as long as $(P_W, \Pi_n^0, \mathrm{E}_n^0(\Pi|W), \mathrm{Var}_n^0(\Pi|W))$ are estimated consistently, the estimate for the parameter of interest $\Psi = \beta$ is consistent.

In the scalar case, our estimating equation is

$$E_n \left[c_0^{-1} j(V)(m(W) - \beta) + D_Y^*(P)(Y, Z, W) - D_A^*(P)(A, Z, W) \right] = 0$$

where the D_Y^*, D_A^* terms to do not depend on β.

For the case of a linear working model, the estimating equation is

$$E_n \left[c_0^{-1} j(V) \begin{pmatrix} 1 \\ W \end{pmatrix} (m(W) - \beta' \begin{pmatrix} 1 \\ W \end{pmatrix}) + D_Y^*(P)(Y, Z, W) - D_A^*(P)(A, Z, W) \right] = 0$$

which can also be solved as a linear equation of β. The terms D_Y^*, D_A^* do not depend on β and are vector-valued here.

Two-stage least squares

The most widely used solution to estimating the effect of a treatment on an outcome in the presence of a confounder and valid instrument is to use a linear model for both the "first-stage" equation $A = \alpha_Z Z + \alpha_W W + \alpha_1 1 + \epsilon_A$ and the "second stage": $Y = \beta_A A + \beta_W W + \beta_1 1 + \epsilon_Y$. When there is a single instrumental variable and treatment, which is the case we study, a solution for scalar $\hat{\beta}$ that is consistent and asymptotically optimal among linear models is $\hat{\beta} = ((Z, W, 1)'(A, W, 1))^{-1}((Z, W, 1)'Y)$. This estimate corresponds to the two-stage least squares solution where one estimates $A^* = E(A|Z, W)$ using a linear model, and then estimates the effect of (A^*, W) on Y using a linear model again (having exogenous variation).

When estimating a vector-valued causal effect, we find $A^* = E(A|Z, W)$ and then do linear regression of Y on cross terms $A^* \times (1, W)$ and covariates $(1, W)$, thus finding a linear treatment effect modifier function $m(W)$ and a linear additive effect function $\theta(W)$. 2SLS is a parametric model and is in general not consistent for estimating our causal parameter of interest.

Ignoring the confounding

We include a "confounded" estimator in each table that ignores the unmeasured confounding between the treatment and outcome, and does not use an instrument. We use a correctly specified parametric model for $m(W)$, $\theta(W)$, and estimate their parameters using $E(Y|W, A) = A \cdot m(W) + \theta(W)$, which will give a biased estimate for $m(W)$ by ignoring the confounding between A and the residual term. The correctly specified model for $m(W)$ converges at a parametric rate, and for large n, we isolate the effect of the bias arising from not using an instrument.

Initial estimates.

For the semiparametric approaches (our three estimators based on TMLE, and estimating equations), initial estimates are formed as follows. We use the empirical distribution of W for P_W and never update this component. For $\mathrm{Var}_n^0(\Pi_n^0(Z, W)|W)$ and $E_n^0(\Pi_n^0(Z, W)|W)$, noting that our instrument Z is binary in the simulations below, we estimate $P(Z = 1|W) = E(Z|W)$ and find the expectation and variance of $\Pi_n^0(Z, W)$ from $P(Z = 1|W)$, instead of directly estimating them as a function of Z, W. Thus we need initial estimates for $E(Z|W), \Pi(Z, W), \theta(W), m(W)$ from the data.

For $\Pi(Z, W)$ in cases where A is binary, and for $E(Z|W)$, we use as candidate learners the following R packages (see the corresponding function specifications in Super Learner): **glm, step, knn, DSA.2, svm, randomForest** (Sinisi and van der Laan 2004). For glm (generalized linear models), step (stepwise model selection using AIC), and svm (support vector machines), we use both linear and second-order terms. In addition, we use cross-validation to find the highest degree of polynomial terms in glm that results in the lowest prediction error, thus using terms of degree higher than two with glm. For

$\Pi(Z, W)$ in cases where A is continuous, we use candidate learners **glm, step, svm, randomForest, nnet** and **polymars**.

For $m(W)$ and $\theta(W)$ which involve continuous outcomes, we use candidate learners **glm, step, svm,** and **polymars**. We need to predict $m(W)$ and $\theta(W)$ so that $\mu(Z, W) = \pi(Z, W) \cdot m(W) + \theta(W)$ retains the structural form. We include $\Pi \times m(W)$ cross-terms as well as $\theta(W)$ terms, having various functional forms for parameterizing $m(W)$, $\theta(W)$.

Results.

In the simulations that follow, we use the following general format for generating data. In accordance with R's notation, the right-hand side of the formulas specify the regressors but leave the link function unspecified. ϵ_{AY} is a confounding term, while the treatment effect modifier function m_W can be highly non-linear.

$$\begin{aligned} W &\sim N(\mu, \Sigma) \\ Z &\sim Binom(p(W)) \\ A &\sim W + Z + \epsilon_{AY} \\ Y &\sim A \cdot m(W) + \theta(W) + \epsilon_{AY} \end{aligned}$$

Nonlinear design 1

We test our estimators in the case of highly nonlinear treatment effect modification $m(W) \sim e^W$ in tables 6.1 and 6.2. As we show, 2SLS can be extremely biased in recovering the correct projection of $m(W)$ unto a linear working model. We use $W \sim N(3, 1)$, $p = .5$ for Z, and a continuous treatment generated as a linear function of its regressor terms.

Scalar parameter. (Table 6.1.) The true effect is 33.23, sample size of $n = 1000$ is relatively small for using an instrumental variable, and 10,000 repetitions are made. The "initial substitution" estimator is formed by substituting the estimates of relevant components into the parameter of interest, which is just $\beta_n^0 = \Psi(Q_n^0) = E_{W,n} m_n^0(W)$ here, or the estimated mean treatment effect. When consistent initial estimates are formed of all components of D^* using Super Learner, we observed a bias of just .0038, and variance of .6990 for the initial substitution estimator. The three new methods all performed very similarly, achieving lower bias than the initial substitution estimator, as well as slightly lower variance. Since all relevant components are consistently specified, the TMLE-based estimators are asymptotically guaranteed to have the lowest possible variance within the class of consistent estimators in our semiparametric model. The same asymptotic guarantees hold for the estimating equations estimator, which achieves similar magnitude bias and slightly higher variance than the TMLE-based estimators. The two-stage least squares (2SLS) estimator, in contrast, achieves not only much higher bias but vastly higher variance than the semiparametric estimators, even though it is a parametric estimator. The highly misspecified linear model that 2SLS fits for the conditional outcome brings about the bias and large finite-sample variance. Finally, the estimate that ignores confounding has a bias of about 21.

In table 6.2, we use an inconsistent initial estimate of $Q(W, R)$, namely, we fit an incorrect linear model $m(W) = b'\binom{1}{W}$. Thus, the substitution estimator essentially functions like 2SLS. The confounded and 2SLS estimators are unchanged. The TMLE-based estimators often show bias removal at the expense of some increase in variance as compared to the unfluctuated initial substitution estimator in the case of misspecification. However we don't see that here with the modest sample size (n=1000), for which the initial substitution estimator has fairly large variance in this simulation. Also, in this case of a scalar parameter, the bias of the initial estimator was quite small (less than 2%). Performing the TMLE fluctuation step causes neither an improvement nor substantial decline in performance here.

Vector-valued parameter. For the projection of $m_0(W)$ unto a linear working model, the true two-dimensional parameter of interest is $[-64.2, 32.3]$.

2SLS solves the following optimization in the second stage: $\arg\min_{\beta_1,\beta_2} \sum_{i=1}^{n}(Y - \Pi(Z, W)\beta_1^T\binom{1}{W} - \beta_2^T\binom{1}{W})^2$. β_1 is output as the parameter of interest. It is easy to check that this can give a very different solution than a semiparametric approach which estimates a function $m(W)$ that can take a variety of functional forms, and then solves $\beta = \arg\min \sum_{i=1}^{n}(m(W) - \beta^T\binom{1}{W})^2$. Specifically, let $\epsilon_\beta(W) = m(W) - \beta^T\binom{1}{W}$ denote the vector of residuals in approximating $m(W)$ by $\beta^T\binom{1}{W}$. Then in the case of a linear $\theta(W)$, 2SLS solves $\arg\min_\beta \sum_{i=1}^{n}(\Pi(Z, W)\epsilon_\beta(W))^2$, while the semiparametric approach solves $\arg\min_\beta \sum_{i=1}^{n}(\epsilon_\beta(W))^2$.

We see in table 6.2 that 2SLS has a mean absolute bias of around 136. A typical value for its estimate is $[-224, 90]$. It is useless for estimating our parameter of interest without knowing the functional form for $m(W)$ a priori. The confounded estimator that is fully correctly specified in its functional forms but ignores confounding has a bias of roughly 10. All the semiparametric approaches achieve very low bias when initial estimates are consistent. Furthermore, they all achieve similar and low variance for a large sample size, as the $n = 10000$ column shows. For the sample sizes in our simulation, the 2SLS estimator is not only extremely biased, it also has larger variance than the semiparametric estimators, due to the large mismatch between the second-stage linear model it fits and the data-generating process.

The right-hand side of table 6.2 shows an incorrect linear fit for $m(W)$ to form an inconsistent initial estimate of $\mu(Z, W)$. The initial substitution estimator works essentially like two-stage least squares in this case. We deliberately start with this enormously biased initial estimator to see if the semiparametric estimators can remove bias sufficiently. Indeed, we see very low finite-sample bias for the three semiparametric consistent estimators. The iterative TMLE-based approach performs best here, with mean absolute bias around just .25 at $n = 10000$ (compared to a mean absolute effect around 48). Furthermore, while the variance of the semiparametric consistent estimators can be an order of magnitude higher than for the initial substitution estimator when $n = 1000$, the variances are at a comparable scale for $n = 10000$.

Scalar effect, nonlinear design 2

In table 6.3, we generate a continuous outcome such that $E(Y|Z, W)$ lies within sharp boundaries covering a much smaller range than Y. TMLE using the logistic fluctuation has been shown to be especially effective with similarly generated data, where the data or conditional outcome falls within sharp cutoffs (Gruber and van der Laan 2010).

We use a 3-dimensional $W \sim N(1, 1)$, $p = .5$ for Z, a binary treatment generated using the binomial link function. The confounding term is $\epsilon_{AY} \sim N(0, 5)$. $m(W)$ and $\theta(W)$ are continuous, and they each have the form $a \cdot plogis(\beta W) + b$, for some constants a, b. Thus, $m(W)$ and $\theta(W)$ fall within some bounds $[b, a + b]$. Furthermore, the parameters are set so that many values for each function are close to the boundaries.

The true effect is 1.00, and we use $n = 1000$. We see that without using an instrument, the estimate is confounded by more than 50%. For the case of consistently specifying all initial estimates, we include the correct parametric form for $E(Y|Z, W)$ in Super Learner's library. In this case the initial substitution estimator has both lowest bias and lowest variance. The logistic fluctuation and estimating equations estimators also do well with relatively low bias and variance, followed by the iterative and linear fluctuation TMLE, and finally, 2SLS has the highest MSE of the unconfounded estimators. In the right hand of table 6.3, we misspecify the initial estimate for $E(Y|Z, W)$ as a second-order polynomial. In this case, TMLE using logistic fluctuation is the clear winner. It achieves an MSE (dominated by the variance) of .34, compared to roughly .45 for the other semiparametric approaches. It also achieves a large reduction in bias for minimal gain in variance compared to the initial substitution estimator.

Vector-valued effect, linear model

In table 6.4, we use a linear model for $m(W)$, so that two-stage least squares with the correctly specified cross terms $\Pi(Z, W) \times W$ estimates $\mu(Z, W)$ consistently. Here we use a 3-dimensional covariate $W \sim N(2, 1)$, Z is binary and of the form $E(Z|W) = \text{plogis}(\alpha' W + \alpha_0)$. Treatment A is also binary and uses the logit link function; $m(W) = \beta^T \binom{1}{W}$.

We see that although 2SLS uses the correct second-stage specification for $E(Y|W, Z)$, it remains slightly biased for all n, with .2 mean absolute bias (about 17%), since $E(A|W, Z)$ uses a nonlinear link function. The confounded estimate has (mean absolute) bias of .34. The semiparametric consistent estimators have much lower bias than 2SLS even for $n = 1000$, with linear fluctuation and estimating equations achieving lower bias than the initial substitution estimator. The table reflects the roughly $\sqrt{n}$ decrease in bias of the consistent estimators and decrease in SD of all estimators. The initial substitution estimator has just slightly higher SD than 2SLS, as the former chooses the correct linear model from a library of methods.

When we use an inconsistent initial estimate for $\mu(Z, W)$: one of the coefficients in β is fixed to an incorrect value and then a linear model is fit (Super Learner is only used for estimating $Pr(Z|W)$, $\Pi(Z, W)$). This makes for a mean absolute bias of roughly 1.5 in the initial substitution estimator (corresponding to an error of 100%). The three semiparametric consistent estimates successfully remove bias; the two TMLE-based approaches have particularly low bias (about 94% of the bias is removed for $n = 10000$). The semiparametric estimates have mean SD's of only around .3 for n=10,000 where mean absolute effect is 1.5. The linear fluctuation TMLE-based estimator performs the best overall, with lowest bias and variance for large samples.

Confidence intervals

Table 6.5 shows 95% confidence intervals corresponding to tables 6.2, 6.4. These are calculated separately for each component of the vector-valued parameter of interest. For the semiparametric estimators, as proved in (van der Laan and Rubin 2006), the following equation holds:

$$\Psi(Q_n^*)-\Psi(Q_0) = \left[P_n-P_0\right] D^*(Q_n^*, g_n^*)-\left[P_n-P_0\right]\text{Proj}(D^*(Q_n^*, g_n^*)|\text{Tang}(g_0))+o_P(\frac{1}{\sqrt{n}})$$

Here $\text{Proj}(D^*(Q_n^*, g_n^*)|\text{Tang}(g_0))$ denotes the projection of the efficient influence curve D^* unto the tangent space of nuisance parameters, $T(g_0)$. It thus follows that a conservative estimator for the variance of $\beta_n^* = \Psi(Q_n^*)$ is the variance of $D^*(Q^*, g^*)$. Note that when all its components are consistently estimated, under regularity conditions, $[P_n - P_0]\ D^*(Q^*, g^*) = [P_n - P_0]\ D^*(Q_0, g_0) + o_P(\frac{1}{\sqrt{n}})$, and thus the semiparametric efficiency bound is achieved. For the three semiparametric consistent estimators, shown at the top of the list in table 6.5, we use the estimated variance of the efficient influence curve $D^*(Q^*, g^*)$ to calculate confidence interval width. For the other three estimators, we simply use the empirical variance. For these cases, we demonstrate that even when "cheating" by accurately knowing the correct width of the confidence intervals, coverage is still very poor due to the bias of the estimators.

We see that for all three semiparametric estimators, the coverage is generally overestimated, as the theory suggests, but is usually not too far from 95%. For the case of consistent initial estimates, coverage is around 96% when estimating a linear treatment effect and closer to 97% when estimating a nonlinear effect. Similar results holds when using misspecified initial estimates; however, estimating equations has poor coverage (in the 80's) due to finite-sample bias. The initial substitution estimator is consistent when the initial estimates of components are; however, it has coverage slightly below 95% even when using the empirical variance to estimate the variance. This could be due to its not being normally distributed. When the initial estimates of components is not consistent, the initial substitution estimator can be heavily biased, and we see 0 coverage for most columns, even using an accurate variance. Likewise the large bias of the confounded and 2SLS estimators for the case of the nonlinear treatment effect causes 0 coverage. When a linear treatment effect is estimated, both the confounded and 2SLS estimators exhibit poor coverage that deteriorates with n. In the case of 2SLS, the bias is due to the mismatch between the linear model and the nonlinear distribution of the conditional treatment $\Pi(Z, W)$.

3.4 Application to a dataset: estimating the effect of parents' education on infant health

We apply our TMLE-based estimators in the context of a program that expanded schooling in Taiwan. In 1968, Taiwan expanded mandatory schooling from 6 years to 9 years, and more than 150 new junior high schools were opened in 1968-1973 to accommodate this program. Prior to this expansion of schools, enrollment in junior high was based on a competitive process in which only part of the population of 12-14 year-old children was accepted. There is significant variation in how much the schooling expansion program affected an individual's access to education based on the individual's birth cohort and county of residence. In counties where there were previously relatively few educated people and spots in school beyond grade 6, many new junior high schools were opened per child. Thus, program intensity as a function of birth cohort and county serves as an instrumental variable that causes exogenous variation in people's educational attainment. This lets one make a consistent estimate of the effect of parents' education on their child's health.

The school expansion program caused junior high enrollment to jump from 62% to 75% within a year in 1968, before leveling off around 84% in 1973.

We use the same dataset as (Chou et al 2010). The treatment variable is either the mother's or the father's education in years (starting from first grade). There are four outcomes we study: low birth weight (< 2500g), neonatal mortality (in the first 27 days after birth), postneonatal mortality (between day 28 and 365), and infant mortality (either neonatal or postneonatal). The instrument is the cumulative number of new junior highs opened in a county by the time a birth cohort reaches junior high, per 1000 children age 12-14 in that year. This serves as a proxy for the intensity of the school expansion program for a particular birth-county cohort. The data is taken by checking every birth certificate for children born in Taiwan between 1978 and 1999. The birth certificates list for both parents their ages, number of years of education, and county of birth (which we use as a usually correct guess for the county in which the parent went to school), as well as the incidence of low birth weight. Birth certificates are matched to death certificates from a similar period using a unique identification number issued for each person born to ascertain if an infant death has taken place. The previous study done on this dataset (Chou et al 2010) used standard OLS and 2SLS, which are sensitive to highly collinear regressors, and as a result separately estimated the effect of father's and mother's education on infant's health. To ease comparison with prior results, we do the same here. Only datapoints where the father was born in [1943-1968], or the mother in [1948-1968] were included in the study. Those points where the parent was at most 12 years old in 1968 constitute the treatment group, and the rest the control group where the instrument Z is 0 (for those who were unaffected by the school reform). This resulted in a sample size of about 6.5 million, of which roughly 4 million were in the treatment group, for either case of parent.

We reestimate (Chou et al 2010)'s scalar effect estimates using our TMLE-based approach. We also give previously unpublished estimates of treatment effect heterogeneity as a function of the parent's and children's birth cohorts.

The usefulness of the semiparametric approaches depend on the $\sigma^2(W) = \mathrm{Var}(\Pi(Z,W)|W)$ term being large (recall $\Pi(Z,W) = E(A|Z,W)$). This term captures the strength of the instrument in predicting the treatment given W, and the variance of the instrument-based estimators blow up when σ^2 is small. Our instrument only depends on the parent's birth cohort and the county, so σ^2 would be 0 if we include both these variables in W. Since most variation in Z is by county (of parent's birth), we do not include county in W, and use as covariates W only parent's and child's birth cohort, coding these as dummy variables. In addition, we remove datapoints where $\sigma^2(W) = 0$, which corresponds to including only points where the parent was born after or in 1956. People born earlier were unaffected by the schooling reform.

We need to check that county (parent's county of birth) does not serve as a confounder causing U_Z, U_Y to be correlated. Using modified outcome Y' (the log-odds ratio for a binary health outcome, see below), we compare the between-county vs the within-county variation. We see that, for any of the 4 health outcomes, and using either mother's or father's county, fixing W, at most 1.1% of the variation in Y' is between-county, but on average only .5%. Thus, we can rule out that confounding from county will effect our estimates. The IV-assumptions in section 3 are satisfied.

Table 3.1 shows summary statistics. Note that for the outcome of post-neonatal mortality, we only include datapoints where the child survived the neonatal period.

We perform our TMLE-based estimates using the noniterative, linear-fluctuation estimator, as this was found to perform well across multiple simulations, and had low bootstrap variance on the data, suggesting a good fit. We use the same library of initial estimates described above in section 3.3, and the empirical distribution for the probability of a county given the birth cohorts, $\Pr(Z|W)$. Since our outcomes are binary with relatively few positives, and the covariates are indicator variables that divide the dataset into cells, we modify our dataset $(W, Z, A, Y) \longrightarrow (W, Z, \bar{A}, Y')$ when forming initial estimates

Table 3.1: Means and SDs of variables.

	Mothers	Fathers
Sample size	4,101,825	4,001,970
Program intensity (R)	0.22	0.22
	(0.11)	(0.11)
Parent's years of schooling	9.93	10.67
	(1.46)	(1.15)
Percentage of low-birthweight births	4.50	4.80
	(1.24)	(1.25)
Neonatal mortality (deaths per thousand births)	2.32	2.33
	(2.38)	(2.38)
Postneonatal mortality (deaths per thousand neonatal survivors)	3.50	3.38
	(2.56)	(2.71)
Infant mortality (deaths per thousand births)	5.81	5.71
	(3.58)	(3.67)

Note: the SDs for the binary outcomes (low birth weight, and mortality) are the SD's for the average rates within each cell (in which county, and parent and child's birth cohorts are fixed). Each cell is weighted by its sample size for the relevant outcome (for example, the total number of births in a cell for infant mortality) in finding the SD.

$\Pi(Z, W)$, $m(W)$, $\theta(W)$. $\bar{A}_i$ is the average value of education A in the i^{th} cell given by the parent and child's birth cohorts and the county (thus, fixing W and Z). Y' is the log-odds ratio given by Cox's modified logistic transformation: $Y_i' = \log \frac{N_i+.5}{D_i-N_i+.5}$, where there are D_i total points in the i-th cell, and N_i of these are 1, for one of the four outcomes of interest.

Tables 6.6-6.7 give estimates of the scalar treatment effects. For the OLS and 2SLS estimates, we include the parent and child's birth cohorts as covariates, with heteroscedasticity-robust standard errors (White's method as implemented in R's sandwich package).

We use the final semiparametric model of the components that TMLE fits ($P_W, \Pi(Z, W)$, etc...), as well as a linear 2SLS model to estimate the number of adverse infant health outcomes prevented by schooling reform. Using our modified log-odds outcome $Y_i' = \log \frac{N_i+.5}{D_i-N_i+.5}$, where i indexes a cell, we estimate the counterfactuals for Y' without the schooling reform, denoted $Y'(Z = 0)$.

We have $Y'(Z = 0) = m(W)\Pi(Z = 0, W) + \theta(W)$, where $\Pi(Z = 0, W)$ estimates the counterfactual $E(A(Z = 0)|W)$. Then we convert from $Y_i'(Z = 0)$ to $N_i(Z = 0)$, which is the (counterfactual) number of adverse outcomes in a cell. $\Delta N = \sum_{cells} N_i(Z = 0) - N_i$ gives the estimated total reduction in an adverse outcome from the schooling reform. We also show the linear 2SLS model's estimate. In this case, $Y'(Z = 0)$ simplifies to $(1, (1, 0, W)'\beta_1, W)^T(\beta_2)$, where β_1, β_2 are the first- and second-stage coefficients, indexing $(1, Z, W)^T$, and $(1, A, W)^T$, respectively.

As tables 6.6-6.7 show, estimates of the scalar effect of (a parent's) education on the log-odds ratio of (infant's) health outcome range from -.2 to -1.0. The estimated percent reduction in adverse outcomes ranges from 1.5% for low birthweight (father's education is treatment, TMLE is the estimator) to 16.7% for neonatal mortality (with mother's education, TMLE estimator). The results imply a significant human benefit from the schooling reform regarding health: our TMLE estimator estimates roughly 1850 infant deaths were spared as an indirect effect of schooling reform.[2] The TMLE estimator finds a significantly greater reduction in adverse outcomes than 2SLS when the outcome is neonatal mortality and mother's education is the treatment, and for infant mortality when father's education is the treatment. TMLE and 2SLS yield similar estimates for the effect for low-birthweight/mother's-education and postneonatal-mortality/father's-education, while TMLE gives a somewhat lower estimate than 2SLS in the remaining two cases. The beneficial effect of father's education on infant and postneonatal mortality was highly significant for either estimation method, while the effect of mother's education on neonatal mortality was highly significant only for the TMLE estimator.

The use of a library of learners (TMLE) instead of a linear parametric model (2SLS) is reflected in the better fit and higher cross-validated R^2 value achieved for both stages. The "first stage" of the method of instruments refers to fitting $\Pi(Z, W)$ in our semiparametric model, and the "second stage" to fitting $\mu(Z, W)$. Especially for the second stage of father's education, there is a large gain in R^2 of .2-.3 from using data-adaptive learning. SuperLearner chooses a least-squares linear model with largest weight in every case; however, our semiparametric model for $E(Y|Z, W) = \Pi(Z, W)m(W) + \theta(W)$ even when $m(W)$, $\theta(W)$ are set to be linear in W is more flexible than the standard linear 2SLS model $E(Y|Z, W) = \beta_A E(A|Z, W) + \beta_W^T \binom{1}{W}$, and we include quadratic terms in W. Support vector machine is also chosen with large weight for both stages, and Random Forest for the first stage. The instrument is slightly

[2]This estimate is made using semiparametric, TMLE-based estimates of the effect of father's education on reducing infant mortality in the treated population.

stronger for predicting mother's education than father's, which might explain the higher first-stage R^2 values for mother's education.

As expected, our semiparametric estimator typically has higher variance than 2SLS; however, this is not always true, as TMLE achieves a better fit, which can make for a lower variance despite the added complexity of choosing from various learners.

We had expected that OLS would be biased from unmeasured confounding between a parent's education and his/her infant's health. One would expect that confounding factors would increase parents' education and decrease adverse health effects, or vice versa, biasing the OLS estimates to overestimate the beneficial effects of education. Surprisingly, we saw that the OLS estimates were smaller in magnitude that either of the instrument-based estimates for several columns in our table. One possible explanation is there might not have been significant unmeasured confounding. Indeed, the Hausman-Wu (F-) test for exogeneity gives low evidence in support of confounding.

In tables 6.8-6.9, we estimate the treatment effect modification, where the parent's or child's membership in a particular birth cohort is a modifier (given by the dummy variables W). As before, we estimate a vector-valued parameter β using a linear working model $\{m_\beta(W) = \beta^T \binom{1}{W} | \beta\}$. Since all covariates in W are binary indicators for birth cohorts, the coefficients in β can be directly compared to one another to reflect treatment effect modification.

The six effect modifiers that are largest in magnitude for each case are shown. The child being born in the late 70's or 80's often corresponded to a substantial increase in the beneficial effects of parent's education. The father being born in 1965, 1967, or 1968 corresponded to increased beneficial effects of his education on his infant's mortality. However, the effect of mother's education on her child's good health was found be diminished for babies born in 1998 or 1999. Virtually all the treatment effect modifiers shown are highly significant for mothers, as well as for fathers when postneonatal mortality is the outcome. The largest magnitude effect modifiers were not necessarily the most statistically significant ones, so the treatment effect modifiers are summarized for each case both as original and as standardized values (effect modifier $\div$ SE). There were roughly 33 total effect modifiers. We see that for some cases, a significant fraction of the effect modifiers had a coefficient of statistically significant magnitude (neonatal mortality for mothers, and low birthweight and postneonatal mortality for fathers).

3.5 Discussion

Our simulations reflect that even compared to a parametric estimator for the scalar effect of interest, such as two-stage least squares, the semiparametric efficient estimates can have both lower bias and far lower variance due to the better fit with relevant components of the data-generating distribution. We also showed that two-stage least squares can be enormously biased when estimating a vector-valued parameter, while TMLE is very effective at removing bias with only a moderate gain in variance in finite samples. Using TMLE with a logistic fluctuation can give the best performance when the conditional mean of the outcome follows sharp cutoffs, and each of the three TMLE-based estimators we describe has datasets on which is it the strongest performer. Finally, using the (estimated) variance of the efficient influence curve to estimate the standard error gives confidence intervals that are just slightly conservative. The confidence intervals based on TMLE can perform better than those based on a conventional semiparametric estimator.

We performed an extensive data analysis estimating the effect of parents' education on their infant's health in the context of a schooling reform in Taiwan. We identified a number of birth cohorts, pertaining to either the parent or the infant, that significantly increased, or decreased, the beneficial effect of education on health.

Several avenues for future work are of interest. One is to work with instrumental variables in the context of more complex causal models, such as when there are multiple instruments and treatments. This may for example occur in the setting of longitudinal data where each time point has an instrument, or in the context that a multivariate instrument is used to control for a multivariate confounded treatment. A number of extensions are of interest along empirical lines as well. For instance, future work could apply our methods to data having a very high-dimensional covariate space W, where V is a tiny subset of W, in finding the effect of the treatment given V.

3.6 Proofs

Proofs of properties of the TMLE-based estimators

Consistency under partial misspecification.

TMLE is constructed so that the efficient influence curve equation holds. We can explicitly write this as a function of the final estimate $\Psi^* = \beta^*$ using the definition of β. Thus we have

$P_nD^*(Q^*, g^*, \beta^*) = 0$ (we drop the n-subscript notation here). Since $P_nD^*(Q^*, g^*, \beta^*)$ converges to $P_0D^*(Q', g', \beta')$, where $\{Q', g', \beta'\}$ are the components in the limiting distribution, when the true parameter of interest $\beta' = \beta_0$ solves $P_0D^*(Q', g', \beta') = 0$, for some case of consistent specification of some of $\{Q', g'\}$, then we have that $\beta^* \longrightarrow \beta_0$ for our TMLE estimators.

Simplifying slightly, we get that $P_0D^*(Q, g, \beta) = 0$ reduces to

$$P_0\ c_0^{-1}h_1(m - m_\beta) \tag{3.10}$$

$$+\ P_0\ c_0^{-1}\frac{h_1}{\sigma^2}(\Pi - E(\Pi))((m_0 - m)\Pi_0 - (\theta_0 - \theta)) \tag{3.11}$$

$$= 0 \tag{3.12}$$

TMLE yields a consistent estimate for $\Psi^* = \beta^*$ under 3 scenarios of partial misspecification of components given below, with the reasoning sketched. Note that for the non-iterative versions of TMLE, only m^0 is updated, and the initial estimates are the same as the final estimates for the other components. For iterative TMLE, it is easy to check that when an initial estimate for a component is consistent, so is the final estimate (i.e. at every step k, $\epsilon_A^k \to 0$ when Π^0 is consistent).

1. Initial estimates Π^0 and $\Pr^0(Z|W)$ are consistent.

   We have $E_0(\ (\Pi_0(Z, W) - E_0(\Pi_0)|W)\Pi_0(Z, W)\ |W) = \sigma^2(W)$ since $\Pi, E\Pi|W$ are correctly specified. Also, since $E_0(\ (\Pi_0(Z, W) - E_0(\Pi_0)|W)\ |W) = 0$, the term involving $(\theta_0(W) - \theta(W))$ is 0 in expectation. Thus, 3.11 reduces to
   $P_0\ c_0^{-1}h_1(m_0 - m)$, so $P_0D^*(Q, g, \beta) = 0$ becomes $P_0\ c_0^{-1}h_1(m_0 - m_\beta) = 0$, and this is solved by $\beta = \beta_0$ by definition of β.

2. Initial estimates m^0 and $\Pr^0(Z|W)$ are consistent.

   The term in 3.11 involving $(m_0 - m)$ is 0 by the consistency of m, and the term involving $(\theta_0 - \theta)$ is also 0 since $E_0(\ \Pi(Z, W) - E_0(\Pi(Z, W)|W)\ |W) = 0$. Thus, we have $P_0\ c_0^{-1}h_1(m_0 - m_\beta) = 0$, which is solved by $\beta = \beta_0$ by definition of β.

3. Initial estimates m^0 and θ^0 are consistent.
   3.11 goes to 0 because both $m_0 - m = 0$, $\theta_0 - \theta = 0$. The rest of the reasoning is the same as above.

Efficiency under correct specification of all relevant components and nuisance parameters.

(See van der Laan and Robins 2003, and van der Laan and Rubin 2006.)

Suppose all initial estimates (Q_n^0, g_n^0) are consistent, and that $\mathrm{Var}_0(D^*(Q_n^*, g_n^*) - D^*(Q_0, g_0)) \in o_P(1)$. Then the final estimate $\Psi(Q_n^*)$ is asymptotically efficient, with

$$\Psi(Q_n^*) - \Psi(Q_0) = \left[P_n - P_0\right] D^*(Q_0, g_0) + o_p(1/\sqrt{n}) \tag{3.13}$$

Sketch of proof: Note that when all initial estimates are consistent, then so are all final estimates (Q_n^*, g_n^*). In the non-iterative case, only $m_n^0(W)$ is updated and $m_n^* \to m_n^0$ when the other components are consistent (see Consistency proof above). Using the definition of the canonical gradient D^* at (Q_n^*, g_0) and calculating a remainder term (see van der Laan and Robins 2003), we have

$$\Psi(Q_n^*) - \Psi(Q_0) = -P_0\ D^*(Q_n^*, g_n^*) + o_p(1/\sqrt{n}) \tag{3.14}$$

Using the key property of TMLE that $P_n D^*(Q_n^*, g_n^*) = 0$, and the fact that $\mathrm{Var}_0(D^*(Q_n^*, g_n^*) - D^*(Q_0, g_0))^2 \in o_P(1)$ when all components $\{Q_n^*, g_n^*\}$ are consistent, and the Donsker class assumption in theorem 1 we get

$$\begin{aligned}
&\Psi(Q_n^*) - \Psi(Q_0)\\
&= \left[P_n - P_0\right] D^*(Q_n^*, g_n^*) + o_p(1/\sqrt{n})\\
&= \left[P_n - P_0\right] D^*(Q_0, g_0) + \left[P_n - P_0\right](D^*(Q_n^*, g_n^*) - D^*(Q_0, g_0)) + o_p(1/\sqrt{n})\\
&= \left[P_n - P_0\right] D^*(Q_0, g_0) + o_p(1/\sqrt{n})
\end{aligned}$$

Efficient influence curve of target parameter

We determine the efficient influence curve of $\Psi : \mathcal{M} \to \mathbb{R}^d$ in a two step process. Firstly, we determine the efficient influence curve in the model in which Π_0 is assumed to be known. Subsequently, we compute the correction term that yields the efficient influence curve in our model of interest in which Π_0 is unspecified.

Efficient influence curve in model in which Π_0 is known.

First, we consider the statistical model $\mathcal{M}(\pi_0) \subset \mathcal{M}$ in which $\Pi_0(Z, W) = E_0(A \mid Z, W)$ is known. For the sake of the derivation of the canonical gradient, let $W \in \mathbb{R}^N$ be discrete with support $\mathcal{W}$ so that we can view our model as a high dimensional parametric model, allowing us to re-use previously established results. That is, we represent the semiparametric regression model as $E_0(Y \mid Z, W) = \Pi_0(Z, W) \sum_w m_0(w) I(W = w) + \theta_0(W)$ so that it corresponds with a linear regression $f_{m_0}(Z, W) = \Pi_0(Z, W) \sum_w m_0(w) I(W = w)$

in which m_0 represents the coefficient vector. Define the N-dimensional vector $h(\Pi_0)(Z,W) = d/dm_0 f_{m_0}(Z,W) = (\Pi_0(Z,W)I(W=w) : w \in \mathcal{W})$. By previous results on the semiparametric regression model, a gradient for the N-dimensional parameter $m(P)$ at $P = P_0 \in \mathcal{M}(\pi_0)$ is given by

$$D^*_{m,\Pi_0}(P_0) = C(\pi_0)^{-1}(h(\Pi_0)(Z,W) - E(h(\Pi_0)(Z,W) \mid W))(Y - f_{m_0}(Z,W) - \theta_0(W)),$$

where $C(\pi_0)$ is a $N \times N$ matrix defined as

$$\begin{aligned} C(\pi_0) &= E_0\{d/dm_0 f_{m_0}(Z,W) - E_0(d/dm_0 f_{m_0}(Z,W) \mid W)\}^2 \\ &= E_0\{(I(W=w)\{\Pi_0(Z,W) - E_0(\Pi_0(Z,W) \mid W\} : w\}^2 \\ &= \mathrm{Diag}\left(E_0\{I(W=w)\{\Pi_0(Z,W) - E_0(\Pi_0(Z,W) \mid W=w)\}^2\} : w\right) \\ &= \mathrm{Diag}\left(P_{W,0}(w)E_0\left(\{\Pi_0(Z,W) - E_0(\Pi_0(Z,W) \mid W)\}^2 \mid W=w\right) : w\right). \end{aligned}$$

For notational convenience, given a vector X, we used notation X^2 for the matrix $XX^\top$. We also used the notation $\mathrm{Diag}(x)$ for the $N \times N$ diagonal matrix with diagonal elements defined by vector x. Thus, the inverse of $C(\pi_0)$ exists in closed form and is given by:

$$C(\pi_0)^{-1} = \mathrm{Diag}\left(\frac{1}{P_{W,0}(w)E_0(\{\Pi_0(Z,W) - E_0(\Pi_0(Z,W) \mid W)\}^2 \mid W=w)} : w\right).$$

This yields the following formula for the efficient influence curve of m_0 in model $\mathcal{M}(\pi_0)$:

$$\begin{array}{l} D^*_{m,\Pi_0,w}(P_0) = \frac{1}{P_{W,0}(w)E_0(\{\Pi_0(Z,W) - E_0(\Pi_0(Z,W) \mid W)\}^2 \mid W=w)} \\ I(W=w)(\Pi_0(Z,W) - E_0(\Pi_0(Z,W) \mid W))(Y - \Pi_0(Z,W)m_0(W) - \theta_0(W)), \end{array}$$

where $D^*_{m,\Pi_0}(P_0)$ is $N \times 1$ vector with components $D^*_{m_0,\Pi_0,w}(P_0)$ indexed by $w \in \mathcal{W}$. We can further simplify this as follows:

$$\begin{array}{l} D^*_{m,\Pi_0,w}(P_0)(W,Z,Y) = \frac{1}{P_{W,0}(w)E_0(\{\Pi_0(Z,W) - E_0(\Pi_0(Z,W) \mid W)\}^2 \mid W=w)} \\ I(W=w)(\Pi_0(Z,w) - E_0(\Pi_0(Z,W) \mid W=w))(Y - \Pi_0(Z,w)m_0(w) - \theta_0(w)). \end{array}$$

This gradient equals the canonical gradient of m_0 in this model $\mathcal{M}(\pi_0)$, if $E_0((Y - E_0(Y \mid \Pi_0, W))^2 \mid Z, W)$ is only a function of W. For example, this would hold if $E(U_Y^2 \mid Z,W) = E_0(U_Y^2 \mid W)$. This might be a reasonable assumption for an instrumental variable Z. For the sake of presentation, we work with this gradient due to its relative simplicity. and the fact that it still equals the actual canonical gradient under this assumption.

We have that $\psi_0 = \phi(m_0, P_{W,0})$ for a mapping

$$\phi(m_0.P_{W,0}) = \arg\min_\beta E_0 \sum_a h(a,V)a^2\left(m_0(W) - m_\beta(V)\right)^2,$$

defined by working model $\{m_\beta : \beta\}$. Let $d\phi(m_0, P_{W,0})(h_m, h_W) = \frac{d}{dm_0}\phi(m_0, P_{W,0})(h_m)+ \frac{d}{dP_{W,0}}\phi(m_0, P_{W,0})(h_W)$ be the directional derivative in direction (h_m, h_W). The gradient of $\Psi : \mathcal{M}(\Pi_0) \to \mathbb{R}^d$ is given by $D^*_{\psi,\Pi_0}(P_0) = \frac{d}{dm_0}\phi(m_0, P_{W,0})D^*_{m,\Pi_0}(P_0)+ \frac{d}{dP_{W,0}}\phi(m_0, P_{W,0})IC_W$, where $IC_W(O) = (I(W = w) - P_{W,0}(w) : w)$. We note that $\beta_0 = \phi(m_0, P_{W,0})$ solves the following $d \times 1$ equation

$$U(\beta_0, m_0, P_{W,0}) \equiv E_0 \sum_a h(a, V)a^2 \frac{d}{d\beta_0} m_{\beta_0}(V)(m_0(W) - m_{\beta_0}(V)) = 0.$$

By the implicit function theorem, the directional derivative of $\beta_0 = \phi(m_0, P_{W,0})$ is given by

$$\begin{aligned} d\phi(m_0, P_{W,0})(h_m, h_W) = -\left\{\tfrac{d}{d\beta_0}U(\beta_0, m_0, P_{W,0})\right\}^{-1} \\ \left\{\tfrac{d}{dm_0}U(\beta_0, m_0, P_{W,0})(h_m) + \tfrac{d}{dP_{W,0}}U(\beta_0, m_0, P_{W,0})(h_W)\right\}. \end{aligned}$$

We need to apply this directional derivative to $(h_m, h_W) = (D^*_{m,\Pi_0}(P_0), IC_W)$. Recall we assumed that m_β is linear in β. We have

$$c_0 \equiv -\frac{d}{d\beta_0}U(\beta_0, m_0) = E_0 \sum_a h(a, V)a^2 \left\{\frac{d}{d\beta_0} m_{\beta_0}(V)\right\}^2,$$

which is a $d \times d$ matrix. Note that if $m_\beta(V) = \sum_j \beta_j V_j$, then this reduces to

$$c_0 = E_0 \sum_a h(a, V)a^2 \vec{V}\vec{V}^\top,$$

where $\vec{V} = (V_1, \ldots, V_d)$. We have

$$\frac{d}{dP_{W,0}}U(\beta_0, m_0, P_{W,0})(h_W) = \sum_w h_W(w) \sum_a h(a, v)a^2 \frac{d}{d\beta_0} m_{\beta_0}(v)(m_0(w)-m_{\beta_0}(v)).$$

Thus, the latter expression applied to $IC_W(O)$ yields $c_0^{-1}D^*_W(P_0)$, where

$$D^*_W(P_0) \equiv \sum_a h(a, V)a^2 \frac{d}{d\beta_0} m_{\beta_0}(V)(m_0(W) - m_{\beta_0}(V)).$$

In addition, the directional derivative $\frac{d}{d\epsilon}\, U(\beta_0, m_0 + \epsilon h_m, P_{W,0})\big|_{\epsilon=0}$ in the direction of the function h_m is given by

$$\frac{d}{dm_0}U(\beta_0, m_0, P_{W,0})(h_m) = E_0 \sum_a h(a, V)a^2 \frac{d}{d\beta_0} m_{\beta_0}(V)h_m(W).$$

We conclude that

$$d\phi(m_0, P_{W,0})(h_m, h_W) = D^*_W(P_0) + c_0^{-1}\left\{E_0 \sum_a h(a,V)a^2 \frac{d}{d\beta_0} m_{\beta_0}(V) D^*_{m,W}(P_0)\right\}.$$

We conclude that the canonical gradient of $\Psi : \mathcal{M}(\Pi_0) \to \mathbb{R}^d$ is given by

$$\begin{array}{l} D^*_{\psi,\Pi_0}(P_0)(O) = D^*_W(P_0)(O) \\ +c_0^{-1} E_0 \sum_a h(a,V)a^2 \frac{d}{d\beta_0} m_{\beta_0}(V) D^*_{m,W}(P_0) \\ = c_0^{-1} \sum_a h(a,V)a^2 \frac{d}{d\beta_0} m_{\beta_0}(V)(m_0(W) - m_{\beta_0}(V)) \\ +c_0^{-1} \sum_a h(a,V)a^2 \frac{d}{d\beta_0} m_{\beta_0}(V) \frac{1}{E_0(\{\Pi_0(Z,W) - E(\Pi_0(Z,W)|W)\}^2|W)} \\ (\Pi_0(Z,W) - E_0(\Pi_0(Z,W) \mid W))(Y - \Pi_0(Z,W)m_0(W) - \theta_0(W)). \end{array}$$

We state this result in the following lemma and also state a robustness result for this efficient influence curve.

Lemma 3 *The efficient influence curve of $\Psi : \mathcal{M}(\Pi_0) \to \mathbb{R}^d$ is given by*

$$\begin{array}{l} D^*_{\psi,\Pi_0}(P_0) = c_0^{-1} \sum_a h(a,V)a^2 \frac{d}{d\beta_0} m_{\beta_0}(V)(m_0(W) - m_{\beta_0}(V)) \\ +c_0^{-1} \sum_a h(a,V)a^2 \frac{d}{d\beta_0} m_{\beta_0}(V) \frac{1}{E_0(\{\Pi_0(Z,W) - E(\Pi_0(Z,W)|W)\}^2|W)} \\ (\Pi_0(Z,W) - E_0(\Pi_0(Z,W) \mid W))(Y - \Pi_0(Z,W)m_0(W) - \theta_0(W)). \end{array}$$

Assume the linear working model $m_\beta(V) = \beta\vec{V}$. Let $h_1(V) = \sum_a h(a,V)a^2\vec{V}$. We have that for all θ, (ρ_0 below refers to $\Pr(Z|W)$):

$$P_0 D^*_{\psi,\Pi_0}(g_0, m, \theta) = 0 \text{ if } E_0 h_1(V)(m - m_0)(W) = 0,$$

or, equivalently, if $\psi \equiv \Psi(m, P_{W,0}) = \Psi(m_0, P_{W,0}) = \psi_0$.

Efficient influence curve in model in which Π_0 is unknown

We will now derive the efficient influence curve in model $\mathcal{M}$ in which Π_0 is unknown, which is obtained by adding a correction term $D_\pi(P_0)$ to the above derived $D^*_{\psi,\Pi_0}(P_0)$. The correction term $D_\pi(P_0)$ that needs to be added to D^*_{ψ,Π_0} is the influence curve of $P_0\{D^*_{\psi,\Pi_0}(\pi_n) - D^*_{\psi,\Pi_0}(\pi_0)\}$, where $D^*_{\psi,\Pi_0}(\pi) = D^*_{\psi,\Pi_0}(\beta_0, \theta_0, m_0, \rho_0, \pi)$ is the efficient influence curve in model $\mathcal{M}(\pi_0)$, as derived above with π_0 replaced by π, and π_n is the nonparametric NPMLE of π_0. Let $h_1(V) \equiv \sum_a h(a,v)a^2 \frac{d}{d\beta_0} m_{\beta_0}(v)$. Let $\pi(\epsilon) = \pi + \epsilon\eta$. We plug in for η the influence curve of the NPMLE $\Pi_n(z,w)$, which is given by

$$\eta(z,w) = \frac{I(Z=z, W=w)}{P_0(z,w)}(A - \Pi(Z,W)).$$

We have

$$\begin{aligned}
&D_\pi(P_0) = \tfrac{d}{d\epsilon} P_0 D^*_\psi(\pi(\epsilon))\Big|_{\epsilon=0} \\
&= P_0 c_0^{-1} h_1(V) \left\{ -2\tfrac{E_0((\pi - E(\pi|W))(\eta - E(\eta|W))|W)}{E_0((\pi - E(\pi|W))^2|W)} \right. \\
&\qquad\qquad \left. (\pi - E(\pi \mid W)(Y - \pi m_0 - \theta_0) \right\} \\
&+ P_0 c_0^{-1} h_1(V) \left\{ \tfrac{(\eta - E(\eta|W))(Y - \pi m_0 - \theta_0)}{E_0((\pi - E(\pi|W))^2|W)} \right\} \\
&- P_0 c_0^{-1} h_1(V) \left\{ \tfrac{(\pi - E(\pi|W))\eta m_0}{E_0((\pi - E(\pi|W))^2|W)} \right\}.
\end{aligned}$$

By writing the expectation w.r.t. P_0 as an expectation of a conditional expectation, given Z, W, and noting that $E(Y - \pi_0 m_0 - \theta_0 \mid Z, W) = 0$, it follows that the first two terms equal zero. Thus,

$$D_\pi(P_0) = -P_0 c_0^{-1} h_1(V) \left\{ \tfrac{(\pi - E_0(\pi|W))\eta m_0}{E_0((\pi - E_0(\pi|W))^2|W)} \right\}.$$

This yields as correction term:

$$\begin{aligned}
&D_\pi(P_0) = -(A - \Pi_0(Z, W)) \int_{z,w} P_0(z, w) c_0^{-1} h_1(V) \left\{ \frac{(\pi - E(\pi|W)) \frac{I(Z=z, W=w)}{P_0(z,w)} m_0}{E_0((\pi - E(\pi|W))^2|W)} \right\} \\
&= -(A - \Pi_0(Z, W)) c_0^{-1} h_1(V) \left\{ \tfrac{(\pi(Z,W) - E(\pi(Z,W)|W)) m_0(W)}{E_0((\pi(Z,W) - E_0(\pi(Z,W)|W))^2|W)} \right\}.
\end{aligned}$$

This proves the following lemma.

Lemma 4 *The efficient influence curve of $\Psi : \mathcal{M} \to \mathbb{R}^d$ is given by*

$$\begin{aligned}
&D^*(P_0) = D^*_W(P_0) \\
&+ c_0^{-1} \tfrac{h_1(V)}{\sigma^2(\rho_0, \pi_0)(W)} (\pi_0(Z, W) - E_0(\pi_0(Z, W) \mid W))(Y - \pi_0(Z, W) m_0(W) - \theta_0(W)) \\
&- c_0^{-1} \tfrac{h_1(V)}{\sigma^2(\rho_0, \pi_0)(W)} \{(\pi_0(Z, W) - E_0(\pi_0(Z, W) \mid W)) m_0(W)\} (A - \pi_0(Z, W)) \\
&\equiv D^*_W(P_0) + C_Y(\rho_0, \pi_0)(Z, W)(Y - \pi_0(Z, W) m_0(W) - \theta_0(W)) \\
&\qquad\qquad\qquad\qquad\qquad\qquad - C_A(\rho_0, \pi_0, m_0)(A - \pi_0(Z, W)) \\
&\equiv D^*_W(P_0) + D^*_Y(P_0) - D^*_A(P_0),
\end{aligned}$$

where

$$\begin{aligned}
\sigma^2(\rho_0, \pi_0)(W) &= E_0(\{\Pi_0(Z, W) - E(\Pi_0(Z, W) \mid W)\}^2 \mid W) \\
h(\rho_0, \pi_0)(W) &= c_0^{-1} \frac{h_1(V)}{\sigma^2(\rho_0, \pi_0)(W)} \\
C_Y(\rho_0, \pi_0)(Z, W) &= h(\rho_0, \pi_0)(W)(\pi_0(Z, W) - E_{\rho_0}(\pi_0(Z, W) \mid W)) \\
C_A(\rho_0, \pi_0, m_0)(Z, W) &= C_Y(\rho_0, \pi_0)(Z, W) m_0(W).
\end{aligned}$$

Double robustness of efficient influence curve: We already showed $P_0D^*(\pi_0, \rho_0, m, \theta) = 0$ if $\phi(m, P_{W,0}) = \phi(m_0, P_{W,0})$. If $\phi(m, P_{W,0}) = \phi(m_0, P_{W,0})$ (i.e., $\psi = \psi_0$), then,

$$P_0D^*(\pi, \rho_0, m, \theta) = P_0\frac{h_1}{\sigma^2(\rho_0, \pi)}(\pi - P_{\rho_0}\pi)(\pi_0 - \pi)(m_0 - m),$$

where we used notation $P_{\rho_0}h = E_{\rho_0}(h(Z, W) \mid W)$ for the conditional expectation operator over Z, given W. This is thus second order in $(m-m_0)(\pi-\pi_0)$. In particular, it equals zero if $m = m_0$ or $\pi = \pi_0$. We can thus also state the following double robustness result: if $m = m_0$, then $P_0D^*(\pi, d, m_0, \theta) = 0$ if $d = \rho_0$ or if $\pi = \pi_0$.

Efficient influence curve of target parameter when assuming a parametric form for effect of treatment as function of covariates

We now assume $m_0 = m_{\alpha_0}$ for some model $\{m_\alpha : \alpha\}$, which implies the semiparametric regression model $E_0(Y \mid Z, W) = \Pi_0(Z, W)m_{\beta_0}(W) + \theta_0(W)$. Let $f_\beta(Z, W) = \Pi_0(Z, W)m_\beta(W)$. Let $m_\alpha(W) = \alpha^\top W^*$, where W^* is k-dimensional vector of functions of W. Note that α is d-dimensional and $\frac{d}{d\alpha}m_\alpha(W) = W^*$.

Efficient influence curve in model in which Π_0 is known.

First, we consider the statistical model $\mathcal{M}(\pi_0) \subset \mathcal{M}$ in which $\Pi_0(Z, W) = E_0(A \mid Z, W)$ is known. Define the k-dimensional vector

$$h(\Pi_0)(Z, W) = d/\alpha_0 m_{\alpha_0}(Z, W) = \Pi_0(Z, W)d/d\alpha_0 m_{\alpha_0}(W) = \Pi_0(Z, W)W^*.$$

By previous results on the semiparametric regression model, a gradient for the k-dimensional parameter $\alpha(P)$ at $P = P_0 \in \mathcal{M}(\pi_0)$ is given by

$$D^*_{\alpha,\Pi_0}(P_0) = C(\pi_0)^{-1}(h(\Pi_0)(Z, W) - E(h(\Pi_0)(Z, W) \mid W))(Y - f_{\alpha_0}(Z, W) - \theta_0(W)),$$

where $C(\pi_0)$ is a $k \times k$ matrix defined as

$$\begin{aligned} C(\pi_0) &= E_0\{d/d\alpha_0 f_{\alpha_0}(Z, W) - E_0(d/d\alpha_0 f_{\alpha_0}(Z, W) \mid W)\}^2 \\ &= E_0\{(W^*W^{*\top}\{\Pi_0(Z, W) - E_0(\Pi_0(Z, W) \mid W\}^2\}. \end{aligned}$$

Let $C(\pi_0)^{-1}$ be the inverse of $C(\pi_0)$.

This gradient equals the canonical gradient of α_0 in this model $\mathcal{M}(\pi_0)$, if $E_0((Y - E_0(Y \mid \Pi_0, W))^2 \mid Z, W)$ is only a function of W. For example, this would hold if $E(U_Y^2 \mid Z, W) = E_0(U_Y^2 \mid W)$. This might be a reasonable assumption for an instrumental variable Z. For the sake of presentation, we work with this gradient due to its relative simplicity. and the fact that it still equals the actual canonical gradient under this assumption.

We have that $\psi_0 = \phi(\alpha_0, P_{W,0})$ for a mapping

$$\phi(\alpha_0.P_{W,0}) = \arg\min_\beta E_0 \sum_a h(a,V)a^2 \left(m_{\alpha_0}(W) - m_\beta(V)\right)^2,$$

defined by working model $\{m_\beta : \beta\}$. Let $d\phi(\alpha_0, P_{W,0})(h_\alpha, h_W) = \frac{d}{d\alpha_0}\phi(\alpha_0, P_{W,0})(h_\alpha) + \frac{d}{dP_{W,0}}\phi(\alpha_0, P_{W,0})(h_W)$ be the directional derivative in direction (h_β, h_W). The gradient of $\Psi : \mathcal{M}(\Pi_0) \to \mathbb{R}^d$ is given by $D^*_{\alpha,\Pi_0}(P_0) = \frac{d}{d\alpha_0}\phi(\alpha_0, P_{W,0})D^*_{\alpha,\Pi_0}(P_0) + \frac{d}{dP_{W,0}}\phi(\alpha_0, P_{W,0})IC_W$, where $IC_W(O) = (I(W = w) - P_{W,0}(w) : w)$ is the influence curve of the empirical distribution of W. We note that $\beta_0 = \phi(\alpha_0, P_{W,0})$ solves the following $d \times 1$ equation

$$U(\beta_0, \alpha_0, P_{W,0}) \equiv E_0 \sum_a h(a,V)a^2 \frac{d}{d\beta_0} m_{\beta_0}(V)(m_{\alpha_0}(W) - m_{\beta_0}(V)) = 0.$$

By the implicit function theorem, the directional derivative of $\beta_0 = \phi(\alpha_0, P_{W,0})$ is given by

$$\begin{array}{l} d\phi(\alpha_0, P_{W,0})(h_\alpha, h_W) = -\left\{\frac{d}{d\beta_0}U(\beta_0, \alpha_0, P_{W,0})\right\}^{-1} \\ \left\{\frac{d}{d\alpha_0}U(\beta_0, \alpha_0, P_{W,0})(h_\alpha) + \frac{d}{dP_{W,0}}U(\beta_0, \alpha_0, P_{W,0})(h_W)\right\}. \end{array}$$

We need to apply this directional derivative to $(h_\alpha, h_W) = (D^*_{\alpha,\Pi_0}(P_0), IC_W)$. Recall we assumed that m_β is linear in β. We have

$$c_0 \equiv -\frac{d}{d\beta_0}U(\beta_0, \alpha_0, P_{W,0}) = E_0 \sum_a h(a,V)a^2 \left\{\frac{d}{d\beta_0}m_{\beta_0}(V)\right\}^2,$$

which is a $d \times d$ matrix. Note that if $m_\beta(V) = \sum_j \beta_j V_j$, then this reduces to

$$c_0 = E_0 \sum_a h(a,V)a^2 \vec{V}\vec{V}^\top,$$

where $\vec{V} = (V_1, \ldots, V_d)$. We have

$$\frac{d}{dP_{W,0}}U(\beta_0, \alpha_0, P_{W,0})(h_W) = \sum_w h_W(w) \sum_a h(a,v)a^2 \frac{d}{d\beta_0}m_{\beta_0}(v)(m_{\alpha_0}(w) - m_{\beta_0}(v)).$$

Thus, the latter expression applied to $IC_W(O)$ yields the contribution $c_0^{-1}D^*_W(P_0)$, where

$$D^*_W(P_0) \equiv \sum_a h(a,V)a^2 \frac{d}{d\beta_0} m_{\beta_0}(V)(m_{\alpha_0}(W) - m_{\beta_0}(V)).$$

In addition,

$$\frac{d}{d\alpha_0} U(\beta_0, \alpha_0, P_{W,0}) = E_0 \sum_a h(a,V)a^2 \frac{d}{d\beta_0} m_{\beta_0}(V) \frac{d}{d\alpha_0} m_{\alpha_0}(W).$$

We conclude that

$$\begin{aligned}
&d\phi(\alpha_0, P_{W,0})(h_\alpha, h_W) = \\
&D^*_W(P_0) + c_0^{-1}\left\{E_0 \textstyle\sum_a h(a,V)a^2 \frac{d}{d\beta_0} m_{\beta_0}(V) \frac{d}{d\alpha_0} m_{\alpha_0}(W) D^*_{\alpha,\Pi_0}(P_0)\right\}.
\end{aligned}$$

We conclude that the canonical gradient of $\Psi : \mathcal{M}(\Pi_0) \to \mathbb{R}^d$ is given by

$$\begin{aligned}
D^*_{\psi,\Pi_0}(P_0) &= D^*_W(P_0)(O) \\
&\quad + c_0^{-1}\left\{E_0 \sum_a h(a,V)a^2 \frac{d}{d\beta_0} m_{\beta_0}(V) \frac{d}{d\alpha_0} m_{\alpha_0}(W)\right\} D^*_{\alpha,\Pi_0}(P_0)(O) \\
&= D^*_W(P_0)(O) + \\
&\quad c_0^{-1}\left\{E_0 h_1(V)\vec{V}\vec{W}^{*\top}\right\} C(\pi_0)^{-1}(h(\Pi_0)(Z,W) - E(h(\Pi_0)(Z,W) \mid W)) \times \\
&\quad (Y - f_{\alpha_0}(Z,W) - \theta_0(W)).
\end{aligned}$$

We state this result in the following lemma and also state a robustness result for this efficient influence curve.

Lemma 5 *Let $h_1(V) = \sum_a h(a,V)a^2\vec{V}$. The efficient influence curve of $\Psi : \mathcal{M}(\Pi_0) \to \mathbb{R}^d$ is given by*

$$\begin{aligned}
&D^*_{\psi,\Pi_0}(P_0) = c_0^{-1}h_1(V)\tfrac{d}{d\beta_0} m_{\beta_0}(V)(m_{\alpha_0}(W) - m_{\beta_0}(V)) \\
&+c_0^{-1}\left\{E_0 h_1(V)\vec{V}\vec{W}^{*\top}\right\} C(\pi_0)^{-1}(h(\Pi_0)(Z,W) - E(h(\Pi_0)(Z,W) \mid W))\times \\
&(Y - f_{\alpha_0}(Z,W) - \theta_0(W)).
\end{aligned}$$

We have that

$$P_0 D^*_{\psi,\Pi_0}(d, m_{\alpha_0}, \theta) = 0, \textit{ if either } d = \rho_0 \textit{ or } \theta = \theta_0.$$

Efficient influence curve in model in which Π_0 is unknown

We will now derive the efficient influence curve in model $\mathcal{M}$ in which Π_0 is unknown, which is obtained by adding a correction term $D_\pi(P_0)$ to the above derived $D^*_{\psi,\Pi_0}(P_0)$. The correction term $D_\pi(P_0)$ that needs to be added to D^*_{ψ,Π_0} is the influence curve of $P_0\{D^*_{\psi,\Pi_0}(\pi_n) - D^*_{\psi,\Pi_0}(\pi_0)\}$, where $D^*_{\psi,\Pi_0}(\pi) = D^*_{\psi,\Pi_0}(\beta_0, \theta_0, \alpha_0, \rho_0, \pi)$ is the efficient influence curve in model $\mathcal{M}(\pi_0)$, as derived above with π_0 replaced by π, and π_n is the nonparametric NPMLE of π_0. Let $h_1(V) \equiv \sum_a h(a,v)a^2 \frac{d}{d\beta_0} m_{\beta_0}(v)$. Let $\pi(\epsilon) = \pi + \epsilon\eta$. We plug in for η the influence curve of the NPMLE $\Pi_n(z,w)$, which is given by

$$\eta(z,w) = \frac{I(Z=z, W=w)}{P_0(z,w)}(A - \Pi(Z,W)).$$

We have

$$\begin{aligned} D_\pi(P_0) &= \left. \frac{d}{d\epsilon} P_0 D^*_\psi(\pi(\epsilon)) \right|_{\epsilon=0} \\ &= -\left\{ P_0 c_0^{-1} h_1(V) \vec{V} W^{*\top} \right\} C(\pi_0)^{-1} P_0 \left\{ W^* W^{*\top} (\pi_0 - E(\pi_0 \mid W)) \eta(Z,W) \right\}. \end{aligned}$$

This yields as correction term:

$$\begin{aligned} &D_\pi(P_0)(O) = -(A - \Pi_0(Z,W)) \\ &\left\{ P_0 c_0^{-1} h_1(V) \vec{V} W^{*\top} \right\} C(\pi_0)^{-1} \left\{ W^* W^{*\top} (\pi_0(Z,W) - E(\pi_0 \mid W)) \right\}. \end{aligned}$$

This proves the following lemma.

Lemma 6 *The efficient influence curve of $\Psi : \mathcal{M} \to \mathbb{R}^d$ is given by*

$$\begin{aligned} &D^*(P_0) = D^*_W(P_0) \\ &+ c_0^{-1} \left\{ E_0 h_1(V) \vec{V} \vec{W}^{*\top} \right\} C(\pi_0)^{-1} W^* (\Pi_0 - E(\Pi_0(Z,W) \mid W))(Y - f_{\alpha_0}(Z,W) - \theta_0(W)) \\ &- \left\{ P_0 c_0^{-1} h_1(V) \vec{V} W^{*\top} \right\} C(\pi_0)^{-1} \left\{ W^* W^{*\top} (\pi_0(Z,W) - E(\pi_0 \mid W)) \right\} (A - \Pi_0(Z,W)) \\ &\equiv D^*_W(P_0) + C_Y(\rho_0, \pi_0)(Z,W)(Y - \pi_0(Z,W) m_{\alpha_0}(W) - \theta_0(W)) \\ &\qquad\qquad - C_A(\rho_0, \pi_0, m_0)(A - \pi_0(Z,W)) \\ &\equiv D^*_W(P_0) + D^*_Y(P_0) - D^*_A(P_0), \end{aligned}$$

where

$$\begin{aligned} &C_Y(\rho_0, \pi_0)(Z,W) = c_0^{-1} \left\{ E_0 \textstyle\sum_a h(a,V) a^2 \vec{V} \vec{W}^{*\top} \right\} \times \\ &\qquad\qquad C(\pi_0)^{-1} (h(\Pi_0)(Z,W) - E(h(\Pi_0)(Z,W) \mid W)) \\ &C_A(\rho_0, \pi_0, m_0)(Z,W) = \left\{ P_0 c_0^{-1} h_1(V) \vec{V} W^{*\top} \right\} C(\pi_0)^{-1} \left\{ W^* W^{*\top} (\pi_0(Z,W) - E(\pi_0 \mid W)) \right\}. \end{aligned}$$

Double robustness of efficient influence curve: We already showed $P_0 D^*(\pi_0, d, \alpha_0, \theta) = 0$ if $d = \rho_0$ or $\theta = \theta_0$. We also have that $P_0 D^*(\pi, \rho_0, \alpha_0, \theta) = 0$ for all θ and π.

The TMLE is analogue to the TMLE presented for the nonparametric model for $m_0(W)$.

Chapter 4

Optimal dynamic intent-to-treat.

4.1 Model and problem.

We consider the problem of estimation and inference under optimal dynamic treatment, in the context of an instrumental-variables model. As before, we have iid data $(W, Z, A, Y) \sim \mathcal{M}$, where $\mathcal{M}$ is a semiparametric model. Z is assumed to be a valid instrument for identifying the effect of treatment A on outcome Y, when one has to account for unmeasured confounding. $V \subseteq W$ is an arbitrary fixed subset of the baseline covariates. In this chapter, the unknown optimal dynamic "treatment" rule $d(V)$ actually refers to an intervention on the instrument Z. Thus from here on, we refer to $d(V)$ as the optimal dynamic assignment to treatment, or optimal dynamic intent-to-treat. The optimal mean counterfactual outcome is attained by setting $Z = d(V)$. In applications, instrument Z is often a randomized encouragement mechanism, or randomized assignment to treatment which may or may not be followed. In other cases, Z is not perfectly randomized but nevertheless promotes or discourages individuals in receiving treatment.

There are no restrictions on the type of data. However, the case of categorical or continuous Z requires special attention and is dealt with in section 4.3. Further, we let $c_T(Z, W)$ be a cost function that gives the total cost associated with assigning an individual with covariates W to a particular Z value. We can think of $c_T(Z, W)$ as the sum of $c_Z(Z, W)$, a cost incurred directly from setting Z, and $E_{A|W,Z} c_A(A, W)$, an average cost incurred from the actual treatment A. For instance, for the get-out-the-vote campaign, a binary instrument denotes whether an individual receives a phone call, and binary treatment A represents whether the individual actually received the intervention of speaking to a campaigner. Then $c_Z(Z=1,W)$ is the cost of making a phone call, and $c_A(A,W)$ is the generally higher cost of speaking to a callee. In this work, we assume a known function $c_T(Z,W)$. [1] We have cost constraint $Ec_Z(Z,W) \leq K$, for a fixed cost K.

Notation. We further assume wlog that intent-to-treat $Z=0$ has lower cost for all V: $E_{W|V}c_T(0,W) \leq E_{W|V}c_T(1,W)$. [2] Let $\underline{K}_0 \triangleq E_W c_T(0,W)$ be the total cost of not assigning any individuals to intent-to-treat, and $K_{T,0} \triangleq E_W c_T(1,W)$ be the total cost of assigning everyone, and we assume a nontrivial constraint $K \leq K_{T,0}$. We have P_W, $\Pr_{V|W}$, $\rho(W) \triangleq \Pr(Z=1|W)$, $\mu(W,Z) \triangleq E(Y|W,Z)$ defined as before. We also define $\mu_b(V) \triangleq E_{W|V}\big[\mu(Z=1,W) - \mu(Z=0,W)\big]$, which gives the mean difference in outcome between setting $Z=1$ and $Z=0$ given V. Similarly, $c_b(V) \triangleq E_{W|V}\big[c_T(Z=1,W) - c_T(Z=0,W)\big]$.

Causal model. We assume Z is exogenous, given W. Assume our usual causal model, with data generated as $W=f_W(U_W)$, $Z=f_Z(W,U_Z)$, $A=f_A(W,Z,U_A)$, $Y=f_Y(W,Z,U_Y)$. Then we have:

Assumption 2 *(Randomization of Z.) We assume that $U_Z \perp U_Y|W$.*

This implies $E(Y(Z)|W) = E(Y|W,Z)$.

In this model, instrument Z functions exactly as a usual unconfounded treatment variable A. We can view our problem of estimating the mean counterfactual outcome under optimal dynamic intent-to-treat, intervening on instrument Z, as the problem of estimating a mean counterfactual outcome under optimal dynamic treatment. The latter problem is very similar to one tackled by (Luedtke and van der Laan 2016a); however, we allow for a non-unit cost function, making this problem a generalization of that one. We present this general result on optimal dynamic treatments in the context of instrumental variables here in keeping with our main theme on IV-based models, and also to emphasize that non-compliance to treatments and unmeasured confounding

[1] It is straightforward to extend this model to incorporate uncertainty in $E(A|W,Z)$ for calculating $c_T(Z,W)$, and thus estimating $c_T(W,Z)$ from the data, given fixed functions c_Z, c_A. There is a correction term that gets added to the D^*_c component of the efficient influence curve.

[2] We are only making this assumption for the sake of easing notation. We can forgo this assumption by introducing notation, i.e. $Z=l(V)$ is the lower cost intent-to-treat value for a stratum defined by covariates V.

are very real possibilities in typical treatment-effect data. Variable A is unnecessary for estimation in our model and is ignored in the rest of this chapter.

Causal parameter of interest.

$$\Psi(P_0) \triangleq \text{Max}_d\ E_{P_0} Y(Z = d(V))\text{s.t. } E_{P_0}[c_T(Z = d(V), W)] \leq K$$

Statistical model. We assume the statistical model $\mathcal{M}$ consisting of all distributions $\mathcal{P}$ of $O = (W, Z, Y)$. P_W and $\mu(W, Z)$ are unspecified, and the distribution for the instrument $\Pr(Z|W) = \rho(W)$ may or may not be specified (Z is often a fully randomized group assignment).

Statistical target parameter. We have the following statistical target parameter:

$$\Psi_0 = E_{P_0}\mu_0(W, Z = d_0(V)) \tag{4.1}$$

where d_0 is the optimal intent-to-treat rule:
$d_0 = \text{argmax}_d\ E_{P_0}\mu_0(W, Z = d(V))$ s.t. $E_{P_0}[c_T(Z = d(V), W)] \leq K$

4.2 Assumptions for identifiability and pathwise differentiability of Ψ_0.

We use notation $d_0 = d_{P_0}$, $\tau_0 = \tau_{P_0}$, etc.

A1) Randomization of Z, as given in section 4.1: $E(Y(Z)|W) = E(Y|W, Z)$.

These next three assumptions are needed to ensure pathwise differentiability and prove the form of the canonical gradient (theorem 2).

A2) Positivity assumption: $0 < \rho_0(W) < 1$.

A3) There is a neighborhood of η_0 where $S_0(x)$ is Lipschitz continuous, and a neighborhood of $S_0(\eta_0) = K - \underline{K}_0$ where $S_0^{-1}(y)$ is Lipschitz continuous. We have $|\eta_1 - \eta_2| < \epsilon \Rightarrow |S_0(\eta_1) - S_0(\eta_2)| < c_y\epsilon$ for $\eta_1,\ \eta_2$ in a δ_x-neighborhood of η_0; and $|y_1 - y_2| < \epsilon \Rightarrow |S_0^{-1}(y_1) - S_0^{-1}(y_2)| < c_x\epsilon$, for $y_1,\ y_2$ in a δ_y-neighborhood of $K - \underline{K}_0$, for some constants $c_x,\ c_y$.

A4) $Pr_0(T_0(V) = \tau) = 0$ for all τ in a neighborhood of τ_0.

Note that A3) implies that $S_0^{-1}(K - \underline{K}_0)$ exists. Note also that A3) actually implies $Pr_0(T_0(V) = \eta) = 0$ for η in a neighborhood of η_0, and thus A3) implies A4) when $\eta_0 > 0$ and $\tau_0 = \eta_0$. This is discussed in the paper where we need to use this fact.

Need for A4) (Guarantee of non-exceptional law).
If A4) does not hold and there is positive probability of individuals being at the threshold for being treated or not under the optimal rule, then the solution $d(V)$ is not unique, and Ψ_0 is no longer pathwise differentiable. Further, the problem of finding an optimal *deterministic* solution in this case is the NP-hard knapsack problem, although it is considered to among the easier problems in this class. Thus, we restrict attention to so-called non-exceptional laws where A4) holds. It is easy to see that in this case, the optimal $d(V)$ over the broader set of non-deterministic decision rules is a deterministic rule. (Luedtke and van der Laan 2016b) derive an online martingale estimator for the optimal counterfactual outcome in case of an exceptional law and show root-n confidence intervals. As they describe, one expects A4) to be a reasonable assumption in practice under an arbitrary constraint $\underline{K}_0 < K < K_{T,0}$ that allows for only a strict subset of the population to be treated. In contrast, in the unconstrained version of the problem of estimating mean optimal outcome under dynamic treatment, the corresponding version of this assumption is that there is zero probability of individuals having exactly zero treatment effect. That seems to be a more problematic assumption.

4.3 Closed-form solution for optimal rule d_0 in the case of binary treatment.

Recall that wlog we think of $Z = 0$ as the 'baseline' intent-to-treat (ITT) value having lower cost. We define a scoring function $T(V) = \frac{\mu_b(V)}{c_b(V)}$ for ordering subgroups (given by V) based on the effect of Z per unit cost. In the optimal intent-to-treat policy, all groups with the highest $T(V)$ values deterministically have Z set to 1, up to cost K and assuming $\mu_b \geq 0$. We write $T_P(V)$ to make explicit the dependence on $P_W, \mu(Z, W)$ from distribution P.

Define a function $S_P : [-\infty, +\infty] \to \mathbb{R}$ as

$$S_P(x) = E_V[I(T_P(V) \geq x)(c_b(V)]$$

In other words, $S_P(x)$ gives the expected (additional above-baseline) cost of setting $Z = 1$ for all subgroups having $T_P(V) \geq x$. We use $S_0(\cdot)$ to denote S_{P_0} from here on. Let
$\underline{K}_P = E_P c_T(Z = 0, W)$ be the cost of treating everyone with the baseline $Z = 0$. Also define $K_{T,P} = E_P c_T(1, W)$ to be the total cost of treating everyone with $Z = 1$.

Define cutoff η_P as

$$\eta_P = S^{-1}(K - \underline{K}_P)$$

The assumptions guarantee that $S^{-1}(K - \underline{K}_P)$ exists and η_P is well-defined. η is set so that there is a total cost K of treating everyone having $T(V) \geq \eta$ with $Z = 1$. Further let:

$$\tau_P = \max\{\eta_P, 0\}$$

Thus, τ gives the cutoff for the scoring function $T(V)$, so the optimal rule is

$$d_P(V) = 1 \text{ iff } T_P(V) \geq \tau_P$$

Lemma 7 *Assume A2)-A4). Then the optimal decision rule d_0 for parameter Ψ_0 as defined in equation 4.1 is the deterministic solution*
$d_0(V) = 1$ iff $T_0(V) \geq \tau_0$, with T_0, τ_0 as defined above.

The proof is given in section 4.8.

Below, we describe modifications to the optimal solution for d_0 when Z is continuous or categorical.

Continuous intent-to-treat Z.

Z can be continuous if we have a linear cost function. Suppose $Z \in [z_{\min}, z_{\max}]$. W can do a linear transformation to $\tilde{Z} \in [0, 1]$. Then the problem is to pick an optimal $d(V) = \tilde{Z} \in [0, 1]$ for every V where $\tilde{Z} = 0$ represents no treatment, and positive values the magnitude of treatment. We make the simplifying assumption that $c_Z(\cdot)$ has the linear form $c_T(\tilde{Z}, W) = \tilde{Z} \cdot c_Z(W)$. The form and proof of the closed-form solution parallel that given above for the case of binary Z, with $d(V) = 1$ for all V such that the expected marginal effect of assignment to $\tilde{Z} = 1$ per marginal cost is above some cutoff. Thus, all groups are either treated at the maximum level $Z = z_{\max}$, or not treated with $Z = z_{\min}$. This is again assuming there is no issue with exceptional laws and assumption A4) holds.

The scoring function for ordering subgroups is $T(V) = E_{W|V} \frac{\mu(W,\tilde{Z}=1)-\mu(W,\tilde{Z}=0)}{c_Z(W)}$. We have function $S_P(x) = E_V[I(T_P(V) \geq x)E_{W|V}c_Z(W)]$. From here on, function threshold η, threshold τ, and decision rule $d(V)$ are defined exactly as for the case of a binary Z. The proof of optimality is also the same.

Categorical intent-to-treat Z.

Returning to the example of the get-out-the-vote campaign, some of the campaigns described have categorical intent-to-treat variables Z [cite]. For instance, a potential voter can receive a phone call, receive a text message, or be left untreated. Our model allows for discrete choices of Z. Suppose we have $|\mathcal{Z}|$ intent-to-treat values $\mathcal{Z} = \{z_0, ..., z_{|\mathcal{Z}|-1}\}$. z_0 represents the baseline assignment. The setting of categorical intent-to-treat with cost functions has a significantly more complicated solution than the binary or continuous cases given above. However, when we consider the special case of uniform cost, where the budget constraint only restricts $d(V)$ so that fraction p of the population receive an ITT setting of $Z = 1$, we again have a similar simple closed-form solution as before.

Categorical intent-to-treat with uniform cost.

Let $z'(V)$ denote $\text{argmax}_{1..|\mathcal{Z}|} E_{W|V}\mu(z', W)$, in other words, the optimal choice of Z for V in the unconstrained case. It is easy to see that when a fixed fraction of the population can be treated with any $Z \neq z_0$, then in the optimal solution has $d(V) = z'(V)$ or $d(V) = z_0$ (a subgroup is either not treated, or treated with its optimal treatment). Now we choose all subgroups with the highest mean outcomes $E_{W|V}\mu(z', W)$ to treat. Let $S : x \longrightarrow \Pr_V(E_{W|V}\mu(z'(V), W) \geq x)$ be a survival function. Then define

$$\eta = S^{-1}(p)$$

$$\tau = \max\{\eta, 0\}$$

and decision rule

$$d(V) = 1 \text{ iff } E_{W|V}\mu(z'(V), W) \geq \tau$$

The proof is essentially the same as for the above cases of binary or continuous treatment.

Categorical intent-to-treat with cost function $c(z, W)$.

The solution can be understood easily as an iterative procedure. We briefly describe it informally first. Start by setting all $d(V)$ to the lowest cost Z assignment given V, thus minimizing $E_{W|V}c(z, W)$. Then pick a subgroup V and treatment Z' such that the marginal gain in expected outcome per marginal gain in cost is maximized, when setting $d(V) = Z'$. For that subgroup, set $d(V) = Z'$, replacing the previous assignment to the lowest cost treatment. Continue to pick a subgroup V and treatment Z' such that the marginal gain per marginal cost is maximized, when switching $d(V)$ to Z'. Here the marginal gain and cost are calculated with respect to the current treatment $Z = d(V)$ for each subgroup. We proceed with this procedure until reaching our allocated budget, successively switching to Z assignments having higher cost and higher outcome than before, chosen in order of decreasing marginal gain per marginal cost. Standard optimization theory guarantees that this procedure is optimal for maximizing a mean outcome under a single positive budget constraint.

Formally, we first define a baseline (lowest cost) Z-value for each subgroup as $Z.0(V) = \text{argmin}_{j\in(0..|\mathcal{Z}|-1)} E_{W|V} c_T(z_j, W)$. Next, we define a series of ordering functions $Z.i(V), i \in [1, |\mathcal{Z}|-1]$. $Z.i(V)$ represents the choice of treatment in $\mathcal{Z} = \{z_0, ..., z_{|\mathcal{Z}|-1}\}$ having the i^{th} highest marginal gain per marginal cost, for stratum defined by V. This function can be defined inductively as:

$$Z.i(V) = \text{argmax}_{k\in 0..|\mathcal{Z}|-1} \left\{ \frac{E_{W|V}(\mu(z_k, W) - \mu(z_{Z.(i-1)(V)}, W))}{E_{W|V}(c(z_k, W) - c(z_{Z.(i-1)(V)}, W))}, \right.$$

$$\left. \text{s.t. } k \neq Z.j(V) \text{ for } j < i, E_{W|V}[\mu(z_k, W) - \mu(z_{Z.(i-1)(V)}, W)] > 0 \right\}$$

If the set in the rhs above is empty, then define $Z.i(V) = Z.(i-1)(V)$. Now define a scoring function for each subgroup based on its marginal gain in outcome per marginal cost. The optimal ordering of treatments $Z.i(V)$ defined above establishes what treatment the marginal values are calculated with respect to.

$$T(i, V) = E_{W|V} \frac{\mu(z_{Z.i(V)}, W) - \mu(z_{Z.(i-1)(V)}, W)}{c(z_{Z.i(V)}, W) - c(z_{Z.(i-1)(V)}, W)}$$

The optimal rule is to treat each subgroup with the treatment associated with the largest cost and gain, such that the marginal cost and gain of that treatment is above some threshold τ. Define such a rule as $q(\tau, V) = \max\{i : T(i, V) \geq \tau\}$. The threshold is chosen so we have K expected cost of this procedure:

$$\text{Set } \tau \text{ s.t. } E_V E_{W|V} c_T(z_{Z.i(V)}, W) = K, \text{ for } i(V) = q(\tau, V)$$

Note that increasing the threshold τ above monotonically decreases the indices $q(\tau, V)$ and thus decreases the cost, so we can solve the equation above. Finally, the optimal rule is given by:

$$d(V) = z_{Z.i(V)}, \text{ for } i(V) = q(\tau, V)$$

4.4 Canonical gradient for Ψ_0.

For $O \in \mathcal{O} = (W, Z, A, Y)$, and deterministic assignment-to-treatment rule $\tilde{d}$, define

$$D_1(\tilde{d}, P)(O) \triangleq \frac{I(Z = \tilde{d}(V))}{\rho_P(W)}(Y - \mu_P(Z, W)) \tag{4.2}$$

$$D_2(\tilde{d}, P)(O) \triangleq \mu_P(\tilde{d}(V), W) - E_P\mu_P(\tilde{d}(V), W) \tag{4.3}$$

Then $D^*_{\tilde{d}}(P)(O) = D_1(\tilde{d}, P)(O) + D_2(\tilde{d}, P)(O)$ is the efficient influence curve of $\Psi_{\tilde{d}}(P)$.

When d is a fixed stochastic rule, it follows from the linearity of the pathwise derivative mapping that the efficient influence curve of $\Psi_d(P)$ is given by $D^*_d(P)(O) = E_U[D_1(d, P)(O) + D_2(d, P)(O)]$. Let $D_3(d, \tau, P) = -\tau(E_U c_T(d(V), W) - K)$.

Define

$$D^*(d, \tau, P)(O) \triangleq D^*_d(P)(O) + D_3(d, \tau, P)$$

Let $D_0 \triangleq D^*(d_0, \tau_0, P_0)$.

Theorem 2 *Assume A1)-A4) above. Then Ψ is pathwise differentiable at P_0 with canonical gradient $D_0 = D^*(d_0, \tau_0, P_0)$.*

This is proved in section 4.8.

4.5 TMLE.

The relevant components for estimating $\Psi(Q) = E_W\mu(Z = d(V), W)$ are $Q = (P_W, \mu(Z, W))$. Decision rule d is also part of Ψ, but it is a function of P_W, $\mu(Z, W)$. The nuisance parameter is $g = \rho(W)$. As usual, convert Y to the unit interval via a linear transformation $Y \rightarrow \tilde{Y}$, so that $\tilde{Y} = 0$ corresponds to $Y_{\min}$ and $\tilde{Y} = 1$ to $Y_{\max}$. We assume $Y \in [0, 1]$ from here.

1. Use the empirical distribution $P_{W,n}$ to estimate P_W. Make initial estimates of $\mu_n(Z, W)$ and $g_n = \rho_n(W)$ using any strategy desired. Data-adaptive learning using Super Learner is recommended.

2. The empirical estimate $P_{W,n}$ gives an estimate of $Pr_{V,n}(V) = E_{W,n}I(F_V(W) = V)$, $\underline{K}_n = E_{W,n}c(0, W)$, $K_{T,n} = E_{W,n}c(1, W)$, and $c_{b,n}(V) = E_{W,n|V}(c_T(1, W) - c_T(0, W))$.

3. Estimate $\mu_{b,0}$ as $\mu_{b,n}(V) = E_{W,n|V}(\mu_n(1,W) - \mu_n(0,W))$.

4. Estimate $T_0(V)$ as $T_n(V) = \frac{\mu_{b,n}(V)}{c_{b,n}(V)}$.

5. Estimate $S_0(x)$ using $S_n(x) = E_{V,n}[I(T_n(V) \geq x)(c_b(V)]$.

6. Estimate η_0 as η_n using $\eta_n = S_n^{-1}(K - \underline{K}_n)$
   and $\tau_n = \max\{0, \eta_n\}$.

7. Estimate the decision rule as $d_n(V) = 1$ iff $T_n(V) \geq \tau_n$.

8. Now fluctuate the initial estimate of $\mu_n(Z,W)$ as follows: For $Z \in [0,1]$, define covariate $H(Z,W) \triangleq \frac{I(d_n(V)=Z)}{g_n(W)}$. Run a logistic regression using:

Outcome:	$(Y_i : i = 1, \ldots, n)$
Offset:	$(\text{logit}\mu_n(Z_i, W_i), i = 1, \ldots, n)$
Covariate:	$(H(Z_i, W_i) : i = 1, \ldots, n)$

   Let ϵ_n represent the level of fluctuation, with
   $\epsilon_n = \text{argmax}_\epsilon \frac{1}{n}\sum_{i=1}^n [\mu_n(\epsilon)(Z_i, W_i)\log Y_i + (1 - \mu_n(\epsilon)(Z_i, W_i))\log(1 - Y_i)]$
   and $\mu_n(\epsilon)(Z,W) = \text{logit}^{-1}(\text{logit}\mu_n(Z,W) + \epsilon H(Z,W))$.

9. Set the final estimate of $\mu(Z,W)$ to $\mu_n^*(Z,W) = \mu_n(\epsilon_n)(Z,W)$.

10. Finally, form final estimate of $\Psi_0 = \Psi_{d_0}(P_0)$ using the plug-in estimator

$$\Psi^* = \Psi_{d_n}(P_n^*) = \frac{1}{n}\sum_{i=1}^n \mu_n^*(Z = d_n(V_i), W_i)$$

We have used the notation $\Psi_d(P)$ refers to mean outcome under decision rule $d(V)$.

- Showing that $P_n D^*(d_n, \tau_n, P_n^*) = 0$.

Using the log likelihood loss function $L(Q_n(\epsilon|g_n), g_n, (O_1, ..., O_n)) \triangleq P_n\left[\mu_n(\epsilon)\log Y + (1 - \mu_n(\epsilon))\log(1 - Y)\right]$ and logistic fluctuation $\mu_n(\epsilon)$ as specified above, we have
$\frac{d}{d\epsilon}L(Q_n(\epsilon|g_n), g_n, (O_1, ..., O_n))|_{\epsilon=0} = P_n H_n(Y - \mu_n)$.
Thus, at $\mu_n^* = \mu_n(\epsilon_n)$, we have $\frac{d}{d\epsilon}L(Q_n(\epsilon|g_n), g_n, (O_1, ..., O_n))|_{\epsilon=0} = 0$, so $P_n D_1(d_n, P_n^*) = 0$ for the first term of the canonical gradient.

It is easy to see that $P_n D_2(d_n, P_n^*) = 0$ when we are using the empirical distribution $P_{W,n}$. We have

$P_n D_3(d_n, \tau_n, P_n^*) = 0$ for the third term of the canonical gradient as well, because $E_{V,n} E_{W,n|V} c_T(d(V), W) = K$, unless $\tau_n = 0$. (This is described in the proof of optimality of the closed-form solution in section 4.8.)

4.6 Theoretical results: efficiency, double robustness, and inference for Ψ_0.

Conditions for efficiency of Ψ_0.

These six conditions are needed to prove asymptotic efficiency (theorem 3). As discussed in section 4.6, when all relevant components and nuisance parameters ($P_{W,n},\ \rho_n,\ \mu_n$) are consistent, then C3) and C4) hold, while C6) holds by construction of the TMLE estimator.

C1) $\rho_0(W)$ satisfies the strong positivity assumption: $Pr_0(\delta < \rho_0(W) < 1 - \delta) = 1$ for some $\delta > 0$.

C2) The estimate $\rho_n(W)$ satisfies the strong positivity assumption, for a fixed $\delta > 0$ with probability approaching 1, so we have $Pr_0(\delta < \rho_n(W) < 1 - \delta) \to 1$.

Define second-order terms as follows:

$$R_1(d, P) \triangleq E_{P_0}\Big[\Big(1 - \frac{Pr_{P_0}(Z = d|W)}{Pr_P(Z = d|W)}\Big)\mu_P(Z = d, W) - \mu_0(Z = d, W)\Big]$$

$$R_2(d, \tau, P) \triangleq E_{P_0}\Big[(d - d_0)(\mu_{b,0}(V) - \tau_0 c_{b,0}(V))\Big]$$

Let $R_0(d, \tau, P) = R_1(d, P) + R_2(d, \tau, P)$.

C3) $R_0(d_n, \tau_0, P_n^*) = o_{P_0}(n^{-\frac{1}{2}})$.

C4) $P_0[(D^*(d_n, \tau_0, P_n^*) - D_0)^2] = o_{P_0}(1)$.

C5) $D^*(d_n, \tau_0, P_n^*)$ belongs to a P_0-Donsker class with probability approaching 1.

C6) $\frac{1}{n}\sum_{i=1}^{n} D^*(d_n, \tau_0, P_n^*)(O_i) = o_{P_0}(n^{-\frac{1}{2}})$.

Sufficient conditions for lemma 15.

E1) GC-like property for $c_b(V)$, $\mu_{b,n}(V)$:
$\sup_V |(E_{W,n|V} - E_{W,0|V})c_{b,T}(W)| = \sup_V(|c_{b,n}(V) - c_{b,0}(V)|) = o_{P_0}(1)$

E2) $\sup_V |E_{W,0|V}\mu_{b,n}(W) - E_{W,0|V}\mu_{b,0}(W)| = o_{P_0}(1)$
This is needed for the proof that $d_n(V) = d_0(V)$ with probability approaching 1.

E3) $S_n(x)$, defined as $x \to E_{V,n}[I(T_n(V) \geq x)c_{b,n}(V)]$ is a GC-class.

E4) Convergence of ρ_n, μ_n to ρ_0, μ_0, respectively, in $L^2(P_0)$ norm at a $O(n^{-1/2})$ rate in each case. This is needed in several places.

When all relevant components and nuisance parameters are consistent, as is the case when theorem 3 below holds and our estimator is efficient, we also expect conditions E1)-E4) to hold.

Theoretical properties of Ψ_n^*.

Theorem 3 (Ψ^* is asymptotically linear and efficient.) *Assume assumptions A1)-A4), and conditions C1)-C6). Then $\Psi^* = \Psi(P_n^*) = \Psi_{d_n}(P_n^*)$ as defined by the TMLE procedure is a RAL estimator of $\Psi(P_0)$ with influence curve D_0, so*

$$\Psi(P_n^*) - \Psi(P_0) = \frac{1}{n}\sum_{i=1}^{n} D_0(O_i) + o_{P_0}(n^{-\frac{1}{2}}).$$

Further, Ψ^* is efficient among all RAL estimators of $\Psi(P_0)$.

Inference. Let $\sigma_0^2 = Var_{P_0}D_0$. By theorem 3 and the central limit theorem, $\sqrt{n}(\Psi(P_n^*) - \Psi(P_0))$ converges in distribution to a $N(0, \sigma_0^2)$ distribution. Let $\sigma_n^2 = \frac{1}{n}\sum_{i=1}^{n} D^*(d_n, \tau_n, P_n^*)(O_i)^2$ be an estimate of σ_0^2.

Lemma 8 *Under the assumptions C1), C2), and conditions E1)-E4), we have $\sigma_n \longrightarrow_{P_0} \sigma_0$. Thus, an asymptotically valid 2-sided $1-\alpha$ confidence interval is given by*

$$\Psi^* \pm z_{1-\frac{\alpha}{2}}\frac{\sigma_n}{\sqrt{n}}$$

where $z_{1-\frac{\alpha}{2}}$ denotes the $(1-\frac{\alpha}{2})$-quantile of a $N(0,1)$ r.v.

Double robustness of Ψ_n^*.

Theorem 3 demonstrates consistency and efficiency when all relevant components and nuisance parameters are consistently estimated. Another important issue is under what cases of partial misspecification we still get a consistent estimate of Ψ_0, albeit an inefficient one. The set of relevant components and nuisance parameters is P_W, $\rho(W) = \Pr(Z = 1|W)$, and $\mu(W, Z) = E(Y|W, Z)$. The empirical distribution $P_{W,n}$ always converges to $P_{W,0}$. Our TMLE-based estimate Ψ^* is a consistent estimate of Ψ_0 under misspecification of $\rho_n(W)$ in the initial estimates, but not under misspecification of $\mu_n(W, Z)$. However, it turns out there is still an important double robustness property. If we consider $\Psi^* = \Psi_{d_n}(P_n^*)$ as an estimate of $\Psi_{d_n}(P_0)$, where the optimal decision rule $d_n(V)$ is estimated from the data, then we have that Ψ^* is double robust to misspecification of ρ_n or μ_n in the initial estimates.

Lemma 9 (Ψ^* is a double robust estimator of $\Psi_{d_n}(P_0)$.) *Assume assumptions A1)-A4) and conditions C1)-C2). Also assume the following version of C4):* $Var_0(D_1(d_n, P_n^*) + D_2(d_n, P_n^*)) < \infty$.
Then $\Psi^* = \Psi(d_n, P_n^*)$ *is a consistent estimator of* $\Psi_{d_n}(P_0)$ *when either* μ_n *is specified correctly, or* ρ_n *is specified correctly.*

To prove this lemma, note from section 4.4 and equation 4.2 that the canonical gradient for parameter $\Psi_{d_n}(P_0)$, in our semiparametric model using estimated decision rule d_n, equals $D_1(d_n, P) + D_2(d_n, P)$. We have remainder term $R_1(d_n, P)$ as defined in the assumptions:
$\Psi_{d_n}(P_n^*) - \Psi_{d_n}(P_0) = -P_0\left[D_1^*(d_n, P_n^*) + D_2^*(d_n, P_n^*)\right] + R_1(d_n, P_n^*)$.
TMLE solves the efficient influence curve equation over the D_1^*, D_2^* terms of D^*. R_1 is a second-order term in ρ, μ with $R_1(d_n, P_n^*) \in O_{P_0}(n^{-\frac{1}{2}})$ when either ρ or μ is specified consistently, while the $(P_n - P_0)[D_1 + D_2]$ term converges to zero under the finite variance assumption.

Discussion of conditions for theorem 3.

Condition C3. This is satisfied if both $R_1(d, P)$ and $R_2(d, K_T)$ are in $o_{P_0}(n^{-\frac{1}{2}})$. $R_1(d, P)$ takes the form of a typical double robust term that is small (generally $O_{P_0}(n^{-\frac{1}{2}})$) when either $\rho_n(W)$ or μ_n is estimated well, and second-order $o_{P_0}(n^{-\frac{1}{2}})$) when both $\rho_n(W)$ and μ_n are estimated well. The Cauchy-Schwarz inequality gives an upper bound for this term as the product of the $L^2(P_0)$ rates of convergence of these two components. If the conditional distribution of the instrument $\rho(W)$ is known, then $\rho_n(W) = \rho_0(W)$ and this component becomes 0.

Ensuring that $R_2(d_n, \tau_0, P_n^*) = o_{P_0}(n^{-\frac{1}{2}}))$ requires more work. The results below are similar to (Audibert and Tsybakov 2007) and (Luedtke and van der Laan 2016a). We make the following margin assumption for some $\alpha > 0$:

$$Pr_0(0 < |T_0(V) - \tau_0| \leq t) \lesssim ct^\alpha \text{ for all } t > 0 \tag{4.4}$$

where $\lesssim$ denotes less than or equal up to a multiplicative constant. If we have a Lipschitz continuity condition for $S_0(x)$ (like that given locally in A3)) holding not just in an interval around τ_0 but for all x, then we can take $\alpha = 1$.

Theorem 4 *If (4.4) holds for some $\alpha > 0$, then*
(i) $|R_2(d_n, \tau_0, P_n^*)| \lesssim \|(T_n(V) - \tau_n) - (T_0(V) - \tau_0)\|_{2,P_0}^{(3\alpha+2)/(2(1+\alpha))}$
(ii) $|R_2(d_n, \tau_0, P_n^*)| \lesssim \|(T_n(V) - \tau_n) - (T_0(V) - \tau_0)\|_{\infty,P_0}^{1+\alpha}$

Taking $\alpha = 1$, this theorem implies that $R_2(d_n, \tau_0, P_n^*) = o_{P_0}(n^{-\frac{1}{2}}))$ if either $\|(T_n(V)-\tau_n)-(T_0(V)-\tau_0)\|_{2,P_0}$ is $o_{P_0}(n^{-\frac{4}{10}})$ or $\|(T_n(V)-\tau_n)-(T_0(V)-\tau_0)\|_{\infty,P_0}$ is $o_{P_0}(n^{-\frac{1}{4}})$.

Condition C4). This is implied by the $L^2(P_0)$ consistency of $\mu_{b,n}$ and ρ_n and convergence of $d_n(V)$ to $d_0(V)$ with probability approaching 1. (The latter is discussed below.)

Condition C5). This condition places restrictions on how data adaptive the estimators of $\mu_{b,n}(W)$ and $\rho_n(W)$ can be. Section 2.10 of (van der Vaart and Wellner 1996) gives conditions under which the estimates of $\mu_{b,0}$ and ρ_0 belonging to Donsker classes (and using the empirical estimate for $\mathrm{Pr}_n(W)$) implies that $D^*(d_n, \tau_0, P_n^*)$ belongs to a Donsker class.

Condition C6). In the TMLE section of this paper, we showed that the empirical mean of the first two terms of the efficient influence curve is 0: $P_n D_1(d_n, P_n^*) = 0$, $P_n D_2(d_n, P_n^*) = 0$. We also showed for the third term that $P_n D_3(d_n, \tau_n, P_n^*) = 0$. It is easy to see that since $\tau_n \to_{P_0} \tau_0$, we have $P_n D_3(d_n, \tau_0, P_n^*) \to_{P_0} o(1)$.

Discussion of conditions for lemma 15.

To see that σ_n converges to σ_0, note that $D^*(d, \tau, P)(O_i)$ depends on the following components: $\{P_W, \rho_P(W), \mu_P(Z, W), d_P, \tau_P\}$. The following is sufficient for convergence of $D^*(d_n, \tau_n, P_n^*)^2$ to D_0^2:

- convergence of τ_n to τ_0 (proved below)
- convergence of $Pr_n(W)$ to $P_{W,0}$ (guaranteed by the fact that we use empirical distribution $P_{W,n}$ for $Pr_n(W)$)

- convergence of ρ_P, μ_P to ρ_0, μ_0, respectively, in $L^2(P_0)$ norm (condition E4)).
- $d_n(V) = d_0(V)$ with probability approaching 1. This is equivalent to $T_n(V) \geq \tau_n \iff T_0(V) \geq \tau_0$ w.p. approaching 1. The convergence of τ_n to τ_0, the uniform convergence of $T_n(V)$ to $T_0(V)$, and A4) guarantee this. The proof that $\tau_n \to \tau_0$ is given below, and contains the proof of uniform convergence of $T_n(V)$ to $T_0(V)$.

4.7 Simulations

Setup.

We use two main data-generating functions:

Dataset 1 (categorical Y).

Data is generated according to:

$$\begin{aligned}
U_{AY} &\sim \text{Bernoulli}(1/2) \\
W1 &\sim \text{Uniform}(-1, 1) \\
W2 &\sim \text{Bernoulli}(1/2) \\
Z &\sim \text{Bernoulli}(\alpha) \\
A &\sim \text{Bernoulli}(W1 + 10 \cdot Z + 2 \cdot U_{AY} - 10) \\
Y &\sim \text{Bernoulli}((1 - A) * (\text{plogis}(W2 - 2 - U_{A,Y})) + (A) * (\text{plogis}(W1 + 4))
\end{aligned}$$

$U_{A,Y}$ is the confounding term. For the simulations where $V \subset W$, we take $V = (1(W1 \geq 0) + -1(W1 < 0), W2)$. Finally, we have $c_T(Z = 1, W) = 1$, $c_T(Z = 0, W) = 0$ for all W here.

Dataset 2 (continuous Y.)

We use 3-dimensional W and distribution

$$\begin{aligned}
U_{AY} &\sim \text{Normal}(0,1)\\
W &\sim \text{Normal}(\mu_\beta, \Sigma)\\
Z &\sim \text{Bernoulli}(0.1)\\
A &\sim -2\cdot W1 + W2^2 + 4\cdot W3\cdot Z + U_{AY}\\
Y &\sim 0.5\cdot W1\cdot W2 - W3 + 3\cdot A\cdot W2 + U_{AY}
\end{aligned}$$

When $V \subset W$, we use either V equals $W1$ rounded to the nearest 0.2, or alternately, V is $W3$ rounded to the nearest 0.2. We also have $c_T(0,W) = 0$ for all W, and $c_T(1,W) = 1 + b\cdot W1$. Parameters μ_β, Σ, and b vary.

Forming initial estimates.

We use the empirical distribution $P_{W,n}$ for the distribution of W. For learning μ_n, we use Super Learner, with the following libraries of learners (the names of learners are as specified in the SuperLearner package):

For continuous Y: glm, step, randomForest, nnet, svm, polymars, rpart, ridge, glmnet, gam, bayesglm, loess, mean.

For categorical Y: glm, step, svm, step.interaction, glm.interaction, nnet.4, gam, randomForest, knn, mean, glmnet, rpart.

Further, we included different parameterizations of some of the learners given above: for nnet, size=2, 3, 4, or 5. For randomForest, ntree=100, 300, 500 1000. For knn, k=10, 30, 50, 100, 200, 300.

Finally, for learning ρ_n, we use a correctly specified logistic regression, regressing Z on W (except for simulation (C) as described below).

Estimators used.

We report results on three estimators throughout this section. One is the TMLE estimator described in section 4.5. We also implemented cross-validated TMLE (CV-TMLE) (see (Zheng and van der Laan 2011)). CV-TMLE uses N-fold cross-validation in choosing the optimal ϵ to minimize loss in the fluctuation step. The relevant components and nuisance parameters are fit on a training set, in each of the N-fold splits of the data, while ϵ is calculated on the hold-out set. This procedure sometimes enhances performance over standard TMLE by avoiding possible overfitting of ϵ to the initial estimates. Finally, our third estimator is the initial substitution estimator. This estimator evaluates the mean outcome using the same initial estimates of relevant components and the nuisance parameter as TMLE. It gives the plug-in estimate

$\Psi^*(P_{W.n}, \mu_n)$. We use machine learning to semi-parametrically learn the components in nearly every case in these simulations. Thus, the initial substitution estimator gives a comparison of TMLE to a straightforward semiparametric, machine learning-based approach that doesn't come with the guarantees of TMLE.

1000 repetitions are done of each simulation by default.

Simulation (A): using a large library of learning algorithms.

Tables 6.10 and 6.11 show the behavior of our estimators when machine learning is used to consistently estimate μ_n. 6.10 deals with categorical Y. In this case, all three estimators achieve very low bias, with or without the TMLE fluctuation step. The σ_n^2 column gives $\mathrm{Var}_n D^*(d_n, \tau_n, P_n^*)$, the estimated variance of the efficient influence curve. This is a consistent estimate of the variance of the TMLE-based estimators, in this case where efficiency holds. We see that all estimators have very low variance that converges to σ_n^2 by $n = 1000$. Although the initial substitution estimator is not guaranteed to be efficient, it displayed similar variance to TMLE, probably because the initial estimates were so accurate that the TMLE fluctuation was miniscule. Coverage of 95% confidence intervals is also displayed, with intervals calculated as $\Psi_n^* \pm 1.96 \frac{\sigma_n}{\sqrt{n}}$, as in lemma 15. The coverage is given in parentheses for the initial substitution estimator, as σ_n^2 is not necessarily the right variance. The TMLE estimators show better coverage, even though in this example, the width of the confidence intervals was accurate for all estimators for $n \geq 1000$. This may be due to the asymptotic linearity property of the TMLE-based estimators, ensuring that they follow a normal distribution as n becomes large.

Y is continuous in table 6.11. The TMLE estimators convincingly outperform the initial substitution estimator in both bias and variance here. While all estimators are consistent, only the TMLE estimators are guaranteed to be efficient, and we see a significant improvement in variance, as well as effective bias reduction. The estimated asymptotic variance σ_n^2 approximates the variance seen in the TMLE estimators fairly well for $n \geq 1000$. The coverage of confidence intervals for TMLE seems to converge to 95% more slowly than for the previous case of categorical Y.

Simulation (B): $V = W$ or $V \subset W$, cost constraint is more or less constraining.

We check the behavior of our estimators across various characteristics of the ODT optimization problem: when we allow V to be a strict subset of covariates W, vs when $V = W$; and under different sized budgets for assigning individuals to intent-to-treat (table 6.12). μ and ρ are estimated consistently here. We see that the TMLE-based estimators typically have an order of magnitude lower bias, and lower variance than the initial substitution estimator. In every single case, they manifest quite low bias on the order of one by $n = 1000$. This simulation also demonstrates the significant gain in mean outcome achieved when the decision rule $d(V)$ is allowed to depend on $V = W$, vs only a subset $V \subset W$.

Simulation (C): double robustness under partial misspecification.
We verify the consistency of Ψ_n^* under two cases of partial misspecification: when μ_n is misspecified, and when ρ_n is misspecified. In each case, the other nuisance parameter and relevant components are estimated consistently using data-adaptive learning. As described in section 4.6, $\Psi_n^* = \Psi_{d_n}^*$ is a double robust estimator of $\Psi_{d_n}(\Psi_0)$, but not necessarily of Ψ_0. Hence in every repetition of sampling a dataset and forming estimates, the target parameter $\Psi_{d_n}(\Psi_0)$ was recalculated using the current estimate of optimal rule d_n. For calculating confidence intervals for $\Psi_{d_n}(\Psi_0)$, the width was estimated using $\sigma_n = \sqrt{\left(\text{Var}_n D_{d_n}^*(P_n^*)\right)}$, where $D_{d_n}^*(P_0)$ is the efficient influence curve of $\Psi_{d_n}(P_0)$ as defined in section 4.4.

In table 6.13, the initial estimate for μ_n is grossly misspecified as $\mu_n = \text{mean}(Y)$. This creates a discrepancy of 15-20 points between $\Psi_{d_n}(P_0)$ and Ψ_0. The initial substitution estimator retains a bias of around -20 in estimating $\Psi_{d_n}(P_0)$, while TMLE demonstrates practically zero bias by $n = 1000$. TMLE is not efficient in this setting of partial misspecification. It has significantly larger variance than the initial substitution estimator for smaller sample sizes, but the variances are similar by $n = 4000$. σ_n was 201.2, 97.8, and 31.9, for $n = 250,\ 1000,\ 4000$, respectively. It provides a conservative (over)-estimate of variance for confidence intervals, as described in section 3.3 (ch. 3). We see that TMLE's coverage converges to just above 95%. On the other hand, coverage is very low for the initial substitution estimator due to its bias. This is despite the fact that the intervals are too wide in this case.

We also confirmed the robustness of Ψ_n^* to misspecification of ρ, when μ is consistently specified. In this case, $\Psi_{d_n}(P_0) = \Psi_0$, and both the TMLE and initial substitution estimators are consistent. TMLE is once again not guaranteed to be efficient. TMLE and the initial substitution estimator were found to perform very similarly in both bias and variance across different sample sizes. For instance, at $n = 1000$, we have bias of -7.23, variance of 43.78 for TMLE, and -7.13, 52.69 for the initial substitution estimator.

Simulation (D): quality of the estimate of d_n vs the true mean outcome attained under rule d_n.

We study how more accurate estimation of the decision rule d_n can lead to a higher objective obtained. The objective maximized here is the mean outcome under rule d_n, where d_n must satisfy a cost constraint. We use the known true distributions for $P_{W,0}$ and μ_0 in calculating the value of mean outcome under d_n as $\Psi_{d_n}(P_0) = E_{P_0}\mu_0(W, Z = d_n(V))$. The highest the true mean outcome can be under a decision rule that satisfies $E_{P_0}c_T(W, Z = d(V)) \leq K$ is Ψ_0 using optimal rule $d = d_0$. Therefore, the discrepancy between $\Psi_{d_n}(P_0)$ and Ψ_0 gives a measure of how inaccurate estimation of the decision rule diminishes the objective.

We compare $\Psi_{d_n}(P_0)$ when estimating μ_n using the usual large library of learners; when using a smaller library of learners consisting of (mean, loess, nnet.size=3, nnet.size=4, nnet.size=5); and finally when we set $\mu_n =$ mean(Y). d_n is estimated as usual (note that it is the same between the initial substitution, and TMLE-based estimates). Table 6.14 confirms the importance of forming a good fit with the data for achieving a high mean outcome. For $K = .2$ when roughly 20% of the population could be assigned $Z = 1$, the mean outcome was only a few points below the true optimal mean outcome Ψ_{d_0} when using the full library of learners (158.9 vs 162.8). However, it was about 15 points lower when using a much smaller library of learners. In fact, even when using machine learning with several nonparametric methods in the case of the smaller library, the objective $\Psi_{d_n}(P_0)$ attained wasn't far from that attained with the most uninformative $\mu_n =$ mean(Y). Very similar results hold for the less constrained case of $K = .8$.

Simulation (E): histograms for Ψ_n^*, Ψ_n^0.

Figures 4.1 and 4.2 give histograms for the TMLE estimate Ψ_n^* and initial substitution estimator Ψ_n^0. A large library of learners is used to consistently estimate μ_n. We see that not only is there significant bias for Ψ_n^0 at $n = 1000$, but the histogram looks far less like a normal curve than that for Ψ_n^*. The fluctuation step of TMLE appears very successful in normalizing the histogram despite the complex data-adaptive learning involved.

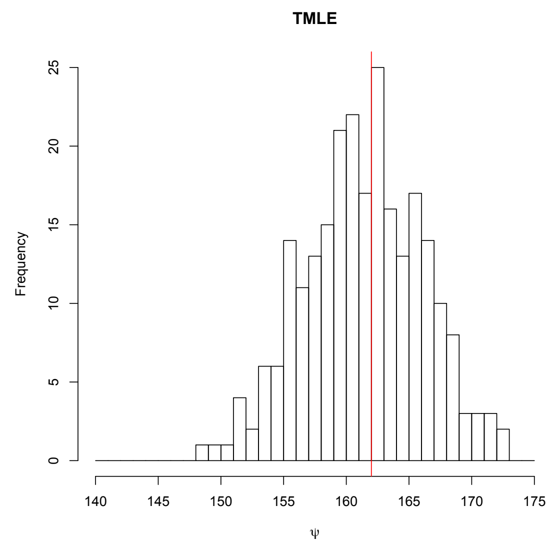

Figure 4.1: (Histogram of TMLE estimates Ψ_n^*. Y is continuous, μ_n and ρ_n are consistently specified using machine learning, and $n = 1000$. The red line depicts Ψ_0.

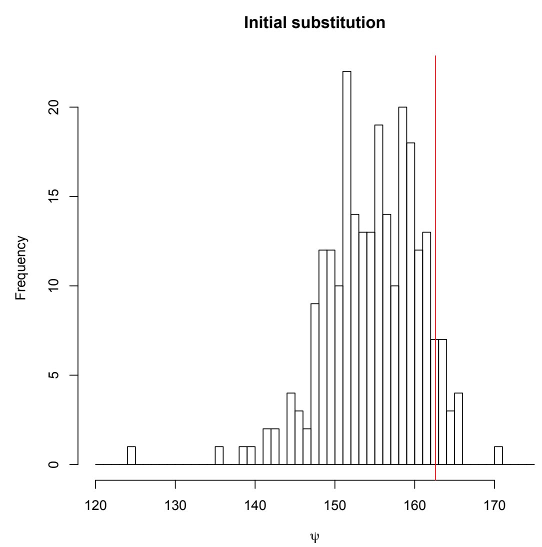

Figure 4.2: Histogram of initial estimates Ψ_n^0. Y is continuous, μ_n and ρ_n are consistently specified using machine learning, and $n = 1000$. The red line depicts Ψ_0.

4.8 Proofs.

Proof of optimal closed-form solution d_0.

We allow the possibility of a nondeterministic optimal treatment rule. In the $E_{P\times U}$ expressions below, U refers to possible randomization over $d(V)$. First, note that we can rewrite the objective as
$E_{P\times U}[Y(Z = d(V)) - Y(Z = 0)] + E_P Y(Z = 0)$ and can similarly rewrite the cost. Thus our parameter is equivalent to

$$\text{Max}_d \; E_{P\times U}[Y(Z = d(V)) - Y(Z = 0)]$$

$$\text{s.t.} \quad E_{P\times U}[c_T(Z = d(V), W) - c_T(Z = 0, W)] \leq K - \underline{K}$$

Secondly, note that the optimal decision rule in the *unconstrained* case would be to set $d(V) = 1$ if
$E_P[Y(Z = 1) - Y(Z = 0)|V] > 0$, and $d(V) = 0$ otherwise. Thus, we treat all subgroups with probability 1 where the average treatment effect is positive, and no other subgroups. This is true because $\text{Max}_{d(X)} \; E_X f(d(X), X)$ has the solution $d(X) = \text{argmax}_d f(d, X)$. Now if the closed-form solution is $d(V) = 1$ for all V, that means there is money in the budget to treat everyone. Also, since $T(V) \geq \tau \geq 0$ for all V, we have
$E_P[Y(Z = 1) - Y(Z = 0)|V] = E_{W|V}\mu(Z = 1, W) - \mu(Z = 0, W) > 0$
for all V. Thus, treatment is beneficial for all subgroups, so clearly the closed-form solution is optimal. In another scenario, the cutoff for $T(V)$ for treated groups is $\tau = 0$, and not all groups are treated. In this case, all subgroups where the average treatment effect (ATE) $E_{W|V}\mu(Z = 1, W) - \mu(Z = 0, W) \geq 0$ are treated, so again, the closed-form solution is optimal.

The only scenario where there is more work to do is when the cutoff for $T(V)$ for treated groups is $\tau = \eta > 0$, and there isn't enough money in the budget to treat all subgroups having positive ATE. (Note that η is the minimum threshold for $T(V)$ so that treating groups with $T(V) \geq \eta$ remains in the budget.) Let $d(V)$ refer to the decision rule given by the closed-form solution, achieving objective $E_{P\times U}[Y(Z = d(V)) - Y(Z = 0)] = C$. In this case, the budget constraint must be tight: Note that assumption A3) guarantees that $S(\eta') > S(\eta) = K - \underline{K}$ for $0 < \eta' < \eta$, which means there is some set $\mathcal{V}$ with positive support and $0 < T(V) < \eta$ and $d(V) = 0$ for $V \in \mathcal{V}$. Then it would be feasible within the budget and increase the objective if V where treated with nonzero probability. Thus, $E_{P\times U}[c_T(Z = d(V), W) - c_T(Z = 0, W)] = K - \underline{K}$.

Suppose now that $d(V)$ is not the optimal decision rule; then there is another rule $d'(V)$, such that $E_{P\times U}[Y(Z = d'(V)) - Y(Z = 0)] > C$. We can assume that the constraint is tight for d', otherwise we can pick another d' with objective at least C and average cost $K - \underline{K}$. Let $V \in \mathcal{V}$ denote the subgroups such that $T(V) > \tau$; thus $d(V) = 1$ for $V \in \mathcal{V}$, $d(V) = 0$ for $V \notin \mathcal{V}$. We must have $d'(V) > 0$ for some $V \notin \mathcal{V}$; otherwise d' would not have higher objective than d. Since the cost of assigning subgroups to non-baseline treatments is the same for d and d', we must have that the additional cost $\tilde{K}$ of assigning $V \notin \mathcal{V}$ with d' must equal the savings in cost over $V \in \mathcal{V}$. Thus,

$$\begin{aligned}\tilde{K} = E_{V\in\mathcal{V}}E_{W|V\times U}[c_T(Z=1,W) - c_T(Z=d'(V),W)] =\\ E_{V\in\mathcal{V}}E_{W|V}(1-d'(V))[c_T(Z=1,W) - c_T(Z=0,W)] =\\ E_{V\notin\mathcal{V}}E_{W|V\times U}[c_T(Z=d'(V),W) - c_T(Z=0,W)] =\\ E_{V\notin\mathcal{V}}E_{W|V}d'(V)[c_T(Z=1,W) - c_T(Z=0,W)]\end{aligned}$$

Then

$$\begin{aligned}0 &< E_{P\times U}[Y(Z=d'(V)) - Y(Z=0)] - E_{P\times U}[Y(Z=d(V)) - Y(Z=0)]\\ &= E_V E_{W|V\times U}[\mu(Z=d'(V),W) - \mu(Z=0,W)] -\\ &\quad E_V E_{W|V\times U}[\mu(Z=d(V),W) - \mu(Z=0,W)]\\ &= E_{V\in\mathcal{V}}E_{W|V\times U}[\mu(Z=d'(V),W) - \mu(Z=1,W)] +\\ &\quad E_{V\notin\mathcal{V}}E_{W|V\times U}[\mu(Z=d'(V),W) - \mu(Z=0,W)]\\ &= E_{V\in\mathcal{V}}E_{W|V}(-(1-d'(V)))[\mu(Z=1,W) - \mu(Z=0,W)] +\\ &\quad E_{V\notin\mathcal{V}}E_{W|V}d'(V)[\mu(Z=1,W) - \mu(Z=0,W)]\\ &\leq E_{V\in\mathcal{V}}(-(1-d'(V)))E_{W|V}[c_T(Z=1,W) - c_T(Z=0,W)]\cdot\eta +\\ &\quad E_{V\notin\mathcal{V}}d'(V)E_{W|V}[c_T(Z=1,W) - c_T(Z=0,W)]\cdot\eta\\ &= -\tilde{K}\cdot\eta + \tilde{K}\cdot\eta = 0\end{aligned}$$

Thus, we have a contradiction, so the given closed-form solution of $d(V)$ is optimal.

Derivation of canonical gradient.

Much of this is adapted from (Luedtke and van der Laan 2016a).

The pathwise derivative of $\Psi(Q)$ is defined as $\frac{d}{d\epsilon}\Psi(Q(\epsilon))|_{\epsilon=0}$ along paths $\{P_\epsilon : \epsilon\} \subset \mathcal{M}$. Here Q represents relevant components P_W, $\mu(A,W)$, and the paths are chosen so that

$$\begin{aligned} dP_{W,\epsilon} &= (1+\epsilon H_W(W))dP_W, \\ &\quad \text{where } EH_W(W)=0 \text{ and } C_W \triangleq \sup_W |H_W(W)| < \infty; \\ d\mu_\epsilon(Z,W) &= (1+\epsilon H_Y(Y|Z,W))d\mu(Z,W), \\ &\quad \text{where } EH_Y(Y|Z,W)=0 \text{ and } C_1 \triangleq \sup_{W,Z,Y} |H_Y(Y|Z,W)| < \infty; \end{aligned}$$

Note that we can assume upper and lower bounds for the cost $0 < C_L < c_{b,0}(V) < C_U$; if a treatment has no additional cost over the baseline treatment for some subgroups, we can treat those subgroups and solve the decision problem over the remaining subgroups having a cost-benefit tradeoff.

- Proof that $D_1(d,P)(O) + D_2(d,P)(O)$ is the efficient influence curve of $\Psi_d(P)$ for fixed deterministic rule d.

This is the same derivation as for the canonical gradient of the mean causal effect in a model without an instrumental variable or dynamic treatment, see for example (van der Laan and Rubin 2006).

- Extending D_0 to an unknown optimal rule $d(V)$.

Assumption A4) implies that the optimal rule $d(V)$ is almost surely deterministic. Let $d(V)$ represent the deterministic rule given by the closed-form solution.

We have that
$\Psi_d = \int \mu(d(V),W)dW = \int_W d(V)(\mu(1,W) - \mu(0,W))dW + E_W\mu(0,W)$
$= \int_W d(V)\mu_b(V)dW + E_W\mu(0,W)$. Thus,

$$\begin{aligned} \Psi(P_\epsilon) - \Psi(P_0) &= \int_W (E_U d_\epsilon(V) - d_0(V))\mu_{b,\epsilon} dP_{W,\epsilon} \\ &+ \int_W d_0(V)(\mu_{b,\epsilon} dP_{W,\epsilon} - \mu_{b,0} dP_{W,0}) \\ &+ E_{P_\epsilon}\mu_\epsilon(0,W) - E_{P_0}\mu_0(0,W) \\ &= \int_W (E_U d_\epsilon(V) - d_0(V))(\mu_{b,\epsilon} - \tau_0(c_{b,\epsilon}(V))\ ) dP_{W,\epsilon} \end{aligned} \tag{4.5}$$

$$+\tau_0 \int_V (E_U d_\epsilon(V) - d_0(V))(c_{b,\epsilon}(V)) dP_{W,\epsilon} \tag{4.6}$$

$$+\Psi_{d_0}(P_\epsilon) - \Psi_{d_0}(P_0). \tag{4.7}$$

In the derivations below, note that we often leave the distribution over $V = F_V(W)$ implicit and specify P_W, which induces a distribution P_V.

- Showing that $S_\epsilon(\eta) - S_0(\eta) = O(\epsilon)$ for η in a δ_x-neighborhood of η_0 (δ_x as given in assumption A3)).

Recall that $T_P(V) = \frac{\mu_{b,P}(V)}{c_{b,P}(V)}$ and $S_P(x) = E_V[I(T_P(V) \geq x)c_{b,P}(V)]$.

First note that $|\underline{K}_\epsilon - \underline{K}_0| = \int_W c_T(0,W)dP_{W,\epsilon} - dP_{W,0}$ $\leq \epsilon \int_W c_T(0,W)H_W(W)dP_{W,0} \leq \epsilon \cdot C_W \sup_W c_T(0,W)$, so $|\underline{K}_\epsilon - \underline{K}_0| \leq C_2$ for some constant C_2.

Next, note that $\mu_{b,\epsilon}(W) - \mu_{b,0}(W) = (1+\epsilon H_Y(Y|Z=1,W))\mu_0(Z=1,W)$ $-(1+\epsilon H_Y(Y|Z=0,W))\mu_0(Z=0,W) - \mu_0(Z=1,W) + \mu_0(Z=0,W)$ $= \epsilon \cdot (H_Y(Y|Z=1,W)\mu_0(Z=1,W) - H_Y(Y|Z=0,W)\mu_0(Z=0,W))$ $\leq \epsilon \cdot C_3$, for a constant C_3, assuming a bounded treatment effect. Defining $F_V(W) = V$ to represent the V-value corresponding to W, we have

$$|\mu_{b,\epsilon}(V) - \mu_{b,0}(V)|$$
$$\leq \Big| \int_W \mu_{b,\epsilon}(W)I(F_V(W)=V)dP_{W,\epsilon} - \int_W \mu_{b,0}(W)I(F_V(W)=V)dP_{W,\epsilon}\Big|$$
$$+\Big| \int_W \mu_{b,0}(W)I(F_V(W)=V)dP_{W,\epsilon} - \int_W \mu_{b,0}(W)I(F_V(W)=V)dP_{W,0}\Big|$$

$\leq \epsilon \cdot C_3 + \epsilon \int_W \mu_{b,0}(W)I(F_V(W)=V)H_W(W)dP_{W,0} \leq \epsilon \cdot C_4$ for constant C_4. It is also straightforward to see that $|c_{b,\epsilon}(V) - c_{b,0}(V)| \leq C_5 \cdot \epsilon$ for constant C_5.

Thus, we have

$$|T_\epsilon(V) - T_0(V)| = \Big|\frac{\mu_{b,\epsilon}(V)}{c_{b,\epsilon}(V)} - \frac{\mu_{b,0}(V)}{c_{b,\epsilon}(V)} + \frac{\mu_{b,0}(V)}{c_{b,\epsilon}(V)} - \frac{\mu_{b,0}(V)}{c_{b,0}(V)}\Big|$$
$$\leq \epsilon \cdot C_6 + \Big|\frac{\mu_{b,0}(V)}{c_{b,0}(V) + \epsilon C_7 c_{b,0}(V)} - \frac{\mu_{b,0}(V)}{c_{b,0}(V)}\Big|$$
$$\leq \epsilon \cdot C_6 + \Big|\frac{\mu_{b,0}(V)}{c_{b,0}(V)} \cdot (\frac{1}{1+\epsilon C_7} - 1)\Big|$$

$\leq \epsilon \cdot C_8$, where the last line is from the Taylor expansion $\frac{1}{1+\epsilon} = 1 - \epsilon + o(\epsilon)$, and the fact that $\mu_{b,0}$ is upper bounded, $c_{b,0}$ upper and lower bounded.

Finally, we have

$$|S_\epsilon(\eta) - S_0(\eta)|$$

$$= \Big| \int_W \big[I(T_\epsilon(V) \geq \eta)c_{b,\epsilon}(V)\big]dP_{W,\epsilon} - \int_W \big[I(T_0(V) \geq \eta)c_{b,0}(V)\big]dP_{W,0}\Big|$$
$$\leq \Big| \int_W \big[I(T_\epsilon(V) \geq \eta)c_{b,\epsilon}(V)\big]dP_{W,\epsilon} - \int_W \big[I(T_\epsilon(V) \geq \eta)c_{b,\epsilon}(V)\big]dP_{W,0}\Big|$$
$$+\Big| \int_W \big[I(T_\epsilon(V) \geq \eta)c_{b,\epsilon}(V)\big]dP_{W,0} - \int_W \big[I(T_\epsilon(V) \geq \eta)c_{b,0}(V)\big]dP_{W,0}\Big|$$

$$+\Big| \int_W \big[I(T_\epsilon(V) \geq \eta)c_{b,0}(V)\big]dP_{W,0} - \int_W \big[I(T_0(V) \geq \eta)c_{b,0}(V)\big]dP_{W,0}\Big|. \quad (4.8)$$

We have that $\int_W \big[I(T_0(V) \geq \eta+\epsilon C_8)c_{b,0}(V)\big]dP_{W,0} \leq \int_W \big[I(T_\epsilon(V) \geq \eta)c_{b,0}(V)\big]dP_{W,0} \leq \int_W \big[I(T_0(V) \geq \eta - \epsilon C_8)c_{b,0}(V)\big]dP_{W,0} \Longrightarrow$
$S_0(\eta + \epsilon C_8) \leq \int_W \big[I(T_\epsilon(V) \geq \eta)c_{b,0}(V)\big]dP_{W,0} \leq S_0(\eta - \epsilon C_8)$, so line (4.8) becomes $K - \underline{K}_0 + 2 \cdot C_8 c_y \epsilon - (K - \underline{K}_0)$ by assumption A3).

Finally, we get $|S_\epsilon(\eta) - S_0(\eta)| \leq \epsilon C_W(C_L + \epsilon) + \epsilon C_5 + 2 \cdot C_8 c_y \epsilon = C_S \epsilon$ for constant C_S.

- Showing that $\tau(\epsilon) - \tau_0 = O(\epsilon)$.

Recall constants $c_x,\ c_y, \delta_x,\ \delta_y$ used in assumption A3). Set $c_1 \triangleq C_S \cdot c_x$. Pick ϵ small enough so that: $\epsilon(c_1 \cdot c_y + C_2) < \delta_y$; $c_x(C_2 + C_S)\epsilon < \delta_x$. Let $y_\epsilon = K - \underline{K}_\epsilon$ and $y_0 = K - \underline{K}_0 = S_0(\eta_0)$.

The derivation above that $|S_\epsilon(\eta) - S_0(\eta)| \leq C_S\epsilon$ gives us $|y_\epsilon - y_0| \leq C_2\epsilon$. Set $\eta' = S_0^{-1}(y_\epsilon)$, and note by assumption A3) that $|\eta' - \eta_0| \leq c_x C_2 \epsilon$. Define $\eta_1 \triangleq \eta' - c_1\epsilon$ and $\eta_2 \triangleq \eta' + c_1\epsilon$. Note that $\eta_1,\ \eta_2$ are in a δ_x-neighborhood of η_0. Thus we have $S_0(\eta_1) \leq y_\epsilon + c_1 c_y \epsilon$, $S_0(\eta_2) \geq y_\epsilon - c_1 c_y \epsilon$. So $S_0(\eta_1),\ S_0(\eta_2)$ are in a δ_y-neighborhood of y_0, so by assumption A3) $|\eta_1 - \eta'| \geq c_1\epsilon \Rightarrow S_0(\eta_1) \geq y_\epsilon + c_1/c_x\epsilon = y_\epsilon + C_S\epsilon$, and similarly, we have $S_0(\eta_2) \leq y_\epsilon - c_1/c_x\epsilon = y_\epsilon - C_S\epsilon$. Using $|S_\epsilon(\eta) - S_0(\eta)| \leq C_S\epsilon$, we have $S_\epsilon(\eta_1) \geq y_\epsilon$, similarly $S_\epsilon(\eta_2) \leq y_\epsilon$. By the monotonicity of $S_\epsilon(x),\ S_\epsilon^{-1}(y)$, we have that $\eta_1 \leq \eta_\epsilon \leq \eta_2$ where $S_\epsilon(\eta_\epsilon) = y_\epsilon$.

Finally, we have $|\eta_\epsilon - \eta_0| \leq |\eta_0 - \eta'| + |\eta_\epsilon - \eta'| \leq c_x C_2 \epsilon + c_1 \epsilon = O(\epsilon)$. Thus, $|\tau_\epsilon - \tau_0| = \max(0, O(\epsilon)) = O(\epsilon)$.

- Showing that line (5.21) in the expansion of $\Psi(P_\epsilon) - \Psi(P_0)$ is $o(\epsilon)$.

We know that

$$T_0(V) - \tau_0 + O(\epsilon) \leq T_\epsilon(V) - \tau_\epsilon \leq T_0(V) - \tau_0 + O(\epsilon)$$

By assumption A4), it follows that there exists some $\delta > 0$ s.t. $\sup_{|\epsilon|<\delta} Pr_0(T_\epsilon(V) = \tau_\epsilon) = 0$. By the absolute continuity of $P_{W,\epsilon}$ with respect to $P_{W,0}$, $\sup_{|\epsilon|<\delta} Pr_{P_\epsilon}(T_\epsilon(V) = \tau_\epsilon) = 0$. It follows that, for all small enough ϵ and almost all U, $d_\epsilon(U,V)$ is the deterministic decision rule $d_\epsilon(U,V) = I(T_\epsilon(V) > \tau_\epsilon)$. Hence,

$$
\begin{aligned}
&\int_W (E_U d_\epsilon(V) - d_0(V))(\mu_{b,\epsilon} - \tau_0 c_{b,\epsilon}(V)) dP_{W,\epsilon} \\
&= \Big| \int_W (I(\frac{\mu_{b,\epsilon}}{c_{b,\epsilon}} > \tau_\epsilon) - I(\frac{\mu_{b,0}}{c_{b,0}} > \tau_0))(\mu_{b,\epsilon} - \tau_0 c_{b,\epsilon}) dP_{W,\epsilon} \Big| \\
&\leq \int_W |I(\mu_{b,\epsilon} > \tau_\epsilon c_{b,\epsilon}) - I(\mu_{b,0} > \tau_0 c_{b,0})|(|\mu_{b,0} - \tau_0 c_{b,\epsilon}| + C_4|\epsilon|) dP_{W,\epsilon} \\
&\leq \int_W I(|\mu_{b,0} - \tau_0 c_{b,0}| \leq |\mu_{b,0} - \tau_0 c_{b,0} - \mu_{b,\epsilon} + \tau_\epsilon c_{b,\epsilon}|) \\
&\qquad \times (|\mu_{b,0} - \tau_0 c_{b,0}| + |\tau_0| C_5 \epsilon + C_5|\epsilon|) dP_{W,\epsilon} \\
&= \int_W \Big[I(0 < |\mu_{b,0} - \tau_0 c_{b,\epsilon}| \leq |\mu_{b,0} - \mu_{b,\epsilon} + \tau_0(c_{b,\epsilon} - c_{b,0}) + (\tau_\epsilon - \tau_0) c_{b,\epsilon}|) \\
&\qquad \times (|\mu_{b,0} - \tau_0 c_{b,0}| + \tilde{C}|\epsilon|) \Big] dP_{W,\epsilon} \\
&\leq O(\epsilon) \int_W I(0 < |\mu_{b,0} - \tau_0 c_{b,0}| \leq O(\epsilon)) dP_{W,\epsilon} \\
&\leq O(\epsilon)(1 + C_W|\epsilon|) Pr_0(0 < |\mu_{b,0} - \tau_0 c_{b,0}| \leq O(\epsilon))
\end{aligned}
$$

Above $\tilde{C}$ is a new constant. We use constants and bounds from the proof that $S_\epsilon(\eta) - S_0(\eta) = O(\epsilon)$, the fact that $c_{b,0}$ is upper and lower bounded, and that $T_0(V)$ and consequently τ_0 are upper bounded. The last line is $o(\epsilon)$ because $\Pr(0 < X \leq \epsilon) \to 0$ as $\epsilon \to 0$ for any random variable X. Dividing by ϵ and taking the limit as $\epsilon \to 0$ yields zero.

- D_0 is as given in theorem 2, case 1: $\tau_0 = 0$.

The first line of $\Psi(P_\epsilon) - \Psi(P_0)$ is $o_P(\epsilon)$ according to the previous derivation, and second line is 0 in this case, so D_0 as given in theorem 2.

- D_0 is as given in theorem 2, case 2: $\tau_0 > 0$.

The second line of $\Psi(P_\epsilon) - \Psi(P_0)$ expands to

$$
\tau_0 \Big[\int_W E_U d_\epsilon(V) c_{b,\epsilon}(V) dP_{W,\epsilon} - \int_W d_0(V) c_{b,0}(V) dP_{W,0} \Big]
$$

$$-\tau_0\Big[\int_W d_0(V)c_{b,\epsilon}(V)dP_{W,\epsilon} - \int_W d_0(V)c_{b,0}(V)dP_{W,0}\Big]$$

We have $\eta_\epsilon > 0$ since $\eta_\epsilon - \eta_0 = O(\epsilon)$.

In the proof of the optimal closed-form solution, we observed that in this case of $\eta_0 > 0$, we have $\int_W d_P(V)c_{b,P}(V)dP_W = K - \underline{K}_P$ for any distribution P. Thus, the first line is $-\underline{K}_\epsilon + \underline{K}_0 = -\int_W c_T(0,W)dP_{W,\epsilon} - dP_{W,0} = -\epsilon\int_W c_T(0,W)H_W(W)dP_{W,0}$, where $H_W(W)$ is a score function. Dividing by ϵ and taking the limit, the contribution to the canonical gradient is from the first line is $-\tau_0 \cdot (c_T(0,W) - \underline{K}_0)$.

Noting that $\int_W d_0(V)c_{b,\epsilon}(V)dP_{W,\epsilon} = \int_W d_0(V)[c_\epsilon(1,W) - c_\epsilon(0,W)]dP_{W,\epsilon}$, we get for the second line $-\tau_0\int_W d_0(V)[c_T(1,W) - c_T(0,W)]\epsilon H_W(W)dP_{W,0}$ for score function $H_W(W)$, so dividing by ϵ and taking the limit, and noting that $\int_W d_0(V)c_{b,0}(V)dP_{W,0} = K - \underline{K}_0$, we get the contribution to the canonical gradient is $-\tau_0(d_0(V)(c_T(1,W) - c_T(0,W)) - (K - \underline{K}_0))$. We add the contribution from the first line to get the term $-\tau_P(E_U c_T(d(V),W) - K)$ of D_0 as given in theorem 2.

Proof of theorem 3 (asymptotic linearity and efficiency).

Lemma 10 (Expression for higher-order remainder terms.) *Let P_0, P be distributions which satisfy the positivity conditions C1), C2), and for which Y is bounded in probability. We have that*
$\Psi_d(P) - \Psi(P_0) = -P_0[D^*(d,\tau_0,P)] + R_0(d,\tau_0,P)$.

- Proof of lemma 10.

$$\Psi_d(P) - \Psi(P_0) + P_0[D^*(d,\tau_0,P)]$$

$$= \Psi_d(P) - \Psi_d(P_0) + \sum_{j=1}^{2} P_0[D_j(d,P)]$$

$$+\Psi_d(P_0) - \Psi_{d_0}(P_0) - \tau_0 E_{P_0}[c_T(d(V),W) - K]$$

Standard calculations show that the first term on the right equals $R_1(d,P)$. The second term equals $E_{P_0}[(d-d_0)\mu_{b,0} - \tau_0[c_T(d(V),W) - c_T(d_0(V),W)]]$, noting that $E_{P_0}(c_T(d_0(V),W)) = K$. Since $c_T(d(V),W) = d(V)c_{b,0}(V) + c_T(Z=0,W)$, we get that the second term on the right equals $R_2(d,\tau_0,P)$.

- Proof of Theorem 3.

$$
\begin{aligned}
& \Psi(P_n^*) - \Psi(P_0) \\
= \; & -P_0 D^*(d, \tau_0, P) + R_0(d, \tau_0, P) \\
= \; & (P_n - P_0) D^*(d, \tau_0, P) + R_0(d, \tau_0, P) + o_{P_0}(n^{-\frac{1}{2}}) \\
= \; & (P_n - P_0) D_0 + (P_n - P_0)[D^*(d, \tau_0, P) - D_0] + R_0(d, \tau_0, P) + o_{P_0}(n^{-\frac{1}{2}})
\end{aligned}
$$

The first line is from lemma 10, and the second line is from C6). The middle term on the last line is $o_{P_0}(n^{-\frac{1}{2}})$ from C1), C2), C4), and C5), and $R_0(d, P, K_T)$ is $o_{P_0}(n^{-\frac{1}{2}})$ from C3). This proves the claim about asymptotic linearity. Standard semiparametric theory gives the result about regularity and efficiency, see for instance (Bickel et al 1993).

Proof of theorem 4.

Define B_n to be the function $v \to T_n(V) - \tau_n$, and B_0 to be $v \to T_0(V) - \tau_0$. We omit the dependence of $B_n,\ B_0$ on V. For any $t > 0$, we have

$$
\begin{aligned}
R_2(d_n, \tau_0, P_n^*) \; & \leq P_0[|(d_n - d_0) B_0 \cdot c_{b,0}|] \\
& \leq C_U \cdot P_0[I(d_n \neq d_0)|B_0|] \\
& = C_U \cdot P_0[I(d_n \neq d_0)|B_0| I(0 < |B_0| \leq t)] \\
& + C_U \cdot P_0[I(d_n \neq d_0)|B_0| I(|B_0| > t)] \\
& \leq C_U \cdot P_0[|B_n - B_0| I(0 < |B_0| \leq t)] \\
& + C_U \cdot P_0[|B_n - B_0| I(|B_n - B_0| > t)] \\
& \leq C_U \big[\|B_n - B_0\|_{2,P_0} Pr_0(0 < |B_0| \leq t)^{1/2} + \\
& \qquad \sqrt{\|B_n - B_0\|_{2,P_0}^2 P_0 I(|B_n - B_0| > t)} \big] \\
& \leq C_U \big[\|B_n - B_0\|_{2,P_0} C_0^{1/2} t^{\alpha/2} + \sqrt{\|B_n - B_0\|_{2,P_0}^2 \frac{P_0|B_n - B_0|}{t}} \big] \\
& \leq C_U \big[\|B_n - B_0\|_{2,P_0} C_0^{1/2} t^{\alpha/2} + \frac{\|B_n - B_0\|_{2,P_0}^{3/2}}{t^{1/2}} \big]
\end{aligned}
$$

where C_U is an upper bound on the cost function $c_{b,0}$, and C_0 in the penultimate line is the constant implied by equation 4.4. The third from last line is from the Cauchy-Schwarz inequality and the next to last line is from the Markov inequality. The first part of theorem 4 follows by plugging $t = \|B_n - B_0\|_{2,P_0}^{1/(1+\alpha)}$ into the upper bound above.

We also have that

$$
\begin{aligned}
|R_2(d_n, \tau_0, P_n^*)| &\leq P_0[I(d_n \neq d_0)|B_0|C_U] \\
&\leq C_U \cdot P_0[I(0 < |B_0| \leq |B_n - B_0|)|B_0|] \\
&\leq C_U \cdot P_0[I(0 < |B_0| \leq \|B_n - B_0\|_{\infty,P_0})|B_0|] \\
&\leq C_U \|B_n - B_0\|_{\infty,P_0} Pr_{P_0}(0 < |B_0| \leq \|B_n - B_0\|_{\infty,P_0})
\end{aligned}
$$

By equation 4.4, and plugging in $t = \|B_n - B_0\|_{\infty,P_0}$, it follows that $R_2(d_n, \tau_0, P_n^*) \lesssim \|B_n - B_0\|_{\infty,P_0}^{1+\alpha}$.

Proof that $\tau_n \to \tau_0$ for lemma 15.

We show $\eta_n \to \eta_0$, and then the consistency of τ_n follows by the continuous mapping theorem.

- $\sup_V |T_n(V) - T_0(V)| = o_{P_0}(1)$ when $K_{T,0} > K$.

  (It would suffice to show $E_0|T_n(V) - T_0(V)| = o_{P_0}(1)$ for this part of the proof, and hence $\tau_n \to \tau_0$. However, we need that $T_n(V) \to T_0(V)$ uniformly over V in order to have that $d_n(V) = d_0(V)$ with probability approaching 1, so we prove the stronger result.)

Note that $P_{W,n|V}$ denotes the empirical distribution $P_{W,n}$ given V. Assume conditions E1)-E4) given in section 4.6 hold.

Invoking E1) and E2), we have

$$
\begin{aligned}
&\sup_V |T_n(V) - T_0(V)| = \sup_V \Big| \frac{\mu_{b,n}(V)}{c_{b,n}(V)} - \frac{\mu_{b,0}(V)}{c_{b,0}(V)} \Big| \\
&= \sup_V \Big| \frac{P_{W,n|V}\mu_{b,n}}{P_{W,n|V}c_{b,n}} - \frac{P_{W,0|V}\mu_{b,0}}{P_{W,0|V}c_{b,n}} \Big| \\
&\longrightarrow_{P_0} \sup_V \frac{|P_{W,n|V}\mu_{b,n} - P_{W,0|V}\mu_{b,0})|}{|P_{W,0|V}c_{b,n}|} \\
&\leq \sup_V \frac{|P_{W,n|V}\mu_{b,n} - P_{W,0|V}\mu_{b,n}|}{|P_{W,0|V}c_{b,n}|} + \sup_V \frac{|P_{W,0|V}\mu_{b,n} - EP_{W,0|V}\mu_{b,0}|}{|P_{W,0|V}c_{b,n}|} \\
&\leq \frac{1}{C_L} \sup_V |P_{W,n|V}\mu_{b,n} - P_{W,0|V}\mu_{b,n}| + \frac{1}{C_L} \sup_V |P_{W,0|V}\mu_{b,n} - P_{W,0|V}\mu_{b,0}| \\
&\longrightarrow_{P_0} o(1).
\end{aligned}
$$

In getting to the next to last line, we used the lower bound on cost $c_b(W) \geq C_L$.

- Next we show that $S_n(\eta) \longrightarrow S_0(\eta)$.

Let η be in the neighborhood of η_0 so that assumptions A3), A4) hold.

$$\begin{aligned}
&\big|S_n(\eta) - S_0(\eta)\big| \\
&= \big|P_n[I(T_n \geq \eta)c_{b,n}] - P_0[I(T_0 \geq \eta)c_{b,0}]\big| \\
&\leq \big|P_0[I(T_n \geq \eta)c_{b,n} - I(T_0 \geq \eta)c_{b,0}]\big| + \big|(P_n - P_0)[I(T_n \geq \eta)c_{b,n}]\big| \\
&\leq \big|P_0[I(T_n \geq \eta)c_{b,n} - I(T_0 \geq \eta)c_{b,0}]\big| + o_{P_0}(1)
\end{aligned}$$

The last line is because $S_n(\eta) = E_{V,n}[I(T_n(V) \geq \eta)c_{b,n}(V)]$ is a GC-class by condition E3).

Note that assumption A3) guarantees that, for η in a neighborhood of η_0, $P_0(T_0 = \eta) = 0$: otherwise, given the non-zero cost of treating any subgroup, we would have strictly larger cost of treating groups with $T_0(V) < \eta$ than the cost of treating groups with $T_0(V) > \eta$, so $S(x)$ would have infinite slope at η. Let $Z_n(\eta, V) \triangleq (I(T_n(V) \geq \eta) - I(T_0(V) \geq \eta))^2$. Then we have for all $q > 0$:

$$\begin{aligned}
&\big|P_0[I(T_n \geq \eta)c_{b,n} - I(T_0 \geq \eta)c_{b,0}]\big| \\
&\leq P_0\big|[I(T_n \geq \eta)c_{b,0} - I(T_0 \geq \eta)c_{b,0}]\big| + P_0\big|I(T_n \geq \eta)(c_{b,0} - c_{b,n})\big| \\
&\leq C_U P_0 Z_n(\eta, \cdot) + o_{P_0}(1) \\
&= C_U\big[P_0 Z_n(\eta, \cdot)I(|T_0 - \eta)| > q) + P_0 Z_n(\eta, \cdot)I(|T_0 - \eta| \leq q)\big] + o_{P_0}(1) \\
&= C_U\big[P_0 Z_n(\eta, \cdot)I(|T_n - T_0| > q) + P_0 Z_n(\eta, \cdot)I(0 < |T_0 - \eta| \leq q)\big] + o_{P_0}(1) \\
&\leq C_U\big[P_0(|T_n - T_0| > q) + P_0(0 < |T_0 - \eta| \leq q)\big] + o_{P_0}(1) \\
&\leq C_U\big[\frac{E_0|T_n(V) - T_0(V)|}{q} + P_0(0 < |T_0 - \eta| \leq q)\big] + o_{P_0}(1)
\end{aligned}$$

The third line follows from the upper bound C_U on $c_{b_0}(V)$, and on the assumption about $\sup_V |c_{b,n}(V) - c_{b,0}(V)|$. The fifth line is from the guarantee that $P_0(T_0(V) = \eta) = 0$, described above. Finally, the last line is from Markov's inequality.

Since the above derivation holds for any $q > 0$, and we proved that $E_0|T_n(V) - T_0(V)| = o_{P_0}(1)$, we can construct a sequence $q_n \longrightarrow 0$ such that $\frac{E_0|T_n(V) - T_0(V)|}{q_n} \longrightarrow_{P_0} 0$, and $P_0(0 < x \leq q_n) = 0$ as $q_n \longrightarrow 0$ for any r.v. x. Finally, we have $|S_n(\eta) - S_0(\eta)| = o_{P_0}(1)$.

- $\eta_n \longrightarrow \eta_0$.

Fix $\epsilon > 0$ small enough so that $\eta_0 + / - \epsilon$ is a neighborhood for the purposes of A3). Note that $\underline{K}_n = E_{W,n} c_T(Z = 0, W)$ converges to $\underline{K}_0 = P_{W,0} c_T(Z = 0, W)$. Thus, there is some N_1 s.t. for $n \geq N_1$, $|(K - \underline{K}_n) - (K - \underline{K}_0)| \leq \frac{\epsilon}{2c_x}$, where c_x is the constant from assumption A3) (Lipschitz continuity of S_0^{-1}). Let $\eta' = S_0^{-1}(K - \underline{K}_n)$. We have $|\eta_0 - \eta'| \leq \frac{\epsilon}{2}$ by A3). Let $\eta_l = \eta' - \frac{\epsilon}{2}$, $\eta_r = \eta' + \frac{\epsilon}{2}$. Then $S_0(\eta_l) > K - \underline{K}_n$, $S_0(\eta_r) < K - \underline{K}_n$. For n bigger than some N_2, $S_n(\eta_l) > K - \underline{K}_n$, $S_n(\eta_r) < K - \underline{K}_n$. Thus, $\eta_l \leq \eta_n = S_n^{-1}(K - \underline{K}_n) \leq \eta_r$ by the monotonicity of the $S(x)$ function. So $|\eta_n - \eta_0| \leq |\eta_n - \eta'| + |\eta' - \eta_0| \leq \epsilon$.

Finally, we get $\tau_n \longrightarrow \tau_0$.

Chapter 5

Optimal dynamic treatments.

5.1 The model.

We now turn our attention to the problem of estimating the mean counterfactual outcome under an optimal dynamic treatment regime, when an instrumental variable is needed to make consistent estimates of treatment effect. This allows us to answer questions like: what would be the mean outcome under an optimal treatment regime, if we could truly achieve compliance with treatment? In the case of the get-out-the-vote campaign, an interesting question is how effective would the optimal assignment of phone calls be in mobilizing voters, if people could actually be assigned to participate in the phone call. Once again, we have iid data $(W, Z, A, Y) \sim \mathcal{M}$, where $\mathcal{M}$ is a semiparametric model, and Z an instrument for identifying the effect of treatment A on outcome Y, when one has to account for unmeasured confounding. $V \subseteq W$ is an arbitrary fixed subset of the baseline covariates. In this chapter, the unknown optimal dynamic treatment rule $d(V)$ gives the intervention $A = d(V)$ to make on treatment variable A such that the optimal value of mean counterfactual outcome Y is attained.

We restrict our attention to a binary instrument. If Z is continuous or categorical, we can still use it for an instrumental variables model if it has two values $Z = z_1$, $Z = z_2$ such that those two values induce variation in A as per assumption 3, part 3. Treatment A can be binary, categorical, or continuous. The solutions sketched out in sections 4.3 for the case of a continuous treatment and linear cost function, or for a categorical treatment, are equally relevant here. We assume outcome Y is binary or continuous. Let $c_A(A, W)$ denote the cost function giving the cost of treating an individual having covariates Z with treatment A. When A is continuous, we assume the treatment variable represents the magnitude of treatment in $[A_{min}, A_{max}]$, and we make the simplifying assumption that we have a linear cost function $c_A(A,W) = c_A(W)A$. To make our notation consistent, we assume continuous treatment values are converted to span the unit interval. Finally, we have a cost constraint $Ec_A(A,W) \leq K$, for a fixed cost K.

Notation. We again assume wlog that treatment $A = 0$ is the baseline treatment with lower cost for all V: $E_{W|V}c_A(0,W) \leq E_{W|V}c_A(1,W)$. [1] Let $\underline{K}_0 \triangleq E_W c_A(0,W)$ be the cost of not assigning any individuals to treatment, and $K_{T,0} \triangleq E_W c_A(1,W)$ be the total cost of assigning everyone, and we assume a nontrivial constraint $K \leq K_{T,0}$. We have P_W, $\Pr_{V|W}$, $\rho(W) \triangleq \Pr(Z=1|W), \Pi(W,Z) \triangleq E(A|W,Z)$, $\mu(W,Z) \triangleq E(Y|W,Z)$ defined as before, and notation $c_b(V) \triangleq E_{W|V}\big[c_T(Z=1,W) - c_T(Z=0,W)\big]$. We also use notation $m(V) \triangleq E_{W|V}\big[m(Z=1,W) - m(Z=0,W)\big]$, where m is the causal effect function defined in the causal assumptions.

Causal model. We assume the same causal model as chapter 3. These assumptions guarantee that $E(Y(A=a))$ equals $E_W m(W)a + \theta(W)$ for identifiable functions m, θ.

Recall our structural equation notation, $W = f_W(U_W), Z = f_Z(W,U_Z), A = f_A(W,Z,U_A), Y = f_Y(W,Z,A,U_Y)$ where $U = (U_W,U_Z,U_A,U_Y) \sim P_{U,0}$ is an exogenous random variable, and f_W, f_Z, f_A, f_Y may be unspecified or partially specified (for instance, we might know that the instrument is randomized). U_Y is possibly confounded with U_A.

Assumption 3 ***Assumptions ensuring that Z is a valid instrument:***

1. ***Exclusion restriction.*** *Z only affects outcome Y through its effect on treatment A. Thus, $f_Y(W,Z,A,U_Y) = f_Y(W,A,U_Y)$.*

2. ***Exogeneity of the instrument.*** *$E(U_Y|W,Z) = 0$ for any W,Z.*

3. *Z* ***induces variation in*** *A.* *$Var_0[E_0(A|Z,W)|W] > 0$ for all W.*

Structural equation for outcome Y:

4. *$Y = Am(W) + \theta(W) + U_Y$ for continuous Y, and*
$\Pr(Y=1|W,A,\tilde{U}_Y) = Am(W) + \theta(W) + \tilde{U}_Y$ for binary Y,

[1] Again we can forgo this assumption by introducing notation, i.e. $A = l(V)$ is the lower cost treatment value for a stratum defined by covariates V.

where $U_Y = (\tilde{U}_Y, U'_Y)$ *for an exogenous r.v.* U'_Y[2] *, and* m*,* θ *are unspecified functions.*

Assumptions 2 and 4 yield that, whether Y is binary or continuous,

$$E(Y|W,Z) = m_0(W)\Pi_0(W,Z) + \theta_0(W)$$

We have a saturated model for $f_Y(W,Z,U_Y)$ when A is binary. It should also be noted that we don't require the instrument to be randomized with respect to treatment ($U_Z \perp\!\!\!\perp U_A|\ W$ is not necessary).

Causal parameter of interest.

$$\Psi(P_0) \triangleq \text{Max}_d\ E_0 Y(A = d(V)) \text{ s.t. } E_0[c_A(A = d(V), W)] \leq K \tag{5.1}$$

Identifiability. $m(W)$ is identified as $\big[(\mu(W,Z=1) - \mu(W,Z=0))/(\Pi(W,Z=1) - \Pi(W,Z=0))\big]$. $\theta(W)$ is identified as $\big[\mu(W,Z) - \Pi(W,Z) \cdot m(W)\big]$. Lemma 1 (ch. 3) states this as a formal identifiability result.

Statistical model. The above stated causal model implies the statistical model $\mathcal{M}$ consisting of all distributions $\mathcal{P}$ of $O = (W,Z,A,Y)$ satisfying $E_P(Y|W,Z) = m_P(W) \cdot \Pi_P(W,Z) + \theta_P(W)$. Here m_P and θ_P are unspecified functions and $\Pi_P(W,Z) = E_P(A|W,Z)$ such that $Var_P(\Pi_P(Z,W)|W) > 0$ for all W. Note that the regression equation $E_P(Y|W,Z) = m_P(W) \cdot \Pi_P(W,Z) + \theta_P(W)$ is always satisfied for some choice of $m(W)$, $\theta(W)$ when Z is binary. The distribution for the instrument $\rho(W)$ may or may not be known, and we generally think of all other components P_W, Π, m, θ as unspecified.

Statistical target parameter.

Lemma 11 *The causal parameter given in 5.1 is identified by the statistical target parameter:*

$$\Psi_0 = E_{P_{W,0}} m_0(W) d_0(V) + \theta_0(W) \tag{5.2}$$

[2]The U'_Y term is an exogenous r.v. whose purpose is for sampling binary Y. Let U'_Y be a Unif[-.5,.5] r.v. (we set it to have 0 mean to conform to assumption 2.) Then $Y = 1((U'_Y + .5) < \tilde{f}_Y(W,Z,A,\tilde{U}_Y))$

Note that optimal decision rule d_0 is a function of $m_0, P_{W,0}$. For Ψ_0 we also use the notation $\Psi(P_{W,0}, m_0, \theta_0)$, or alternately $\Psi(P_{W,0}, \Pi(W,Z), \mu(W,Z))$, using the above identifiability results.

This lemma follows from our causal assumptions:
$\Psi(P_0) = EY(A = d_0(V)) = E_W E_{U_Y|W} EY(A = d_0(V)|W, U_Y)$
The right hand side becomes $E_W E_{U_W|Y}(m(W)d_0(V) + \theta(W) + U_Y)$ for a continuous Y, and $E_W E_{U_W|Y}(m(W)d_0(V) + \theta(W) + \tilde{U}_Y)$ for a binary Y.

5.2 Assumptions for identifiability and pathwise differentiability.

We use notation $d_0 = d_{P_0}$, $\tau_0 = \tau_{P_0}$, etc.

A1) The four causal assumptions denoted Assumption 3 above are satisfied. (Z is a valid instrument for identifying the causal effect of A on Y, and a marginal structural model guarantees identifiability.)

A2) Positivity assumption: $0 < \rho_0(W) < 1$.

A3) There is a neighborhood of η_0 where $S_0(x)$ is Lipschitz continuous, and a neighborhood of $S_0(\eta_0) = K - \underline{K}_0$ where $S_0^{-1}(y)$ is Lipschitz continuous. We have $|\eta_1 - \eta_2| < \epsilon \Rightarrow |S_0(\eta_1) - S_0(\eta_2)| < c_y\epsilon$ for $\eta_1,\ \eta_2$ in a δ_x-neighborhood of η_0; and $|y_1 - y_2| < \epsilon \Rightarrow |S_0^{-1}(y_1) - S_0^{-1}(y_2)| < c_x\epsilon$, for $y_1,\ y_2$ in a δ_y-neighborhood of $K - \underline{K}_0$, for some constants $c_x,\ c_y$.

A4) $Pr_0(T_0(V) = \tau) = 0$ for all τ in a neighborhood of τ_0.

Note that A3) implies that $S_0^{-1}(K - \underline{K}_0)$ exists.

Condition A4) ensures a non-exceptional law, so that Ψ_0 is pathwise differentiable. The optimal rule $d(V)$ over the class of non-deterministic rules is a deterministic rule in this case (Luedtke and van der Laan, 2016a).

5.3 Closed-form solution of optimal rule for the case of binary treatment.

The closed-form solution is the same as in chapter 4, with the two main modifications that: 1) replace intervention variable Z with A, and 2) replace $\mu_b(W)$ with $m(W)$. These latter quantities represent the effect on Y of applying the intervention vs the baseline treatment (at Z or A).

Again let $d(V) = 0$ denote the baseline treatment having lower cost given V. Define a scoring function $T(V) = \frac{m(V)}{c_b(V)}$ for ordering subgroups (given by V) based on the effect of treatment per unit cost. In the optimal treatment rule, all groups with the highest $T(V)$ are assigned deterministically, which under our assumptions has cost exactly K and there is a unique solution. We write $T_P(V)$ to make explicit the dependence on $P_W, m(W)$ from distribution P.

Again we have a function $S_P(x) = E_{P_V}[I(T_P(V) \geq x)(c_{b,P}(V)]$ which gives the expected (additional above-baseline) cost of treating all subgroups having $T_P(V) \geq x$. Cutoff η and τ for T_V are defined as before, as is the optimal decision rule $d(V)$:

$$\eta_P = S_P^{-1}(K - \underline{K}_P)$$

$$\tau_P = \max\{\eta_P, 0\}$$
$$d_P(V) = 1 \text{ iff } T_P(V) \geq \tau_P$$

Lemma 12 *Assume A2)-A4). Then the optimal decision rule d_0 for parameter Ψ_0 as defined in 5.2 is the deterministic solution*
$d_0(V) = 1$ iff $T_0(V) \geq \tau_0$, with $T_0,\ \tau_0$ as defined above.

The proof that this is the unique optimal solution is the same as in chapter 4. The extensions given in section 4.3 to a continuous intent-to-treat variable Z under a linear cost function, and to a categorical Z, apply here as well.

5.4 Efficient influence curve $D^*(\Psi_0)$.

Lemma 13 *Let*

$$J_0(Z,W) = \frac{I(Z=1)}{\rho_0(W)} + \frac{\left(\frac{I(Z=1)}{\rho_0(W)} - \frac{I(Z=0)}{1-\rho_0(W)}\right)\left(d(V) - \Pi_0(W,Z=1)\right)}{\Pi_0(W,Z=1) - \Pi_0(W,Z=0)}$$

The efficient influence curve $D^*(\Psi_0)$ is

$$D^*(\Psi_0) = -\tau_0 E_{P_0}[c_T(d(V), W) - K] \tag{5.3}$$
$$+m_0(W)d(V) + \theta_0(W) - \Psi_0 \tag{5.4}$$
$$-J_0(Z, W)m_0(W)\big[A - \Pi_0(W, Z)\big] \tag{5.5}$$
$$\tag{5.6}$$
$$+J_0(Z, W)\big[Y - (m_0(W)\Pi_0(W, Z) - \theta_0(W))\big] \tag{5.7}$$

We also write $D^*(d_0, \tau_0, P_0)$. For convenience, denote lines (1)-(4) of D^* above as D^*_c, D^*_W, D^*_A, and D^*_Y, respectively.

The derivation for this is given in section 5.8.

Checking that D^* is in tangent space T_{P_0}.

We need to check that $D^*(\Psi_0)$ is in the tangent space $T_{P,0}$ of the semiparametric statistical model given in 5.1. We denote the subspaces as $T_{P,0} = T_Y + T_A + T_Z + T_W$, where T_Y is the tangent space of the $\Pr(Y|W, Z, A)$ component in our model $\mathcal{M}$, and so forth for T_A, T_Z, T_W. Note that components D^*_W, D^*_c are mean-zero functions of W, so $D^*_W + D^*_c$ is clearly in T_W given our nonparametric model for $\Pr(W)$.

Further note that for binary Z, any function $h(Z, W)$ can be written as $h_1(W)Z + h_2(W)$, so we can write $J_0(Z, W)$ in this form.

Y is continuous.

We need to verify that there are score functions s_P spanning $D^*_Y + D^*_A$ at model P with respect to paths $\{P_\epsilon\}$ s.t. $P, \{P_\epsilon\} \subseteq \mathcal{M}$, using statistical model $\mathcal{M}$ as given in 5.1. Let $\Pr(W)$, $\rho(W) = \Pr(Z|W)$, $\Pi(Z, W)$, $m(W)$, and $\theta(W)$ be arbitrary functions. The first D^*_W component is clearly in the tangent space of $\Pr_0(W)$. Let $\Pr(Y|W, Z, A) = \text{pnorm}(r(W, A, Y))$, where we have residual $r(W, A, Y) = Y - m(W)A - \theta(W)$. Thus, Y is normally distributed, with no confounding term in the semiparametric regression model. Finally, let $\Pr_\epsilon(Y|W, Z, A) = \text{pnorm}(r_\epsilon(W, Z, A, Y))$, where $r_\epsilon(W, Z, A, Y) = Y - m(W)A - \theta(W) - \epsilon(\Pi(W, Z)h_1(W) + h_2(W))$. Note that $\Pr_\epsilon(Y|W, Z, A)$ stays in our model, because we have $Y = (m(W) + \epsilon h_1(W))A + (\theta(W) + \epsilon h_2(W)) + \epsilon(\Pi(Z, W) - A)h_1(W)$, where the last term is mean 0 given W, Z. Then we have $s_P = (h_1(W)Z + h_2(W))(Y - m(W)A - \theta(W))$, so $(D^*_Y + D^*_A)$ is in the tangent space of $P(Y|W, Z, A)$.

Y is binary.

Fix arbitrary $\Pr(W)$, $\Pr(Z|W)$, $\Pi(W, Z)$, $m(W)$, and $\theta(W)$. We first define a fluctuation P_ϵ on $E(A|W, Z) = \Pi(W, Z)$. If A is binary, we have

$\Pr(A = 1|W, Z) = \Pi(Z, W)$, and use fluctuation $\Pi_\epsilon(W, Z) = \Pi(W, Z) + \epsilon\Pi(W, Z)(1 - \Pi(W, Z))(h_1(W)Z + h_2(W))m(W)$. If A is continuous, we use the model $A = \Pi(W, Z) + r_A(W, Z)$, where $r_A(W, Z)$ is a $N(0, 1)$ residual term, and we use fluctuation $A = \Pi(W, Z) + \epsilon m(W)(h_1(W)Z + h_2(W)) + r_A(W, Z)$ for residual term $r_A(W, Z) \sim N(0, 1)$. In either case, it is straightforward to check that the score with respect to P_ϵ is $(h_1(W)Z + h_2(W))m(W)(A - \Pi(W, Z))$.

We now define another fluctuation P_ϵ on $\Pr(Y|W, A, Z)$. We use model $\Pr(Y|W, Z, A) = m(W)A + \theta(W) + m(W)(\Pi(W, Z) - A) = m(W)\Pi(W, Z) + \theta(W) \triangleq \mu(W, Z)$. Note that $\Pr(Y|W, Z, A) \subseteq \mathcal{M}$. Set P_ϵ so that $\Pr(Y|W, Z, A) = m_\epsilon(W)A + \theta_\epsilon(W) + U_Y = m_\epsilon(W)\Pi(W, Z) + \theta_\epsilon(W)$. We have that $\mu(W, Z)$ is fluctuated by $\epsilon(\mu(W, Z)(1 - \mu(W, Z))(h_1(W)Z + h_2(W))) \triangleq \epsilon\mu'(W, Z)$. Note that since Z is binary, $\mu'(W, Z)$ can be written as $\mu'(W, Z) = l_1(W)Z + l_2(W)$. Since $\Pi(W, Z = 1) \neq \Pi(W, Z = 0)$ for any W and Z is binary, simple algebra shows we can write $\mu'(W, Z)$ as $m'(W)\Pi(W, Z) + \theta'(W)$ for some m', θ'. Thus, we have $\Pr_\epsilon(Y|W, Z, A) = (m(W) + \epsilon m'(W))\Pi(Z, W) + (\theta(W) + \epsilon\theta'(W))$, and so clearly $P_\epsilon \subseteq \mathcal{M}$. It follows that $s_P = (h_1(W)Z + h_2(W))(Y - m(W)\Pi(W, Z) - \theta(W))$.

Note that if $\mu(Z, W)$ equals 0 or 1 above, then there is a potential issue with the fluctuation moving outside permissible bounds $[0, 1]$. However, in this case D^* vanishes as the $(Y - \mu(Z, W))$ term is 0, and the fluctuation $\mu'(Z, W)$ has value 0. Similar comments apply to binary A, $\Pi(W, Z) = 0$ or 1.

Finally, it follows that $D^*_Y + D^*_A$ is in tangent space $T_Y + T_A$.

5.5 Remainder term.

We calculate the remainder term $R(P, P_0) \triangleq \Psi(P) - \Psi(P_0) + P_0(D^*(P))$ and verify that it is second-order. $J(Z, W)$ is the coefficient as defined above. We have

$$R_0(P, P_0) = \tag{5.8}$$
$$\Psi_d(P) - \Psi(P_0) + P_0[D^*(d, \tau_0, P)] \tag{5.9}$$
$$= \Psi_d(P) - \Psi_d(P_0) + P_0 D^*_W(d, P) + P_0 D^*_A(d, P) + P_0 D^*_Y(d, P) \tag{5.10}$$
$$+\Psi_d(P_0) - \Psi_{d_0}(P_0) - \tau_0 E_{P_0}[c_T(d(V), W) - K] \tag{5.11}$$

We can write $R_0(P, P_0) = R_d(P, P_0) + R_c(P, P_0)$, where R_d, R_c are lines 5.10 and 5.11, respectively. R_d is the remainder term when estimating Ψ treating the estimated decision rule as correct, while R_c is the remainder term arising from the estimation of d.

We have

$$R(P,P_0) = R_d(P,P_0) + P_0\big[(d-d_0)(m_0(V) - \tau_0 c_{b,0}(V))\big] \tag{5.12}$$

with

$$\begin{aligned}
&R_d(P,P_0) = \qquad (5.13)\\
&P_0\left[\frac{\rho_0-\rho}{\rho}\Big((m_0-m)\Pi_0(W,1) + (\theta_0-\theta)\Big)\Big(1+\frac{d-\Pi(W,1)}{\Pi(W,1)-\Pi(W,0)}\Big)\right]\\
&+P_0\left[\frac{\rho-\rho_0}{1-\rho}\Big((m_0-m)\Pi_0(W,0) + (\theta_0-\theta)\Big)\Big(\frac{-(d-\Pi(W,1))}{\Pi(W,1)-\Pi(W,0)}\Big)\right]\\
&+P_0\Bigg[(m_0-m)\Big((\Pi_0(W,1)-\Pi(W,1)) - (\Pi_0(W,0)-\Pi(W,0))\Big)\\
&\qquad\times\Big(\frac{d-\Pi(W,1)}{\Pi(W,1)-\Pi(W,0)}\Big)\Bigg]\\
&+P_0\big((m_0-m)(\Pi_0(W,1)-\Pi(W,1))\big)
\end{aligned}$$

We can alternatively write $R(P,P_0)$ as a function of the estimated distribution $R_0(d,\tau_0,P)$.

Lemma 14 *(Expression for higher-order remainder terms.) Let P_0, P be distributions which satisfy the positivity assumptions C1), C2), and for which Y is bounded in probability. Then*

$$\Psi(P) - \Psi(P_0) = -P_0[D^*(d,\tau_0,P)] + R_0(d,\tau_0,P)$$

with $R_0(d,\tau_0,P)$ as given in equations 5.12, 5.13.

5.6 TMLE-based estimators.

We propose two different TMLE-based estimators for Ψ_0. The first estimator is analogous to the iterative estimator and the second to the non-iterative logistic estimator given in chapter 3. The important advantage of the iterative one is that it involves a standard, numerically well-behaved and easily understood, likelihood maximization operation at each step. The second estimator, on the other hand, has the advantages that 1) there is a one-step solution with no need to reach convergence iteratively, and 2) the estimate μ respects the bounds of Y found in the data.

Iterative estimator.

The relevant components for estimating $\Psi(Q) = E_W[m(W)d(V) + \theta(W)]$ are $Q = (P_W, m, \theta)$. The nuisance parameters are $g = (\rho, \Pi)$. $d(V)$ and τ can be thought of as functions of P_W, m here. Let
$h_1(W) \triangleq \frac{1}{\rho(W)(\Pi(W,1)-\Pi(W,0))} + \frac{d(V)-\Pi(W,1)}{(\Pi(W,1)-\Pi(W,0))^2}\frac{1}{\rho(W)(1-\rho(W))}$. Also, let
$h_2(W) \triangleq \frac{1}{\rho}\big[1 - \frac{\Pi(W,1)}{\Pi(W,1)-\Pi(W,0)} + \frac{d-\Pi(W,1)}{\Pi(W,1)-\Pi(W,0)}(1 - \frac{\Pi(W,1)}{\Pi(W,1)-\Pi(W,0)}\frac{1}{1-\rho})\big]$.
Then we have that $D^*_Y = (h_1\Pi + h_2)(Y - m\Pi - \theta)$.

If A is not binary, convert A to the unit interval via a linear transformation $A \to \tilde{A}$ so that $\tilde{A} = 0$ corresponds to $A_{\min}$ and $\tilde{A} = 1$ to $A_{\max}$. We assume $A \in [0, 1]$ from here.

1. Use the empirical distribution $P_{W,n}$ to estimate P_W. Make initial estimates of $Q = \{m_n(W),\ \theta_n(W)\}$ and $g_n = \{\rho_n(W), \Pi_n(W, Z)\}$ using any strategy desired. Data-adaptive learning using Super Learner is recommended.

2. The empirical estimate $P_{W,n}$ gives an estimate of $Pr_{V,n}(V) = E_{W,n}I(F_V(W) = V)$, $\underline{K}_n = E_{W,n}c(0, W)$, $K_{T,n} = E_{W,n}c(1, W)$, and $c_{b,n}(V) = E_{W,n|V}(c_T(1, W) - c_T(0, W))$.

3. Estimate $m_n(V)$ as $E_{W,n|V}m(W)$.

4. Estimate $T_0(V)$ as $T_n(V) = \frac{m_n(V)}{c_{b,n}(V)}$.

5. Estimate $S_0(x)$ using $S_n(x) = E_{V,n}[I(T_n(V) \geq x)(c_{b,n}(V))]$.

6. Estimate η_0 as using $\eta_n = S_n^{-1}(K - \underline{K}_n)$

   and $\tau_n = \max\{0, \eta_n\}$.

7. Estimate the decision rule as $d_n(V) = 1$ iff $T_n(V) \geq \tau_n$.

ITERATE STEPS 8)-9) UNTIL CONVERGENCE:

8. Fluctuate the initial estimate of $m_n(W),\ \theta_n(W)$ as follows: Using $\mu_n(W, Z) = m_n(W)\Pi_n(W, Z) + \theta_n(W)$, run an OLS regression:

$$\begin{aligned} \text{Outcome:} &\quad (Y_i : i = 1, \ldots, n) \\ \text{Offset:} &\quad (\mu_n(W_i, Z_i), i = 1, \ldots, n) \\ \text{Covariate:} &\quad (h_1(W_i)\Pi(W_i, Z_i) + h_2(W_i) : i = 1, \ldots, n) \end{aligned}$$

Let ϵ_n represent the level of fluctuation, with
$\epsilon_n = \text{argmax}_\epsilon \frac{1}{n}\sum_{i=1}^n (Y_i - \mu_n(\epsilon)(W_i, Z_i))^2$ and
$\mu_n(\epsilon)(W,Z) = \mu_n(W,Z) + \epsilon(h_1(W)\Pi(W,Z) + h_2(W))$.

Note that $\mu_n(\epsilon) = (m_n + \epsilon h_1)\Pi + (\theta_n + \epsilon h_2)$ stays in the semiparametric regression model.

Update m_n to $m_n(\epsilon)$, θ_n to $\theta_n(\epsilon)$.

9. Now fluctuate the initial estimate of $\Pi_n(Z, W)$ as follows: Use covariate $J(Z, W)$ as defined in lemma 13. Run a logistic regression using:

$$\begin{aligned} \text{Outcome:} &\quad (A_i : i = 1, \dots, n) \\ \text{Offset:} &\quad (\text{logit}\Pi_n(W_i, Z_i), i = 1, \dots, n) \\ \text{Covariate:} &\quad (K(W_i, Z_i)m(W_i) : i = 1, \dots, n) \end{aligned}$$

Let ϵ_n represent the level of fluctuation, with
$\epsilon_n = \text{argmax}_\epsilon \frac{1}{n}\sum_{i=1}^n [\Pi_n(\epsilon)(W_i, Z_i)\log A_i + (1-\Pi_n(\epsilon)(W_i, Z_i))\log(1-A_i)]$
and $\Pi_n(\epsilon)(W,Z) = \text{logit}^{-1}(\text{logit}\Pi_n(Z,W) + \epsilon K(W,Z)m(W))$.
Update Π_n to $\Pi_n(\epsilon)$. Also update $h_1(W)$, $h_2(W)$ to reflect the new Π_n.

---

10. Finally, form final estimate of $\Psi_0 = \Psi_{d_0}(P_0)$ using the plug-in estimator

$$\Psi^* = \Psi_{d_n}(P_n^*) = \frac{1}{n}\sum_{i=1}^n m_n^*(W_i) \cdot d_n(V_i) + \theta_n^*(W_i)$$

- Showing that $P_n D(d_n, \tau_0, P_n^*) = 0$.

The usual calculations (given in 3.2) show that for the linear fluctuation $m_(\epsilon)$, $\theta_n(\epsilon)$ with quadratic loss function, $\frac{d}{d\epsilon}L(Q_n(\epsilon|g_n), g_n, (O_1, ..., O_n))|_{\epsilon=0} = J_n(Z,W)(Y - \mu_n)$, so we have $P_n D_Y^* = 0$ upon convergence. Similarly, for the logistic fluctuation $\Pi_n(\epsilon)$ with logistic loss function, we have $\frac{d}{d\epsilon}L(Q_n(\epsilon|g_n), g_n, (O_1, ..., O_n))|_{\epsilon=0} = J_n(Z,W)m_n(A - \Pi_n)$, and $P_n D_A^* = 0$ upon convergence. As usual $P_n D_W^* = 0$ when using the empirical distribution $P_{W,n}$. Lastly, we have $P_n D_c(d_n, \tau_0, P_n^*) = 0$ for the third term of the canonical gradient as well, because $E_{V,n}E_{W,n|V}c_T(d_n(V), W) = K$, unless $\tau_0 = 0$. (This is described in the proof of optimality of the closed-form solution.)

Non-iterative, logistic estimator.

As in chapter 3, several variants of a non-iterative TMLE-based estimator are possible. We present here one based on a logistic fluctuation and logistic loss function, as this enables us to match the observed bounds of Y. (Gruber and van der Laan 2010) have shown performance gains of this feature when the data is sparse. This estimator essentially parallels the logistic estimator given in chapter 3, so we present it briefly. We fluctuate $m_n(W) \to m_n(\epsilon)(W)$ so that the efficient influence curve equation is directly solved:
$P_n D^*(Q_n^* = \{m_n^0(\epsilon), \theta_n, P_{W,n}\}, g_n = \{\rho_n, \Pi_n\}) = 0$.

Assume Y has been converted via a linear transformation $Y \to \tilde{Y}$, so that 0 corresponds to the minimum value and 1 to the maximum. We use notation $\tilde{Y}, \tilde{m}, \tilde{\theta}$ to reflect the change of scale. Knowing that $EY(A=1|W) - EY(A=0|W) = m_0(W)$, we can think of $\tilde{m}_0$ as being in $[-1,1]$. We bound $\tilde{m}_n^0$ to be in $[-1,1]$. We use fluctuation

$$\tilde{m}_n^0(\epsilon)(W) = 2 \times \text{logistic}(\text{logit}(\frac{\tilde{m}_n^0(W)+1}{2}) + \epsilon) - 1 \tag{5.14}$$

This corresponds to the mapping $f(\epsilon) = \text{logistic}(\text{logit}(f) + \epsilon)$ where f is $\tilde{m}_n^0$ scaled to be in $[0,1]$.

Thus, we get that $D^*(Q^*, g_n) = 0$ reduces to

$$\frac{1}{n}\sum_{i=1}^{n} \left[J_n(W,Z)(\tilde{Y} - A \cdot \tilde{m}_n^0(\epsilon)(W) - \tilde{\theta}_n(W))\right] = 0 \tag{5.15}$$

Since $E_0(\tilde{Y} - A \cdot \tilde{m}_0(W) - \tilde{\theta}_0(W)|W,Z) = 0$, the equation above has a solution ϵ for any reasonable estimates (Q_n^0, g_n). We can use bisection as a computationally simple method to find ϵ. Once one solves for ϵ and finds $\tilde{m}_n^* = \tilde{m}_n^0(\epsilon)$, one converts back to the original scale for outcome Y, by setting $m_n^* = \tilde{m}_n^* \cdot (\max(Y) - \min(Y))$. Then the parameter of interest is evaluated by finding $\Psi(d_n, P_n^*)$.

5.7 Theoretical results: efficiency, double robustness, and valid inference for Ψ_n^*.

Conditions for efficiency.

These six conditions are needed to prove asymptotic efficiency (theorem 5). When all relevant components and nuisance parameters are consistently specified, C3) and C4) hold, while C6) holds by construction of the TMLE estimator.

C1) $\rho_0(W)$ satisfies the strong positivity assumption: $Pr_0(\delta < \rho_0(W) < 1-\delta) = 1$ for some $\delta > 0$.

C2) The estimate $\rho_n(W)$ satisfies the strong positivity assumption, for a fixed $\delta > 0$ with probability approaching 1, so we have $Pr_0(\delta < \rho_n(W) < 1-\delta) \to 1$.

Remainder terms $R_0(d, \tau, P)$ are defined in section 5.5.

C3) $R_0(d_n, \tau_0, P_n^*) = o_{P_0}(n^{-\frac{1}{2}})$.

C4) $P_0[(D(d_n, \tau_0, P_n^*) - D_0)^2] = o_{P_0}(1)$.

C5) $D(d_n, \tau_0, P_n^*)$ belongs to a P_0-Donsker class with probability approaching 1.

C6) $\frac{1}{n}\sum_{i=1}^{n} D(d_n, \tau_0, P_n^*)(O_i) = o_{P_0}(n^{-\frac{1}{2}})$.

Sufficient conditions for lemma 15.

The following four conditions are sufficient for lemma 15. When all relevant components and nuisance parameters converge at a $O(n^{-\frac{1}{2}})$ rate, asymptotic efficiency as given in theorem 5 holds, and we expect all these conditions below to also hold, giving lemma 15.

E1) GC-like property for $c_b(V)$, $m_n(V)$: $\sup_V |(E_{W,n|V} - E_{W,0|V})m_n(W)| = o_{P_0}(1)$, similarly for $c_b(V)$.

E2) $\sup_V |E_{W,0|V}\mu_{b,n}(W) - E_{W,0|V}\mu_{b,0}(W)| = o_{P_0}(1)$
(This is needed for the proof that $d_n(V) = d_0(V)$ with probability approaching 1.)

E3) $S_n(x)$, defined as $x \to E_{V,n}[I(T_n(V) \geq x)c_{b,n}(V)]$ is a GC-class.

E4) Convergence of ρ_n, Π_n, m_n, θ_n to ρ_0, Π_0, m_0, θ_0, respectively, in $L^2(P_0)$ norm at a $O(n^{-1/2})$ rate (this is needed in several places).

Theoretical properties of Ψ_n^*.

Theorem 5 (Ψ^* is asymptotically linear and efficient.) *Assume assumptions A1)-A4) and conditions C1)-C6). Then* $\Psi^* = \Psi(P_n^*) = \Psi_{d_n}(P_n^*)$ *as defined by the TMLE procedure is a RAL estimator of* $\Psi(P_0)$ *with influence curve* $D^*(\Psi_0)$*, so*

$$\Psi(P_n^*) - \Psi(P_0) = \frac{1}{n}\sum_{i=1}^{n} D^*(\Psi_0)(O_i) + o_{P_0}(n^{-\frac{1}{2}}).$$

Further, Ψ^* *is efficient among all RAL estimators of* $\Psi(P_0)$.

Inference. Let $\sigma_0^2 = Var_{P_0} D^*(\Psi_0)$. By theorem 5 and the central limit theorem, $\sqrt{n}(\Psi(P_n^*) - \Psi(P_0))$ converges in distribution to a $N(0, \sigma_0^2)$ distribution. Let $\sigma_n^2 = \frac{1}{n}\sum_{i=1}^{n} D^*(d_n, \tau_n, P_n^*)(O_i)^2$ be an estimate of σ_0^2.

Lemma 15 *Under the conditions C1)-C2), and E1)-E4) we have* $\sigma_n \longrightarrow_{P_0} \sigma_0$. *Thus, an asymptotically valid 2-sided* $1-\alpha$ *confidence interval is given by*

$$\Psi^* \pm z_{1-\frac{\alpha}{2}} \frac{\sigma_n}{\sqrt{n}}$$

where $z_{1-\frac{\alpha}{2}}$ denotes the $(1-\frac{\alpha}{2})$-quantile of a $N(0,1)$ r.v.

Proofs.

The proof of theorem 5, and the proof that $\tau_n \to \tau_0$ as needed for lemma 15, are essentially the same as the corresponding proofs in chapter 4 section 4.8, with μ_b replaced by m.

Double robustness of Ψ_n^*.

We have a double robustness result for Ψ_0. As in chapter 4, Ψ^* is not a double robust estimator of Ψ_0: component $m(W)$ must always be consistently specified as a necessary condition for consistency of Ψ^*. However, if we consider $\Psi^* = \Psi_{d_n}(P_n^*)$ as an estimate of $\Psi_{d_n}(P_0)$, where the optimal decision rule $d_n(V)$ is estimated from the data, then we have that Ψ^* is double robust:

Lemma 16 (Ψ^* is a double robust estimator of $\Psi_{d_n}(P_0)$.) *Assume A1)-A4) and C1)-C2). Also assume* $Var_0(D_d(d_n, P_n^*)) < \infty$.
Then $\Psi^* = \Psi(d_n, P_n^*)$ *is a consistent estimator of* $\Psi_{d_n}(P_0)$ *when either:*

- m_n *and* θ_n *are consistent*
- ρ_n *and* Π_n *are consistent*
- m_n *and* ρ_n *are consistent*

Above D_d refers to $D_Y + D_A + D_W$, the portions of the efficient influence curve that are orthogonal to variation in decision rule d. This lemma is straightforward to prove. First, observe that
$\Psi_{d_n}(P_n^*) - \Psi_{d_n}(P_0) = -P_0\left[D_d^*(d_n, P_n^*)\right] + R_d(d_n, P_n^*)$.
By our TMLE procedure for Ψ^*, $P_n D_d(d_n, P_n^*) = 0$, so the rhs becomes
$(P_n - P_0)D_d^* + R_d$, for a second-order remainder term as given in section 5.5, which is $O_{P_0}(n^{-\frac{1}{2}})$ when either of the consistency conditions hold.

Discussion of conditions for theorem 5.

Condition C3. This is satisfied when both $R_d(d_n, P_n^*)$ and $R_c(d_n, \tau_0, P_n^*)$ are $o_{P_0}(n^{-\frac{1}{2}})$. As given in section 5.5, R_d is a double robust term that is second order when all components are estimated well. The proof that $R_c = o_{P_0}(n^{-\frac{1}{2}})$ is the same as in 4.6.

Condition C4-C6. Similar comments to those given in section 4.6 apply here.

Discussion of conditions for lemma 15.

To see that σ_n converges to σ_0, note that $D(d, \tau, P)(O_i)$ depends on the following components: $\{P_W, K_{T,P}, \rho_P, \Pi_P, m_P, \theta_P, d_P, \tau_P\}$. The following is sufficient for convergence of $D^*(d_n, \tau_n, P_n^*)^2$ to D_0^*:

- convergence of τ_n to τ_0 (proof is the same as in chapter 4)
- convergence of $P_{W,n}$ to $P_{W,0}$ and convergence of $K_{T,n}$ to $K_{T,0}$ (guaranteed by the fact that we use empirical distribution $P_{W,n}$ for $Pr_n(W)$)
- convergence of ρ_P, Π_P, m_P, and θ_P to ρ_0, Π_0, m_0, θ_0 respectively, in $L^2(P_0)$ norm (assumption E4)).

– $d_n(V) = d_0(V)$ with probability approaching 1. This is equivalent to $T_n(V) \geq \tau_n \iff T_0(V) \geq \tau_0$ w.p. approaching 1. The convergence of τ_n to τ_0, the uniform convergence of $T_n(V)$ to $T_0(V)$, and A4) guarantee this. The uniform convergence of $T_n(V)$ to $T_0(V)$ is proved in the proof that $\tau_n \to \tau_0$.

Testing for unmeasured confounding.

An instrumental variable-based estimator gives a consistent estimate of the treatment effect, even when unmeasured confounding exists. However, it is typically of much larger variance than a "direct" estimate of causal effect that is based on the assumption of no confounding and attempts to identify the counterfactual value of $Y(A)$ using $E(Y|A,W)$. If in fact there is no unmeasured confounding, the two approaches will converge to the same correct estimate of causal effect, and one can use the typically more precise direct estimate. On the other hand, if there is confounding, the two estimates do not converge, and we can use this difference of estimates as a statistic to test for confounding. In some applications, detecting unmeasured confounding might be a very useful end in itself. Such a test for unmeasured confounding is standard in the econometrics literature, by the name of Hausman-Wu test, although it is usually described using parametric models. We briefly describe a test in the semiparametric, TMLE framework.

When there is no unmeasured confounding, one can bypass the use of an instrumental variable to estimate the mean outcome under optimal dynamic treatment. In this case, the counterfactual outcome under treatment A is equal to the conditional mean given A: $EY(A=a) = E_W E(Y|W, A=a)$. Thus, we can write the parameter of interest Ψ as

$$\Psi = E_W EY(A = d(V)) =$$
$$E_W\big[(E(Y|A=1,W) - E(Y|A=0,W)) \cdot d(V) + E(Y|A=0,W)\big]$$

A TMLE-based estimator can be derived that is quite similar to that given in chapter 4, but replacing variable Z with A: one forms initial estimates of $E_{P_n}(Y|A,W)$ and $\Pr_n(A|W)$, estimates $d_n(V)$ from those, and performs a single-step fluctuation. This estimator has similar properties to the instrumental variable-based estimator presented here (double robust estimator of Ψ_{d_n}, efficient under consistent specification), but can estimate Ψ_0 with significantly lower variance in case there is no unmeasured confounding.

Let Ψ_{IV} denote the TMLE instrumental variable-based estimator described for causal effect Ψ_0, and Ψ_{NC} the TMLE estimator for Ψ_0 that assumes no confounding. Let $P^*_{n,IC}$, $P^*_{n,NC}$ denote the final estimates of the data generating distributions, in calculating $\Psi_{IV},\ \Psi_{NC}$, respectively. D_{IV} and D_{NC} denote the respective efficient influence curves. If we have that all relevant components and nuisance parameters for Ψ_{IV} are consistently specified, and $D_{IV}(P^*_{n,IV})$ belongs to a Donsker class, then the final estimate Ψ_{IV} is asymptotically linear, with asymptotic variance $Var(D_{IV}(P^*_{n,IC}))$. We have the analogous statement for Ψ_{NC}. Then $\Psi_{IV}-\Psi_{NC}$ is asymptotically linear with asymptotic variance $Var\big[D_{IV}(P^*_{n,IV})-D_{NC}(P^*_{n,NC})\big]$. This gives the following theorem.

Theorem 6 *(Test for unmeasured confounding.) Let $\Psi_{IV},\ \Psi_{NC}$ be as defined above. Assume $D_{IV}(P^*_{n,IV}),\ D_{NC}(P^*_{n,NC})$ belong to P_0-Donsker classes with probability approaching 1, and that all relevant components and nuisance parameters of $D_{IV},\ D_{NC}$ are consistently specified. Define a test statistic*
$\hat{T}=\sqrt{n}(\Psi_{IV}-\Psi_{NC})/\sqrt{Var_n(D_{IV}(P^*_{n,IV})-D_{NC}(P^*_{n,NC}))}$
Then under the null hypothesis of no unmeasured confounding, we have that $\hat{T}$ converges to a $N(0,1)$ random variable.

5.8 Proofs

Derivation for efficient influence curve $D^*(\Psi_0)$

We first derive the efficient influence curve $D^*_{d,\Pi}$, that assumes optimal decision rule d is fixed and Π is known. Next we extend the derivation to the case where Π is estimated non-parametrically and obtain D^*_d. Finally, we derive the correction term D^*_c that needs to be added to account for the need to estimate d.

$D^*_{d,\Pi}$ when $\Pi(W,Z)$, treatment rule $d(V)$ are known.

We derive the efficient influence curve for $\Psi_{\Pi,d}=\Psi_{\Pi,d}(P_{W,0},\mu_0)$ in model $\mathcal{M}$ given above, where we assume $\Pi=E(A|W,Z)$ is known and rule $d(V)$ is fixed. To ease notation, we do not always state the dependence on $\Pi(W,Z)$ and $d(V)$ in the notation from here on. For $\Psi(Q)=\Psi(P_W,\mu)$, we have relevant components P_W,μ, and the canonical gradient is the pathwise derivative of $\Psi(Q)$, in the tangent space of model $\mathcal{M}$. This is defined as $\frac{d}{d\epsilon}\Psi(Q(\epsilon))|_{\epsilon=0}$

along paths $\{P_\epsilon : \epsilon\} \subseteq \mathcal{M}$. The paths are represented here as

$$
\begin{aligned}
dP_{W,\epsilon} &= (1+\epsilon H_W(W))dP_W, \\
& \quad \text{where } EH_W(W) = 0 \text{ and } C_W \triangleq \sup_W |H_W(W)| < \infty; \\
d\mu_\epsilon(Z,W) &= (1+\epsilon H_Y(Y|Z,W))d\mu(Z,W), \\
& \quad \text{where } EH_Y(Y|Z,W) = 0 \text{ and } C_1 \triangleq \sup_{W,Z,Y} |H_Y(Y|Z,W)| < \infty;
\end{aligned}
$$

Let $\Psi_m \triangleq E_{P_W} m(W)d(V)$ and $\Psi_\theta \triangleq E_{P_W}\theta(W)$. By linearity of the operation of taking a pathwise derivative, we have that $D^*(\Psi) = D^*(\Psi_m) + D^*(\Psi_\theta)$, where $D^*(\Psi_m)$, $D^*(\Psi_\theta)$ are the efficient influence curves of the respective parameters in model $\mathcal{M}$.

We have

$$
\Psi_m = \int_{Y,A,W} Y \left(\frac{\frac{I(Z=1)}{\rho(W)} - \frac{I(Z=0)}{1-\rho(W)}}{\Pi(W,Z=1) - \Pi(W,Z=0)} \right) d(V)\ dP_{Y|Z,W} dP_{Z|W} dP_W
$$

Then proceeding with the delta method, we have that

$$
\begin{aligned}
&\Psi_m(P_\epsilon) - \Psi_m(P_0) = \\
&\int_{Y,A,W} Y \left(\frac{\frac{I(Z=1)}{\rho_0(W)} - \frac{I(Z=0)}{1-\rho_0(W)}}{\Pi(W,Z=1) - \Pi(W,Z=0)} \right) d(V)\ dP_{Y,\epsilon|Z,W} dP_{Z,0|W} dP_{W,\epsilon} \\
&- \int_{Y,A,W} Y \left(\frac{\frac{I(Z=1)}{\rho_0(W)} - \frac{I(Z=0)}{1-\rho_0(W)}}{\Pi(W,Z=1) - \Pi(W,Z=0)} \right) d(V)\ dP_{Y,0|Z,W} dP_{Z,0|W} dP_{W,0} \\
&= \epsilon \int_{Y,A,W} Y \left(\frac{\frac{I(Z=1)}{\rho_0(W)} - \frac{I(Z=0)}{1-\rho_0(W)}}{\Pi(W,Z=1) - \Pi(W,Z=0)} \right) d(V) H_Y(Y|Z,W)\ dP_{Y,0|Z,W} dP_{Z,0|W} dP_{W,0} \\
&+ \epsilon \int_{Y,A,W} Y \left(\frac{\frac{I(Z=1)}{\rho_0(W)} - \frac{I(Z=0)}{1-\rho_0(W)}}{\Pi(W,Z=1) - \Pi(W,Z=0)} \right) d(V) H_W(W)\ dP_{Y,0|Z,W} dP_{Z,0|W} dP_{W,0} + O(\epsilon^2)
\end{aligned}
$$

In the final line, the first term implies that $Y(\frac{\frac{I(Z=1)}{\rho_0(W)} - \frac{I(Z=0)}{1-\rho_0(W)}}{\Pi(W,Z=1)-\Pi(W,Z=0)})d(V)$ is the contribution to the gradient of Ψ_m from fluctuation $\mu_\epsilon(W,Z)$. The second term reduces to
$\epsilon \int_W (\mu_0(Z=1,W) - \mu_0(Z=0,W))/(\Pi(Z=1,W) - \Pi(Z=0,W))d(V)H_W(W)dP_{W,0}$
so it follows that $m_0(W)d(V)$ is the contribution to the gradient from fluctuation $P_{W,\epsilon}$. Finally, setting the mean to zero, we have the efficient influence

curve

$$\begin{aligned} D^*(\Psi_m) = \ & d(V)\Big(\frac{\frac{I(Z=1)}{\rho_0(W)} - \frac{I(Z=0)}{1-\rho_0(W)}}{\Pi(W,Z=1)-\Pi(W,Z=0)}\Big)(Y-\mu_0(W,Z)) \\ & + m_0(W)d(V) - E_W(m_0(W)d(V)) \end{aligned}$$

Using the identifiability of $\theta(W)$ as $E(Y|W,Z=1) - \Pi(W,Z=1)m(W)$, we have that

$$\Psi_\theta = \int_{Y,A,W} Y\left(\frac{I(Z=1)}{\rho(W)} - \frac{\big(\frac{I(Z=1)}{\rho(W)} - \frac{I(Z=0)}{1-\rho(W)}\big)\Pi(W,Z=1)}{\Pi(W,Z=1)-\Pi(W,Z=0)}\right) dP_{Y|Z,W}dP_{Z|W}dP_W$$

Proceeding as in the derivation for Ψ_m, we get

$$\begin{aligned} D^*(\Psi_\theta) = \ & \Big(\frac{I(Z=1)}{\rho_0(W)} - \frac{\big(\frac{I(Z=1)}{\rho_0(W)} - \frac{I(Z=0)}{1-\rho_0(W)}\big)\Pi(W,Z=1)}{\Pi(W,Z=1)-\Pi(W,Z=0)}\Big)(Y-\mu_0(W,Z)) \\ & + \theta_0(W) - E_W(\theta_0(W)) \end{aligned}$$

Combining, we get

$$\begin{aligned} D^*_{d,\Pi}(\Psi_0) = \ & J_0(W,Z)(Y-\mu_0(W,Z)) \\ & + m_0(W)d(V) + \theta_0(W) - \Psi_0 \end{aligned} \tag{5.16}$$

D^*_d when optimal treatment rule $d(V)$ is known, but Π is estimated non-parametrically.

Allowing non-parametric Π_ϵ, we assume

$$\begin{aligned} d\Pi_\epsilon(Z,W) \ & = (1+\epsilon H_A(A|Z,W))d\Pi(Z,W), \\ & \quad \text{where } EH_A(A|Z,W) = 0 \text{ and } C_2 \triangleq \sup_{W,Z,A}|H_A(A|Z,W)| < \infty; \end{aligned}$$

We need to adjust $D^*_{d,\Pi}$ above for the usual setting where $\Pi(W,Z)$ is not known but is estimated from the data through a NPMLE. Let $D^*_{d,\Pi}(\Pi')$ denote the efficient influence curve $D^*_{d,\Pi}(\Psi_0)$ in model $\mathcal{M}(\Pi_0)$ as given above, but with Π replaced with Π', so $D^*_{d,\Pi}(\Pi') = D^*_{d,\Pi}(P_W, \rho, \Pi', \mu)$. Let Π_n denote the non-parametric NPMLE estimate (ie empirical distribution) for $E(A|W,Z)$. The correction term $D_C(P_0)$ to add to $D^*_{d,\Pi}$ is the influence curve of $P_0(D^*_{d,\Pi_0}(\Pi_n) - D^*_{d,\Pi_0}(\Pi_0)$, which is the influence curve of $P_0 D^*_{d,\Pi_0}(\Pi_n)$.

Let $\Pi(\epsilon) = \Pi_0 + \epsilon\eta$. When using the NPMLE for Π_n, we have that $\eta = \frac{d}{d\epsilon}\Pi(P_0 + \epsilon(O - P_0))$. Plugging in the influence curve of a NPMLE, we have

$$\eta(z, w) = \frac{I(Z = z, W = w)}{\Pr(z, w)}(A - \Pi_0(Z, W))$$

Then we have

$$\begin{aligned}
&D_C(P_0) = \frac{d}{d\epsilon} P_0 D^*_{d,\Pi_0}(\Pi_\epsilon)|_{\epsilon=0} \\
&= P_0\left[J(w, z)(y - \mu_0(w, z))\right] \\
&+P_0\left[\Big(\frac{I(z = 1)}{\rho_0(w)} + \frac{\left(\frac{I(z=1)}{\rho_0(w)} - \frac{I(z=0)}{1-\rho_0(w)}\right)\left(d(v) - \Pi_0(w, 1)\right)}{\Pi_0(w, 1) - \Pi(w, 0)}\Big)(-m_0(w))\eta\right]
\end{aligned}$$

using shorthand $\Pi_0(w, 1) = \Pi_0(w, z = 1)$ and $J(w, z)$ as the appropriate function of w, z. The $P_0[J(w, z)(y - \mu_0(w, z))]$ term becomes 0. Plugging in for $\eta(z, w)$ and taking the expectation over $\Pr_0(w, z)$, we get

$$D_C(P_0) = \Bigg[\Big(\frac{I(Z = 1)}{\rho_0(W)} + \frac{\left(\frac{I(Z=1)}{\rho_0(W)} - \frac{I(Z=0)}{1-\rho_0(W)}\right)\left(d(V) - \Pi_0(W, 1)\right)}{\Pi_0(W, 1) - \Pi_0(W, 0)}\Big) \tag{5.17}$$

$$\times(-m_0(W))(A - \Pi_0(Z, W))\Bigg] \tag{5.18}$$

$$\tag{5.19}$$

Finally we get that the $-m(W)\Pi(Z, W)$ component in the canonical gradient D^*_d changes to $-m(W)A$.

$$\begin{aligned}
D^*_d(\Psi_0) &= m_0(W)d(V) + \theta_0(W) - \Psi_0 \\
&+J_0(W, Z)(Y - (m_0(W)A - \theta_0(W)))
\end{aligned} \tag{5.20}$$

Proof that D^*_c is the right correction term.

Above we derived the efficient influence curve $D^*_d(\Psi_0)$ when the optimal decision rule d was assumed known. To extend $D^*_d(\Psi_0)$ to unknown optimal rule $d_0(V)$, we first note that assumption A4) implies that the optimal rule is almost surely deterministic. Let $d_P(V)$ denote the optimal deterministic rule given by the closed form solution, given the distribution assumed, which depends on the $\{P_W, \underline{K}, m\}$ components of distribution P. Let $d_0(P)$ denote the true optimal dynamic treatment, corresponding to the closed form solution under the true data generating distribution P_0. Then we have that the correction term that needs to be added to D_d^* is $D_c^* = -\tau_0 E_{P_0}[c_T(d(V), W) - K]$.

Note that P_W induces a distribution over $V = F_V(W)$. We have that $\Psi_d = \int_W d(V)m(W) + \theta(W)dW$. Thus,

$$\begin{aligned}
\Psi(P_\epsilon) - \Psi(P_0) &= \int_W (E_U d_\epsilon(V) - d_0(V)) m_\epsilon dP_{W,\epsilon} \\
&+ \int_W d_0(V)(m_\epsilon dP_{W,\epsilon} - m_0 dP_{W,0}) \\
&+ E_{P_\epsilon}\theta_\epsilon(W) - E_{P_0}\theta_0(W) \\
&= \int_W (E_U d_\epsilon(V) - d_0(V))(m_\epsilon - \tau_0(c_{b,\epsilon}(V))\ )dP_{W,\epsilon} \qquad (5.21) \\
&+ \tau_0 \int_V (E_U d_\epsilon(V) - d_0(V))(c_{b,\epsilon}(V))dP_{W,\epsilon} \qquad (5.22) \\
&+ \Psi_{d_0}(P_\epsilon) - \Psi_{d_0}(P_0). \qquad (5.23)
\end{aligned}$$

From here on, we have essentially the same derivation as in 4.8 that D_c^* as given is the correction term that needs to be added to D^* to compensate for the variation resulting from estimating d. What needs to be changed is that m replaces μ_b and θ replaces $\mu(Z = 0, W)$ throughout. The one additional detail that needs to be added to that proof is to show that $m_\epsilon(W) - m_0(W) = \epsilon \cdot C_3$ for some constant C_3. This is needed for the proof that $S_\epsilon(\eta) - S_0(\eta) = O(\epsilon)$ for η in a δ_x-neighborhood of η_0 . We have $|m_\epsilon(W) - m_0(W)| = |\frac{\mu_\epsilon(W,1) - \mu_\epsilon(W,0)}{\Pi_\epsilon(W,1) - \Pi_\epsilon(W,0)} - \frac{\mu_0(W,1) - \mu_0(W,0)}{\Pi_0(W,1) - \Pi_0(W,0)}| =$
$|\frac{\mu_\epsilon(W,1) - \mu_\epsilon(W,0)}{\Pi_\epsilon(W,1) - \Pi_\epsilon(W,0)} - \frac{\mu_0(W,1) - \mu_0(W,0)}{\Pi_\epsilon(W,1) - \Pi_\epsilon(W,0)}(1 + \epsilon(H_A(W,1) - H_A(W,0))|$
$\leq |\frac{(\mu_0(W,1) - \mu_0(W,0))\epsilon(H_Y(W,1) - H_Y(W,0)) - (\mu_0(W,1) - \mu_0(W,0))\epsilon(H_A(W,1) - H_Y(W,0)) + O(\epsilon^2)}{\Pi_\epsilon(W,1) - \Pi_\epsilon(W,0)}| \leq C_3\epsilon$

for some constant C_3, noting that we can assume $\Pi(1, W) - \Pi(0, W)$ is bounded away from 0, because Z is a valid instrument that induces variation in A, as guaranteed by our causal model. We note that we know H_A, H_Y, and μ are all bounded.

Derivation for remainder term

We give a sketch of how lemma 14 is derived.

Note that $\Psi_d(P) = P(md + \theta)$ and $\Psi_d P_0 = P_0(m_0 d + \theta_0)$. For components

$D^*_Y(d,P)$, $D^*_A(d,P)$ we have:

$$P_0\big[D^*_Y(d,P)+D^*_A(d,P)\big] = E_0\big[(E_0(Y|W,A,Z)-m(W)A-\theta(W))J(Z,W)\big]$$
$$= P_0\big[((m_0-m)\Pi_0+(\theta_0-\theta))K\big]$$

$$= P_0\left[\frac{\rho_0}{\rho}\Big((m_0-m)\Pi_0(W,1)+(\theta_0-\theta)\Big)\Big(1+\frac{d-\Pi(W,1)}{\Pi(W,1)-\Pi(W,0)}\Big)\right]$$
$$+\left[\frac{1-\rho_0}{1-\rho}\Big((m_0-m)\Pi_0(W,0)+(\theta_0-\theta)\Big)\Big(\frac{-(d-\Pi(W,1))}{\Pi(W,1)-\Pi(W,0)}\Big)\right]$$

$$= P_0\Big[\Big((m_0-m)\Pi_0(W,1)+(\theta_0-\theta)\Big)\Big(1+\frac{d-\Pi(W,1)}{\Pi(W,1)-\Pi(W,0)}\Big)\Big]$$
$$+\Big[\Big((m_0-m)\Pi_0(W,0)+(\theta_0-\theta)\Big)\Big(\frac{-(d-\Pi(W,1))}{\Pi(W,1)-\Pi(W,0)}\Big)\Big] + \text{2nd order terms}$$

$$= P_0\Big((m_0-m)(\Pi_0(W,1)-\Pi_0(W,0))\Big)\Big(\frac{d-\Pi(W,1)}{\Pi(W,1)-\Pi(W,0)}\Big)$$
$$+P_0\big((m_0-m)\Pi_0(W,1)+(\theta_0-\theta)\big) + \text{2nd order terms}$$

$$= P_0\big((m_0-m)(d-\Pi(W,1))\big) + P_0\big((m_0-m)\Pi_0(W,1)+(\theta_0-\theta)\big) + \text{2nd order terms}$$
$$= P_0\big((m_0-m)d+(\theta_0-\theta)\big) + \text{2nd order terms}$$

where the fourth equality is from expanding $\frac{\rho_0}{\rho}$ into $1-\frac{\rho_0-\rho}{\rho}$, likewise $\frac{1-\rho_0}{1-\rho}$. The final two equalities follow from expanding Π_0 into $\Pi+(\Pi_0-\Pi)$.

We also have that $P_0D^*_W(d,P) = P_0(md+\theta)-P(md+\theta)$.

Combining, we get

$$R_d(P,P_0) =$$
$$\Psi_d(P)-\Psi_d(P_0)+P_0(D^*_d)(P)$$
$$= P_0\big((m-m_0)d+(\theta-\theta_0)+(m_0-m)d+(\theta_0-\theta)\big) + \text{2nd order terms}$$
$$= \text{2nd order terms}$$

For the sake of brevity, we didn't write out the second-order terms accumulating in the derivations above, but it is straightforward to do so and obtain $R_d(d,P)$ as given in 5.10.

To get $R_c(d,P)$, first note that line 5.11 reduces to $E_{P_0}[(d-d_0)m_0 - \tau_0[c_T(d(V),W)-c_T(d_0(V),W)]]$. Note that $E_{P_0}(c_T(d_0(V),W)) = K$ (this is justified in section 4.8). Since $c_T(d(V),W) = d(V)c_{b,0}(V)+c_T(Z=0,W)$, we get that line 5.11 equals $P_0[(d-d_0)(m_0(V)-\tau_0c_{b,0}(V))]$.

Chapter 6

Tables

Table 6.1: Performance of estimators in estimating a **scalar causal effect, nonlinear design 1**. The initial estimator for $E(Y|Z,W)$ is either consistently specified or misspecified, and all other nuisance parameters are consistenly specified. Sample size is 1000, and 10,000 repetitions were made. The true effect is 33.23.

CONSISTENTLY SPECIFIED

	Estimator	**Bias**	**Var**	**MSE**
New methods	Iterative	.0016	.6103	.6103
	Linear fluctuation	.0015	.6189	.6189
	Logistic fluctuation	.0015	.6189	.6189
Non-parametric	Estimating equations	$-.0016$	.7834	.7834
	Initial substitution estimator	.0038	.6990	.6990
	Confounded	20.97	0.000	439.7
	Two-stage least squares	$-.3904$	52.74	52.89

$E(Y|W,Z)$ IS MISSPECIFIED

	Estimator	**Bias**	**Var**	**MSE**
New methods	Iterative	.3157	117.7	117.8
	Linear fluctuation	.6214	78.27	78.65
	Logistic fluctuation	.8193	82.99	83.66
Non-parametric	Estimating equations	$-.2088$	35.14	35.18
	Initial substitution estimator	$-.3941$	54.07	54.22
	Confounded	20.97	0.000	439.7
	Two-stage least squares	$-.3904$	52.74	52.89

Table 6.2: Performance of estimators in estimating **vector-valued** causal effect, when the **treatment effect is nonlinear (design 1)**. Causal parameter β to estimate is projection of effect unto linear function of covariates $\{m_\beta(W) = \beta^T W|\beta\}$. The true effect is $[-64.2, 32.3]$.

	MEAN ABSOLUTE BIAS OF ESTIMATORS					
	Consistent specification			$E(Y\|Z,W)$ is misspecified		
	n=1000	n=3000	n=10000	n=1000	n=3000	n=10000
Iterative	.0683	.0219	.0191	1.350	.6043	.2552
Linear fluctuation	.3025	.0247	.0056	8.773	2.589	.6587
Estimating equations	.0128	.0084	.0119	1.521	1.013	.4110
Initial substitution estimator	.6478	.0595	.0473	136.0	136.3	136.1
Two-stage least squares	136.6	136.4	136.6	136.6	136.4	136.6
Confounded	10.93	10.15	10.72	10.93	10.15	10.72

	MEAN ABSOLUTE STD DEV OF ESTIMATORS					
	Consistent specification			$E(Y\|Z,W)$ is misspecified		
	n=1000	n=3000	n=10000	n=1000	n=3000	n=10000
Iterative	11.34	3.782	1.860	34.59	12.00	4.851
Linear fluctuation	8.192	3.016	1.494	86.78	17.34	5.212
Estimating equations	5.029	2.741	1.492	19.03	10.15	5.954
Initial substitution estimator	4.861	2.709	1.565	11.37	6.743	3.789
Two-stage least squares	11.12	6.235	3.694	11.12	6.235	3.694
Confounded	.0021	.0009	.0005	.0021	.0009	.0005

Table 6.3: Performance of estimators in estimating a **scalar causal effect, nonlinear design 2**, where $E(Y|W,Z)$ follows **sharp cutoffs**. The initial estimator for $E(Y|Z,W)$ is either consistently specified or misspecified, and all other nuisance parameters are consistenly specified. Sample size is 1000, and 10,000 repetitions were made. The true effect is 1.00.

CONSISTENTLY SPECIFIED

	Estimator	**Bias**	**Var**	**MSE**
New methods	Iterative	$-.0853$	.2226	.2299
	Linear fluctuation	$-.0827$	.2198	.2266
	Logistic fluctuation	.0307	.1645	.1654
Non-parametric	Estimating equations	$-.0643$	.1508	.1549
	Initial substitution estimator	.0202	.1196	.1200
	Confounded	.5735	.0170	.3459
	Two-stage least squares	.0926	.2792	.2878

$E(Y|W,Z)$ IS MISSPECIFIED

	Estimator	**Bias**	**Var**	**MSE**
New methods	Iterative	$-.0703$	.4498	.4547
	Linear fluctuation	$-.0414$	.4561	.4578
	Logistic fluctuation	.0487	.3396	.3420
Non-parametric	Estimating equations	$-.0636$	.4492	.4532
	Initial substitution estimator	.0865	.3870	.3945
	Confounded	.5735	.0170	.3459
	Two-stage least squares	.0926	.2792	.2878

Table 6.4: Performance of estimators in estimating **vector-valued** causal effect, when the **treatment effect is linear**. The true effect is $[0, 1, 2, 3]$. Causal parameter β to estimate is coefficient α, where the treatment effect is $m(W) = \alpha^T W$.

	MEAN ABSOLUTE BIAS OF ESTIMATORS					
	Consistent specification			$E(Y\|Z, W)$ is misspecified		
	n=1000	n=3000	n=10000	n=1000	n=3000	n=10000
Iterative	.0705	.0417	.0071	.1540	.1302	.0980
Linear fluctuation	.0049	.0034	.0015	.1101	.1134	.0861
Estimating equations	.0062	.0027	.0020	.4194	.3803	.2374
Initial substitution estimator	.0090	.0117	.0029	1.546	1.499	1.503
Two-stage least squares	.2446	.2324	.2443	.2446	.2324	.2443
Confounded	.3484	.3432	.3430	.3484	.3432	.3430

	MEAN ABSOLUTE STD DEV OF ESTIMATORS					
	Consistent specification			$E(Y\|Z, W)$ is misspecified		
	n=1000	n=3000	n=10000	n=1000	n=3000	n=10000
Iterative	1.044	.5967	.1927	1.038	.6305	.3421
Linear fluctuation	.5746	.3067	.1799	.5528	.3410	.2268
Estimating equations	.6356	.3944	.1618	.5371	.3549	.2989
Initial substitution estimator	.5413	.3140	.1713	.4296	.2514	.1345
Two-stage least squares	.5104	.2906	.1580	.5104	.2906	.1580
Confounded	.1188	.0657	.0359	.1188	.0657	.0359

Table 6.5: Mean coverage of **95% confidence intervals**. The coverage is calculated for each dimension of the parameter of interest and the average taken. For the top three estimators in each table, the empirical variance of the efficient influence curve $\mathrm{Var}(D^*(Q_n^*, g_n^*))$ is used to calculate the standard error. For the other estimators, we give the unfair advantage of using the accurate variance in calculating the confidence intervals.

LINEAR TREATMENT EFFECT	Consistent specification			$E(Y\|Z,W)$ is misspecified		
	n=1000	n=3000	n=10000	n=1000	n=3000	n=10000
Iterative	96.8	96.4	96.2	96.1	96.0	95.2
Linear fluctuation	96.6	96.1	96.3	95.9	95.1	94.7
Estimating equations	96.4	96.0	96.3	89.4	88.8	90.3
Initial substitution estimator	94.6	94.8	94.8	5.98	0	0
Two-stage least squares	92.6	87.5	67.7	92.6	87.5	67.7
Confounded	19.8	0.22	0	19.8	0.22	0

NONLINEAR TREATMENT EFFECT	Consistent specification			$E(Y\|Z,W)$ is misspecified		
	n=1000	n=3000	n=10000	n=1000	n=3000	n=10000
Iterative	97.5	97.1	96.7	97.3	97.2	96.9
Linear fluctuation	97.2	96.5	96.8	96.9	95.9	96.6
Estimating equations	96.4	95.7	96.2	96.8	96.3	97.0
Initial substitution estimator	94.2	94.6	94.3	0	0	0
Two-stage least squares	0	0	0	0	0	0
Confounded	0	0	0	0	0	0

Table 6.6: Estimates of effect of mother's education on infant's health, where the latter is given by the log-odds ratio of a binary outcome. Significant effects are marked with (∗) and highly significant effects with (∗∗).

	Mother's education		
	Low birthweight	Neonatal mortality	Postneonatal mortality
TMLE, linear fluctuation			
Mean effect	-.266 (.163)*	-1.04 (.193)**	-.358 (.255)
First stage CV-R^2	.814	.805	.798
Second stage CV-R^2	.539	.561	.453
Change in outcome	-.041%	-.387	-.222
Percent change in outcome	-7.07%	-16.7%	-6.54%
Two-stage least squares			
Mean effect	-.212 (.175)	-.265 (.124)*	-.529 (.229)*
First stage CV-R^2	.778	.747	.760
Second stage CV-R^2	.426	.531	.376
Change in outcome	-.027%	-.078	-.255
Percent change in outcome	-4.58%	-3.42%	-7.14%
F-test for weak IV	15.5	12.1	13.7
Hausman-Wu test	1.64	.738	.684
OLS			
Mean effect	-.177 (.009)	-.381 (.011)	-.434 (.010)
R^2	.428	.538	.424

Table 6.7: Estimates of effect of father's education on infant's health, where the latter is given by the log-odds ratio of a binary outcome. Significant effects are marked with (∗) and highly significant effects with (∗∗).

	Father's education		
	Low birthweight	Neonatal mortality	Postneonatal mortality
TMLE, linear fluctuation			
Mean effect	-.126 (.179)	-.632 (.105)**	-.569 (.242)**
First stage CV-R^2	.801	.784	.751
Second stage CV-R^2	.616	.622	.626
Change in outcome	-.001	-.342	-.166
Percent change in outcome	-1.54%	-5.99%	-4.75%
Two-stage least squares			
Mean effect	-.298 (.160)*	-.480 (.182)**	-.602 (.254)**
First stage CV-R^2	.749	.752	.716
Second stage CV-R^2	.318	.378	.395
Change in outcome	-.001	-.161	-.178
Percent change in outcome	-2.03%	-2.70%	-5.29%
F-test for weak IV	11.1	9.22	9.69
Hausman-Wu test	.892	.186	.432
OLS			
Mean effect	-.223 (.006)	-.345 (.010)	-.419 (.011)
R^2	.345	.400	.415

Table 6.8: Additive treatment effect modifiers for the effect of **mother's** education on infant's health. Significant effects are marked with (∗) and highly significant effects with (∗∗). The standardized effect modifiers refer to (treatment effect modifier ÷ SE).

	Low birthweight	Postneonatal mortality
Strongest additive treatment effect modifiers	$YOB$1982: -1.24 (.472)** $YOB$1985: -.608 (.162)** $YOB$1988: -.506 (.167)** $YOB$1989: -.497 (.168)** $YOB$1987: -.494 (.166)** $YOB$1998: +.494 (.188)**	$YOB$1980 +1.22 (.430)** $YOB$1982 -.916 (.317)** $YOB$1989 -.886 (.227)** $YOB$1998 +.795 (.233)** $YOB$1999 +.667 (.342) $YOB$1987 -.664 (.202)**
Summary statistics	Q1 Q2 Mean Q3 -.237 -.072 -.148 -.010	Q1 Q2 Mean Q3 -.358 -.104 -.111 .047
Summary statisics, standardized effect modifiers	Q1 Q2 Mean Q3 -1.78 -.511 -.833 -.056	Q1 Q2 Mean Q3 -1.75 -.721 -.628 .372

Table 6.9: Additive treatment effect modifiers for the effect of **father's** education on infant's health. Significant effects are marked with ($*$) and highly significant effects with ($**$). The standardized effect modifiers refer to (treatment effect modifier $\div$ SE).

	Low birthweight	Postneonatal mortality
Strongest additive treatment effect modifiers	$YOB$1980: -3.98 (8.93) $YOB$1979: +3.54 (1.73)* $YOB$1978: -2.39 (.962)* $YOB$1981: -1.43 (.437)** $YOB$1985: -1.13 (.239)** $YOB$1982: -.880 (.385)*	$YOB$1979 -5.103 (.297)** $YOB$1978 -3.27 (1.14)** $YOB$1981 -1.34 (.241)** $COH$1968: -1.29 (.427)** $COH$1967: -1.26 (.429)** $YOB$1983: -1.11 (.778)
Summary statistics	Q1 Q2 Mean Q3 -.792 -.351 -.504 -.307	Q1 Q2 Mean Q3 -.931 -.419 -.636 -.055
Summary statisics standardized effect modifiers	Q1 Q2 Mean Q3 -3.16 -.634 -1.59 -.276	Q1 Q2 Mean Q3 -2.17 -1.39 -1.19 -.198

Table 6.10: (Simulation A.) Consistent estimation using machine learning, categorical Y. $\Psi_0 = 0.3456$.

	N=250				
Estimator	Ψ^*	**Bias**	**Var**	σ_n^2	**Cover.**
TMLE	.3545	.0089	.0071	.0010	88.3
CV-TMLE	.3541	.0085	.0017	.0010	90.6
Init. Substit.	.3427	-.0029	.0067	.0010	(87.9)

	N=1000				
Estimator	Ψ^*	**Bias**	**Var**	σ_n^2	**Cover.**
TMLE	.3485	.0029	.0003	.0003	93.3
CV-TMLE	.3497	.0041	.0002	.0003	96.8
Init. Substit.	.3344	-.0112	.0003	.0003	(88.3)

	N=4000				
Estimator	Ψ^*	**Bias**	**Var**	σ_n^2	**Cover.**
TMLE	.3467	.0011	.0001	.0001	95.0
CV-TMLE	.3498	.0002	.0001	.0001	94.7
Init. Substit.	.3429	-.0027	.0001	.0001	(93.3)

Table 6.11: (Simulation A.) Consistent estimation using machine learning, continuous Y. $\Psi_0 = 288.8$.

	N=250				
Estimator	Ψ^*	**Bias**	**Var**	σ_n^2	**Cover.**
TMLE	277.8	-11.05	221.0	247.5	81.7
CV-TMLE	283.8	-5.03	506.3	265.1	83.2
Init. Substit.	260.5	-28.34	548.9	303.6	(64.3)

	N=1000				
Estimator	Ψ^*	**Bias**	**Var**	σ_n^2	**Cover.**
TMLE	285.1	-3.72	39.20	37.91	86.8
CV-TMLE	285.9	-2.86	36.89	35.65	88.9
Init. Substit.	275.6	-13.29	120.75	41.141	(50.3)

	N=4000				
Estimator	Ψ^*	**Bias**	**Var**	σ_n^2	**Cover.**
TMLE	286.8	-2.16	7.68	8.18	91.7
CV-TMLE	287.5	-1.26	8.51	9.19	93.1
Init. Substit.	281.4	-7.44	24.69	8.36	(38.3)

Table 6.12: (Simulation B.) Varying parameters of the optimization problem. $\Psi_0 = 162.6$ when $V = W$, $K = .2$; $\Psi_0 = 141.3$ when $V \subset W$, $K = .2$; $\Psi_0 = 288.9$ when $V = W$, $K = .8$; and $\Psi_0 = 201.9$ when $V \subset W$, $K = .8$. Sample size is $n = 1000$.

CONTINUOUS Y

	Estimator	Ψ^*	**Bias**	**Var**
$V = W,\ K = .2$	TMLE	161.0	-1.64	25.17
	CV-TMLE	162.2	-.48	33.28
	Init. Substit.	153.2	-9.48	20.99
$V \subset W,\ K = .2$	TMLE	139.5	-1.79	4.21
	CV-TMLE	138.6	-2.74	7.64
	Init. Substit.	135.6	-5.65	9.38
$V = W,\ K = .8$	TMLE	287.6	-1.42	14.34
	CV-TMLE	288.9	-.06	19.5
	Init. Substit.	278.5	-10.52	17.78
$V \subset W,\ K = .8$	TMLE	200.2	-1.86	43.2
	CV-TMLE	200.9	-1.07	40.7
	Init. Substit.	192.6	-9.34	72.2

Table 6.13: (Simulation C.) Robustness to partial misspecification, μ is misspecified. $\Psi_0 = 163.0$.

	N=250				
Estimator	Ψ^*	$\left(\Psi^* - \Psi_{d_n}(P_0)\right)$	$\left(\Psi^* - \Psi_0\right)$	**Var**	**Cover.**
TMLE	137.8	-10.71	-25.19	555.4	74.2
Init. Substit.	123.4	-25.09	-39.57	153.5	(84.8)

	N=1000				
Estimator	Ψ^*	$\left(\Psi^* - \Psi_{d_n}(P_0)\right)$	$\left(\Psi^* - \Psi_0\right)$	**Var**	**Cover.**
TMLE	144.5	-0.19	-18.54	145.0	92.5
Init. Substit.	124.3	-21.40	-38.75	52.86	(70.7)

	N=4000				
Estimator	Ψ^*	$\left(\Psi^* - \Psi_{d_n}(P_0)\right)$	$\left(\Psi^* - \Psi_0\right)$	**Var**	**Cover.**
TMLE	143.9	-0.03	-19.13	24.08	96.1
Init. Substit.	123.8	-20.01	-39.17	22.48	(40.6)

Table 6.14: (Simulation D.) True mean outcome $\Psi_{d_n}(P_0)$, under rule d_n. $\Psi_0 = 162.8$ when $K = .2$, and $\Psi_0 = 289.1$ when $K = .8$. Sample size is $n = 1000$.

	K=.2		K=.8	
Learning μ_n	$\Psi_{d_n}(P_0)$	**Var**	$\Psi_{d_n}(P_0)$	**Var**
Large library	158.9	8.14	286.4	9.32
Small library	148.3	49.45	267.9	16.28
No fitting	142.2	12.83	264.1	10.30

Bibliography

[1] A Abadie. "Semiparametric instrumental variable estimation of treatment response models". In: *Journal of Econometrics, Vol. 113, p.213-263* (2003).

[2] J.D. Angrist, G.W. Imbens, and D.B. Rubin. "Identification of causal effects using instrumental variables". In: *Journal of the American Statistical Association, Vol. 91, p. 444-471* (1996).

[3] Joshua Angrist. "Treatment effect heterogeneity in theory and practice". In: *The Economic Journal, Vol. 114, p. C52-C83* (2004).

[4] Joshua Angrist and Alan Krueger. "Does compulsory school attendance affect schooling and earnings". In: *The Quarterly Journal of Economics, Vol. 106, No. 4, p. 979-1014* (1991).

[5] Kevin Arceneaux, Alan Gerber, and Donald Green. "Comparing experimental and matching methods using a large-scale voter mobilization experiment". In: *Political Analysis, Vol. 14, p. 37-62* (2006).

[6] Mike Baiocchi et al. "Near/far matching: a study design approach to instrumental variables". In: *Health Services and Outcomes Research Methodology, Vol. 12, No. 4, p. 237-253* (2012).

[7] Alexander Balke and Judea Pearl. "Bounds on treatment effects from studies with imperfect compliance". In: *Journal of the American Statistical Association, Vol. 92, No. 439, p. 1172-1176* (1997).

[8] Peter J. Bickel et al. *Efficient and Adaptive Estimation for Semiparametric Models*. Johns Hopkins Studies in the Mathematical Sciences, 1st Edition, 1993.

[9] M. Alan Brookhart, Jeremy A. Rassen, and Sebastian Schneeweiss. "Instrumental variable methods in comparative safety and effectiveness research". In: *Pharmacoepidemiology and Drug Safety, Vol. 19, Issue 6, p. 537-554* (2010).

[10] M. Alan Brookhart and Sebastian Schneeweiss. "Preference-based instrumental variable methods for the estimation of treatment effects". In: *International Journal of Biostatistics: Vol.3(1), p.1-14* (2007).

[11] Bibhas Chakraborty, Eric Laber, and Ying-Qi Zhao. "Inference for optimal dynamic treatment regimes using an adaptive m-out-of-n bootstrap scheme". In: *Biometrics, Vol. 69, No. 3, p. 714-723* (2013).

[12] Bibhas Chakraborty and E E Moodie. *Statistical Methods for Dynamic Treatment Regimes*. Springer, Berlin Heidelberg New York, 2013.

[13] J. Cheng et al. "Efficient nonparametric estimation of causal effects in randomized trials with noncompliance". In: *Biometrika: No. 96, p.1-9* (2009).

[14] Shin-Yi Chou et al. "Parental education and child health: evidence from a natural experiment in Taiwan". In: *American Economic Journal: Applied Economics, Volume 2, Issue 1, p. 33-61* (2010).

[15] Paul Clarke and Frank Windmeijer. "Identification of causal effects on binary outcomes using structural mean models". In: *Biostatistics* (2010).

[16] Paul Clarke and Frank Windmeijer. "Instrumental variables estimators for binary outcomes". In: *Center for Market and Public Organization Working Paper Series, No. 10/239* (2010).

[17] Graham Dunn and Richard Bentall. "Modelling treatment effect heterogeneity in randomized controlled trials of complex interventions (psychological treatments)". In: *Statistics in Medicine, Vol. 26, p. 4719-4745* (2007).

[18] Susan Gruber and Mark J. van der Laan. "A targeted maximum likelihood estimator of a causal effect on a bounded continuous outcome". In: *International Journal of Biostatistics: Vol. 6, Issue 1, Article 26* (2010).

[19] Miguel Hernan and James Robins. "Instruments for causal inference: an epidemiologist's dream". In: *Epidemiology, Vol. 17, No. 4* (2006).

[20] Han Hong and Denis Nekipelov. "Semiparametric efficiency in nonlinear LATE models". In: *Quantitative Economics, Vol. 1, Issue 2* (2010).

[21] Kosuke Imai and Aaron Strauss. "Estimation of heterogeneous treatment effects from randomized experiments, with application to the optimal planning of the get-out-the-vote campaign". In: *Political Analysis, Vol. 19, p. 1-19* (2011).

[22] G.W. Imbens and J.D. Angrist. "Identification and estimation of local average treatment effects". In: *Econometrica, Vol. 62, p. 467-475* (1994).

[23] Richard Karp. *Reducibility Among Combinatorial Problems.* Springer, New York Berlin Heidelberg, 1972.

[24] Maximilian Kasy. "Semiparametrically efficient estimation of conditional instrumental variable parameters". In: *International Journal of Biostatistics, Vol. 5, Issue 1, Article 2* (2009).

[25] Mark J. van der Laan, Eric C. Polley, and Alan E. Hubbard. "Super Learner". In: *Statistical Applications in Genetics and Molecular Biology: Vol. 6, Issue 1, Article 25* (2007).

[26] Mark J. van der Laan and James M Robins. *Unified Methods for Censored Longitudinal Data and Causality. Springer Verlag: New York.* 2003.

[27] Mark J. van der Laan and Daniel Rubin. "Targeted maximum likelihood learning". In: *International Journal of Biostatistics: Vol. 2, Issue 1, Article 11* (2006).

[28] MJ van der Laan, AE Hubbard, and NP Jewell. "Estimation of treatment effects in randomized trials with non-compliance and a dichotomous outcome." In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology), Vol. 69.* (2007).

[29] MJ van der Laan and Sherri Rose. *Targeted Learning: Causal Inference for Observational and Experimental Data. New York, Springer.* 2011.

[30] Alexander Luedtke and Mark van der Laan. "Optimal individualized treatments in resource-limited settings". In: *International Journal of Biostatistics, Vol. 12, No. 1, p. 283-303* (2016a).

[31] Alexander Luedtke and Mark van der Laan. "Statistical inference for the mean outcome under a possibly non-unique optimal treatment strategy". In: *Annals of Statistics, Vol. 44, No. 2, p. 713-742* (2016b).

[32] Whitney Newey. "Semiparametric efficiency bounds". In: *Journal of the Applied Econometrics, Vol. 5, No. 2, p. 99-135* (1990).

[33] Joseph P. Newhouse and Mark McClellan. "Econometrics in outcomes research: the use of instrumental variables". In: *Annual Review of Public Health, Vol. 19, p. 17-34* (1998).

[34] El Ogburn, A Rotnizky, and JM Robins. "Doubly robust estimation of the local average treatment effect curve". In: *Journal of the Royal Statistical Society, Series B, Vol. 77, No. 2, p. 373-396* (2014).

[35] Judea Pearl. *Causality: Models, Reasoning, and Inference.* Cambridge University Press, 2000.

[36] Maya Petersen et al. "History-adjusted marginal structural models for estimating time-varying effect modification". In: *American Journal of Epidemiology, Vol. 166, No. 9* (2007).

[37] JM Robins. "Correcting for non-compliance in randomized trials using structural nested mean models." In: *Communications in Statistics: Vol. 23, p.2379-2412* (1994).

[38] JM Robins. "Marginal structural models and causal inference in epidemiology." In: *Epidemiology, Vol. 11, p.550-560* (2000).

[39] JM Robins. "Optimal structural nested models for optimal sequential decisions." In: *Proc. Second Seattle Symp. Biostat., Vol. 179, p. 189-326* (2004).

[40] D Rubin. "Estimating causal effects of treatments in randomized and nonrandomized studies." In: *Journal of Educational Psychology: Vol. 66, No. 6, p.689* (1974).

[41] SE Sinisi and MJ van der Laan. "Loss-based cross-validation deletion/substitution/addition algorithms in estimation." In: *Technical report, UC Berkeley Division of Biostatistics Working Paper Series No. 143* (2004).

[42] Douglas Staiger and James Stock. "Instrumental variables regression with weak instruments". In: *Econometrica, Vol. 65, No. 3, p. 557-586* (1997).

[43] J Stock, J Wright, and M Yogo. "A survey of weak instruments and weak identification in generalized method of moments." In: *Journal of the American Statistical Association, Vol. 20, Issue 4, p. 518-529* (2002).

[44] Z Tan. "Marginal and nested structural models using instrumental variables". In: *Journal of the American Statistical Association, Vol. 105, p.157-169* (2010).

[45] Z Tan. "Regression and weighting methods for causal inference using instrumental variables". In: *Journal of the American Statistical Association, Vol. 101, p.1607-1618* (2006).

[46] Joseph Terza, David Bradford, and Clara E Dismuke. "The use of linear instrumental variables methods in health services research and health economics: a cautionary note". In: *Health Services Research: Vol. 43, Issue 3, p.1102-1120* (2008).

[47] Boriska Toth and Mark van der Laan. "Efficient targeted estimation using instrumental variables". In: *Proceedings of the Joint Statistical Meetings (JSM)* (2011).

[48] Boriska Toth and Mark van der Laan. "TMLE for marginal structural models based on an instrument". In: *UC Berkeley Division of Biostatistics Working Papers Series* (2016).

[49] Boriska Toth and Mark van der Laan. "Towards optimal semiparametric inference using instrumental variables". In: *Proceedings of the Joint Statistical Meetings (JSM)* (2012).

[50] SD Uysal. "Doubly robust IV estimation of the local average treatment effects." In: *(Available from http://www.ihs.ac.at/vienna/resources/Economics/Papers/Uysal_paper.pdf .)* (2011).

[51] AW van der Vaart. *Asymptotic Statistics*. Cambridge University Press, 2000.

[52] B Zhang et al. "A robust method for estimating optimal treatment regimes". In: *Biometrics, Vol. 68, p. 1010-1018* (2012).

[53] Wenjing Zheng and Mark van der Laan. *Cross-validated targeted minimum-loss-based estimation*. In book: Targeted Learning: Causal Inference for Observational and Experimental Data, 2011.