Optimal Structural Nested Models for Optimal Sequential Decisions

James M. Robins

Departments of Epidemiology and Biostatistics, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115, email: robins@hsph.harvard.edu

ABSTRACT: I describe two new methods for estimating the optimal treatment regime (equivalently, protocol, plan or strategy) from very high dimesional observational and experimental data: (i) g-estimation of an *optimal* double-regime structural nested mean model (drSNMM) and (ii) g-estimation of a standard single regime SNMM combined with sequential dynamic-programming (DP) regression. These methods are compared to certain regression methods found in the sequential decision and reinforcement learning literatures and to the regret modelling methods of Murphy (2003). I consider both Bayesian and frequentist inference. In particular, I propose a novel "Bayes-frequentist compromise" that combines honest subjective non- or semiparametric Bayesian inference with good frequentist behavior, even in cases where the model is so large and the likelihood function so complex that standard (uncompromised) Bayes procedures have poor frequentist performance.

1 Introduction

The goal of this paper is to describe several methods for estimating the optimal treatment regime (equivalently, protocol, plan or strategy) from observational (i.e. nonexperimental) and randomized studies when a data on high dimensional time-dependent response (i.e. covariate) processes are available. The first method is based on doubly robust locally semiparametric efficient (dr-lse) g-estimation of the so-called blip (i.e. treatment effect) function of an *optimal* double-regime structural nested mean model (drSNMM). The second method is based on a dynamic-programming (DP) -like regression model applied to g-estimates of the blip (treatment effect) function of a standard single regime SNMM. I shall refer to the models required by this method as DP-regression SNMMs.

I introduced *standard* single-regime structural nested mean model (srSNMMs) and proposed dr-lse g-estimation of their blip function in Robins

(1994,1997,2000). *Standard* srSNMMs model the effect of a final blip of active treatment (versus zero treatment) at time m before following the zero treatment regime from time $m+1$ to end of follow-up. Here "zero" treatment refers to a substantively meaningful baseline level of treatment such as "no treatment". Double-regime SNMMs and optimal drSNMMs are introduced here. A drSNMM models the effect of a final blip of active treatment (versus zero treatment) at time m before following a given prespecified regime from *time* $m+1$ to end of follow-up. An optimal drSNMM is a drSNMM in which the regime followed from time $m+1$ onwards is the optimal treatment regime. In Sections 3 and 4 I show drSNMMs and optimal drSNMMs are but a minor technical generalization of the standard srSNMMs of Robins (1994).

My methods were motivated by but differ from but those proposed by Susan Murphy in her seminal paper on this topic. Section 6 compares and contrasts Murphy's methods with mine. I show that Murphy's semiparametric regret model is a particular *nonstandard* single-regime SNMM, not only in terms of the observed data likelihood, but also as a counterfactual model. A *nonstandard* srSNMM differs from a standard srSNMM only in that the data analyst's definition of the "zero" level of treatment at any time m may vary both with time and with past treatment and covariate history, and thus may have little or no consistent substantive meaning. From a mathematical and statistical point of view, standard and non-standard srSNMMs are identical; they only differ in the substantive meaning of the treatment effect encoded in their blip functions. It follows, as conjectured by Murphy, that my prior work on SNMMs indicates how to extend her results to include (i) sensitivity analysis and instrumental variable methodologies that allow for unmeasured confounders (Robins, Rotnitzky, Sharfstein, 1999a, Sec 8.1b,8.2b; Robins, Greenland, Hu 1999d, Sec.2d.5; Section 7 below) (ii) continuous time treatments (Robins,1998ab; Section 8 below), (iii) locally semiparametric efficient doubly- robust (lse-dr) estimation (Robins,1994, 2000a; Sections 3-4 below), and (iv) an asymptotic distribution-free test of the g-null hypothesis that the mean response is the same for all regimes (Robins, 1994,1997; Sections 3-4 below) [provided the treatment probabilities are known (as in a randomized study) or can be correctly modelled.].

In addition I show in Section 6.1 that the proposed optimal drSNMMs have two advantages over the semiparametric regret models proposed by Murphy (2003). First, when the time interval between treatment innovations is not too short, they admit closed formed estimators of the optimal treatment regime for non-continuous (e.g. dichotomous) treatments without requiring smoothing or (differentiable) approximation of indicator functions. Second, and more importantly, I believe it is easier to specify substantively meaningful (e.g. biologically meaningful) optimal drSNMMs than semiparametric regret models, because regrets are not effect measures about which scientists have clear substantive opinions amenable to easy modelling. I believe this last statement to be true not only for the sequential decision problem that is the subject of this paper but also for the simpler single time decision problem.

Section 6.2 compares and contrasts optimal regime SNMMs with DP-regression SNMMs. Optimal drSNMMs enjoy a certain robustness property not shared by the DP-regression SNMMs. Specifically, when the treatment probabilities are known (as in a randomized study) or can be correctly modelled, optimal drSNMMs quite generally allow consistent estimation of the optimal treatment regime without restricting or modelling the joint distribution of the time -dependent covariate processes. In contrast, the DP-regression component of a DP-regression SNMM model is a model for aspects of the joint distribution of this covariate process. I show that misspecification of the DP-regression model results in inconsistent estimation of the optimal treatment regime, except when the g-null hypothesis of no treatment effect holds. However, I do not consider this lack of robustness to be a shortcoming of the DP-regression SNMM compared to the optimal drSNMM methodology for the following reason. To obtain a consistent estimate of the optimal regime $\overline{d}_{op}$ under an optimal drSNMM, the optimal drSNMM must be correct. But, due to the high dimension of the covariate processes, this is not possible whenever the g-null hypothesis is false. Indeed, no method can provide a consistent estimator for $\overline{d}_{op}$ under such a high dimensional alternative, even when the treatment probabilities are known. The relevant question then is whether, based on our substantive subject specific knowledge, do we expect to obtain less biased estimates of $\overline{d}_{op}$ by specifying a DP-regression SNMM or by specifying an optimal drSNMM? Even when we possess a great deal of accurate substantive knowledge, this is often a very difficult question to answer. As a consequence I would recommend estimation of the optimal regime based on both models and then either (i) checking whether the optimal regimes and estimated expected utilities (or confidence sets for the optimal regime and for their expected utility) computed under the two different models agree at least qualitatively or (ii) choosing among the estimated optimal regimes by comparing independent estimates of their expected utilities by employing the cross-validation methods described in Section 9.

Overview: To fix ideas consider the HIV infected patients receiving their health care at a large HMO. Suppose that each patient is seen weekly beginning shortly after their time of infection with the virus. At each visit clinical and laboratory data are obtained. On the basis of a patient's treatment, clinical, and laboratory history, the patient's physician decides whether to treat with anti-virals in the next week, and, if so, the particular drugs and dosages to prescribe. Physicians wish to maximize the quality-adjusted survival time of their patients. Because there is no single agreed upon treatment protocol, different physicians make different treatment decisions in the face of the same patient history. Suppose that, after several years, the HMO wishes to estimate from data collected to date an optimal treatment protocol that, when implemented uniformly by all HMO-physicians, will maximize quality adjusted survival time. A treatment protocol or regime is a rule (function) that at the beginning of each week m takes as input a patients treatment, laboratory, and clinical history up to m and deterministically prescribes the

dosage of each available antiviral drug to be taken during the week. Our goal is to develop methods for estimating the optimal treatment protocol by utilizing the variability in treatment found in the HMO data. This variability exists because different physicians have made different treatment decisions in the face of similar patient histories.

A key identifying assumption that we shall assume until Section 7 is the assumption of no unmeasured confounders, which is also referred to as the assumption of sequential randomization. It says that, in the HMO data, among patients with a common measured (i.e. recorded for data analysis) treatment, laboratory, and clinical history up to any time m, those receiving a particular dose of treatment in week m do not differ on unmeasured (unrecorded) determinants of quality-adjusted survival. This assumption will be true if all determinants of quality-adjusted survival that are used by physicians to determine the dosage of treatment at m have been recorded in the data base. For example, since physicians tend to withhold anti-viral treatment from subjects with very low white blood count (WBC), and in untreated subjects, low white blood count is a predictor of survival, the assumption of no unmeasured confounders would be false if data on WBC history was not recorded. It is a primary goal of the HMO epidemiologists to record data on a sufficient number of covariates to ensure that the assumption of no unmeasured confounders will be at least approximately true.

The assumption of no unmeasured confounders is the fundamental condition that will allow us to draw causal inferences and to estimate the effect of interventions from observational data. It is precisely because it cannot be guaranteed to hold in an observational study and is not empirically testable that it is so very hazardous to draw causal inferences from observational data. On the other hand, the assumption of no unmeasured confounders is guaranteed to be true in a sequential randomized trial. A sequential randomized trial is a trial in which, at each time m, the dose of treatment is chosen at random by the flip of a (possibly multi-sided) coin, with the coin probabilities depending on past measured laboratory, clinical and treatment-history . It is because physical randomization guarantees the assumption of sequential randomization that most people accept that valid causal inferences can be obtained from a randomized trial. In Section 7, we discuss how the consequences of violations of the assumption of no unmeasured confounders can be explored through sensitivity analysis.

The problem of determining an optimal treatment strategy is a sequential decision problem in that the treatment to be prescribed at each week m is decided based on updated information. As discussed by Murphy (2003), Sutton and Barto (1998), Cowell et. al. (1999), and Bertsekas and Tsitsiklis (1996), this same problem is of concern in many other disciplines and subdisciplines including the disciplines of Markov Decision Processes, Multi-stage or Sequential Decision Analysis, Influence Diagrams, Decision Trees, Dynamic Programming, Partially Observed Markov Decision Processes, and Reinforcement Learning. Susan Murphy (2003) showed that under the assumption of

no unmeasured confounders, the mathematical problem we are attempting to solve is identical to that treated in these other disciplines. Thus the estimation methods developed in these disciplines could be used in place of the proposed methodologies to estimate the optimal treatment regime. However, we argue that in principle for many, if not most, of the very high dimensional sequential decision problems faced in clinical medicine, the estimation methods of Murphy and myself are better than any previous approach for the following reason. In biomedical studies, it is often the case that the treatment being evaluated has no effect on survival but may cause mild to moderately severe side effects. If so, the treatment regime that maximizes the quality-adjusted survival is to withhold all treatment. In section 2, I argue that in this setting, previous methods, other than Murphy's and my own, can, with probability near 1, estimate an optimal treatment regime that inappropriately recommends treatment for certain patients. The advantage of Murphy and my methods is intimately connected with recent work on the foundations of statistics in very high dimensional models (Robins and Ritov, 1997; Robins Rotnitzky, van der Laan, 2000, Robins and Rotnitzky,2001). This work shows that in high dimensional, sequential randomized trials with known randomization probabilities (i.e. known conditional probabilities of treatment at each time m), any method, whether Bayesian or frequentist, that satisfies the likelihood principle must ignore the known randomization probabilities. However any method of estimation that ignores these probabilities will, with high probability, incorrectly estimate, as optimal, regimes that prescribe active drug, even under the null hypothesis of no drug effect. In contrast, methods such as Murphy's and mine, that violate the likelihood principle by using these probabilities, will correctly estimate as optimal the regime that withholds all drug. In this paper the question of optimal experimental design is not adressed. Rather we take the data as given and develop analytic methods for estimation of optimal or near optimal decision rules.

Organization of the Paper: In Section 2, we formalize our problem and describe its relation with the problem treated in the sequential decision literature. In Section 3, we define drSNMMs and srSNMMs. and construct dr-lse estimators of the model parameters. In Section 4, we define optimal drSNMMs and construct dr-lse estimators. In Section 5 , we consider how to use our estimates of an optimal drSNMM to make optimal decisions. We consider both frequentist and Bayesian approaches. Because, for an optimal drSNMM, the expectation of our dr-lse estimating functions may not be differentiable with respect to the model parameters, standard methods of confidence interval construction can fail. In Section 5 we describe a novel method for the construction of frequentist confidence intervals in this setting. Furthermore, in Section 5, we show that, due to the curse of dimesionality, standard Bayesian methods fail when analyzing high dimensional data arising from either a sequential randomized trial or an observational study data. We therefore propose a new type of Bayes-Frequentist compromise procedure that allows a valid Bayesian analysis of high dimensional data by reducing the data to the locally effi-

cient doubly robust frequentist estimating function for the data, which is then viewed as a stochastic process whose index is the parameter vector of our optimal drSNMM. In Section 6 we compare and contrast drSNMMs with Murphy's regret models and DP-regression SNMMs. In Section 7, we allow for unmeasured confounding variables and propose sensitivity analysis and instrumental variable methodologies. Section 8 briefly considers extensions to the continuous time setting in which the treatment and covariate processes can jump at random times. Results in Sections 1-8 rely on two assumptions that will never be strictly correct: the first that our optimal drSNMM is correct and the second that either (but not necessarily both) a low dimensional model for the conditional law of treatment or a low dimensional model for the mean of the counterfactual utility given the past is correct. In section 9, we relax both assumptions, although not simultaneously. To accomplish this, we use recent results of van der Laan and Dudoit (2003) on model selection via cross-validation and of Robins and van der Vaart (2003,2004) on adaptive non-parametric confidence intervals and inference based on higher order influence functions. [Robins and van der Vaart (2004) consider relaxing both assumptions simultaneously.] In Appendix 1, we provide detailed calculations for a specific example in order to clarify the inferential consequences of using estimating functions with the non- differentiable expectations. Finally, Appendix 2 and 3 contain the deferred proofs of several theorems.

2 The Data and Analysis Goals

2.1 Observational Studies and Sequential Randomized Trials:

In an observational study or sequential randomized trial we assume that we observe n i.i.d. copies O_i, $i = 1, ..n$, of the random vector $O = \left(\overline{L}_{K+1}, \overline{A}_{K+1}\right)$ where $\overline{A}_{K+1} = \{A_0, \ldots, A_{K+1}\}$ are temporally ordered treatment variables given at non-random times $t_0, \ldots, t_{K+1}$, $\overline{L}_{K+1} = \{L_0, L_1, \ldots, L_{K+1}\}$ are responses (i.e. covariates) with L_m temporally subsequent to A_{m-1} but prior to A_m, and we have represented random variables by capital letters. Both A_m and L_m may be multivariate. For example in a study of HIV infected subjects, $A_m = (A_{m1}, A_{m2}, A_{m3})^T$ might be the three vector of doses of protease inhibitor, nonnucleoside reverse transcriptase inhibitor (RTI) and nucleoside RTI received in the interval $(t_m, t_{m+1}]$, while L_m might be a vector with components of white count, red count, CD4 count, level of serum HIV RNA, indicators for each HIV associated opportunistic infection, weight, height, blood pressure, etc. recorded at time t_m. [In particular L_m would include the indicator D_m of survival to t_m.] In typical applications, the number of time steps K would be in the range of 10 to 500. If the time subscript is absent, we take the history indicated by the overbar through the end of the study. For example, $\overline{L} = \overline{L}_{K+1}$. Without loss of generality, we will take $A_{K+1} = 0$ with probability one, as A_{K+1} cannot causally influence $\overline{L}_{K+1}$. Thus, in a

mild abuse of notation, we can write $\overline{A} = \overline{A}_K$. For any random vector Z_k, we use the corresponding lower case letter z_k to denote a realization of Z_k and the corresponding calligraphic letter $\mathcal{Z}_k$ to denote the support set (i.e. possible values of) of Z_k. We write $\overline{Z}_k = \{Z_0, Z_1, \ldots, Z_k\}$ and, by convention, set $\overline{Z}_{-1} \equiv Z_{-1} \equiv \overline{z}_{-1} \equiv z_{-1} \equiv 0$ with probability one.

We assume there is a known function $y(\cdot)$ of the observed data O whose expectation $E[y(O)]$ we would like to maximize. That is $Y = y(O)$ can be thought of as a utility and our goal is to maximize expected utility. For example, $y(O) = \sum_{m=1}^{K+1} D_m w_m\left(\overline{L}_m, \overline{A}_m\right)$ is quality-adjusted years of life where $D_m = 1$ if alive at t_m, 0 otherwise and $w_m(\cdot,\cdot)$ is an agreed upon function of past treatment and covariate history quantifying the quality of life at t_m. If $w_m(\cdot,\cdot) \equiv 1$, $y(O)$ is total years of life during the study period. [Note that a subject who survived until end of follow-up at time t_{K+1}, will presumably survive a considerable time past end of follow-up. This implies that the weight function $w_{K+1}(\cdot,\cdot)$ should dominate the function $w_m(\cdot,\cdot)$, $m < K+1$, especially for values of $\overline{L}_{K+1}, \overline{A}_K$ that suggest the subject is healthy and is expected to live a long time after the study comes to an end.]

A treatment regime or strategy $\overline{p} = (p_0, ..., p_K)$ is a collection of conditional densities $p_m = p_m\left(a_m|\overline{l}_m, \overline{a}_{m-1}\right)$ for the density of treatment at t_m given past treatment and covariate history. A given treatment strategy $\overline{p}$ is a deterministic strategy, say $\overline{d} = (d_0, ..., d_K)$, if for all m, a_m is a deterministic function $d_m = d_m\left(\overline{l}_m, \overline{a}_{m-1}\right)$ of $\left(\overline{l}_m, \overline{a}_{m-1}\right)$. That is $\overline{p} = \overline{d}$ means $p_m\left(a_m|\overline{l}_m, \overline{a}_{m-1}\right) = 1$ if $a_m = d_m\left(\overline{l}_m, \overline{a}_{m-1}\right)$. Let $\overline{p}_{obs} = (p_{0,obs}, ..., p_{K,obs})$ be the set of conditional treatment densities that generated the observed data. Then we can write the density $f_{obs} = f_{obs}(o)$ of the law that generated the observed data with respect to a dominating measure $\mu(\cdot)$ as the product of the conditional response densities f_{res} and the conditional treatment densities $f_{tr,\overline{p}_{obs}}$. Specifically we write, $f_{obs} \equiv f_{\overline{p}_{obs}} \equiv f_{res} f_{tr,\overline{p}_{obs}}$ where for any $\overline{p}$

$$
\begin{aligned}
f_{\overline{p}}(o) &= f_{res}(o)\, f_{tr,\overline{p}}(o), \qquad (2.1)\\
f_{res} = f_{res}(o) &= \prod_{m=0}^{K+1} f\left[l_m \mid \overline{l}_{m-1}, \overline{a}_{m-1}\right],\\
f_{tr,\overline{p}} = f_{tr,\overline{p}}(o) &= \prod_{m=0}^{K} p_m\left(a_m|\overline{l}_m, \overline{a}_{m-1}\right)
\end{aligned}
$$

Note that knowing $f_{res}(o)$ is the same as knowing each $f\left[l_m \mid \overline{l}_{m-1}, \overline{a}_{m-1}\right]$, $m = 0, ..., K+1$, since the $f\left[l_m \mid \overline{l}_{m-1}, \overline{a}_{m-1}\right]$ are densities; for example $f(l_0) = f\left[l_0 \mid \overline{l}_{0-1}, \overline{a}_{0-1}\right] = \int \cdots \int f_{res}(o) \prod_{m=1}^{K+1} d\mu(l_m)$. In an observational study neither f_{res} nor $f_{tr,\overline{p}_{obs}}$ are known and must therefore be estimated from the observed data $O_i, i = 1, ..., n$. In a sequential randomized trial f_{res} is unknown but $f_{tr,\overline{p}_{obs}}$ is known by design since the randomization probabilities $p_{m,obs}\left(a_m|\overline{l}_m, \overline{a}_{m-1}\right)$ are chosen by the investigators .

Note $f_{\overline{p}}(o) = f_{\overline{p}} = f_{res} f_{tr,\overline{p}}$ is a law in which the treatment regime $\overline{p}$ replaces the observed regime $\overline{p}_{obs}$. Actually $f_{\overline{p}}$ is not well-defined unless $\overline{p}$ is a feasible regime defined as follows.

Definition: A regime $\overline{p}$ is feasible if $f_{\overline{p}}\left(\overline{l}_m, \overline{a}_{m-1}\right) > 0$ implies the support of A_m under the conditional law $p_m\left(\cdot|\overline{l}_m, \overline{a}_{m-1}\right)$ is contained in the support of $p_{m,obs}\left(\cdot|\overline{l}_m, \overline{a}_{m-1}\right)$. We let $\overline{\mathcal{P}}$ be the set of all feasible regimes and let $\overline{\mathcal{D}}$ be the set of all deterministic feasible regimes.

If $\overline{p}$ is not feasible, then there must exist $\left(\overline{l}_m, \overline{a}_m\right)$ such that $f_{\overline{p}}\left(\overline{l}_m, \overline{a}_m\right) > 0$ but $f_{\overline{p}_{obs}}\left(\overline{l}_m, \overline{a}_m\right) = 0$. Thus $f_{\overline{p}}\left(l_{m+1}|\overline{l}_m, \overline{a}_m\right)$ is not a function of $f_{obs} \equiv f_{\overline{p}_{obs}}$ and thus is not non-parametrically identified. For any given $\overline{p} \in \overline{\mathcal{P}}$, $f_{\overline{p}}$ is identified from the observed data, since $f_{\overline{p}}$ is a function of f_{res} and f_{res} is a functional of the joint density $f_{\overline{p}_{obs}}$ of the observed data. Until Section 3.2, we restrict attention to feasible regimes. In Section 3.2, we make additional modelling assumptions that allow borrowing of information across regimes. As a consequence, the restriction to feasible regimes becomes unnecessary.

Let $f_{int}^{\overline{p}}(o)$ be the density of O that would have been observed if, contrary to fact, all subjects had followed regime $\overline{p}$. The subscript "int" denotes that $f_{int}^{\overline{p}}$ is the density of the data that would have been observed under an intervention that forced all subjects to follow regime $\overline{p}$. Our interest in $f_{\overline{p}}$ derives from the fact that, as formally discussed in section 3, under a sequential randomization assumption, $f_{int}^{\overline{p}}(o) = f_{\overline{p}}(o)$ for all $\overline{p} \in \overline{\mathcal{P}}$, which implies $f_{int}^{\overline{p}}(o)$ is identified from the observed data. Let $J_{int}^{\overline{p}} = E_{int,\overline{p}}[Y]$ and $J^{\overline{p}} = E_{\overline{p}}[Y]$ be the expectation of $Y = y(O)$ under $f_{int}^{\overline{p}}$ and $f_{\overline{p}}$, respectively. Also for a deterministic strategy $\overline{d}$, define $J^{\overline{d}}$ to be $J^{\overline{p}}$ for $\overline{p} = \overline{d}$. Note $J^{\overline{p}_{obs}}$ is the mean of Y in the observed study. If our goal is to treat a new patient, exchangeable with the n patients in the study, with the regime that maximizes expected utility, we wish to treat with a (possibly nonunique) regime $\overline{p}_{op}$ that maximizes $J_{int}^{\overline{p}}$ over $\overline{p} \in \overline{\mathcal{P}}$. Under sequential randomization, this is equivalent to maximizing $J^{\overline{p}}$. It is clear that we can always take $\overline{p}_{op}$ to be a deterministic strategy $\overline{d}_{op}$, because one can always match a random strategy with a deterministic one. That is, $\sup_{\overline{p} \in \overline{\mathcal{P}}} J^{\overline{p}} = \sup_{\overline{d} \in \overline{\mathcal{D}}} J^{\overline{d}}$. Thus, under sequential randomization, our goal is to find $\overline{d}_{op}$ maximizing $J^{\overline{d}}$.

2.2 The Sequential Decision Literature:

The standard sequential decision literature with a finite time horizon $K+1$ deals with the following problem. A set $\overline{\mathcal{P}}$ of strategies and a product of conditional densities $f_{res}(o) = \prod_{m=0}^{K+1} f\left[l_m \mid \overline{l}_{m-1}, \overline{a}_{m-1}\right]$ are given such that $f_{\overline{p}} = f_{res} f_{tr,\overline{p}}$ is well defined for all $\overline{p} \in \overline{\mathcal{P}}$. The goal is again to find $\overline{d}_{op} \in \overline{\mathcal{D}} \subset \overline{\mathcal{P}}$ maximizing $J^{\overline{d}}$ over $\overline{d} \in \overline{\mathcal{D}}$. See Sutton and Barto (1998) and Bertsekas and Tsitsiklis (1996). Thus the problem treated in this literature is exactly as above except now (i) f_{res} is known rather than estimated from the

data and (ii) there is no data. Thus the problem is purely computational rather than statistical. The dynamic programming algorithm of Bellman (1957) is the classic method of computing $\overline{d}_{op}$. However in high dimensional problems the classical dynamic programming algorithm cannot be implemented because it depends on the calculation of conditional expectations that require the evaluation of very high dimensional integrals and thus only approximate solutions are available. These approximate solutions are often based on choosing a known regime $\overline{p}_{obs}$ and simulating data O_i, $i = 1, ..., n$, from $f_{\overline{p}_{obs}}$. The needed conditional expectations are then evaluated by fitting linear or non-linear (e.g. neural network) regression models by least squares. These models are meant to roughly approximate the true functional form of the conditional expectations determined by $f_{res}(o)$ (Bertekas and Tsilikis,1996). We will call this fitting method the *regression method*. Thus in the recent sequential decision literature, the analysts actually create sequential randomized trials to approximately solve an intractable computational problem by simulation. However we will argue that, for the high dimensional problems typically occurring in biomedical studies, the regression method is inferior to dr-lse g-estimation of optimal double-regime SNMMs. Indeed the difficulty with the regression approach is already evident in high dimensional single decision problems.

Difficulty with the regression-method: We will consider the special case where $K = 1$ so the observed data is $o = (l_0, a_0, l_1, a_1, l_2)$. Further, to simplify matters, we suppose A_1 is the constant 0 with probability 1 so we can only consider treating with $a_1 = 0$ at t_1 and so A_1 can be ignored. For convenience we suppress the subscript 0 so $d_0 = d, l_0 = l, a_0 = a$. Our goal is then to find the function $d_{op}(\cdot)$ of l that maximizes $J^{\overline{d}} = \int y(l_2, l_1, l, a)\, dF(l_2 \mid a, l, l_1)\, dF(l)\, dF(l_1 \mid a, l)$. Thus $d_{op}(l) = \arg\max_{a \in \mathcal{A}} E[Y \mid A = a, L = l] = \int y(l_2, l_1, l, a)\, dF(l_2 \mid a, l, l_1) dF(l_1 | a, l)$. Now if l_1 and/or l_2 is very high dimensional the integral is intractable even if as in the sequential decision literature $f(l_1|a, l)$ and $f(l_2 \mid a, l, l_1)$ are known . In that case we choose a known density $p_{obs}(a|l)$ and simulate n iid copies $O_i = (L_{2i}, L_{1i}, A_i, L_i)$ from $f_{res} f_{tr,p_{obs}}(o) = f(l_2|a, l, l_1) f(l_1|a, l) f(l)\, p_{obs}(a|l)$. In the regression method we obtain the OLS estimate $\widehat{\eta}$ of η from the fit of a regression model $E(Y|A, L) = \eta^T w(A, L)$ to the data $Y_i = y(L_{2i}, L_{1i}, L_i, A_i), A_i, L_i$, $i = 1, ..., n$, where $w(A, L)$ is an investigator chosen vector of regressors selected to have $\eta^T w(a, l)$ approximate the true regression function $E(Y|A = a, L = l) = \int y(l_2, l_1, l, a)\, dF(l_2 \mid a, l, l_1) dF(l_1|a, l)$. The regression-estimate of $d_{op}(l)$ is $\widehat{d}_{op}(l) = \arg\max_{a \in \mathcal{A}} \widehat{\eta}^T w(a, l)$. Suppose, for the moment, the utility $y(l_2, l_1, l, a) = l_2$ where L_2 is the biological outcome of interest. In biomedical studies, it is often the case that the treatment being evaluated has no effect on the biological outcome L_2. In that case all treatment regimes d result in the same utility, so the true regression function $E(Y|A = a, L = l)$ does not depend on a. However when l is very high dimensional, the regression model $\eta^T w(A, L)$ is almost guaranteed to be misspecified and, as discussed further below, because of this misspecification, even in large samples $\widehat{\eta}^T w(a, l)$

will nearly always depend on a . It follows that if for some non-zero dose a, $\widehat{\eta}^T w(a,l)$ exceeds $\widehat{\eta}^T w(0,l)$, treatment will be unneccessarily recommended.

Remark : Of course, when $y\,(l_2,l_1,l,a)$ does not depend on a and both $f(l_1|a,l)$ *and* $f(l_2 \mid a,l,l_1)$ are known then a sufficient condition for the null hypothesis to hold is that both the known densities $f(l_1|a,l)$ and $f(l_2 \mid a,l,l_1)$ do not depend on a, which can be checked by inspection. However, surprisingly, as discussed by Robins and Wasserman (1997), it often happens in realistic biomedical settings that both densities depend on a and yet the null hypothesis that the integral $J^{\overline{d}}$ is the same for all $\overline{d} \in \overline{\mathcal{D}}$ is true. Thus whether this null hypothesis holds cannot be checked without evaluation of the integral $J^{\overline{d}}$ but, in high dimensions, this is not possible. Robins and Wasserman (1997) refer to this as the null paradox since the factors in the integral $J^{\overline{d}}$ depend on a but the integral does not. Although such a cancellation may appear miraculous it is quite common in many realistic biomedical settings. Specifically it will occur whenever (i) neither A nor L_1 affects L_2 (ii) treatment A affects L_1 and (iii), as will nearly always be the case, pre-treatment unmeasured health status, say U, is a causal determinant of both L_1 and L_2. As a specific example in HIV patients infected with AZT-resistant HIV, A being treatment with AZT, L_1 being red blood count (RBC) and L_2 being HIV RNA, (i) and (ii) will hold as (i) *neither* AZT nor RBC will influence HIV RNA but (ii) AZT is a direct red blood cell toxin.

Now if the null hypothesis of no treatment effect is false and thus $E\,(Y|A=a,L=l)$ depends on a, then, because of the high dimension of l and logistical limits to the number of observations n we can obtain or simulate, we cannot hope to correctly estimate $d_{op}\,(l)$. Thus some approximation is necessary. But we would like to use a statistical method to do this approximation that, in contrast to the regression method, does not lead us to falsely conclude that $E\,(Y|A=a,L=l)$ depends on a when in fact the null hypothesis is true. We now describe such a method after first giving a more precise treatment of the failings of the regression method.

Let $\gamma\,(l,a) = E\,(Y|A=a,L=l) - E\,(Y|A=0,L=l)$ and $b\,(l) = E\,(Y|A=0,L=l)$, where 0 is a substantively meaningful baseline level of a (e.g. no treatment). Then $E\,(Y|A=a,L=l) = \gamma\,(l,a) + b(l)$. Note $\gamma\,(l,0) = 0$. The optimal strategy depends solely on the function $\gamma\,(l,a)$ since $d_{op}\,(l) = \arg\max_{a\in\mathcal{A}} \gamma\,(l,a)$. Suppose, as an example, we model $\gamma\,(l,a)$ by $\gamma\,(l,a,\psi) = \sum_{j=1}^{s} \psi_j h_j\,(a)\, w_{1j}\,(l)$ and $E\,(Y|A=0,L=l)$ by $b\,(l,\theta) = \sum_{j=1}^{r} \theta_j w_{2j}\,(l)$ where h_j, w_{1j} and w_{2j} are functions chosen by the analyst with $h_j\,(0)=0$, so $\gamma\,(l,0,\psi)=0$ as it must. The model $\gamma\,(l,a,\psi)$ for $\gamma\,(l,a)$ is a simple special case of an optimal drSNMM of Section 4. Thus our model is

$$E\,(Y|A=a,L=l) = \eta^T w(a,l) = \gamma\,(l,a,\psi) + b\,(l,\theta) \tag{2.2}$$

Consider also the larger semiparametric partially linear model

$$E\,(Y|A=a,L=l) = \gamma\,(l,a,\psi) + b\,(l) \tag{2.3}$$

where $b(l)$ is a completely unknown function of l. Note, under the null hypothesis $E(Y|A=a,L=l)=E(Y|L=l)$, model (2.3) is guaranteed to be correctly specified with $\psi=0$ since $E(Y|A=a,L=l)$ does not depend on a if and only if $\psi=(\psi_1,...,\psi_s)^T=0$. In contrast, under the null, the model $\gamma(l,a,\psi)+b(l,\theta)$ is correctly specified only if $b(l,\theta)$ equals $E(Y|A=0,L=l)$ for some θ, which is almost certainly false when L is high dimensional. Because of this model misspecification, the regression method will fail. Specifically the OLS estimate $\widehat{\psi}$ from the fit of the model $\gamma(l,a,\psi)+b(l,\theta)$ will , except for certain exceptional data generating processes, be biased with confidence intervals that in large samples will fail to include zero, even though the null $\psi=0$ is true. [In exceptional cases the OLS estimate $\widehat{\psi}$ of $\psi=0$ remains unbiased under the null ($\psi=0$) even though the model $b(l,\theta)$ for $E(Y|A=0,L=l)$ is misspecified. An example of such an exception is the special case in which, in the definitions of $\gamma(l,a,\psi)$ and $b(l,\theta)$, we have $r=s$, $w_{1j}(l)$=$w_{2j}(l)$ for each j, and A is independent of L under $p_{obs}(a|l)=p_{obs}(a)$.]

If we could obtain a consistent estimator of ψ in the semiparametric model (2.3) we would be guaranteed to not be led astray under the null as, then, (2.3) is a correctly specified model with $\psi=0$. The key to doing so is to use the fact that $p_{obs}(a|l)$ is determined by the simulator or randomizer and thus is known. For example, consider the so-called g-estimator $\widehat{\psi}(s)$ solving $0=\sum_i U_i(\psi,s)$ with $U(\psi,s)=H(\psi)\{s(A,L)-E[s(A,L)|L]\}, H(\psi)=Y-\gamma(L,A,\psi)$, $s(A,L)$ is a vector function of the dimension of ψ chosen by the analyst, and $E[\cdot|L]$ is computed under the known $p_{obs}(a|l)$. Because under the model (2.3), $U(\psi,s)$ has mean zero at the true ψ, it follows that under standard regularity conditions $\widehat{\psi}(s)$ is a consistent asymptotically normal estimate of ψ, i.e., as $n\rightarrow\infty$, $n^{1/2}\left(\widehat{\psi}(s)-\psi\right)$ converges in law to a normal with mean zero and variance that depends on the particular choice of $s(A,L)$. Efficient choices for $s(A,L)$ are discussed in Section *3*. However a goal that may be more important than efficiency is to choose $s(A,L)$ and $p_{obs}(a|l)$ such that $E[s(A,L)|L]$ can be computed in closed form. For example if we choose $s(A,L)=As^*(L)$ and $p_{obs}(a|l)$ to have mean $v(L)$ then $E[s(A,L)|L]=v(L)s^*(L)$. In section 3, we extend this approach to the sequential decision setting.

Remark 2.1: Dimension Reduction and Estimation of ψ Using A Deterministic Design: Suppose A and L are discrete with card($\mathcal{A}$) and card($\mathcal{L}$) levels respectively. It would not be unusual to have the vector L have card($\mathcal{L}$) $=10^8$ in a biomedical application. Then in a saturated model $\gamma(l,a,\psi)$ for $\gamma(l,a)$, ψ is $\{card(\mathcal{A})-1\}$card($\mathcal{L}$) dimensional and its value determines which of the card$(\mathcal{A})^{card(\mathcal{L})}$ treatment regimes is optimal. Our use of an unsaturated model $\gamma(l,a,\psi)$ with ψ of dimension much less than $\{card(\mathcal{A})-1\}$card($\mathcal{L}$) makes the problem tractable at the cost of possible misspecification bias when the null hypothesis is false. If ψ is sufficiently low dimensional and, as in the sequential decision literature, the design parameters of a simulation study are under the analyst's control, it is computationally feasible to estimate ψ without bias under the null hypothesis and without

using g-estimation by employing a deterministic design. We describe 2 such methods. Suppose $\gamma(l,a,\psi) = a\left(\psi_0 + \sum_{j=1}^{s}\psi_j l_j\right)$. where $l = (l_1,...,l_s)^T$. Then we can estimate ψ_0 by simulating O_i $i = 1,...,n$ as above but with $L_i = 0$ for all i and with $A_i = a_{\max} = \max\{a; a \in \mathcal{A}\}$ for $i \leq n/2$ and with $A_i = a_{\min} = \min\{a; a \in \mathcal{A}\}$for $i > n/2$, and estimating the mean of ψ_0 by $\{a_{\max} - a_{\min}\}^{-1}$ times the difference between the averages of Y_i in the first and second half of the sample. We can then estimate $\psi_0 + \psi_j$ and thus ψ_j for each j by repeating the above with L_i the vector that has a 1 in entry j and 0 elsewhere. This approach is unbiased under the null hypothesis. Alternatively we could use a $L-$matched design to control confounding wherein we first take a random sample $O_{1i} = (L_{1i}, A_{1i}, Y_{1i})$ $i = 1,...,n$ of size n and then create a second sample O_{2i} $i = 1,...,n$ where $O_{2i} = (L_{2i} = L_{1i}, A_{2i} = 1 - A_{1i}, Y_{2i})$ with Y_{2i} drawn from the conditional law of Y given (L_{2i}, A_{2i}) and then minimize $\sum_{i=1}^{n}\left\{Y_{2i} - Y_{1i} - (A_{2i} - A_{1i})\left(\psi_0 + \sum_{j=1}^{s}\psi_j L_{ji}\right)\right\}^2$ over ψ.

2.3 A Counterfactual Formulation:

Following Murphy (2002), we will find it useful to have a counterfactual formulation of the problem. We adopt the model of Robins (1986). In this model associated with each treatment history $\overline{a} = \overline{a}_K \in \overline{\mathcal{A}} = \overline{\mathcal{A}}_K$ and each m we have a counterfactual response vector $\overline{L}_{\overline{a},m} = \overline{L}_{\overline{a}_{m-1},m} = (L_{\overline{a},0},, L_{\overline{a},1},..., L_{\overline{a},m}) = \left(L_{0},, L_{\overline{a_0},1},..., L_{\overline{a}_{m-1},m}\right)$ where $L_{\overline{a},j} = L_{\overline{a}_{j-1},j}$ records a subjects outcome at t_j if, possibly contrary to fact, the subject had followed treatment history $\overline{a}_{j-1}$. This notation includes the assumption that the future does not determine the past, since for any two treatment histories $\overline{a}$ and $\overline{a}^*$ that agree through t_{j-1}, $L_{\overline{a},j} = L_{\overline{a}^*,j}$. The counterfactual data is linked to the observed data by the consistency assumption that if a subject's observed treatment history $\overline{A}$ agrees with a treatment history $\overline{a}$ through t_{j-1} then a subject's observed response L_j equals the counterfactual response $L_{\overline{a},j}$ at time $t_{j.}$ That is,

$$If\ \overline{A}_{j-1} = \overline{a}_{j-1},\ \text{then}\ L_j = L_{\overline{a}_{j-1},j} \tag{2.4}$$

In particular, $L_{\overline{a},0}$ equals L_0.

Let D_m be the set of all functions d_m that are the m^{th} component of some $\overline{d} \in \overline{\mathcal{D}}$. Let $\overline{\mathcal{D}}_m$ and $\underline{\mathcal{D}}_m$ be the set of $\overline{d}_m = (d_0,...,d_m)$ *and* $\underline{d}_m = (d_m,...,d_K)$ respectively, each of whose components is in some D_k. We define the counterfactual outcome and treatment histories $\overline{L}_{\overline{d},m} = \overline{L}_{\overline{d}_{m-1},m} = \left(L_{\overline{d},0},,...,L_{\overline{d},m}\right) = \left(L_0, L_{\overline{d_0},1},...,L_{\overline{d}_{m-1},m}\right)$ and
$\overline{A}_{\overline{d},m} = \overline{A}_{\overline{d}_m,m} = \left(A_{\overline{d},0},,...,A_{\overline{d},m}\right) = \left(A_{\overline{d_0},0},...,A_{\overline{d_m},m}\right)$ associated with following regime $\overline{d}$ in terms of the $L_{\overline{a},j}$ recursively as follows: $L_{\overline{d},0} = L_0$ and, for $m = 0,...,K+1$, $A_{\overline{d_j},j} = d_j\left(\overline{L}_{\overline{d}_{j-1},j}, \overline{A}_{\overline{d}_{j-1},j-1}\right)$, where $L_{\overline{d}_{j-1},j} = L_{\overline{a}_{j-1},j}$ with $\overline{a}_{j-1} = \overline{A}_{\overline{d}_{j-1},j-1}$. That is ones counterfactual treatment $A_{\overline{d},j} = A_{\overline{d_j},j}$

under $\overline{d}$ at time t_j is the treatment assignment function $d_j(\cdot,\cdot)$ applied to the past counterfactual history $\left(\overline{L}_{\overline{d}_{j-1},j}, \overline{A}_{\overline{d}_{j-1},j-1}\right)$. Similarly, ones counterfactual outcome $L_{\overline{d},j} = L_{\overline{d}_{j-1},j}$ under $\overline{d}$ at time t_j is ones counterfactual outcome $L_{\overline{a}_{j-1},j}$ associated with the treatment $\overline{a}_{j-1}$ equal to one's counterfactual treatment history $\overline{A}_{\overline{d}_{j-1},j-1}$ through t_{j-1} under $\overline{d}$. This notation incorporates the assumption that the response of a given subject under regime $\overline{d}$ does not depend on the treatments received by any other subject. By definition, the distribution of these counterfactuals are linked to the previously introduced intervention distributions $f_{int}^{\bar{d}}$ by $f_{int}^{\bar{d}}(\bar{l},\bar{a}) = f_{\bar{L}_{\bar{d}},\bar{A}_{\bar{d}}}(\bar{l},\bar{a})$ where, for any random variable X, $f_X(x)$ is the density of X at x and $F_X(x)$ is the distribution function of X at x.

The distribution $F_{int}^{\bar{d}}(\bar{l}) = F_{\bar{L}_{\bar{d}}}(\bar{l})$ of the counterfactual data $\overline{L}_{\overline{d}} = \overline{L}_{\overline{d},K+1}$ for $\overline{d} \in \overline{\mathcal{D}}$ is identified from the distribution of the observables under the following sequential randomization assumption.

$$\text{Sequential Randomization:} \{\overline{L}_{\overline{a}}; \overline{a} \in \overline{\mathcal{A}}\} \coprod A_k \mid \overline{L}_k, \overline{A}_{k-1} \text{ w.p.1}, \quad (2.5)$$
$$for\ k = 0, 1, \ldots, K\ and\ \overline{a} \in \overline{\mathcal{A}}$$

We also refer to (2.5) as the assumption of no unmeasured confounders. Eq. (2.5) states that among subjects with covariate history $\overline{L}_k$ through t_k, and treatment history $\overline{A}_{k-1}$ through t_{k-1}, treatment A_k at time t_k is independent of the set of counterfactual outcomes $\overline{L}_{\overline{a}}$. It can be shown that if (2.5) holds, it also holds with $\{\overline{L}_{\overline{d}}, \overline{d} \in \overline{\mathcal{D}}\}$ substituted for $\{\overline{L}_{\overline{a}}; \overline{a} \in \overline{\mathcal{A}}\}$ (*Robins*, 1986). A sequential randomized trial is a designed study in which treatment level is randomly assigned at each time t_k with the known randomization probabilities $p_k\left(a_k | \overline{L}_k, \overline{A}_{k-1}\right)$, possibly depending on past outcome and treatment history $\left(\overline{L}_k, \overline{A}_{k-1}\right)$. Eq. (2.5) will be true in a sequential randomized trial because $\overline{L}_{\overline{a}}$, like gender or age at enrollment, is a fixed characteristic of a subject unaffected by the randomized treatments A_k actually received and thus is independent of A_k given the determinants $\left(\overline{L}_k, \overline{A}_{k-1}\right)$ of A_k. As stated in the following theorem of Robins (1986), Eqs. (2.5) implies that the density of $f_{int}^{\bar{d}}\left(\overline{l}_{K+1}\right)$ of $\overline{L}_{\overline{d}}$ is a functional (termed the g-computational algorithm functional or g-functional) of the distribution of the observables O and thus is identified.

Theorem 2.1 : Under (2.5), for all $\overline{d} \in \overline{\mathcal{D}}, (i)$ $f_{int}^{\bar{d}}\left(\overline{l}_{K+1}\right)$ is equal to $f_{\overline{d}}\left(\overline{l}_{K+1}\right) = \prod_{m=0}^{K+1} f\left[l_m \mid \overline{l}_{m-1}, \overline{a}_{m-1}\right]$, where, recursively, $a_m = d_m\left(\overline{l}_m, \overline{a}_{m-1}\right)$, for $m = 0, 1, ..., K$, and

$(ii) f_{int}^{\overline{d}}(y) = \int \cdots \int y\left(\overline{l}_{K+1,}, \overline{a}_K\right) \prod_{m=0}^{K} dF\left[l_m \mid \overline{l}_{m-1}, \overline{a}_{m-1}\right]$

$with\ a_m = d_m\left(\overline{l}_m, \overline{a}_{m-1}\right)$

Note: The expression for $f_{\overline{d}}\left(\overline{l}_{K+1}\right)$ given in the theorem is equivalent to the previous definition of $f_{\overline{d}}\left(\overline{l}_{K+1}\right)$, because the treatment density at each m is degenerate, taking the value $a_m = d_m\left(\overline{l}_m, \overline{a}_{m-1}\right)$ with probability 1.

We have two goals. The first goal is to develop useful models for estimating the mean $E\left[Y_{\overline{d}}\right] \equiv E^{\bar{l}}_{int}\left[Y\right] \equiv J^{\overline{d}}_{int}$ of Y under a given regime $\overline{d}$ and to derive easily computed locally-efficient doubly-robust estimators of the model parameters. Note, by Theorem 2.1, under (2.5), $E\left[Y_{\overline{d}}\right]$ equals the identifiable quantity $J^{\overline{a}}$. The second goal is to find the optimal regime(s) maximizing $E\left[Y_{\overline{d}}\right] = J^{\overline{d}}$ over $\overline{d} \in \overline{\mathcal{D}}$ and to estimate $E\left[Y_{\overline{d}}\right]$ for the optimal regime. We turn next to goal one.

3 Regime-Specific Additive Structural Nested Mean Models

3.1 Definition and Characterization:

In this section we show that regime-specific additive structural nested mean models can be used to estimate the mean of $Y_{\overline{d}}$ for a given regime $\overline{d}$.

We begin with some definitions. For any z_k, let $\underline{z}_k = (z_k, ..., z_{K+1})$. Given regimes $\overline{d} = (d_0, ..., d_K)$ and $\overline{d}^*$ and a treatment history $\overline{a}_{k-1}$, we write $Y_{\overline{a}_{k-1}, d^*_k, \underline{d}_{k+1}}$, as the response (utility) Y under the regime $\left(\overline{a}_{k-1}, d^*_k, \underline{d}_{k+1}\right)$, assumed to be $\overline{\mathcal{D}}$, in which the nondynamic regime $\overline{a}_{k-1}$ is followed through t_{k-1}, $\overline{d}^*$ is followed at t_k and the regime $\overline{d}$ is followed from t_{k+1} onwards. Let the "blip" function

$$\gamma^{\overline{d},\overline{d}^*}\left(\overline{l}_m, \overline{a}_m\right) \equiv \gamma^{\underline{d}_{m+1}, d^*_m}\left(\overline{l}_m, \overline{a}_m\right) =$$
$$E\left[Y_{\overline{a}_m, \underline{d}_{m+1}} - Y_{\overline{a}_{m-1}, d^*_m, \underline{d}_{m+1}} | \overline{L}_m = \overline{l}_m, \overline{A}_m = \overline{a}_m\right]$$

be the causal effect on the mean of Y among those with history $\overline{L}_m = \overline{l}_m, \overline{A}_m = \overline{a}_m$ of taking one's observed treatment a_m versus treatment $d^*_m\left(\overline{l}_m, \overline{a}_{m-1}\right)$ at time t_m and then following the regime $\overline{d}$ from t_{m+1} onwards. Note that from its definition $\gamma^{\overline{d},\overline{d}^*}\left(\overline{l}_K, \overline{a}_K\right)$ is the same for all $\overline{d} \in \overline{\mathcal{D}}$ at the final treatment time t_K (since $\underline{d}_{K+1}$ is empty). The functions $\gamma^{\overline{d},\overline{d}^*}\left(\overline{l}_m, \overline{a}_m\right)$ will be useful for estimating $E\left[Y_{\overline{d}}\right]$. The regime $\overline{d}^*$ functions only as a baseline level of treatment and will commonly be chosen to be the regime in which treatment is always withheld. It is for this reason that we give $\overline{d}$ pride of place before $\overline{d}^*$ when writing $\gamma^{\overline{d},\overline{d}^*}$.

Remark 3.1: The preceding definition includes a minor abuse of notation: Technically we should have written $\gamma^{\overline{d},\overline{d}^*}\left(\overline{l}_m, \overline{a}_m\right)$ as $\gamma^{\overline{d},\overline{d}^*}_m\left(\overline{l}_m, \overline{a}_m\right)$ because, for $k \neq m$, $\gamma^{\overline{d},\overline{d}^*}\left(\overline{l}_m, \overline{a}_m\right)$ and $\gamma^{\overline{d},\overline{d}^*}\left(\overline{l}_k, \overline{a}_k\right)$ have different domains and are different functions .

Remark 3.2: If for $k \geq m$, $d_{k+1}\left(\overline{l}_{k+1}, \overline{a}_k\right)$ only depends on the variables $\left(\overline{l}_m, \overline{a}_{m-1}\right)$ prior to a_m, then $\gamma^{\overline{d},\overline{d}^*}\left(\overline{l}_m, \overline{a}_m\right)$ represents the effect of one last blip of treatment a_m (versus $d_m^*(\overline{l}_m, \overline{a}_{m-1})$) at time t_m. However when, for some $k > m$, $d_{k+1}\left(\overline{l}_{k+1}, \overline{a}_k\right)$ depends on $\left(\underline{l}_{m+1}, \underline{a}_m\right)$, the actual treatment of subjects following regime $\left(\overline{a}_m, \underline{d}_{m+1}\right)$ may differ from those following regime $\left(\overline{a}_{m-1}, d_m^*, \underline{d}_{m+1}\right)$ at times subsequent to t_m and thus $\gamma^{\overline{d},\overline{d}^*}\left(\overline{l}_m, \overline{a}_m\right)$ may depend on the effect of treatments at times subsequent to t_m.

The following lemma that is a consequence of Theorem 2.1 states that under sequential randomization $\gamma^{\underline{d}_{m+1},d_m^*}\left(\overline{l}_m, \overline{a}_m\right)$ is also equal to the above counterfactual contrast among all subjects who would have history $\overline{l}_m$ under the counterfactual treatment history $\overline{a}_{m-1}$.

Lemma 3.1: Under the sequential randomization assumption (2.5),

$$\begin{aligned}\gamma^{\overline{d},\overline{d}^*}\left(\overline{l}_m, \overline{a}_m\right) &= \gamma^{\underline{d}_{m+1},d_m^*}\left(\overline{l}_m, \overline{a}_m\right) \\ &= E\left[Y_{\overline{a}_m,\underline{d}_{m+1}}, - Y_{\overline{a}_{m-1},d_m^*,\underline{d}_{m+1}}, |\overline{L}_{\overline{a}_{m-1},m} = \overline{l}_m\right].\end{aligned}$$

Note that sequential randomization in the hypothesis of Lemma 3.1 guarantees that the conditional means of $Y_{\overline{a}_m,\underline{d}_{m+1}}$, and $Y_{\overline{a}_{m-1},d_m^*,\underline{d}_{m+1}}$ among the subset of the population with observed history $\left(\overline{l}_m, \overline{a}_m\right)$ is equal to the means among the larger subset who would have had history $\overline{l}_m$ upon receiving the (possibly counterfactual) treatment $\overline{a}_{m-1}$ through t_{m-1}. It follows from Theorem 2.1 and Lemma 3.1 that the function $\gamma^{\overline{d},\overline{d}^*}\left(\overline{l}_m, \overline{a}_m\right)$ is nonparametrically identified under sequential randomization. The following theorem below states that under sequential randomization the function $\gamma^{\overline{d},\overline{d}^*}\left(\overline{l}_m, \overline{a}_m\right)$ being identically zero for a single regime pair $(\overline{d}, \overline{d}^*)$ is equivalent to the $g-null$ hypothesis

$$E\left[Y_{\overline{d}^{***}}\right] = E\left[Y_{\overline{d}^{**}}\right] \text{ for all } \overline{d}^{***}, \overline{d}^{**} \in \overline{\mathcal{D}} \tag{3.1}$$

of no effect of any feasible treatment regime on the mean of Y. Further the theorem provides formulae for $E\left[Y_{\overline{d}}\right]$ when $\gamma^{\overline{d},\overline{d}^*}\left(\overline{l}_m, \overline{a}_m\right)$ is non-zero. Below we use the notational convention $h(x) \equiv 0$ to denote that the range of the function $h(x)$ is zero. Parts (i)-(ii) are special cases of Theorem 7.1 below. Part (iii) is a special case of Theorem 7.6.

Theorem 3.1: Given sequential randomization (2.5), and regimes $\overline{d}$ and $\bar{d}^* \in \overline{\mathcal{D}}$,

(i) $\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_m\right) \equiv 0$ w.p.1 for $m = 0, ..., K$ if and only if the g-null mean hypothesis (3.1) holds.

(ii)

$$E\left[Y_{\overline{d}}\right] = E\left[H^{\overline{d}}\right] \tag{3.2}$$

where $H^{\overline{d}} = Y + \sum_{m=0}^{K} \varrho^{\overline{d}}\left(\overline{L}_m, \overline{A}_m\right)$ and

$$\begin{aligned}\varrho^{\overline{d}}\left(\overline{L}_m, \overline{A}_m\right) &= \gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_{m-1}, d_m\left(\overline{L}_m, \overline{A}_{m-1}\right)\right) - \gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_m\right) \\ &= E\left[Y_{\overline{A}_{m-1},\underline{d}_m} - Y_{\overline{A}_m,\underline{d}_{m+1}} | \overline{L}_m, \overline{A}_m\right] = -\gamma^{\overline{d},\overline{d}}(\overline{L}_m, \overline{A}_m)\end{aligned}$$

More generally,

$$E\left[Y_{\overline{a}_{k-1},\underline{d}_k}|\overline{L}_k=\overline{l}_k,\overline{A}_k=\overline{a}_k\right]=E\left[H_k^{\overline{d}}|\overline{L}_k=\overline{l}_k,\overline{A}_k=\overline{a}_k\right] \tag{3.3}$$

where $H_k^{\overline{d}}=Y+\sum_{m=k}^{K}\underline{\varrho}^{\overline{d}}\left(\overline{L}_m,\overline{A}_m\right)$.

(iii) For any other $\overline{d}^{**}\in\overline{\mathcal{D}}$,

$$E\left[Y_{\overline{d}^{**}}\right]=$$

$$E\left[Y_{\overline{d}}\right]+\int\cdots\int\left[\sum_{m=0}^{K}\left\{\gamma^{\overline{d},\overline{d}^*}\left(\overline{l}_m,\overline{a}_m\right)-\gamma^{\overline{d},\overline{d}^*}\left(\overline{l}_m,\overline{a}_{m-1},d_m\left(\overline{l}_m,\overline{a}_{m-1}\right)\right)\right\}\right]$$

$$\times\prod_{m=0}^{K}dF\left[l_m\mid\overline{l}_{m-1},\overline{a}_{m-1}\right]$$

$$\textit{where, for } m=0,1,...,K, a_m=d_m^{**}\left(\overline{l}_m,\overline{a}_{m-1}\right)\ .$$

Note that neither $H^{\bar{d}},H_k^{\bar{d}}$ nor $\underline{\varrho}^{\bar{d}}(\bar{l}_m,\bar{a}_m)$ depend on $\bar{d}^*$. It follows from Theorem 3.1(ii) that if the identifiable function $\gamma^{\overline{d},\overline{d}^*}\left(\overline{l}_m,\overline{a}_m\right)$ were known, a $n^{1/2}$- consistent estimate of $E\left[Y_{\overline{d}}\right]$ would be the sample average of the observable random variable $H^{\overline{d}.}$. Intuitively at each time t_m, $H^{\overline{d}}$ removes the effect of observed treatment A_m (compared to treatment $d_m^*\left(\overline{L}_m,\overline{A}_{m-1}\right)$) from Y by subtracting $\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$ and then adds back the effect $\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_{m-1},d_m\left(\overline{L}_m,\overline{A}_{m-1}\right)\right)$ of the treatment $d_m\left(\overline{L}_m,\overline{A}_{m-1}\right)$ specified by the regime $\overline{d}$ (again compared to treatment $d_m^*(\overline{L}_m,\overline{A}_{m-1})$)

Why Eq. (3.2) and (3.3) should be expected to hold is particularly transparent under *additive local rank preservation*, which we now define.

Definition: Let $\gamma^{\overline{d},\overline{d}^*}\left(y,\overline{L}_m,\overline{A}_m\right)$ be the counterfactual conditional quantile-quantile function $F^{-1}_{Y_{\overline{A}_{m-1},d_m^*,\underline{d}_{m+1}}}[\left\{F_{Y_{\overline{A}_m,\underline{d}_{m+1}}}\left(y|\overline{L}_m,\overline{A}_m\right)\right\}|\overline{L}_m,\overline{A}_m]$. We say we have local rank preservation w.r.t. $\left(\overline{d},\overline{d}^*\right)$ if $\gamma^{\overline{d},\overline{d}^*}\left(Y_{\overline{A}_m,\underline{d}_{m+1}},\overline{L}_m,\overline{A}_m\right)=Y_{\overline{A}_{m-1},d_m^*,\underline{d}_{m+1}}$ w.p.1. for all m.

Definition: If $\gamma^{\overline{d},\overline{d}^*}\left(y,\overline{L}_m,\overline{A}_m\right)=y-\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$ w.p1. and we have local rank preservation w.r.t. $\left(\overline{d},\overline{d}^*\right)$, we say we have additive local rank preservation w.r.t. $\left(\overline{d},\overline{d}^*\right)$.

The existence of local rank preservation is non-identifiable (i.e. untestable) since we never observe both $Y_{\overline{A}_m,\underline{d}_{m+1}}$ and $Y_{\overline{A}_{m-1},d_m^*,\underline{d}_{m+1}}$ on any subject with $A_m\neq d_m^*\left(\overline{L}_m,\overline{A}_{m-1}\right)$. However, we can rule out additive local rank preservation under sequential randomization as both $\gamma^{\overline{d},\overline{d}^*}\left(y,\overline{L}_m,\overline{A}_m\right)$ and $\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$ are identifiable so the truth of $\gamma^{\overline{d},\overline{d}^*}\left(y,\overline{L}_m,\overline{A}_m\right)=y-\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$ is subject to empirical test.

Lemma: Suppose additive local rank preservation w.r.t. $\left(\overline{d}, \overline{d}^*\right)$ holds. Then, with probability 1,

$$\begin{aligned}
Y &= Y_{\overline{A}_K}, \\
Y - \gamma^{\underline{d}_{K+1}, d_K^*}\left(\overline{L}_K, \overline{A}_K\right) &= Y_{\overline{A}_{K-1}, d_K^*}, \\
H_K^{\overline{d}} &= Y - \gamma^{\underline{d}_{K+1}, d_K^*}\left(\overline{L}_K, \overline{A}_K\right) + \\
\gamma^{\underline{d}_{K+1}, d_K^*}\left(\overline{L}_K, \overline{A}_{K-1}, d_K\left(\overline{L}_K, \overline{A}_{K-1}\right)\right) &= Y_{\overline{A}_{K-1}, d_K} \\
H_k^{\overline{d}} &= Y + \sum_{m=k}^{K} \varrho^{\overline{d}}\left(\overline{L}_m, \overline{A}_m\right) = Y_{\overline{A}_{k-1}, \underline{d}_k} \\
H_k^{\overline{d}} - \gamma^{\overline{d}, \overline{d}^*}\left(\overline{L}_{k-1}, \overline{A}_{k-1}\right) &= Y + \sum_{m=k}^{K} \varrho^{\overline{d}}\left(\overline{L}_m, \overline{A}_m\right) - \gamma^{\overline{d}, \overline{d}^*}\left(\overline{L}_{k-1}, \overline{A}_{k-1}\right) \\
&= Y_{\overline{A}_{k-2}, d_{k-1}^*, \underline{d}_k}, \\
H_{k-1}^{\overline{d}} &= H_k^{\overline{d}} - \gamma^{\overline{d}, \overline{d}^*}\left(\overline{L}_{k-1}, \overline{A}_{k-1}\right) + \\
\gamma^{\overline{d}, \overline{d}^*}\left(\overline{L}_{k-1}, \overline{A}_{k-2}, d_{k-1}\left(\overline{L}_{k-1}, \overline{A}_{k-2}\right)\right) &= Y_{\overline{A}_{k-2}, \underline{d}_{k-1}}
\end{aligned}$$

In particular, (3.2) and (3.3) hold.

Proof: All results follow directly from the definitions and the assumption of additive local rank preservation w.r.t. $\left(\overline{d}, \overline{d}^*\right)$.

It follows that (3.2) and (3.3) hold under additive local rank preservation w.r.t. $\left(\overline{d}, \overline{d}^*\right)$ even without the assumption of sequential randomization. In fact, Theorem 7.1 below implies the following more general result.

Lemma 3.2 : Eqs. (3.2) and (3.3) hold if

$$\begin{aligned}
&\gamma^{\overline{d}, \overline{d}^*}\left(\overline{L}_m, \overline{A}_{m-1,} d_m\left(\overline{L}_m, \overline{A}_{m-1}\right)\right) \\
&= E\left[Y_{\overline{A}_{m-1}, d_m, \underline{d}_{m+1}} - Y_{\overline{A}_{m-1}, d_m^*, \underline{d}_{m+1}} | \overline{L}_m, \overline{A}_{m-1}, A_m \neq d_m\left(\overline{L}_m, \overline{A}_{m-1}\right)\right].
\end{aligned}$$

The left hand side of the above equation is, by definition,

$$E\left[Y_{\overline{A}_{m-1}, d_m, \underline{d}_{m+1}} - Y_{\overline{A}_{m-1}, d_m^*, \underline{d}_{m+1}} | \overline{L}_m, \overline{A}_{m-1}, A_m = d_m\left(\overline{L}_m, \overline{A}_{m-1}\right)\right]$$

The previous equation will be true if any one of the following are true: the assumption of sequential randomization, the assumption of additive local rank preservation w.r.t. $\left(\overline{d}, \overline{d}^*\right)$, or the regime $\overline{d}$ is the same as the regime $\overline{d}^*$ (since in this latter case the left hand side and right hand side of the equation are both zero).

Since $\gamma^{\overline{d}, \overline{d}^*}\left(\overline{l}_m, \overline{a}_m\right)$ has a high dimensional argument $\left(\overline{l}_m, \overline{a}_m\right)$, nonparametric estimation is not feasible. Therefore we shall consider parametric models for $\gamma^{\overline{d}, \overline{d}^*}\left(\overline{l}_m, \overline{a}_m\right)$.

Definition: A $\overline{d}, \overline{d}^*$ double-regime additive SNMM for the outcome Y specifies that

$$\gamma^{\overline{d},\overline{d}^*}\left(\overline{l}_m, \overline{a}_m\right) = \gamma^{\overline{d}^*}\left(\overline{l}_m, \overline{a}_m; \psi^\dagger\right) \tag{3.4}$$

where $\gamma^{\overline{d}^*}\left(\overline{l}_m, \overline{a}_m; \psi\right)$ is a known function depending on a finite dimensional parameter $\psi \in \Psi \subset R^p$ that (i) satisfies $\gamma^{\overline{d}^*}\left(\overline{l}_m, \overline{a}_m; \psi\right) = 0$ if $a_m = d_m^*\left(\overline{l}_m, \overline{a}_{m-1}\right)$ and (ii) $\gamma^{\overline{d}^*}\left(\overline{l}_m, \overline{a}_m; \psi\right) \equiv 0$ for all m if and only if $\psi = 0$, so under sequential randomization, $\psi^\dagger = 0$ is equivalent to the g-null mean hypothesis (3.1).

An example of such a function is $\gamma^{\overline{d}^*}\left(\overline{l}_m, \overline{a}_m; \psi\right) = a_m\left(1, l_m, a_{m-1}\right)\psi_m = a_m\left(\psi_{1m} + \psi_{2m} l_m + \psi_{3m} a_{m-1}\right)$. Note it is comprised of a main effect term for a_m plus terms representing interactions of a_m with past treatment and covariates. In this example the dimension of $\psi = (\psi_1^T, ..., \psi_K^T)^T$ is $3K$.

Notational Conventions: The function $\gamma^{\overline{d}^*}\left(\overline{l}_m, \overline{a}_m; \psi\right)$ is indexed by $\overline{d}^*$ to indicate that condition (i) in the definition below Equation (3.4) holds. Sometimes to help the reader recall the $\overline{d}, \overline{d}^*$ regime that we are modelling, we will write the function $\gamma^{\overline{d}^*}\left(\overline{l}_m, \overline{a}_m; \psi\right)$ as $\gamma^{\overline{d},\overline{d}^*}\left(\overline{l}_m, \overline{a}_m; \psi\right)$. Similarly we will sometimes write the parameter ψ as $\psi^{\overline{d},\overline{d}^*}$ to indicate it is the parameter of a drSNMM for $\gamma^{\overline{d},\overline{d}^*}\left(\overline{l}_m, \overline{a}_m\right)$. Occasionally we will write $\gamma^{\overline{d}^*}\left(\overline{l}_m, \overline{a}_m; \psi\right)$ simply as $\gamma\left(\overline{l}_m, \overline{a}_m; \psi\right)$ to indicate a generic drSNMM model.

Remark: Robins (1994,1997) proved Theorem 3.1 and then studied estimation of the parameters $\psi^\dagger$ for $(\bar{d}, \bar{d}^*)$ -double-regime additive SNMMs in the special cases in which $\overline{d}$ and $\overline{d}^*$ are (i) the same treatment regime and (ii) this common regime $\overline{d}$ is the identically "zero" regime- that is $d_m(\bar{l}_m, \bar{a}_{m-1}) \equiv 0$ for all m where 0 is a baseline level of treatment selected by the data analyst. Because, as discussed in Robins (1986), we are free to define the "zero" level of treatment a_m for subjects with history $\left(\overline{l}_m, \overline{a}_{m-1}\right)$ to be $d_m\left(\overline{l}_m, \overline{a}_{m-1}\right)$ (even though the substantive meaning of $a_m = 0$ now depends on the $\left(\overline{l}_m, \overline{a}_{m-1}\right)$), it follows that Robins' (1994) additive SNMM model applies whenever property (i) of this remark holds, with (ii) simply a notational convention. When $\bar{d}$ equals $\bar{d}^*$, we refer to our model as a $\bar{d}-$regime additive srSNMM where sr stands for single-regime. If, in addition, $\bar{d}$ is the zero regime and the treatment level 0 has a natural substantive meaning (such as no treatment), we shall refer to the $\overline{0}-$regime srSNMM as a standard srSNMM.

Remark: From the fact that, by definition,

$$\begin{aligned}&\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_{m-1}, d_m\left(\overline{L}_m, \overline{A}_{m-1}\right)\right) - \gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_m\right)\\ &= E\left[Y_{\overline{A}_{m-1},\underline{d}_m} - Y_{\overline{A}_m,\underline{d}_{m+1}} | \overline{L}_m, \overline{A}_m\right]\end{aligned}$$

it follows that if $\gamma^{\overline{d}^*}\left(\overline{L}_m, \overline{A}_m, \psi\right)$ is a correctly specified $\overline{d}, \overline{d}^*$ drSNMM for $\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_m\right)$, then

$$\varrho^{\overline{d}}\left(\overline{L}_m, \overline{A}_m, \psi\right) = \gamma^{\overline{d}^*}\left(\overline{L}_m, \overline{A}_m, \psi\right) - \gamma^{\overline{d}^*}\left(\overline{L}_m, \overline{A}_{m-1}, d_m\left(\overline{L}_m, \overline{A}_{m-1}\right), \psi\right)$$

is a correctly specified $\bar{d}-$regime srSNMM for $\gamma^{\overline{d},\overline{d}}\left(\overline{l}_m, \overline{a}_m\right)$.

3.2 Inefficient Estimation

We now discuss (inefficient) estimation of the parameter $\psi^\dagger$ of a $(\bar{d}, \bar{d}^*)$ - double-regime additive SNMM when, as in a sequential randomized trial, $p_m\left[a_m \mid \overline{A}_{m-1}, \overline{L}_m\right]$is known by design. Efficient estimation and estimation in observational studies with unknown $p_m\left[a_m \mid \overline{A}_{m-1}, \overline{L}_m\right]$ is considered in Section 3.3. Our fundamental tool is the characterization of the true blip function $\gamma^{\overline{d},\overline{d}^*}\left(\overline{l}_m, \overline{a}_m\right)$ given in Theorem 3.2. Let $\gamma^{\overline{d}^*\,*}\left(\overline{l}_m, \overline{a}_m\right)$ denote an arbitrary function of $\left(\overline{l}_m, \overline{a}_m\right)$ satisfying (3.6) below, where the second "$*$" denotes that this is a particular, but arbitrary, function. Let

$$H_m^{\overline{d}}\left(\gamma^{\overline{d}^*\,*}\right) = H_m^{\underline{d}_m}\left(\gamma^{\overline{d}^*\,*}\right) \tag{3.5}$$
$$= Y + \sum_{j=m}^{K}\left\{\gamma^{\overline{d}^*\,*}\left(\overline{L}_j, \overline{A}_{j-1,} d_j\left(\overline{L}_j, \overline{A}_{j-1}\right)\right) - \gamma^{\overline{d}^*\,*}\left(\overline{L}_j, \overline{A}_j\right)\right\}$$

The following is a special case of Theorems 7.1 and 7.2.

Theorem 3.2: Given a regime $\overline{d}$, sequential randomization (2.5), and any function $\gamma^{\overline{d}^*\,*}\left(\overline{L}_m, \overline{A}_m\right)$ satisfying for all m

$$\gamma^{\overline{d}^*\,*}\left(\overline{L}_m, \overline{A}_m\right) = 0 \text{ if } A_m = d_m^*\left(\overline{L}_m, \overline{A}_{m-1}\right) \tag{3.6}$$

the following three conditions are equivalent.

(i) $\gamma^{\overline{d}^*\,*}\left(\overline{L}_m, \overline{A}_m\right) = \gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_m\right)$ w.p. 1,

i.e., $\gamma^{\overline{d}^*\,*}\left(\overline{L}_m, \overline{A}_m\right)$ is the true blip function $\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_m\right)$

(ii) For $m = 0, \ldots, K$,

$$E\left[H_m^{\underline{d}_m}\left(\gamma^{\overline{d}^*\,*}\right) \mid \overline{A}_m, \overline{L}_m\right] \tag{3.7}$$
$$\textit{does not depend on } A_m \textit{ w.p. } 1.$$

(iii) For each function $S_m\left(A_m\right) \equiv s_m\left(A_m, \overline{A}_{m-1}, \overline{L}_m\right)$

$$E\left[H_m^{\underline{d}_m}\left(\gamma^{\overline{d}^*\,*}\right)\left\{S_m\left(A_m\right) - E\left[S_m\left(A_m\right) \mid \overline{A}_{m-1}, \overline{L}_m\right]\right\}\right] \tag{3.8}$$
$$= 0\ .$$

provided the expectation exists.

Remark: In contrast with theorem 3.1, neither the assumption of additive local rank preservation w.r.t. $\left(\overline{d}, \overline{d}^*\right)$ nor having the regime $\overline{d}$ be the same as the regime $\overline{d}^*$can substitute for the assumption of sequential randomization in theorem 3.2.

We now show how to use the Theorem 3.2 to estimate the parameters of a drSNMM. Given a dr SNMM, define

$$H_m^{\overline{d},\overline{d}^*}(\psi) = H_{\underline{m}}^{\underline{d}_m,\underline{d}_m^*}(\psi)$$
$$= Y + \sum_{j=m}^{K} \left\{ \gamma^{\overline{d}^*}\left(\overline{L}_j, \overline{A}_{j-1}, d_j\left(\overline{L}_j, \overline{A}_{j-1}\right), \psi\right) - \gamma^{\overline{d}^*}\left(\overline{L}_j, \overline{A}_j, \psi\right) \right\}$$

and

$$U^{\overline{d},\overline{d}^*}(\psi, s) = \sum_{m=0}^{K} U_m^{\overline{d},\overline{d}^*}(\psi, s) \text{ with} \tag{3.9}$$
$$U_m^{\overline{d},\overline{d}^*}(\psi, s) = H_{\underline{m}}^{\underline{d}_m,\underline{d}_m^*}(\psi)\left\{ S_m(A_m) - E\left[S_m(A_m) \mid \overline{A}_{m-1}, \overline{L}_m \right] \right\}$$

where $s = (s_0, ..., s_K)$ and $S_m(A_m) \equiv s_m\left(A_m, \overline{A}_{m-1}, \overline{L}_m\right)$ is a vector valued function with range the dimension of ψ, chosen by the analyst. Note $U^{\overline{d},\overline{d}^*}(\psi, s)$ can be explicitly computed from the data when $p\left[a_m \mid \overline{A}_{m-1}, \overline{L}_m\right]$ is known, since then the expectation in (3.9) can be calculated. It follows from (3.8) that $U^{\overline{d},\overline{d}}(\psi, s)$ is an unbiased estimating function, i.e., $E\left[U^{\overline{d},\overline{d}^*}\left(\psi^\dagger, s\right)\right] = 0$. Define P_n to be the expectation operator with respect to the empirical distribution so $P_n[Z] = n^{-1}\sum_i Z_i$. Then, by the central limit theorem, $n^{1/2}P_n[U^{\overline{d},\overline{d}^*}(\psi^\dagger, s)]$ is asymptotically normal with mean 0 and covariance $\Sigma(\psi^\dagger, s)$ which can be consistently estimated by $\hat{\Sigma}(\psi^\dagger, s) = P_n[U^{\overline{d},\bar{d}^*}(\psi^\dagger, s)^{\otimes 2}]$. Thus the set of ψ for which $nP_n[U^{\bar{d}\bar{d}^*}(\psi, s)]^T \hat{\Sigma}(\psi, s)^{-1} P_n[U^{\overline{d}\bar{d}^*}(\psi, s)]$ is less than the α upper quantile of a χ^2 distribution on the dimension of ψ degrees of freedom is a large sample $1-\alpha$ confidence interval for $\psi^{\dagger\cdot}$. Further $U^{\overline{d},\overline{d}}(\psi, s)$ is smooth in ψ when $\gamma\left(\overline{L}_m, \overline{A}_m, \psi\right)$ is smooth in ψ and is linear in ψ when $\gamma\left(\overline{L}_m, \overline{A}_m, \psi\right)$ is linear. Suppose $E\left[\partial U^{\overline{d},\overline{d}*}\left(\psi^\dagger, s\right)/\partial\psi\right]$ exists and is invertible. Then under standard regularity conditions there will be a consistent asymptotically normal (CAN) root $\widehat{\psi} = \widehat{\psi}(s)$ of the estimating function $P_n\left[U^{\overline{d},\overline{d}^*}(\psi, s)\right] = 0$ under (2.5). That is $n^{1/2}\left(\widehat{\psi} - \psi^\dagger\right)$ will converge to a normal distribution with zero mean and finite variance; further $P_n\left[H_0^{\underline{d}_0\underline{d}_0^*}\left(\widehat{\psi}\right)\right]$ is a CAN estimator of $E\left[Y_{\overline{d}}\right]$ that does not require knowing or modelling $f\left[l_m \mid \overline{l}_{m-1}, \overline{a}_{m-1}\right]$.

Remark: In fact $P_n\left[H_0^{\underline{d}_0\underline{d}_0^*}\left(\widehat{\psi}\right)\right]$ is a CAN estimator of $E\left[Y_{\overline{d}}\right]$ even if $\overline{d}$ is not a feasible regime, because, by assuming the parametric model (3.4) is correct, the estimator $P_n\left[H_0^{\underline{d}_0\underline{d}_0^*}\left(\widehat{\psi}\right)\right]$ succeeds in borrowing sufficient information from subject's following other regimes to estimate $E\left[Y_{\overline{d}}\right]$ at an $n^{1/2}-$rate.

When $\gamma^{\overline{d}^*}\left(\overline{L}_m, \overline{A}_m, \psi\right)$ is a linear function of $\psi, P_n\left[U^{\overline{d},\overline{d}^*}(\psi, s)\right] = 0$ will, with probability approaching 1, have a unique root. However when $\gamma^{\overline{d}^*}\left(\overline{L}_m, \overline{A}_m, \psi\right)$ is a nonlinear function of ψ, $E\left[U^{\overline{d},\overline{d}^*}(\psi, s)\right] = 0$ may have roots in addition to $\psi^\dagger$, in which case even asymptotically $P_n\left[U^{\overline{d},\overline{d}}(\psi, s)\right] = 0$ will have additional inconsistent roots. In that case it can help to choose functions s_m whose range is of greater dimension than ψ and find the minimizer $\widehat{\psi}(s)$ of the smooth function $\left\{P_n\left[U^{\overline{d},\overline{d}^*}(\psi, s)\right]\right\}^T \times$ $B\{P_n\left[U^{\overline{d},\overline{d}^*}(\psi, s)\right]\}$ for some positive definite matrix B (*e.g.* the identity matrix) [as $P_n\left[U^{\overline{d},\overline{d}^*}(\psi, s)\right] = 0$ will no longer have a solution.] It follows from Theorem 3.2(iii) that in sufficiently large samples, this should solve the multiple root problem provided that one can solve the computational problem of minimizing $\left\{P_n\left[U^{\overline{d},\overline{d}^*}(\psi, s)\right]\right\}^T B\{P_n\left[U^{\overline{d},\overline{d}^*}(\psi, s)\right]\}$.

Remark: The approach based on choosing functions s_m whose range is of greater dimension than ψ will fail when the model (3.4) is even slightly misspecified as then there will in general be no ψ for which $E[U^{\bar{d},\bar{d}^*}(\psi, s)] = 0$. Unfortunately, in practice, the drSNMM (3.4) would never be expected to be precisely correct (unless saturated) [even though it might well be close enough to being correct that the bias of $P_n\left[H_0^{\underline{d}_0}{}^{\underline{d}_0^*}\left(\widehat{\psi}\right)\right]$ as an estimator of $E\left[Y_{\overline{d}}^*\right]$ is of little substantive importance.] Nonetheless we will continue to assume (3.4) is precisely correct until Section 9. In Section 9 we explore the consequences of dropping this assumption.

3.3 Locally Efficient Doubly-Robust Estimation:

We now consider how to obtain a locally semiparametric efficient estimator of the parameter ψ^* of a drSNMM. To do so we introduce a generalization of the estimating function $U(\psi, s)$.

Again let $S_m(A_m) \equiv s_m\left(A_m, \overline{A}_{m-1}, \overline{L}_m\right)$ have range the dimension of ψ. Now given *dim* ψ vector-valued functions $C_m = c_m\left(\overline{A}_m, \overline{L}_m\right)$ chosen by the investigator, define $C = c\left(\overline{L}_K, \overline{A}_K\right) = \sum_{m=0}^{K}\left\{C_m - E\left[C_m|\overline{A}_{m-1}, \overline{L}_m\right]\right\}$ and

$$U^\dagger(\psi, s, c) = U(\psi, s) - C$$

where for notational convenience we have suppressed the $\overline{d}$ and $\overline{d}^*$ superscripts denoting the regimes under consideration. We will now consider estimators of ψ under three different semiparametric models which we denote by (a.1), (a.2), and (a.3). The three models differ only in the a priori knowledge about the conditional treatment laws $p\left[A_m \mid \overline{L}_m, \overline{A}_{m-1}\right]$. Specifically, they are characterized by an $\overline{d}, \bar{d}^*-$regime specific additive SNMM

model $\gamma\left(\overline{l}_m, \overline{a}_m, \psi\right)$, the sequential randomization assumption (2.5), and a model $\left\{p_m\left(a_m \mid \overline{l}_m, \overline{a}_{m-1}; \alpha\right); \alpha \in \alpha\right\}$ for $p_m\left(a_m \mid \overline{l}_m, \overline{a}_{m-1}\right)$ with true parameter $\alpha^\dagger$ where $p_m\left(a_m \mid \overline{l}_m, \overline{a}_{m-1}; \alpha\right)$ is a conditional density known up to α. In model (a.1), $\alpha = \left\{\alpha^\dagger\right\}$ is a singleton, so $p_m\left(a_m \mid \overline{l}_m, \overline{a}_{m-1}\right) = p_m\left(a_m \mid \overline{l}_m, \overline{a}_{m-1}; \alpha^\dagger\right)$ is known as in a sequential randomized trial. In model (a.2), $\alpha \subset R^p$ is a finite-dimensional parameter. In model (a.3), α is an infinite-dimensional set indexing all conditional densities possibly restricted by the requirement that they belong to some prespecified smoothness class.

To avoid misspecification bias we would like to use model (a.3) in analyzing observational data. Due to the high dimension of $\left(\overline{L}_m, \overline{A}_{m-1}\right)$ this is often not possible, and thus in practice, model (a.2) would often be used for dimension reduction. Let $U^\dagger(\psi, \alpha, s, c)$ be $U^\dagger(\psi, s, c)$ with $p_m\left[A_m \mid \overline{L}_m, \overline{A}_{m-1}\right]$ replaced by
$p\left[A_m \mid \overline{L}_m, \overline{A}_{m-1}; \alpha\right]$ in the expectation in (3.9) and let $\widehat{\psi}(s, c; \alpha)$ denote a solution to

$$P_n\left[U^\dagger(\psi, \alpha, s, c)\right] = 0.$$

Finally, in model (a.2), let $\widehat{\alpha}$ solve the parametric partial likelihood score equation

$$P_n\left[S_{part}(\alpha)\right] = 0$$

with $S_{part}(\alpha) = \partial \log \prod_{m=0}^{K} p_m\left[A_m \mid \overline{L}_m, \overline{A}_{m-1}; \alpha\right] / \partial\alpha$, and in model (a.3) let $p_m\left[A_m \mid \overline{L}_m, \overline{A}_{m-1}; \widehat{\alpha}_{smooth}\right]$ denote a high-dimensional non-parametric estimate of $p_m\left[A_m \mid \overline{L}_m, \overline{A}_{m-1}\right]$ possibly obtained by smoothing.

The following theorem 3.3 states results related to the asymptotic properties of estimators of the form $\widehat{\psi}\left(s, c; \alpha^\dagger\right), \widehat{\psi}(s, c; \widehat{\alpha})$ and $\widehat{\psi}(s, c; \widehat{\alpha}_{smooth})$. Theorem 3.3 and 3.4 below are special cases of Theorem 4.3 in Robins and Rotnitzky (2003). We stress that when, as will usually be the case, $\left(\overline{L}_m, \overline{A}_{m-1}\right)$ is high dimensional, $p_m\left[A_m \mid \overline{L}_m, \overline{A}_{m-1}; \widehat{\alpha}_{smooth}\right]$ will be a poor estimate of $p_m\left[A_m \mid \overline{L}_m, \overline{A}_{m-1}\right]$ in the moderate sized samples occurring in practice and thus the asymptotic results for model (a.3) described in the theorem will not be a useful guide to finite sample properties. However knowledge of the asymptotic results for model (a.3) will be useful in understanding the mathematical structure of the problem.

We shall need the following definitions. An estimator $\widehat{\psi}$ of $\psi^\dagger$ is asymptotically linear if $n^{1/2}\left(\widehat{\psi} - \psi^\dagger\right) = n^{1/2} P_n[D] + o_p(1)$ with $E[D] = 0, E\left[D^T D\right] < \infty$, and $o_p(1)$ denotes a random variable converging to zero in probability. D is to referred to as the influence function of $\widehat{\psi}$. If an estimator is asymptotically linear then $n^{1/2}\left(\widehat{\psi} - \psi^\dagger\right)$ is asymptotically normal with mean 0 and variance $E\left[DD^T\right]$. An estimator $\widehat{\psi}$ is regular if its convergence to its limiting distribution is locally uniform in $n^{-1/2}$ neighborhoods of the truth.

Theorem 3.3: Suppose that the sequential randomization assumption (2.5) holds and we have a correctly specified dr SNMM $\gamma\left(\overline{l}_m, \overline{a}_m, \psi\right)$. Suppose

that for all c and for all s such that $\partial E\left\{U\left(\psi,s\right)\right\}/\partial\psi_{|\psi=\psi^{\dagger}}$ exists and is non-singular, $\widehat{\psi}\left(s,c;\widehat{\alpha}_{smooth}\right)$ is a regular asymptotically linear (RAL) estimator in model (a.3) and thus in models (a.2) and (a.1), $\widehat{\psi}\left(s,c;\widehat{\alpha}\right)$ is a RAL estimator in model (a.2) and thus in model (a.1), and $\widehat{\psi}\left(s,c;\alpha^{\dagger}\right)$ is a RAL estimator in model (a.1). Then

(i) The influence function of any RAL estimator of ψ in models (a.1), (a.2) and (a.3) is equal to that of $\widehat{\psi}\left(s,c;\alpha^{\dagger}\right)$, $\widehat{\psi}\left(s,c;\widehat{\alpha}\right)$, and $\widehat{\psi}\left(s,c;\widehat{\alpha}_{smooth}\right)$, respectively for some (s,c).

(ii) For a given s, the asymptotic variance of any RAL estimator of the form $\widehat{\psi}\left(s,c;\alpha^{\dagger}\right)$, $\widehat{\psi}\left(s,c;\widehat{\alpha}\right)$, or $\widehat{\psi}\left(s,c;\widehat{\alpha}_{smooth}\right)$ is minimized at $c^{s}\left(\overline{L}_{K},\overline{A}_{K}\right)=\sum_{m=0}^{K}\left\{C_{m}^{s}-E\left[C_{m}^{s}|\overline{A}_{m-1},\overline{L}_{m}\right]\right\}$ with

$$C_{m}^{s}\equiv E\left[U\left(\psi^{\dagger},s\right)\mid\overline{A}_{m},\overline{L}_{m}\right]$$

Further, under (2.5), $E\left[H_{m}\left(\psi\right)\mid\overline{A}_{j},\overline{L}_{j}\right]=E\left[H_{m}\left(\psi\right)\mid\overline{A}_{j-1},\overline{L}_{j}\right]$ for $j\geq m$. Thus

$$\begin{aligned}
&C_{m}^{s}-E\left[C_{m}^{s}|\overline{A}_{m-1},\overline{L}_{m}\right]\\
&=E\left[U_{m}\left(\psi^{\dagger},s\right)\mid\overline{A}_{m},\overline{L}_{m}\right]-E\left[U_{m}\left(\psi^{\dagger},s\right)|\overline{A}_{m-1},\overline{L}_{m}\right]\\
&=E\left[H_{m}\left(\psi^{\dagger}\right)\left\{S_{m}\left(A_{m}\right)-E\left[S_{m}\left(A_{m}\right)\mid\overline{A}_{m},\overline{L}_{m}\right]\right\}\mid\overline{A}_{m},\overline{L}_{m}\right]\\
&=E\left[H_{m}\left(\psi^{\dagger}\right)|\overline{A}_{m-1},\overline{L}_{m}\right]\left\{S_{m}\left(A_{m}\right)-E\left[S_{m}\left(A_{m}\right)\mid\overline{A}_{m-1},\overline{L}_{m}\right]\right\}.
\end{aligned}$$

Hence

$$U^{\dagger}\left(\psi,s,c^{s}\right)= \tag{3.10}$$

$$\sum_{m=0}^{K}\left\{H_{m}\left(\psi\right)-E\left[H_{m}\left(\psi\right)|\overline{A}_{m-1},\overline{L}_{m}\right]\right\}\left\{S_{m}\left(A_{m}\right)-E\left[S_{m}\left(A_{m}\right)\mid\overline{A}_{m-1},\overline{L}_{m}\right]\right\}\ .$$

Further, in model (a.3) the RAL estimators $\widehat{\psi}\left(s,c;\widehat{\alpha}_{smooth}\right)$ have the same influence function for all c.

(iii) The influence functions of $\widehat{\psi}\left(s,c;\alpha^{\dagger}\right)$, $\widehat{\psi}\left(s,c;\widehat{\alpha}\right)$, and $\widehat{\psi}\left(s,c;\widehat{\alpha}_{smooth}\right)$ are respectively equal to $I^{-1}U^{\dagger}\left(s,c\right)$, $I^{-1}\left\{U^{\dagger}\left(s,c\right)-E\left[U^{\dagger}\left(s,c\right)S_{part}^{T}\right]\left\{E\left[S_{part}^{\otimes2}\right]\right\}^{-1}S_{part}\right\}$, and $I^{-1}U^{\dagger}\left(s,c^{s}\right)$ where $U^{\dagger}\left(s,c\right)=U^{\dagger}\left(\psi^{\dagger},s,c\right)$, $S_{part}=S_{part}\left(\alpha^{\dagger}\right)$, and $I=-E\left[\partial U\left(\psi^{\dagger},s\right)/\partial\psi\right]$. The asymptotic variance of $\widehat{\psi}\left(s,c;\widehat{\alpha}_{smooth}\right)$ is always less than or equal to that of $\widehat{\psi}\left(s,c;\widehat{\alpha}\right)$, which is always less than or equal to that of $\widehat{\psi}\left(s,c;\alpha^{\dagger}\right)$. However, when $c=c^{s}$, all three estimators have the same asymptotic variance.

(iv) The semiparametric variance bound is the same in models (a.1), (a.2) and (a.3) and is equal to the asymptotic variance of $\widehat{\psi}\left(s_{eff},c^{s_{eff}}\right)$. An explicit

closed form expression for s_{eff} is given in Robins (1994). The expression is very complex except in the special case in which

$$var\left(H_m\left(\psi^\dagger\right)|\overline{L}_m,\overline{A}_m\right)=var\left(H_m\left(\psi^\dagger\right)|\overline{L}_m,\overline{A}_{m-1}\right) \tag{3.11}$$

does not depend on A_m for all m. In that case

$$s_{eff,m}\left(\overline{L}_m,\overline{A}_m\right)\equiv S_{eff,m}\left(A_m,\psi^\dagger\right) \text{ with} \tag{3.12}$$
$$S_{eff,m}\left(A_m,\psi\right)=\left\{E\left[\partial H_m\left(\psi\right)/\partial\psi|\overline{L}_m,\overline{A}_m\right]\right\}\left\{var\left(H_m\left(\psi\right)|\overline{L}_m,\overline{A}_{m-1}\right)\right\}^{-1}.$$

Double Robustness: Surprisingly, given (2.5) and a correct drSNMM, the estimator $\widehat{\psi}\left(s,c^s;\widehat{\alpha}\right)$ is a CAN estimator of $\psi^\dagger$ even if the model $p_m\left[A_m\mid\overline{L}_m,\overline{A}_{m-1};\alpha\right]$ for $p_m\left[A_m\mid\overline{L}_m,\overline{A}_{m-1}\right]$ is misspecified because $E\left[U^\dagger\left(\psi^\dagger,s,c^s;\widehat{\alpha}\right)\right]=0$ under misspecification of $p_m\left[A_m\mid\overline{L}_m,\overline{A}_{m-1};\alpha\right]$. However this is of no direct use because $E\left[H_m\left(\psi\right)|\overline{A}_{m-1},\overline{L}_m\right]$ is unknown and thus $U^\dagger\left(\psi^\dagger,s,c^s;\widehat{\alpha}\right)$ is not computable. Therefore we consider a model ς^TW_m for $E\left[H_m\left(\psi^\dagger\right)|\overline{L}_m,\overline{A}_{m-1}\right]$ where $W_m=w_m\left(\overline{A}_{m-1},\overline{L}_m\right)$ is a known vector function of $\left(\overline{A}_{m-1},\overline{L}_m\right)$ and ς is an unknown parameter. Define $\widehat{\psi}\left(s,c^s;\widehat{\alpha},\widehat{\varsigma}\right)$ to be a solution to $P_n\left[U^\dagger\left(\psi,s,c^s;\widehat{\alpha},\widehat{\varsigma}\right)\right]=0$ where

$$\begin{aligned}&U^\dagger\left(\psi,s,c^s;\widehat{\alpha},\widehat{\varsigma}\right) \\ &=\sum_{m=0}^{K}\left\{H_m\left(\psi\right)-\widehat{\varsigma}^T\left(\psi\right)W_m\right\}\left\{S_m\left(A_m\right)-E_{\widehat{\alpha}}\left[S_m\left(A_m\right)\mid\overline{A}_{m-1},\overline{L}_m\right]\right\}\end{aligned} \tag{3.13}$$

is Equation (3.10) with $-E_{\widehat{\alpha}}\left[S_m\left(A_m\right)\mid\overline{A}_{m-1},\overline{L}_m\right]\}$ substituted for $-E\left[S_m\left(A_m\right)\mid\overline{A}_{m-1},\overline{L}_m\right]$ and with $\widehat{\varsigma}^T\left(\psi\right)W_m$ substituted for $E\left[H_m\left(\psi\right)|\overline{A}_{m-1},\overline{L}_m\right]$. Here $\widehat{\varsigma}^T\left(\psi\right)$ solves the OLS estimating equation $P_n\left[\sum_{m=0}^{K}\left(H_m\left(\psi\right)-\varsigma^TW_m\right)W_m\right]=0$. The following theorem and Remark 3.1 below describe the so called "double-robustness" properties of $\widehat{\psi}\left(s,c^s;\widehat{\alpha},\widehat{\varsigma}\right)$ and $U^\dagger\left(\psi^\dagger,s,c^s;\widehat{\alpha},\widehat{\varsigma}\right)$.

Theorem 3.4: Consider the $\overline{d},\overline{d}^*$"union" model characterized by (i) the sequential randomization assumption (2.5), (ii) a correctly specified $(\bar{d},\bar{d}^*)$-double-regime-specific SNMM $\gamma\left(\bar{l}_m,\overline{a}_m,\psi\right)$ and that either (but not necessarily both) the parametric model $p_m\left[A_m\mid\overline{L}_m,\overline{A}_{m-1};\alpha\right]$ for $p_m\left[A_m\mid\overline{L}_m,\overline{A}_{m-1}\right]$ is correct or the regression model ς^TW_m for $E\left[H_m\left(\psi^\dagger\right)|\overline{L}_m,\overline{A}_{m-1}\right]$ is correct. Then , under standard regularity conditions, (i) $n^{-1/2}P_n\left[U^\dagger\left(\psi^\dagger,s,c^s;\widehat{\alpha},\widehat{\varsigma}\right)\right]=n^{-1/2}P_n\left[U^\dagger_{adj}\left(\psi^\dagger,s,c^s;\alpha^*,\varsigma^*\right)\right]+o_p\left(1\right)$, where

$$
\begin{array}{l}
U^{\dagger}_{adj}\left(\psi^{\dagger},s,c^{s};\alpha^{*},\varsigma^{*}\right)= \\
U^{\dagger}\left(\psi^{\dagger},s,c^{s};\alpha^{*},\varsigma^{*}\right)- \\
E\left[\partial U^{\dagger}\left(\psi^{\dagger},s,c^{s};\alpha^{*},\varsigma^{*}\right)/\partial\alpha^{T}\right]E\left[\partial S_{part}\left(\alpha^{*}\right)/\partial\alpha^{T}\right]^{-1}S_{part}\left(\alpha^{*}\right)+ \\
E\left[\partial U^{\dagger}\left(\psi^{\dagger},s,c^{s};\alpha^{*},\varsigma^{*}\right)/\partial\varsigma\left(\psi^{\dagger}\right)^{T}\right]\times \\
E\left[\sum_{m=0}^{K}W_{m}^{\otimes 2}\right]^{-1}\sum_{m=0}^{K}\left(H_{m}\left(\psi^{\dagger}\right)-\varsigma^{*}\left(\psi^{\dagger}\right)W_{m}\right)W_{m}
\end{array}
\tag{3.14}
$$

and $(\alpha^{*},\varsigma^{*})$ is the probability limit of $(\widehat{\alpha},\widehat{\varsigma})$. Since $S_{part}\left(\alpha^{*}\right)$, $\sum_{m=0}^{K}\left(H_{m}\left(\psi^{\dagger}\right)-\varsigma^{*}\left(\psi^{\dagger}\right)W_{m}\right)W_{m}$ and, under the union model, $U^{\dagger}\left(\psi^{\dagger},s,c^{s};\alpha^{*},\varsigma^{*}\right)$ all have mean 0, it follows that, by Slutsky's theorem, (i) $n^{-1/2}P_{n}\left[U^{\dagger}\left(\psi^{\dagger},s,c^{s};\widehat{\alpha},\widehat{\varsigma}\right)\right]$ is asymptotically normal with mean zero and covariance $E\left[U^{\dagger}_{adj}\left(\psi^{\dagger},s,c^{s};\alpha^{*},\varsigma^{*}\right)^{\otimes 2}\right]$, (ii) $\widehat{\psi}\left(s,c^{s};\widehat{\alpha},\widehat{\varsigma}\right)$ is asymptotically linear with influence function

$$
-E\left[\partial U^{\dagger}_{adj}\left(\psi^{\dagger},s,c^{s};\alpha^{*},\varsigma^{*}\right)/\partial\psi\right]^{-1}U^{\dagger}_{adj}\left(\psi^{\dagger},s,c^{s};\alpha^{*},\varsigma^{*}\right)
$$

provided $E\left[\partial U^{\dagger}_{adj}\left(\psi^{\dagger},s,c^{s};\alpha^{*},\varsigma^{*}\right)/\partial\psi\right]$ exists and is invertible and (iii) $\widehat{\psi}\left(s_{eff},c^{s_{eff}};\widehat{\alpha},\widehat{\varsigma}\right)$ attains the semiparametric efficiency bound for the "union" model at the submodel in which (3.11) holds and both the model $p_{m}\left[A_{m}\mid\overline{L}_{m},\overline{A}_{m-1};\alpha\right]$ and the model $\varsigma^{T}W_{m}$ for $E\left[H_{m}\left(\psi^{\dagger}\right)|\overline{L}_{m},\overline{A}_{m-1}\right]$ are correct. Theorem 3.4 suggests an analysis based on the following algorithm.

Algorithm:

Step 1: Specify regimes of interest $\overline{d}$ and $\bar{d}^{*}$.

Step 2:Specify a drSNMM model $\gamma\left(\overline{L}_{m},\overline{A}_{m},\psi\right)$ for $\gamma^{\overline{d},\overline{d}^{*}}\left(\overline{L}_{m},\overline{A}_{m}\right)$

Step 3: For each subject calculate $H_{m}\left(\psi\right)$, $m=0,...,K$ for each possible value of ψ. (Obviously this is not computationally feasible since ψ can take values in $R^{\dim(\psi)}$. Below we discuss how one can reduce the number of ψ for which $H_{m}\left(\psi\right)$ must be evaluated in order to make the algorithm feasible).

Step 4: For each subject estimate $E\left[H_{m}\left(\psi\right)|\overline{L}_{m},\overline{A}_{m-1}\right]$ under a model $\varsigma^{T}W_{m}$ by $\widehat{\varsigma}^{T}\left(\psi\right)W_{m}$ where $\widehat{\varsigma}^{T}\left(\psi\right)$ solves the OLS estimating equation $P_{n}\left[\sum_{m=0}^{K}\left(H_{m}\left(\psi\right)-\varsigma W_{m}\right)W_{m}\right]=0$ and $W_{m}=w_{m}\left(\overline{A}_{m-1},\overline{L}_{m}\right)$ is known vector function of $\left(\overline{A}_{m-1},\overline{L}_{m}\right)$.

Step 5. For each subject estimate $var\left(H_{m}\left(\psi^{\dagger}\right)|\overline{A}_{m-1},\overline{L}_{m}\right)$ by the fitted value $exp\left(\widehat{\varkappa}\left(\psi\right)^{T}B_{m}\right)$ from the nonlinear least squares fit of the regression of the estimated squared residual $\left\{H_{m}\left(\psi\right)-\widehat{\varsigma}\left(\psi\right)W_{m}\right\}^{2}$ on the regression function $exp\left(\varkappa B_{m}\right)$ where $B_{m}=b_{m}\left(\overline{L}_{m},\overline{A}_{m-1}\right)$ is a known vector function.

Step 6. For each subject estimate $E\left[\partial H_{m}\left(\psi\right)/\partial\psi|\overline{L}_{m},\overline{A}_{m}\right]$ under a multivariate regression model ζR_{m} by the fitted value $\widehat{\zeta}\left(\psi\right)R_{m}$ from the multivariate OLS regression of $\partial H_{m}\left(\psi\right)/\partial\psi$ on R_{m} where $R_{m}=r_{m}\left(\overline{A}_{m},\overline{L}_{m}\right)$ is a known vector function and ζ is a matrix of regression coefficients.

Step 7: Specify a parametric model $p_m\left[A_m \mid \overline{L}_m, \overline{A}_{m-1}; \alpha\right]$ and find $\widehat{\alpha}$ solving $P_n\left[S_{part}\left(\alpha\right)\right] = 0$

Step 8: For each subject compute

$\widehat{S}_{eff,m}\left(A_m\right) = \widehat{\zeta}\left(\psi\right) R_m \left\{exp\left(\widehat{\varkappa}\left(\psi\right)^T B_m\right)\right\}^{-1}$ and

$$U^{\dagger}\left(\psi, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}\,, \widehat{\varkappa}, \widehat{\zeta}\right) \tag{3.15}$$
$$= \sum_{m=0}^{K} \left\{H_m\left(\psi\right) - \widehat{\varsigma}(\psi)\, W_m\right\} \left\{\widehat{S}_{eff,m}\left(A_m\right) - E_{\widehat{\alpha}}\left[\widehat{S}_{eff,m}\left(A_m\right) \mid \overline{A}_{m-1}, \overline{L}_m\right]\right\}$$

Step 9: Compute

$$P_n\left[\partial U^{\dagger}\left(\psi, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}\,, \widehat{\varkappa}, \widehat{\zeta}\right) / \partial \widehat{\varsigma}(\psi)^T\right],$$
$$P_n\left[\partial U^{\dagger}\left(\psi, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}\,, \widehat{\varkappa}, \widehat{\zeta}\right) / \partial \widehat{\alpha}^T\right],$$
$$P_n\left[\sum_{m=0}^{K} W_m^{\otimes 2}\right],$$
$$P_n\left[\partial S_{part}\left(\widehat{\alpha}\right) / \partial \alpha^T\right],$$
$$P_n\left[\partial U^{\dagger}\left(\psi, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}\,, \widehat{\varkappa}, \widehat{\zeta}\right) / \partial \psi^T\right]$$

Step 10: For each subject compute $U_{adj}^{\dagger}\left(\psi, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}\,, \widehat{\varkappa}, \widehat{\zeta}\right)$

$$= U^{\dagger}\left(\psi, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}\,, \widehat{\varkappa}, \widehat{\zeta}\right) - \tag{3.16}$$
$$P_n\left[\partial U^{\dagger}\left(\psi, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}\,, \widehat{\varkappa}, \widehat{\zeta}\right) / \partial \widehat{\alpha}^T\right] P_n\left[\partial S_{part}\left(\widehat{\alpha}\right) / \partial \alpha^T\right] S_{part}\left(\widehat{\alpha}\right) +$$
$$P_n\left[\partial U^{\dagger}\left(\psi, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}\,, \widehat{\varkappa}, \widehat{\zeta}\right) / \partial \widehat{\varsigma}(\psi)^T\right] P_n\left[\sum_{m=0}^{K} W_m^{\otimes 2}\right]^{-1} \times$$
$$\sum_{m=0}^{K} \left(H_m\left(\psi\right) - \widehat{\varsigma}(\psi)\, W_m\right) W_m$$

Step 11:Declare the set ψ for which

$$\chi^2_{score}\left(\psi\right) = nP_n[U_{adj}^{\dagger}\left(\psi, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}\,, \widehat{\varkappa}, \widehat{\zeta}\right)^T] \times \tag{3.17}$$
$$\hat{\Sigma}_{adj}\left(\psi, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}\,, \widehat{\varkappa}, \widehat{\zeta}\right)^{-1} P_n[U_{adj}^{\dagger}\left(\psi, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}\,, \widehat{\varkappa}, \widehat{\zeta}\right)]$$

is less than the α upper quantile of a χ^2 distribution on the dimension of ψ degrees of freedom to be a locally efficient large sample $1-\alpha$ score confidence

interval for $\psi^\dagger$ where

$$\hat{\Sigma}_{adj}\left(\psi, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right) = P_n\left[U^\dagger_{adj}\left(\psi, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right)^{\otimes 2}\right].$$

Step 12: Solve $P_n\left[U^\dagger_{adj}\left(\psi, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right)\right] = 0$ to obtain the estimator $\widehat{\psi}\left(\widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right)$.

Estimate $J^{\overline{d}} = E\left[Y^{\overline{d}}\right]$ by $P_n\left[H_0\left(\widehat{\psi}\left(\widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right)\right)\right]$.

Step13: Declare the set ψ for which $\chi^2_{wald}(\psi) =$

$$n\left(\widehat{\psi}\left(\widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right)^T - \psi^T\right) \times$$
$$P_n\left[\partial U^\dagger_{adj}\left(\widehat{\psi}, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right)/\partial\psi^T\right] \times$$
$$\hat{\Sigma}_{adj}\left(\widehat{\psi}, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right)^{-1} P_n\left[\partial U^\dagger_{adj}\left(\widehat{\psi}, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right)/\partial\psi\right] \times$$
$$\left(\widehat{\psi}\left(\widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right) - \psi\right) \tag{3.18}$$

is less than the α upper quantile of a χ^2 distribution on the dimension of ψ degrees of freedom to be a locally efficient large sample $1-\alpha$ Wald confidence interval for $\psi^\dagger$. Here $\widehat{\psi} = \widehat{\psi}\left(\widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right)$

Remark: If inference is to be based on Steps 12 and 13 then we can use Newton-Raphson to obtain the solution in Step 12 and we only need to run the algorithm for the values of ψ required by the Newton-Raphson updates. If Step 11 is to be used for inference then we might use a finite search procedure and run the algorithm for the values of ψ on a finite lattice. Other computational approaches are also possible. It follows from Theorems 3.3 and 3.4 above that the estimator $\widehat{\psi}\left(\widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right)$ is, under the union model, RAL and locally efficient at the submodel in which (3.11) holds and the model $p_m\left[A_m \mid \overline{L}_m, \overline{A}_{m-1}; \alpha\right]$, the model $\varsigma^T W_m$ for $E\left[H_m\left(\psi^\dagger\right)|\overline{L}_m, \overline{A}_{m-1}\right]$, the model $exp\left(\varkappa B_m\right)$ for $var\left(H_m\left(\psi^\dagger\right)|\overline{L}_{m-1}, \overline{A}_m\right)$, and the model ζR_m for $E\left[\partial H_m\left(\psi^\dagger\right)/\partial\psi|\overline{L}_m, \overline{A}_m\right]$ are all correct. Analogously the estimator $P_n\left[H_0\left(\widehat{\psi}\left(\widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right)\right)\right]$ is a RAL estimator of $J^{\overline{d}} = E\left[Y^{\overline{d}}\right]$ in the union model that is efficient at the above submodel. We therefore refer to these estimators as locally semiparametric efficient in the union model at the above submodel. These estimators are also locally semiparametric efficient (at the same submodel) in the more restrictive models (a.1) and (a.2) which, in addition to the assumptions of the union model, respectively assume that $p_m\left[A_m \mid \overline{L}_m, \overline{A}_{m-1}\right]$ is known and the model $p_m\left[A_m \mid \overline{L}_m, \overline{A}_{m-1}; \alpha\right]$ is correct.

Remark: We henceforth refer to $\widehat{\psi}\left(\widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right)$ as a doubly-robust locally semiparametric efficient (dr-lse) estimator. It is doubly robust because it is a RAL estimator in the union model that assumes either a model

for $p_m\left[A_m \mid \overline{L}_m, \overline{A}_{m-1}\right]$ or a model for $E\left[H_m\left(\psi^\dagger\right)|\overline{L}_m, \overline{A}_{m-1}\right]$ be correct. It is locally semiparametric efficient because it is efficient in the union model at the submodel described above.

All the results obtained in the this and the previous section in regards to point and interval estimates of $\psi^\dagger$ (but not of $E\left[Y_{\overline{d}}\right]$) remain true when we replace $H_m^{\underline{d}_m, \underline{d}_m^*}(\psi)$ by the modified version

$$\begin{aligned} H_{\text{mod},m}^{\underline{d}_m, \underline{d}_m^*}(\psi) &= H_m^{\underline{d}_m, \underline{d}_m^*}(\psi) - \gamma^{\overline{d}^*}\left(\overline{L}_m, \overline{A}_{m-1,} d_m\left(\overline{L}_m, \overline{A}_{m-1}\right), \psi\right) \qquad (3.19) \\ &= Y - \gamma^{\overline{d}^*}\left(\overline{L}_m, \overline{A}_m, \psi\right) + \\ &\sum_{j=m+1}^{K}\left\{\gamma^{\overline{d}^*}\left(\overline{L}_j, \overline{A}_{j-1,} d_j\left(\overline{L}_j, \overline{A}_{j-1}\right), \psi\right) - \gamma^{\overline{d}^*}\left(\overline{L}_j, \overline{A}_j, \psi\right)\right\} \end{aligned}$$

to obtain modified estimating functions $U_{\text{mod}}^{\overline{d}, \overline{d}^*}$ and $U_{\text{mod}}^{\overline{d}, \overline{d}^*, \dagger}$.This follows from the fact that the difference between $H_{\text{mod},m}^{\underline{d}_m, \underline{d}_m^*}(\psi)$ and $H_m^{\underline{d}_m, \underline{d}_m^*}(\psi)$ is fixed when we condition on $\overline{L}_m, \overline{A}_{m-1}$. Futher the efficient choice s_{eff} is exactly the same function of the data whether we use $H_{\text{mod},m}^{\underline{d}_m, \underline{d}_m^*}(\psi)$ or $H_m^{\underline{d}_m, \underline{d}_m^*}(\psi)$. The usefulness of this modification will become clear in Section 4.1.1 below.

4 Optimal drSNMM

4.1 Estimation

We now turn attention to estimation of the optimal treatment regime when, as in a sequential RCT , sequential randomization (2.5) holds and $p_m\left[A_m \mid \overline{L}_m, \overline{A}_{m-1}\right]$ is known. We first characterize the optimal regime. Define

$$d_{op,K}\left(\overline{L}_K, \overline{A}_{K-1}\right) = \arg\max_{a_K \in \mathcal{A}_K} E\left[Y_{\overline{A}_{K-1}, a_K} | \overline{L}_K, \overline{A}_{K-1}\right], \qquad (4.1)$$

so $d_{op,K}\left(\overline{L}_K, \overline{A}_{K-1}\right)$ is the optimal treatment at time t_K for someone with observed history $\left(\overline{L}_K, \overline{A}_{K-1}\right)$. If there is more than one value of $a_K \in A_K$ at which $E\left[Y_{\overline{A}_{K-1}, a_K} | \overline{L}_K, \overline{A}_{K-1}\right]$ is maximized, $d_{op,K}\left(\overline{L}_K, \overline{A}_{K-1}\right)$ can be arbitrarily chosen to be any one of the maximizers.

For $m = K-1, ...0$, recursively define

$$d_{op,m}\left(\overline{L}_m, \overline{A}_{m-1}\right) = \arg\max_{a_m \in \mathcal{A}_m} E\left[Y_{\overline{A}_{m-1}, a_m, \underline{d}_{op,m+1}} | \overline{L}_m, \overline{A}_{m-1}\right] \qquad (4.2)$$

where $\underline{d}_{op,m+1} = (d_{op,m+1}, ..., d_{op,K})$ with ties again broken arbitrarily. Thus $d_{op,m}\left(\overline{L}_m, \overline{A}_{m-1}\right)$ is the optimal treatment at time t_m for someone with observed history $\left(\overline{L}_m, \overline{A}_{m-1}\right)$ who is planning to follow regime $\underline{d}_{op,m+1}$ beginning at time t_{m+1}. The following trivial Lemma follows directly from the definitions.

Lemma 4.1:
$d_{op,m} = d_{op,m}\left(\overline{L}_m, \overline{A}_{m-1}\right) = \arg\max_{a_m \in \mathcal{A}_m} \gamma^{\overline{d}_{op}, \overline{d}^*}\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right)$.
We then have the following which is a special case of Theorem 7.4 below.

Theorem 4.1: Under the sequential randomization assumption (2.5), the regime $\underline{d}_{op,m}$ maximizes
$E\left[Y_{\overline{A}_{m-1}, \underline{d}_m} | \overline{L}_m, \overline{A}_{m-1}\right]$ over all regimes $\underline{d}_m \in \underline{\mathcal{D}}_m$ where $\underline{\mathcal{D}}_m$ is the set of all feasible regimes beginning at time t_m. Further $\overline{d}_{op} = \underline{d}_{op,0}$ maximizes $E\left[Y_{\overline{d}}\right]$ over all $\overline{d} \in \overline{\mathcal{D}}$. In addition Theorems 3.1 and 3.2 still hold with $\overline{d}_{op}$ substituted for $\overline{d}$ and with $H_m^{\underline{d}_m}\left(\gamma^{\overline{d}^*}\right) = H_m^{\underline{d}_{op,m}}\left(\gamma^{\overline{d}^*}\right)$ redefined to be

$$H_m^{\underline{d}_{op,m}}\left(\gamma^{\overline{d}^*}\right) = Y + \sum_{j=m}^{K} \left\{\gamma^{\overline{d}^*}\left(\overline{L}_j, \overline{A}_{j-1}, d_j^{**}\left(\overline{L}_j, \overline{A}_{j-1}\right)\right) - \gamma^{\overline{d}^*}\left(\overline{L}_j, \overline{A}_j\right)\right\}$$

with $d_m^{**}\left(\overline{L}_m, \overline{A}_{m-1}\right) = \arg\max_{a_m \in \mathcal{A}_m} \gamma^{\overline{d}^*}\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right)$.

Suppose for a given subject we are unable to intervene prior to time t_m although data on $\overline{L}_m, \overline{A}_{m-1}$ is available. Theorem 4.1 implies that $\underline{d}_{op,m}$ is the optimal treatment plan beginning at time t_m.

Remark: Note that without sequential randomization Theorem 4.1 will not hold, because then, for example, $\underline{d}_{op,K-1} = (d_{op,K-1}, d_{op,K})$ may not maximize $E\left[Y_{\overline{A}_{K-2}, \underline{d}_{K-1}} | \overline{L}_{K-1}, \overline{A}_{K-2}\right]$. To see this consider the subset of subjects with observed history $\left(\overline{l}_{K-1}, \overline{a}_{K-2}\right)$. Note, without sequential randomization, the mean of $Y_{\overline{a}_{K-2}, d_{op,K-1}\left(\overline{l}_{K-1}, \overline{a}_{K-2}\right), a_K}$ among the subgroup with observed history $\left(\overline{l}_K, \overline{a}_{K-2}, a_{K-1} = d_{op,K-1}\left(\overline{l}_{K-1}, \overline{a}_{K-2}\right)\right)$ need not equal the mean among the larger subgroup who would have had history $\left(\overline{l}_K, \overline{a}_{K-2}, a_{K-1} = d_{op,K-1}\left(\overline{l}_{K-1}, \overline{a}_{K-2}\right)\right)$ if, contrary to fact, the entire subset with observed history $\left(\overline{l}_{K-1}, \overline{a}_{K-2}\right)$ had followed following regime $d_{op,K-1}$ and thus received treatment $d_{op,K-1}\left(\overline{l}_{K-1}, \overline{a}_{K-2}\right)$ at t_{K-1}. In that case one cannot use the basic method of backward induction (i.e. dynamic programming) to solve the sequential decision problem. See section 7.2 below for a more complete discussion.

Definition: Given a regime $\overline{d}^*$, an optimal drSNMM model$\gamma^{\overline{d}_{op}, \overline{d}^*}\left(\overline{l}_m, \overline{a}_m\right) \equiv \gamma^{\overline{d}^*}\left(\overline{l}_m, \overline{a}_m; \psi\right)$ is a drSNMM model for $\gamma^{\overline{d}_{op}, \overline{d}^*}\left(\overline{l}_m, \overline{a}_m\right)$.

Definition: Define

$$d_{op,m}\left(\overline{L}_m, \overline{A}_{m-1}, \psi\right) = \arg \max_{a_m \in \mathcal{A}_m} \gamma^{\overline{d}^*}\left(\overline{L}_m, \overline{A}_{m-1}, a_m, \psi\right)$$

Note we shall model $\gamma^{\overline{d}_{op}, \overline{d}^*}\left(\overline{l}_m, \overline{a}_m\right)$ even though $\overline{d}_{op}$ is unknown. However the model itself determines $\overline{d}_{op}$ as a function of the parameter $\psi^\dagger$ via $d_{op,m}\left(\overline{l}_m, \overline{a}_{m-1}\right) = \arg\max_{a_m \in \mathcal{A}_m} \gamma^{\overline{d}^*}\left(\overline{l}_m, \overline{a}_m; \psi^\dagger\right)$. We shall see that under sequential randomization the parameter $\psi^\dagger$ can be estimated at a $n^{1/2}-$rate.

Without loss of generality, we shall usually take the regime $\overline{d}^*$ to be the zero-regime $\overline{0}$.

Example (1): Suppose $\gamma^{\overline{0}}\left(\overline{l}_m, \overline{a}_m, \psi\right) = a_m r\left(m, \overline{l}_m, \overline{a}_{m-1}, \psi_m\right)$ with a the common maximum and 0 the common minimum element in each $\mathcal{A}_m$, r a known function of ψ_m, and $\psi^T = (\psi_0, ..., \psi_K)$. Then $d_{op,m}\left(\overline{l}_m, \overline{a}_{m-1}, \psi\right) = a\left[I\left\{r\left(m, \overline{l}_m, \overline{a}_{m-1}, \psi_m\right) > 0\right\}\right]$. We consider the explicit choice $r\left(m, \overline{l}_m, \overline{a}_{m-1}, \psi_m\right) = (1, l_m, a_{m-1})\,\psi_m = \psi_{1m} + \psi_{2m} l_m + \psi_{3m} a_{m-1}$ in much of the ensuing discussion.

Example (2): Suppose now

$$\gamma^{\overline{0}}\left(\overline{l}_m, \overline{a}_m, \psi\right) = a_m\left[r_1\left(m, \overline{l}_m, \overline{a}_{m-1}, \psi_1\right)\right] + a_m^2\left[r_2\left(m, \overline{l}_m, \overline{a}_{m-1}, \psi_2\right)\right]$$

with each $\mathcal{A}_m$ being the interval $[0, a]$ on the real line and r_1 and r_2 are known functions of ψ_1 and ψ_2. Then

$$d_{op,m}\left(\overline{l}_m, \overline{a}_{m-1}, \psi\right) =$$
$$\left\{I\left[r_2\left(m, \overline{l}_m, \overline{a}_{m-1}, \psi_2\right) < 0\right]\right\} I\left(0 \leq g\left(m, \overline{l}_m, \overline{a}_{m-1}, \psi\right) \leq a\right) g\left(m, \overline{l}_m, \overline{a}_{m-1}, \psi\right)$$
$$+a\left\{I\left[r_2\left(m, \overline{l}_m, \overline{a}_{m-1}, \psi_2\right) \geq 0\right]\right\}\left\{I\left[\begin{array}{l} ar_2\left(m, \overline{l}_m, \overline{a}_{m-1}, \psi_2\right) \\ +r_1\left(m, \overline{l}_m, \overline{a}_{m-1}, \psi_1\right)\end{array}\right] > 0\right\}$$

where $g\left(m, \overline{l}_m, \overline{a}_{m-1}, \psi\right) = -r_1\left(m, \overline{l}_m, \overline{a}_{m-1}, \psi_1\right)/2r_2\left(m, \overline{l}_m, \overline{a}_{m-1}, \psi_2\right)$ and the first term corresponds to a maximum in the interior of $[0, a]$ and the second term corresponds to a maximum at the boundary point a. Here $I[B]$ is the indicator function that takes the value 1 if B is true and 0 otherwise.

Under sequential randomization (2.5) and $p_m\left(a_m | \overline{L}_m, \overline{A}_{m-1}\right)$ known ,we now consider the properties of approximate minimizers of $\left\{P_n\left[U\left(\psi, s\right)\right]\right\}^T B\left\{P_n\left[U\left(\psi, s\right)\right]\right\}$ with $U\left(\psi, s\right) = U^{\overline{d}_{op}, \overline{0}}\left(\psi, s\right)$ as defined in Equation (3.9), each function s_m having range the dimension of ψ, B a positive definite square matrix, and $H_m^{\overline{d}_{op}, \overline{0}}\left(\psi\right)$ defined by

$$H_m^{\overline{d}_{op}, \overline{0}}\left(\psi\right) = Y + \sum_{j=m}^{K}\left\{\gamma^{\overline{0}}\left(\overline{L}_j, \overline{A}_{j-1}, d_{op,j}\left(\overline{L}_j, \overline{A}_{j-1}, \psi\right), \psi\right) - \gamma^{\overline{0}}\left(\overline{L}_j, \overline{A}_j, \psi\right)\right\}.$$

Note that even when as in examples 1 and 2, $\gamma^{\overline{0}}\left(\overline{l}_m, \overline{a}_m, \psi\right)$ is smooth (or even linear) in ψ, $H^{\overline{d}_{op}, \overline{0}}\left(\psi\right)$ and $U^{\overline{d}_{op}, \overline{0}}\left(\psi, s\right)$ will not be everywhere differentiable in ψ (because ψ appears within indicator functions), although both remain continuous in ψ. Because $U^{\overline{d}_{op}, \overline{0}}\left(\psi, s\right)$ is a nonlinear function of $\psi, E\left[U^{\overline{d}_{op}, \overline{0}}\left(\psi, s\right)\right] = 0$ may have roots in addition to $\psi^\dagger$ in which case $\left\{P_n\left[U\left(\psi, s\right)\right]\right\}^T B\left\{P_n\left[U\left(\psi, s\right)\right]\right\}$ will have an inconsistent minimizer even asymptotically. As before if we increase the dimension of the s, then in sufficiently large samples the problem of multiple minimizers should disappear under correct specification, but not if the model $\gamma^{\overline{0}}\left(\overline{l}_m, \overline{a}_m, \psi\right)$ is misspecified.

Define $U^{\dagger \overline{d}_{op},\overline{0}}\left(\psi, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}\,, \widehat{\varkappa}, \widehat{\zeta}\right)$ as above but restricted to the set of ψ at which the required $\psi-$derivatives exist. Refer to this set as the admissible set. [In example 1, these $\psi-$derivatives will fail to exist only for values of ψ satisfying $(1, L_{m,i}, A_{m-1,i})\,\psi_m = 0$ for some study subject i as it is only at these values of ψ for which the contribution of subject i is nondifferentiable. Specifically in that case $d_{op,j}\left(\overline{L}_{j,i}, \overline{A}_{j-1,i}, \psi\right)$ and $H_{j,i}^{\overline{d}_{op},\overline{0}}(\psi), j \leq m$ are non differentiable for that subject.] A unique $n^{1/2}$ – *consistent* member $\widehat{\psi}\left(\widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}\,, \widehat{\varkappa}, \widehat{\zeta}\right)$ of the admissible set solving $P_n\left[U^{\dagger \overline{d}_{op},\overline{0}}\left(\psi, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}\,, \widehat{\varkappa}, \widehat{\zeta}\right)\right] = 0$ will exist with probability approaching one. It is a dr-lse estimator (i.e. it is RAL and lse not only in model (a.1) with $p_m\left(a_m|\overline{L}_m, \overline{A}_{m-1}\right)$ known but also in the larger $\overline{d}_{op}, \overline{0}$ "union" model of Theorem 3.4 in which $p_m\left(a_m|\overline{L}_m, \overline{A}_{m-1}\right)$ is not known), except under the exceptional laws F_O defined as those laws at which, for some $m, \arg\max_{a_m \in \mathcal{A}_m} \gamma^{\overline{0}}\left(\overline{L}_m, \overline{A}_{m-1}, a_m, \psi^\dagger\right)$ is not unique with positive probability. In example 1, these exceptional laws are the laws under which $(1, L_m, A_{m-1})\,\psi_m^\dagger = 0$ with positive probability. Even at these laws this estimator remains $n^{1/2}-$*consistent* for $\psi^\dagger$ although not CAN and thus not RAL. These exceptional laws are discussed in detail in Section 5.1 and Appendix 1.

A Closed-Form Estimator of an Optimal drSNMM

We now show that, if, as in our examples, $\gamma^{\overline{0}}\left(\overline{l}_m, \overline{a}_m, \psi\right)$ is linear in ψ, we can obtain a closed -form $n^{1/2}$-consistent estimator $\widetilde{\psi}$ under the union model of Theorem 3.4 based on solving a sequence of linear estimating functions. Then $P_n\left[H_0^{\overline{d}_{op},\overline{0}}\left(\widetilde{\psi}\right)\right]$ is a closed-form $n^{1/2}$-consistent estimator of $E\left[Y_{\overline{d}_{op}}\right]$ that does not require modelling $f\left[l_m \mid \overline{l}_{m-1}, \overline{a}_{m-1}\right]$. The importance of having a closed-form $n^{1/2}$-consistent estimator of $\psi^\dagger$ is that it avoids the problem with our earlier estimating functions. That is they could have multiple roots, an unknown one of which is consistent. Although the closed-form estimator $\widetilde{\psi}$ is not locally efficient, the one-step update

$$\widetilde{\psi}^{(1)} = \widetilde{\psi} + P_n\left[EIF\left(\widetilde{\psi}\right)\right] \text{ where}$$
$$EIF\left(\widetilde{\psi}\right) = P_n\left[\partial U_{adj}^\dagger\left(\widetilde{\psi}, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}\,, \widehat{\varkappa}, \widehat{\zeta}\right)/\partial\psi^T\right]^{-1} \times$$
$$P_n\left[U_{adj}^\dagger\left(\widetilde{\psi}, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}\,, \widehat{\varkappa}, \widehat{\zeta}\right)\right]$$

is locally efficient in the union model of Theorem 3.4. Note $EIF\left(\widetilde{\psi}\right)$ is an estimator of the efficient influence function for ψ in the union model at a particular submodel. The key to obtaining $\widetilde{\psi}$ is to use $U_{\text{mod}}^{\overline{d}_{op},\overline{0}}(\psi, s) =$

$\sum_{m=0}^{K} U^{\overline{d}_{op},\overline{0}}_{\text{mod},m}(\psi, s_m)$ defined just before Section 4 as a basis for estimation rather than $U^{\overline{d}_{op},\overline{0}}(\psi, s)$.

We first show how to construct $\widetilde{\psi}$ in the case in which $p_m\left(a_m|\overline{L}_m, \overline{A}_{m-1}\right)$ is known, $\psi = (\psi_0, ..., \psi_K)$ and $\gamma^{\overline{0}}\left(\overline{l}_m, \overline{a}_m, \psi\right) = \gamma^{\overline{0}}\left(\overline{l}_m, \overline{a}_m, \underline{\psi}_m\right)$ with $\gamma^{\overline{0}}\left(\overline{l}_m, \overline{a}_m, \underline{\psi}_m\right)$ linear and/or smooth in ψ_m but not necessarily in $\underline{\psi}_{m+1}$. It is crucial that at time t_m, our model $\gamma^{\overline{0}}\left(\overline{l}_m, \overline{a}_m, \psi\right)$ depends on ψ only through $\underline{\psi}_m$. We estimate the ψ_m recursively beginning with ψ_K so that at the step in which we estimate ψ_m, we have already estimated the vector $\underline{\psi}_{m+1}$. Specifically let $\widetilde{\psi}_K$ solve $P_n\left[U^{\overline{d}_{op},\overline{0}}_{\text{mod},K}\left(\psi_K, s_K\right)\right] = 0$ where $U^{\overline{d}_{op},\overline{0}}_{\text{mod},K}\left(\psi_K, s_K\right) = H^{\overline{d}_{op},\overline{0}}_{\text{mod},K}\left(\psi_K\right)\left\{S_K\left(A_K\right) - E\left[S_K\left(A_K\right) \mid \overline{A}_{K-1}, \overline{L}_K\right]\right\}$, and $H^{\overline{d}_{op},\overline{0}}_{\text{mod},K}\left(\psi_K\right) = Y - \gamma^{\overline{0}}\left(\overline{L}_K, \overline{A}_K, \psi_K\right)$. Then, for $m = K-1, ..., 0$, let $\widetilde{\psi}_m$ recursively solve, $P_n\left[U^{\overline{d}_{op},\overline{0}}_{\text{mod},m}\left(\psi_m, \underline{\widetilde{\psi}}_{m+1}, s_m\right)\right] = 0$, where

$$U^{\overline{d}_{op},\overline{0}}_{\text{mod},m}\left(\psi_m, \underline{\widetilde{\psi}}_{m+1}, s_m\right)$$
$$= H^{\underline{d}_{op,m},\overline{0}}_{\text{mod},m}\left(\psi_m, \underline{\widetilde{\psi}}_{m+1}\right)\left\{S_m\left(A_m\right) - E\left[S_m\left(A_m\right) \mid \overline{A}_{m-1}, \overline{L}_m\right]\right\},$$

$$H^{\underline{d}_{op,m},\overline{0}}_{\text{mod},m}\left(\psi_m, \underline{\widetilde{\psi}}_{m+1}\right) = Y - \gamma^{\overline{0}}\left(\overline{L}_m, \overline{A}_m, \psi_m, \underline{\widetilde{\psi}}_{m+1}\right)$$
$$+ \sum_{j=m+1}^{K}\left\{\gamma^{\overline{0}}\left(\overline{L}_j, \overline{A}_{j-1}, d_{op,j}\left(\overline{L}_j, \overline{A}_{j-1}, \underline{\widetilde{\psi}}_j\right), \underline{\widetilde{\psi}}_j\right) - \gamma^{\overline{0}}\left(\overline{L}_j, \overline{A}_j, \underline{\widetilde{\psi}}_j\right)\right\},$$

and $S_m\left(a_m\right) = s_m\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right)$ has range the dimension of ψ_m. Note $E\left[U^{\overline{d}_{op},\overline{0}}_{\text{mod},m}\left(\psi^{\dagger}_m, \underline{\psi}^{\dagger}_{m+1}, s_m\right)\right] = 0$. Further each of these estimating functions are smooth in ψ_m when $\gamma^{\overline{0}}\left(\overline{l}_m, \overline{a}_m, \underline{\psi}_m\right)$ is smooth in ψ_m and will be linear in ψ_m when $\gamma^{\overline{0}}\left(\overline{l}_m, \overline{a}_m, \underline{\psi}_m\right)$ is linear in ψ_m. In the linear case $\widetilde{\psi}(s) = \underline{\widetilde{\psi}}_0(s) = \underline{\widetilde{\psi}}_0 = \left(\widetilde{\psi}_0, .., \widetilde{\psi}_K\right)^T$ is unique and exists in closed form. In the following, we suppress the dependence on s and write $\underline{\widetilde{\psi}}_0(s) = \underline{\widetilde{\psi}}_0 = \left(\widetilde{\psi}_0, .., \widetilde{\psi}_K\right)^T$. With $\gamma^{\overline{0}}\left(\overline{l}_m, \overline{a}_m, \psi\right) = \left(1, l_m, a_{m-1}\right)\psi_m$ of Example 1,

$$\widetilde{\psi}_m = \widetilde{I_m}^{-1} \times P_n$$
$$\left[\begin{array}{c}\left\{S_m\left(A_m\right) - E[S_m\left(A_m\right)|\overline{A}_{m-1}, \overline{L}_m]\right\} \times \\ \left\{Y + \sum_{j=m+1}^{K}\left\{\gamma^{\overline{0}}\left(\overline{L}_j, \overline{A}_{j-1}, d_{op,j}\left(\overline{L}_j, \overline{A}_{j-1}, \widetilde{\psi}_j\right), \widetilde{\psi}_j\right) - \gamma^{\overline{0}}\left(\overline{L}_j, \overline{A}_j, \widetilde{\psi}_j\right)\right\}\right\}\end{array}\right]$$

whenever the three by three derivative matrix $\widetilde{I}_m =$
$P_n\left[\left\{S_m\left(A_m\right) - E[S_m\left(A_m\right)|\overline{A}_{m-1}, \overline{L}_m]\right\} A_m\left(1, L_m, A_{m-1}\right)\right]$ is invertible. Further, when $\gamma^{\overline{0}}\left(\overline{L}_m, \overline{A}_m, \psi_m, \underline{\widetilde{\psi}}_{m+1}\right)$ is linear in ψ_m , $\widetilde{\psi}_m\left(\widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}\ , \widehat{\varkappa}, \widehat{\zeta}\right)$ solving $P_n\left[U^{\dagger, \overline{d}_{op}, \overline{0}}_{\text{mod}, adj, m}\left(\psi_m, \underline{\widetilde{\psi}}_{m+1}, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}\ , \widehat{\varkappa}, \widehat{\zeta}\right)\right] = 0$ is a closed form dr-lse estimator (i.e. it is lse not only in model a.1 with $p_m\left(a_m|\overline{L}_m, \overline{A}_{m-1}\right)$ known but also in the larger $\overline{d}_{op}, \overline{0}$ "union" model of Theorem 3.4 in which $p_m\left(a_m|\overline{L}_m, \overline{A}_{m-1}\right)$ is not known) except at the exceptional laws mentioned above where it remains $n^{1/2}-$ consistent.

Remark on Estimation Under A Deterministic Design: When ψ_m is of sufficiently low dimension and, as in the sequential decision literature, the design parameters of the simulation generating the data under analysis are in the analyst's control, it is computationally feasible to accurately estimate the ψ_m without using the above estimation methods by recursively employing time-specific deterministic designs to obtain estimators $\widehat{\psi}_m, m = K, ..., 0$ as described in Section 2.2. For example at time t_m we can estimate ψ_{1m} by fixing $\left(\overline{L}_{m,i}, \overline{A}_{m-1,i}\right) = 0$ for all i, $i = 1, ..., n$, and simulating Y_i under the subsequent regime $\left\{a_{\max,m}, \underline{d}_{op,m+1}\left(\underline{\widehat{\psi}}_{m+1}\right)\right\}$for $i \leq n/2$ and $\left\{a_{\min,m}, \underline{d}_{op,m+1}\left(\underline{\widehat{\psi}}_{m+1}\right)\right\}$ for $i > n/2$, and taking $\widehat{\psi}_{1m}$ to be $\left\{a_{\max,m} - a_{\min,m}\right\}^{-1}$ times the difference between the averages of Y_i in the first and second half of the sample of size n. One can gain computational efficiency by reusing relevant simulations from times greater than t_m. Alternatively one could use a matched design as described in Remark 2.1 of Section 2.2.

Now suppose that, as in example 1, ψ_m is the same for each m so $\gamma^{\overline{0}}\left(\overline{l}_m, \overline{a}_m, \psi_m\right) = \left(1, l_m, a_{m-1}\right)\psi$. One can still artificially regard the common ψ as $K+1$ separate time-specific parameter vectors ψ_m. We can obtain the $K+1$ estimates $\widetilde{\psi}_m\left(s\right)$ and their (estimated) covariance matrix and combine them by inverse covariance weighting to give a final more efficient estimate $\widetilde{\psi}\left(s\right)$. Indeed if we let $\widetilde{\psi}_m\left(\widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}\ , \widehat{\varkappa}, \widehat{\zeta}\right)$ be as above, the inverse covariance weighted estimator will be a dr-lse of the time-independent parameter ψ, except at the exceptional laws. This efficiency result requires that $\widehat{s}_{eff}$ estimate the optimal choice s_{eff} for the model that does not impose equality of the different ψ_m. Alternatively the estimator $\widetilde{\psi}^{(1)}$ is a dr-lse, but here $\widehat{s}_{eff}$ must estimate s_{eff} for the model that does impose equality of the different ψ_m

Suppose the time $\Delta t = t_{m+1} - t_m$ between treatments is so short that very few subjects have A_{m+1} different from A_m. Then, $\widetilde{\psi}_m$,$m < K$, can have unacceptable small sample bias, even though it is asymptotically unbiased if Δt stays fixed as $n \to \infty$. To see why note that $\widetilde{\psi}_K$ will then be excessively variable because its variance is an increasing function of the inverse of the conditional variance of A_K given $\left(\overline{A}_{K-1}, \overline{L}_K\right)$ that increases to infinity as the

conditional variance of A_{K+1} decreases to zero. The excessive variability in $\widetilde{\psi}_K$ will then lead to severe small sample bias in $\widetilde{\psi}_{K-1}$ because $\widetilde{\psi}_K$ enters the estimating function $U^{\overline{d}_{op}}_{\text{mod},K-1}\left(\psi_{K-1}, \widetilde{\psi}_K, s^*_{K-1}\right)$ in a highly non-linear manner through it's inclusion within an indicator function. Thus even when all the $\psi_m = \psi$ are known to be equal and K is large so there is plenty of information available to estimate ψ, our locally efficient one-step estimator $\widetilde{\psi}^{(1)}$ can be very biased. Thus we have no choice but to solve the non-differentiable vector of estimating functions $U^{\overline{d}_{op}}_{\text{mod}}(\psi, s)$ non-sequentially, which can be computationally difficult and may suffer from the problem of multiple roots . In Section 8 we consider the extreme case of treatment innovations in continuous time, where Δt can be arbitrarily small.

Uniform Asymptotic Confidence Intervals For ψ and $\overline{d}_{op}$ By Inverting Tests

Although the estimators $\widetilde{\psi}$ and $\widehat{\psi}$ of the previous subsections will be $n^{1/2}$-consistent (i.e., $n^{1/2}\left(\widehat{\psi} - \psi^\dagger\right)$ and $n^{1/2}\left(\widetilde{\psi} - \psi^\dagger\right)$ are bounded in probability) at all laws allowed by our model , they will be neither asymptotically unbiased nor asymptotically normal under "exceptional" laws F_O satisfying, for some m, $\arg\max_{a_m \in \mathcal{A}_m} \gamma^{\overline{0}}\left(\overline{L}_m, \overline{A}_{m-1}, a_m, \psi^\dagger\right)$ is not unique with positive probability. Specifically, as discussed in Appendix 1.1, the limiting distributions of $n^{1/2}\left(\widehat{\psi} - \psi^\dagger\right)$ and $n^{1/2}\left(\widetilde{\psi} - \psi^\dagger\right)$ will be non-normal with a non zero mean (and thus will not be CAN) at such exceptional laws. Indeed, as shown in Appendix 1.1, $\psi^\dagger$ is not a regular parameter at these laws in the sense that it is not a differentiable function of a smooth finite dimensional parameterization of the law of O. It follows that, because of their unknown asymptotic bias, we cannot obtain uniformly asymptotically unbiased estimators of $\psi^\dagger$ and thus cannot construct valid (i.e. uniform over the entire model) asymptotic Wald-type confidence intervals for $\psi^\dagger$ centered on any estimator including $\widehat{\psi}$ or $\widetilde{\psi}$ of the previous subsections.

However one can obtain valid uniform confidence intervals for such non-regular parameters by inverting tests. For example, uniformly over all laws F_O allowed by the model with $p_m\left(a_m|\overline{L}_m, \overline{A}_{m-1}\right)$ known the sets $C(1-\alpha)$ and $C_{\text{mod}}(1-\alpha)$ of ψ for which $nP_n[U^{\overline{d}_{op},\overline{0}}(\psi, s)]^T \hat{\Sigma}(\psi, s)^{-1} P_n[U^{\overline{d}_{op},\overline{0}}(\psi, s)]$ and $nP_n[U^{\overline{d}_{op},\overline{0}}_{\text{mod}}(\psi, s)^T]\hat{\Sigma}_{\text{mod}}(\psi, s)^{-1} P_n[U^{\overline{d}_{op},\overline{0}}_{\text{mod}}(\psi, s)]$, respectively, are less than the α upper quantile, say, $\chi_{\alpha,\dim(\psi)}$, of a χ^2 distribution on the dimension of ψ degrees of freedom are uniform large sample $1-\alpha$ confidence interval for $\psi^{\dagger\cdot}$. [The set $C(1-\alpha)$ has no relation to the random variable C defined in Section 3.3.] Further under the $\overline{d}_{op}, \overline{0}$ union model of Theorem 3.4 or the more restrictive models (a.1) and (a.2), the interval $C_{op}(1-\alpha)$ based on $\chi^{2,\overline{d}_{op},\overline{0}}_{score}(\psi)$ of Eq (3.17) will be a uniform asymptotic confidence interval over all laws allowed by the model.

One can also obtain conservative uniform large sample (i.e. asymptotic) $1-\alpha$ confidence intervals for functions of $\psi^{\dagger}$, even discrete functions such as $\overline{d}_{op}$. The set

$$D_{op}(1-\alpha) = \tag{4.3}$$
$$\left\{\overline{d}_{op}(\psi); d_{op,m}(\psi) = \arg\max_{a_m \in \mathcal{A}_m} \gamma^{\overline{d}_{op}\overline{0}}\left(\overline{l}_m, \overline{a}_m; \psi\right),\ \psi \in C_{op}(1-\alpha)\right\}$$

is an example. That is, writing $\overline{d}_{op}\left(\psi^{\dagger}\right)$ as $\overline{d}_{op}(F_O)$ and $D_{op}(1-\alpha)$ as $D_{op,n}(1-\alpha)$ to emphasize the dependence on sample size n, by the definition of an conservative uniform asymptotic confidence interval in Appendix 1.1, we have that given any $\delta > 0$, there exists a $N(\delta)$ such that

$$\text{for all } n > N(\delta) \text{ and all sequences } F_{O,1}, F_{O,2}, ... \text{ in the union model}$$
$$(1-\alpha) - \Pr_{F_{O,n}}\left[\overline{d}_{op}(F_{O,n}) \in D_{op,n}(1-\alpha)\right] < \delta. \tag{4.4}$$

In contrast, by definition, a non-uniform conservative asymptotic $(1-\alpha)$ confidence interval $D_n^*(1-\alpha)$ satisfies: given any $\delta > 0$ and any F_O in the union model, there exists a $N(\delta, F_O)$ depending on F_O, such that

$$\text{for all } n > N(\delta, F_O) \tag{4.5}$$
$$(1-\alpha) - \Pr_{F_O}\left[\overline{d}_{op}(F_O) \in D_n^*(1-\alpha)\right] < \delta$$

Thus only the conservative uniform asymptotic confidence interval has the property that there exists some finite sample size $N(\delta)$ such that for $n > N(\delta)$, the interval $D_{op,n}(1-\alpha)$ contains the true optimal regime $\overline{d}_{op}$ with probability greater than $(1-\alpha-\delta)$. In contrast a conservative non-uniform asymptotic $(1-\alpha)$ confidence interval is not guaranteed to contain the true $\overline{d}_{op}$ with probability greater than $(1-\alpha-\delta)$ at any sample size. For this reason we are only interested in uniform asymptotic confidence intervals. Now because the volume of $C_{op}(1-\alpha)$ is asymptotically equivalent to the volume of the Wald intervals based on the dr-lse estimator. $\widehat{\psi}\left(\widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right)$ at the unexceptional laws, we doubt it will be possible to obtain an interval for $\overline{d}_{op}\left(\psi^{\dagger}\right)$ that is substantially narrower (say by a factor of more than 2 under any law in the union model) than $D_{op,n}(1-\alpha)$. However the confidence interval $D_{op,n}(1-\alpha)$ for the optimal regime $\overline{d}_{op}$ may at a given sample size n be too wide (i.e., contain too many candidate regimes $\overline{d}_{op}(\psi)$) to provide reliable optimal decisions. In that case we must make a choice among these $\overline{d}_{op}(\psi)$. We discuss this issue further below and in particular discuss a Bayesian solution based on informative priors for the unknown parameters.

It is useful to understand when $D_{op,n}(1-\alpha)$ may contain many candidate regimes $\overline{d}_{op}(\psi)$. In a slight abuse of notation we say the confidence interval $C_{op,n}(1-\alpha)$ contains a law F_O if $C_{op,n}(1-\alpha)$ contains a value ψ such that

$\gamma^{\overline{0}}\left(\overline{L}_m, \overline{A}_{m-1}, a_m, \psi\right)$ equals the blip function $\gamma^{\overline{d}_{op}\overline{0}}\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right)$ associated with F_O. At a fixed sample size n, even if the data were generated under a (unknown) nonexceptional law $F_{ne,O}$ and n is very large, the confidence interval $C_{op,n}(1-\alpha)$ may contain one or more exceptional laws (i.e a law $F_{e,O}$ under which $\arg\max_{a_m \in \mathcal{A}_m} \gamma^{\overline{d}_{op}\overline{0}}\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right)$ is not unique with positive probability). It then follows that the interval $C_{op,n}(1-\alpha)$ will (generally) also contain another non-exceptional law $F^*_{ne,O}$ such that $\overline{d}_{op}(F_{ne,O})$ differs from $\overline{d}_{op}\left(F^*_{ne,O}\right)$, *as the exceptional laws form the boundary separating non-exceptional laws corresponding to different optimal treatment regimes.*

Thus $C_{op,n}(1-\alpha)$ may often contain both $F_{ne,O}$ and $F^*_{ne,O}$ and thus $D_{op,n}(1-\alpha)$ will often contain both $\overline{d}_{op}(F_{ne,O})$ and $\overline{d}_{op}\left(F^*_{ne,O}\right)$. Because both $F_{ne,O}$ and $F^*_{ne,O}$ are non-exceptional laws, we know that under $F_{ne,O}$, the maximal expected utility $E[Y_{\overline{d}_{op}(F_{ne,O})}]$ is strictly greater than $E[Y_{\overline{d}_{op}(F^*_{ne,O})}]$ while under $F^*_{ne,O}$, $E[Y_{\overline{d}_{op}(F^*_{ne,O})}]$ is strictly greater than $E[Y_{\overline{d}_{op}(F_{ne,O})}]$. Now since, without further information we remain uncertain whether $F_{ne,O}$ or $F^*_{ne,O}$ or neither generated the data, we do not know whether to choose $\overline{d}_{op}(F_{ne,O})$ or $\overline{d}_{op}\left(F^*_{ne,O}\right)$ or neither. Such further information may be available as informative priors for unknown parameters.

Note that because (i) the data were generated under $F_{ne,O}$ and (ii) $\overline{d}_{op}(F_{ne,O})$ differs from $\overline{d}_{op}\left(F^*_{ne,O}\right)$, it is true that our non-uniform asymptotics guarantees that there exists some sample size $\widetilde{n} = \widetilde{n}\left(F_{ne,O}, F^*_{ne,O}\right), \widetilde{n} > n$, such that $D_{op,n}(1-\alpha)$ will exclude $\overline{d}_{op}\left(F^*_{ne,O}\right)$ with very high probability. But of course this fact is of no use when the actual sample size is n. Thus, we see that a "uniform asymptotics" in contrast to a "non-uniform asymptotics" correctly recognizes that even when data is generated under an unknown non-exceptional law $F_{ne,O}$, at (even a large) fixed sample size n, we often cannot rule out the hypothesis that the data were generated either under an exceptional law $F_{e,O}$ or under a non-exceptional law $F^*_{ne,O}$ with $\overline{d}_{op}\left(F^*_{ne,O}\right)$ different from $\overline{d}_{op}(F_{ne,O})$.

5 Locally Efficient Optimal Treatment Decisions For Individual Patients

In this section we will restrict attention to dichotomous $(0,1)$ treatments, so that any model for $\gamma^{\overline{d}_{op},\overline{0}}\left(\overline{l}_m, \overline{a}_m\right)$ takes the form $\gamma^{\overline{0}}\left(\overline{l}_m, \overline{a}_m, \psi\right) = a_m r\left(m, \overline{l}_m, \overline{a}_{m-1}, \psi\right)$ of example 1 of Section 4.1 with $a_m \in \{0,1\}$. Thus $d_{op,m}\left(\overline{l}_m, \overline{a}_{m-1}, \psi\right) = I\left\{r\left(m, \overline{l}_m, \overline{a}_{m-1}, \psi\right) > 0\right\}$. Suppose, after the study is completed, a new patient appears in his physicians office at time t_m since diagnosis with past data $\left(\overline{l}_m, \overline{a}_{m-1}\right)$. One might suppose that for this patient on day t_m, the parameter of immediate interest is $\psi_{int} = r\left(m, \overline{l}_m, \overline{a}_{m-1}, \psi\right)$, as the patient should be treated if $\psi^{\dagger}_{int} = r\left(m, \overline{l}_m, \overline{a}_{m-1}, \psi^{\dagger}\right)$ is positive and not treated if negative. [Here 'int'

abbreviates 'interest' and not 'intervention']. If this supposition were correct we might wish to obtain locally efficient tests of the single null hypothesis $H_0 : \psi^{\dagger}_{int} \leq 0$ versus the alternative $H_1 : \psi^{\dagger}_{int} > 0$ without any consideration given to simultaneous inference or the multiple testing problem that results from testing similar hypotheses at many times and for many patients with different $\left(\overline{l}_m, \overline{a}_{m-1}\right)$ histories. However the sign of $\psi^{\dagger}_{int}$ only determines the optimal treatment decision on day t_m for this patient if in fact the patient follows the optimal regime from t_{m+1} onwards. But because of uncertainty we cannot be assured that the subject will follow the optimal regime subsequent to time t_m unless $m = K$, as t_K is the final treatment time (even if we assume, as we shall do for the time being, that all uncertainty is attributable to sampling variability and not also to model mispecification, to unmeasured confounding, or to lack of compliance.) Nonetheless we shall study the interesting mathematical issue of how to obtain locally efficient tests of the single null hypothesis $H_0 : \psi^{\dagger}_{int} \leq 0$ versus the alternative $H_1 : \psi^{\dagger}_{int} > 0$, in the prescence of exceptional laws, as the results will be useful later.

5.1 A Frequentist Approach to Locally Efficient Tests of $H_0 : \psi^{\dagger}_{int} \leq 0$

It follows from Equation (3.19) that we can rewrite $H^{\overline{d}_{op},\overline{0}}_{\text{mod},m}(\psi) =$

$$Y - A_m r\left(m, \overline{L}_m, \overline{A}_{m-1}, \psi\right) + \sum_{j=m+1}^{K} \left\{ I\left\{ r\left(j, \overline{L}_j, \overline{A}_{j-1}, \psi_j\right) > 0 \right\} - A_j \right\} r\left(j, \overline{L}_j, \overline{A}_{j-1}, \psi\right),$$

and from Equation (3.9) and (3.10) that we can rewrite $U^{\dagger,\overline{d}_{op},\overline{0}}_{\text{mod}}\left(\psi, s_{eff}, c^s_{eff}\right) =$

$$\sum_{m=0}^{K} \left\{ H^{\overline{d}_{op},\overline{0}}_{\text{mod},m}(\psi) - E\left[H^{\overline{d}_{op},\overline{0}}_{\text{mod},m}(\psi) \,|\overline{A}_{m-1}, \overline{L}_m \right] \right\} \times \left\{ S_{eff,m}(1,\psi) - S_{eff,m}(0,\psi) \right\} \left\{ A_m - E\left[A_m \mid \overline{A}_{m-1}, \overline{L}_m \right] \right\}$$

In the remainder of this section we suppress the $\overline{d}_{op}, \overline{0}$ superscript and assume $r\left(m, \overline{L}_m, \overline{A}_{m-1}, \psi\right)$ is twice continuously differentiable with respect to ψ with probability one. Differentiability of $U^{\dagger}_{\text{mod}}\left(\psi^{\dagger}, s_{eff}, c^s_{eff}\right)$ at $\psi^{\dagger}$ with probability one will turn out to be of crucial concern. We now show that $U^{\dagger}_{\text{mod}}\left(\psi^{\dagger}, s_{eff}, c^s_{eff}\right)$ and (indeed $U^{\dagger}_{\text{mod}}\left(\psi^{\dagger}, s, c^s\right)$ for any s) is differentiable at $\psi^{\dagger}$ w.p.1 except for laws F_O under which for some (j, m) ,with non-zero probability all of the following hold, (i) $r\left(j, \overline{L}_j, \overline{A}_{j-1}, \psi^{\dagger}\right) = 0$ (ii) $m < j$ and (iii) $B_{mj}\left(\psi^{\dagger}\right) = 1$. Here $B_{mj}\left(\psi^{\dagger}\right)$ is the indicator of the event that $r\left(j, \overline{L}_j, \overline{A}_{j-1}, \psi\right)$ depends on at least one component of $\left(\underline{A}_m, \underline{L}_{m+1}\right)$ for

some ψ in every open neighborhood of $\psi^\dagger$. To obtain this result note (i) $H_{\text{mod},m}(\psi) - E\left[H_{\text{mod},m}(\psi)\,|\overline{A}_{m-1}, \overline{L}_m\right]$ will be a function of $I\left\{r\left(j, \overline{L}_j, \overline{A}_{j-1}, \psi\right) > 0\right\}$ unless $B_{mj}(\psi) = 0$ for all $j > m$ and (ii) if, for all ψ in a neighborhood of $\psi^\dagger, B_{mj}(\psi) = 0$ w.p1. for all $(m, j), j > m$, then $U^\dagger_{\text{mod}}\left(\psi, s_{eff}, c^s_{eff}\right)$ does not have ψ occurring within an indicator function and is thus smooth at $\psi^\dagger$. A law F_O satisfying (i)-(iii) is said to be an exceptional law. In general any law F_O under which the g-null hypothesis $\psi^\dagger = 0$ holds will be an exceptional law, except if the model $r\left(m, \overline{l}_m, \overline{a}_{m-1}, \psi\right)$ rules out all treatment interactions apriori by specifying that $r\left(m, \overline{l}_m, \overline{a}_{m-1}, \psi\right) = r(m, \psi)$ does not depend on $\overline{l}_m, \overline{a}_{m-1}$ for all ψ, in which case $U^\dagger_{\text{mod}}\left(\psi, s_{eff}, c^s_{eff}\right)$ is smooth in ψ.

Remark: The non-doubly robust statistic $U_{\text{mod}}(\psi, s)$, in contrast to $U^\dagger_{\text{mod}}(\psi, s, c^s)$, will be nondifferentiable at $\psi^\dagger$ with positive probability when, for some m, $r\left(m, \overline{L}_m, \overline{A}_{m-1}, \psi^\dagger\right) = 0$ with non-zero probability. This is equivalent to saying that $\arg_{a_m \in \mathcal{A}_m} \gamma^{\overline{0}}\left(\overline{L}_m, \overline{A}_{m-1}, a_m, \psi^\dagger\right)$ is not unique with positive probability. Thus, when earlier, we were considering tests and estimators based on $U_{\text{mod}}(\psi, s)$, we referred to any law F_O satisfying only condition (i) as an exceptional law.

We now reparametrize so that the parameter ψ_{int} is substituted for some component of ψ on which $r\left(m, \overline{l}_m, \overline{a}_{m-1}, \psi\right)$ functionally depends, so now, in a slight abuse of notation,

$\psi = \left(\psi_{int}, \psi^T_{-int}\right)^T$ where ψ_{-int} are the unchanged components of the original ψ.

Write $U^\dagger_{\text{mod}\,adj}\left(\psi, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right)$ as

$$\left(U^\dagger_{\text{mod}\,adj,int}\left(\psi, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right), U^\dagger_{\text{mod}\,adj,-int}\left(\psi, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right)^T\right)^T$$

Define $S_{int,eff}\left(\psi, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right) =$

$$\begin{aligned}
&U^\dagger_{\text{mod}\,adj,int}\left(\psi, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right) - \\
&E[\partial U^\dagger_{\text{mod}\,adj,int}\left(\psi, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right)/\partial \psi^T_{-int}] \times \\
&E[\partial U^\dagger_{\text{mod}\,adj,-int}\left(\psi, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right)/\partial \psi^T_{-int}]^{-1} \times \\
&U^\dagger_{\text{mod}\,adj,-int}\left(\psi, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right)
\end{aligned}$$

The (locally) efficient score $S_{int,eff}\left(\psi^\dagger, \alpha^*, \varsigma^*, \varkappa^*, \zeta^*\right)$ for ψ_{int} with the other components of ψ_{-int} treated as nuisance parameters is the probability limit of $S_{int,eff}\left(\psi^\dagger, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right)$ in the $\overline{d}_{op}, \overline{0}$ union model of Theorem 3.4, provided F_O is not an exceptional law. The locally efficient score for ψ_{int} is undefined

at exceptional laws, because ψ_{int} is a non-regular parameter under these laws. See appendix A1.1.

We now construct a conservative $1-\alpha$ uniform asymptotic confidence interval for $\psi_{int}^{\dagger}$ that quite generally will be narrower than the so-called $1-\alpha$ projection interval $\{\psi_{int}; \psi_{int} = \psi_{int}(\psi) \text{ for some } \psi \in C_{op}(1-\alpha)\}$ based on the interval $C_{op}(1-\alpha)$ for the vector ψ. In the following $\varnothing$ denotes the null set and $C_{-int,}(1-\varepsilon, \psi_{int})$ is a uniform large sample $1-\varepsilon$ joint confidence interval for ψ_{-int} when assuming that $\psi_{int}^{\dagger}$ is known apriori to be equal to ψ_{int}. For example, $C_{-int,}(1-\varepsilon, \psi_{int})$ could be the interval $C_{\text{mod}}(1-\varepsilon)$ of the last subsection given by $U_{\text{mod}}^{\overline{d}_{op},\overline{0}}(\psi, s)^T \hat{\Sigma}_{\text{mod}}(\psi, s)^{-1} U_{\text{mod}}^{\overline{d}_{op},\overline{0}}(\psi, s)$ is less than the ε upper quantile of a χ^2 distribtion on $\dim(\psi_{-int}) = \dim(\psi) - 1$ d.f. except that one component (say the last) of $U_{\text{mod}}^{\overline{d}_{op},\overline{0}}(\psi, s)$ has been eliminated so $U_{\text{mod}}^{\overline{d}_{op},\overline{0}}(\psi, s)$ is now of $\dim(\psi) - 1$ to reflect the fact that ψ_{int} is regarded as known.

Theorem 5.1: . Let the interval $C_{-int,}(1-\varepsilon, \psi_{int})$ be a uniform large sample $1-\varepsilon$ joint confidence interval for ψ_{-int}, when assuming that $\psi_{int}^{\dagger}$ is known apriori to be equal to ψ_{int}. Under regularity conditions sketched in the proof below, the region

$$\left\{ \begin{array}{c} \psi_{int}; C_{-int}(1-\varepsilon, \psi_{int}) \neq \varnothing \\ inf_{\psi_{-int} \in C_{-int,}(1-\varepsilon,\psi_{int})} \left| \frac{n^{1/2} P_n\left[S_{int,eff}\left((\psi_{int},\psi_{-int}),\widehat{\alpha},\widehat{\varsigma}\ ,\widehat{\varkappa},\widehat{\zeta}\right)\right]}{P_n\left[S_{int,eff}\left((\psi_{int},\psi_{-int}),\widehat{\alpha},\widehat{\varsigma}\ ,\widehat{\varkappa},\widehat{\zeta}\right)^2\right]^{1/2}} \right| < z_{\alpha/2} \end{array} \right\} \tag{5.1}$$

is a conservative $1-\alpha-\varepsilon$ uniform asymptotic confidence region for $\psi_{int}^{\dagger}$ over all laws allowed by the model. The interval's coverage and length are asymptotically equivalent to the asymptotic coverage of $(1-\alpha)$ and the length of the locally optimal interval $\left\{ \psi_{int}; \left| \frac{n^{1/2} P_n\left[S_{int,eff}\left(\left(\psi_{int},\psi_{-int}^{\dagger}\right),\alpha^*,\varsigma^*\ ,\varkappa^*,\zeta^*\right)\right]}{E\left[S_{int,eff}\left(\left(\psi_{int},\psi_{-int}^{\dagger}\right),\alpha^*,\varsigma^*\ ,\varkappa^*,\zeta^*\right)^2\right]^{1/2}} \right| < z_{\alpha/2} \right\}$ at all laws but the exceptional laws. Similarly for $z_{\alpha} > 1$, the test that rejects when

$$I\left\{ \begin{array}{c} C_{-int}(1-\varepsilon, \psi_{int}) \neq \varnothing \text{ or} \\ inf_{\psi_{-int} \in C_{-int,}(1-\varepsilon,\psi_{int})} \frac{n^{1/2} P_n\left[S_{int,eff}\left((0,\psi_{-int}),\widehat{\alpha},\widehat{\varsigma}\ ,\widehat{\varkappa},\widehat{\zeta}\right)\right]}{P_n\left[S_{int,eff}\left((0,\psi_{-int}),\widehat{\alpha},\widehat{\varsigma}\ ,\widehat{\varkappa},\widehat{\zeta}\right)^2\right]^{1/2}} > z_{\alpha} \end{array} \right\}$$

is a conservative uniform asymptotic $\alpha + \varepsilon - level$ test of $H_0 : \psi_{int}^{\dagger} \leq 0$ versus $H_1 : \psi_{int}^{\dagger} > 0$ whose asymptotic level will be α and whose asymptotic local power will be equal to that of the locally optimal test $I\left\{ \frac{n^{1/2} P_n\left[S_{int,eff}\left(\left(0,\psi_{-int}^{\dagger}\right),\alpha^*,\varsigma^*\ ,\varkappa^*,\zeta^*\right)\right]}{P_n\left[S_{int,eff}\left(\left(0,\psi_{-int}^{\dagger}\right),\alpha^*,\varsigma^*\ ,\varkappa^*,\zeta^*\right)^2\right]^{1/2}} > z_{\alpha} \right\}$ under all laws but the exceptional laws.

Remark: When a conservative uniform asymptotic confidence interval for a vector ψ is available, the method given in Theorem 5.1 is a quite general

method of obtaining a conservative uniform asymptotic confidence interval for a subvector ψ_{int} of ψ that will in general be narrower than the projection interval. The narrowness of the interval and the power of the associated test depend critically on the fact that, when $\psi_{int} = \psi^{\dagger}_{int}$, the statistic's numerator, here $n^{1/2}P_n\left[S_{int,eff}\left(\left(\psi_{int},\psi_{-int}\right),\widehat{\alpha},\widehat{\varsigma}\ ,\widehat{\varkappa},\widehat{\zeta}\right)\right]$, has an expectation that, for n sufficiently large, varies less (to a smaller order) as ψ_{-int} varies over a $1-\varepsilon$ confidence interval $C_{-int}\left(1-\varepsilon,\psi^{\dagger}_{int}\right)$ than as ψ_{int} varies around $\psi^{\dagger}_{int}$ over an interval of length $O\left(n^{-1/2}\right)$. This general approach would allow us to construct confidence intervals for ψ_{int} that can do better than projection intervals even when ψ_{-int} is infinite dimensional and/or ψ_{int} is only estimable at non square-root n rates.

Proof: At any law that is not an exceptional law, ψ_{int} and ψ_{-int} are regular parameters. At any such law we assume that, with probability one, $U^{\dagger}_{\mathrm{mod}\,adj}\left(\psi,\widehat{s}_{eff},c^{\widehat{s}_{eff}},\widehat{\alpha},\widehat{\varsigma}\ ,\widehat{\varkappa},\widehat{\zeta}\right)$ has bounded second derivatives with respect to ψ_{-int}. [Note this assumption cannot hold at an exceptional law since at these laws $U^{\dagger}_{\mathrm{mod}\,adj}\left(\psi,\widehat{s}_{eff},c^{\widehat{s}_{eff}},\widehat{\alpha},\widehat{\varsigma}\ ,\widehat{\varkappa},\widehat{\zeta}\right)$, although continuous, will not be even once differentiable.] Further, when $\psi_{int} = \psi^{\dagger}_{int}$, the value $\widehat{\psi}_{\mathrm{inf},-int}$ where the infimum is attained will be uniformly $n^{1/2}-$*consistent* for $\psi^{\dagger}_{-int}$ as all members of the set $C_{-int,}\left(1-\varepsilon,\psi_{int}\right)$ are uniformly $n^{1/2}-$*consistent* over all laws in the model. It then follows from the following lemma 5.1 and some standard limit arguments using contiguity theory that $\frac{n^{1/2}P_n\left[S_{int,eff}\left(\left(\psi_{int},\widehat{\psi}_{\mathrm{inf},-int}\right),\widehat{\alpha},\widehat{\varsigma}\ ,\widehat{\varkappa},\widehat{\zeta}\right)\right]}{P_n\left[S_{int,eff}\left(\left(\psi_{int},\widehat{\psi}_{\mathrm{inf},-int}\right),\widehat{\alpha},\widehat{\varsigma}\ ,\widehat{\varkappa},\widehat{\zeta}\right)^2\right]^{1/2}} = \frac{n^{1/2}P_n\left[S_{int,eff}\left(\left(\psi_{int},\psi^{\dagger}_{-int}\right),\alpha^*,\varsigma^*\ ,\varkappa^*,\zeta^*\right)\right]}{E\left[S_{int,eff}\left(\left(\psi_{int},\psi^{\dagger}_{-int}\right),\alpha^*,\varsigma^*\ ,\varkappa^*,\zeta^*\right)^2\right]^{1/2}}+$ $o_p\left(1\right)$ when $\psi_{int} = \psi^{\dagger}_{int} + kn^{-1/2}$ for any constant k.

Suppose now we are at an exceptional law and $\psi_{int} = \psi^{\dagger}_{int}$. Let $\widehat{\psi}_{\mathrm{inf},-int}$ be the minimizer of (5.1). We know with probability $1-\varepsilon$ that $C_{-int,}\left(1-\varepsilon,\psi_{int}\right)$ contains $\psi^{\dagger}_{-int}$. Thus with uniform probability $1-\varepsilon$,

$$\left|\frac{n^{1/2}P_n\left[S_{int,eff}\left(\left(\psi_{int},\widehat{\psi}_{\mathrm{inf},-int}\right),\widehat{\alpha},\widehat{\varsigma}\ ,\widehat{\varkappa},\widehat{\zeta}\right)\right]}{P_n\left[S_{int,eff}\left(\left(\psi_{int},\widehat{\psi}_{\mathrm{inf},-int}\right),\widehat{\alpha},\widehat{\varsigma}\ ,\widehat{\varkappa},\widehat{\zeta}\right)^{\times 2}\right]^{1/2}}\right| \leq \left|\frac{n^{1/2}P_n\left[S_{int,eff}\left(\left(\psi_{int},\psi^{\dagger}_{-int}\right),\widehat{\alpha},\widehat{\varsigma}\ ,\widehat{\varkappa},\widehat{\zeta}\right)\right]}{P_n\left[S_{int,eff}\left(\left(\psi_{int},\psi^{\dagger}_{-int}\right),\widehat{\alpha},\widehat{\varsigma}\ ,\widehat{\varkappa},\widehat{\zeta}\right)^{\times 2}\right]^{1/2}}\right|.$$

But $\left|\frac{n^{1/2}P_n\left[S_{int,eff}\left(\left(\psi_{int},\psi^{\dagger}_{-int}\right),\widehat{\alpha},\widehat{\varsigma}\ ,\widehat{\varkappa},\widehat{\zeta}\right)\right]}{P_n\left[S_{int,eff}\left(\left(\psi_{int},\psi^{\dagger}_{-int}\right),\widehat{\alpha},\widehat{\varsigma}\ ,\widehat{\varkappa},\widehat{\zeta}\right)^{\times 2}\right]^{1/2}}\right| < z_{\alpha/2}$ with uniform probability $1-\alpha$ as $n\to\infty$. Thus as $n\to\infty$, $\left|\frac{n^{1/2}P_n\left[S_{int,eff}\left(\left(\psi_{int},\widehat{\psi}_{\mathrm{inf},-int}\right),\widehat{\alpha},\widehat{\varsigma}\ ,\widehat{\varkappa},\widehat{\zeta}\right)\right]}{P_n\left[S_{int,eff}\left(\left(\psi_{int},\widehat{\psi}_{\mathrm{inf},-int}\right),\widehat{\alpha},\widehat{\varsigma}\ ,\widehat{\varkappa},\widehat{\zeta}\right)^{\times 2}\right]^{1/2}}\right| < z_{\alpha/2}$ with uniform probability at least $1-\varepsilon-\alpha$.

Lemma 5.1: Suppose $E_{\psi_1,\psi_2}\left[U_a\left(\psi_1,\psi_2\right)\right] = E_{\psi_1,\psi_2}\left[U_b\left(\psi_1,\psi_2\right)\right]$ and $\partial^2 U_a\left(\psi_1,\psi_2\right)/\partial^2\psi_2$ and $\partial^2 U_b\left(\psi_1,\psi_2\right)/\partial^2\psi_2$ are continuous and bounded with probability 1. Define

$$U_1(\psi_1,\psi_2)$$
$$= U_a(\psi_1,\psi_2) - P_n\left[\partial U_a(\psi_1,\psi_2)/\partial\psi_2\right]\left\{P_n\left[\partial U_b(\psi_1,\psi_2)/\partial\psi_2\right]\right\}^{-1} U_b(\psi_1,\psi_2)$$

Then under mild regularity conditions P_{ψ_1,ψ_2},$n^{1/2}P_n\left[U_1\left(\psi_1,\widehat{\psi}_2\right)\right]$
$= n^{1/2}P_n\left[U_1(\psi_1,\psi_2)\right] + o_p(1)$ whenever $n^{1/2}\left(\widehat{\psi}_2 - \psi_2\right)$ is uniformly bounded in probability.

Proof: see Appendix 2

Remark: Note that, even if $U_a(\psi_1,\psi_2)$ and $U_b(\psi_1,\psi_2)$ are non differentiable, Lemma 5.1 remains true under weak regularity conditions if the derivatives $m_a(\psi_1,\psi_2) = \partial E_{\psi_1,\psi_2}\left[U_a(\psi_1,\psi_2^*)\right]/\partial\psi_2^*|_{\psi_2^*=\psi_2}$ and $m_b(\psi_1,\psi_2) = \partial E_{\psi_1,\psi_2}\left[U_b(\psi_1,\psi_2^*)\right]/\partial\psi_2^*|_{\psi_2^*=\psi_2}$ exist and $U_1(\psi_1,\psi_2)$ is redefined to be

$$U_1(\psi_1,\psi_2)$$
$$= U_a(\psi_1,\psi_2) - \widetilde{m}_a(\psi_1,\psi_2)\left\{\widetilde{m}_b(\psi_1,\psi_2)\right\}^{-1} U_b(\psi_1,\psi_2)$$

where $\widetilde{m}_a(\psi_1,\psi_2)$ and $\widetilde{m}_b(\psi_1,\psi_2)$ are based on "numerical derivatives" of $P_n\left[U_a(\psi_1,\psi_2)\right]$ and $P_n\left[U_b(\psi_1,\psi_2)\right]$ with step sizes of $O\left(n^{-1/2}\right)$ whenever $U_a(\psi_1,\psi_2)$ and $U_b(\psi_1,\psi_2)$ are non differentiable. This still does not help at exceptional laws since, under exceptional laws, the derivative of $E\left[S_{int,eff}\left(\left(\psi_{int},\psi_{-int}^{\dagger}\right),\alpha^*,\varsigma^*,\varkappa^*,\zeta^*\right)\right]$ with respect to $\psi_{-int}^{\dagger}$ does not exist.

Remark 5.1: $\left\{\psi_{int};\left|\frac{n^{1/2}P_n\left[S_{int,eff}\left(\left(\psi_{int},\widehat{\psi}_{-int}\right),\widehat{\alpha},\widehat{\varsigma},\widehat{\varkappa},\widehat{\zeta}\right)\right]}{P_n\left[S_{int,eff}\left(\left(\psi_{int},\widehat{\psi}_{-int}\right),\widehat{\alpha},\widehat{\varsigma},\widehat{\varkappa},\widehat{\zeta}\right)^{\times 2}\right]^{1/2}}\right| < z_{\alpha/2}\right\}$ for $\widehat{\psi}_{-int}$ an arbitrary $n^{1/2}$-consistent estimator of $\psi_{-int}^{\dagger}$ need not be a 1-$\alpha-\varepsilon$ uniform asymptotic confidence interval for $\psi_{int}^{\dagger}$ for any given ε because, at the exceptional laws, we do not obtain guaranteed coverage of 1-$\alpha-\varepsilon$ unless we use the minimizer $\widehat{\psi}_{\inf,-int}$ of (5.1). Similiarly $I\left\{\frac{n^{1/2}P_n\left[S_{int,eff}\left(\left(0,\widehat{\psi}_{-int}\right),\widehat{\alpha},\widehat{\varsigma},\widehat{\varkappa},\widehat{\zeta}\right)\right]}{P_n\left[S_{int,eff}\left(\left(0,\widehat{\psi}_{-int}\right),\widehat{\alpha},\widehat{\varsigma},\widehat{\varkappa},\widehat{\zeta}\right)^{\times 2}\right]^{1/2}} > z_{\alpha}\right\}$ need not be a conservative asymptotic $\alpha+\varepsilon-level$ test under an exceptional law. Furthermore had we taken the infimum over all ψ_{-int} (rather than just over ψ_{-int} in $C_{-int,}(1-\varepsilon,\psi_{int})$), the resulting test

$$I\left\{inf_{\psi_{-int}}\frac{n^{1/2}P_n\left[S_{int,eff}\left(\left(0,\psi_{-int}\right),\widehat{\alpha},\widehat{\varsigma},\widehat{\varkappa},\widehat{\zeta}\right)\right]}{P_n\left[S_{int,eff}\left(\left(0,\psi_{-int}\right),\widehat{\alpha},\widehat{\varsigma},\widehat{\varkappa},\widehat{\zeta}\right)^{\times 2}\right]^{1/2}} > z_{\alpha}\right\}$$

with $z_{\alpha} > 1$, though conservative, may have power zero, as it is possible that *the* test statistic may equal zero with probability one, at both local and global alternatives to $\psi_{int.}^{\dagger} < 0$ whether the data were generated under an exceptional or nonexceptional law. To see why the power can be poor if we

take the infimum over all ψ_{-int}, consider the following simple example of a model with an exceptional law.

Example: Suppose we observe n iid copies of
$X \tilde{} N(\beta, 1), Y \tilde{} N(\psi - I(\beta > 0)\beta, 1)$, X and Y independent. Then ψ is in the role of ψ_{int}, β is in the role of ψ_{-int} and ψ is a regular parameter except at the exceptional laws that have $\beta = 0$. If $\beta \neq 0$, the efficient score based on one observation for ψ at $\psi = 0$ is $I(\beta \leq 0) Y + I(\beta > 0)(Y + X)$. Thus our .05 level test of $\psi > 0$ that takes the infimum over all β rejects only if $I\left\{min\left(n^{1/2} Y_{av}, n^{1/2}(Y_{av} + X_{av})/\sqrt{2}\right) > 1.64\right\} = 1$. Suppose the law generating the data has $\beta = -100n^{-1/2}$, $\psi = 10n^{-1/2}$ so Y has mean $10n^{-1/2}$, X has mean $-100n^{-1/2}$ and $Y + X$ has mean $-90n^{-1/2}$, and $X - Y$ has mean $-110n^{-1/2}$. Then the test will fail to accept the true alternative $\psi > 0$ with probability essentially one (i.e it has power zero). On the other hand the efficient score for β with $\psi = 0$ is $I(\beta \leq 0)(X - \beta) + I(\beta > 0)(X - Y - 2\beta)$ so $C_\beta(1 - \varepsilon, \psi = 0) =$

$$\left\{ \begin{array}{c} \beta; \beta \leq 0 \text{ and } X_{av} - n^{-1/2} z_\varepsilon < \beta < X_{av} + n^{-1/2} z_\varepsilon, \text{ or} \\ \beta > 0 \text{ and } \left[\frac{(X_{av} - Y_{av}) - n^{-1/2} z_\varepsilon \sqrt{2}}{2}\right] < \beta < \left[\frac{(X_{av} - Y_{av}) + n^{-1/2} z_\varepsilon \sqrt{2}}{2}\right] \end{array} \right\}.$$

Now, if we choose ε large enough that $z_\varepsilon < 100/\sqrt{2}$, then $C_\beta(1 - \varepsilon, \psi = 0)$ is, with probability essentially one, the set
$\left\{\beta; \beta \leq 0 \text{ and } X_{av} - n^{-1/2} z_\varepsilon < \beta < X_{av} + n^{-1/2} z_\varepsilon\right\}$. Thus the test of $\psi > 0$ that takes the infimum only over all $\beta \in C_\beta(1 - \varepsilon, \psi = 0)$ rejects whenever $I\left\{n^{1/2} Y_{av} > 1.64\right\} = 1$ and thus has power essentially 1. If we take the infimum over all β, the power of zero is not only against local alternatives or near exceptional laws as can be seen by noting that again we have power zero if $\beta = -100$, $\psi = 10$ or $\beta = -100, \psi = 10n^{-1/2}$.

We can also use this example to understand why we needed to add the condition $C_{-int}(1 - \varepsilon, \psi_{int}) \neq \varnothing$ to our interval and testing procedures. Suppose the true data generating process has $\beta = 100n^{-1/2}$ and $\psi = 300n^{-1/2}$. Then with probability essentially one, the interval $C_\beta(1 - \varepsilon, \psi = 0)$ will be empty for any reasonably small ε since X has mean $100n^{-1/2}$ and $X - Y$ has mean $100n^{-1/2} - 200n^{-1/2} = -100n^{-1/2}$. The emptiness of the confidence interval for β can be taken as evidence of model misspecification due to assuming, for the purpose of hypothesis testing, that ψ is known apriori to be equal to 0. Thus $\psi = 0$ should, as is appropriate, not be included in the confidence interval and also be rejected when doing two-sided testing. For a one sided test we may not wish to reject as we may not know which side of $\psi = 0$ we are on.

The above results leave open the question of what $\alpha - level$ or more precisely $\alpha + \varepsilon - level$ to choose to make the best treatment decision. We shall see there is no easy answer. Suppose we choose ε to be much smaller than α so we can ignore the distinction between α and $\alpha + \varepsilon$. The classical thoughtless frequentist tradition is to choose $\alpha = .05$, but there is really no justification for this convention. Further as patients are seen at many different times t_m

from diagnosis with different histories $\left(\overline{l}_m, \overline{a}_{m-1}\right)$, we will be testing the null hypothesis $\psi_{int} \leq 0$ about many time-history-specific parameters ψ_{int}. How should we adjust the $\alpha-level$ of our tests to account for the multiple hypotheses being tested? How should we account for the fact that with the exception of the last treatment occassion K, we cannot be certain we shall succeed in following the optimal regime in the future? Perhaps a Bayesian approach is best. We next explore such an approach.

5.2 Approximate Bayesian Inference:

To begin we will study the following two decision rules (i) and (ii) in which we treat a patient with past data $\left(\overline{l}_m, \overline{a}_{m-1}\right)$ on day t_m if (i) the posterior probability that $\psi_{int} = r\left(m, \overline{l}_m, \overline{a}_{m-1}, \psi\right)$ is positive exceeds 1/2 (i.e. the posterior median is positive) or (ii) the posterior mean of $\psi_{int} = r\left(m, \overline{l}_m, \overline{a}_{m-1}, \psi\right)$ is positive. In our complex high-dimensional model, exact Bayesian inference that respects the semiparametric nature of the problem is not computationally feasible nor, as we shall argue, would it be desirable even if computationally feasible. Therefore, we consider an approximate (i.e. asymptotic) Bayesian analysis. Our analysis will face 2 separate problems. First how do we obtain an approximate posterior for the model parameter $\psi^\dagger$ and thus for $r\left(m, \overline{l}_m, \overline{a}_{m-1}, \psi^\dagger\right)$. Our second problem will be to consider whether either of the two decision rules (i) and (ii) is the optimal Bayes decision rule and, if not, how might we approximate the optimal Bayes rule. The first problem we shall consider now. As for the second, we shall later show that neither decision rule may be the Bayes rule.

Let $\mathbf{t}$ be the $\dim(\psi)$ vector with each component equal to t. For any $\dim(\psi)$ random variable V, define the the $t-truncated$ version

$$V_t = VI\left[V^T V < dim\left(\psi\right) t^2\right] + \mathbf{t} I\left[V^T V > dim\left(\psi\right) t^2\right]$$

of V to the variable that takes the value $\mathbf{t}$ whenever the norm of V exceeds $dim\left(\psi\right)^{1/2} t$. In this section, our approach will be to pretend that rather than observing the data O_i, $i = 1, ..., n$, we *only* observed the $t-truncated$ version $\widehat{Z}_t\left(\cdot\right) = Z_t\left(\cdot, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right)$ of the multivariate stochastic process $\widehat{Z}\left(\cdot\right)$ indexed by ψ

where

$$\begin{aligned}
&\widehat{Z}\left(\cdot\right) = Z\left(\cdot, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right) \\
&= \widehat{\Sigma}^{-1/2}\left(\cdot, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right) n^{1/2} P_n\left[U^\dagger_{adj}\left(\cdot, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right)\right].
\end{aligned}$$

We now argue that for large n and for $t = t\left(n\right)$ going to infinity slowly enough with n, the distribution associated with the stochastic process $\widehat{Z}_t\left(\cdot\right)$ under a law F_O with parameter $\psi^\dagger = \psi^\dagger\left(F_O\right)$ depends on F_O only through $\psi^\dagger$ with a "density" proportional to that of the t-truncated version of a dim

(ψ) $MVN(0,I)$ random variable evaluated at the argument $\widehat{Z}_t\left(\psi^\dagger\right)$. That is $f\left(\widehat{Z}_t(\cdot)\,|F_O\right)$ is proportional to $\phi_{t-Normal}\left(\widehat{Z}_t\left(\psi^\dagger\right)\right)$ where $\phi_{t-Normal}(u)$ is the density of a $t-$truncated MVN$(0,1)$ random variable. Thus with p the dimension of $\psi^\dagger$ we can approximate the likelihood $f\left(\widehat{Z}_t(\cdot)\,|F_O\right)$with $(2\pi)^{-p/2}\exp\left(-\widehat{Z}_t\left(\psi^\dagger\right)^T\widehat{Z}_t\left(\psi^\dagger\right)/2\right)$ uniformly in F_O, whenever $\widehat{Z}_t\left(\psi^\dagger\right)^T\widehat{Z}_t\left(\psi^\dagger\right) < \dim(\psi)\,t^2$. To see why, write Z_t as $Z_{n,t(n)}$ and $\widehat{Z}_t$ as $\widehat{Z}_{n,t(n)}$. Then, the approximation follows from the fact that, under regularity conditions, it can be shown that (i) $\widehat{Z}_{n,t(n)}(\cdot)$ is uniformly asymptotically degenerate, i.e. given any $\epsilon>0$, there exists $N(\epsilon)$ such that for $n>N(\epsilon)$

$$\sup_{F_O} pr_{F_O}\left[\left|\frac{f\left(\widehat{Z}_{n,t(n)}(\cdot)\,|F_O\right)}{f\left(\widehat{Z}_{n,t(n)}(\psi^\dagger)\,|F_O\right)}-1\right|>\epsilon\right]<\epsilon$$

where $\psi^\dagger=\psi^\dagger(F_O)$, (ii) $\widehat{Z}_{n,t(n)}\left(\psi^\dagger\right)$ can be approximated by the t-truncated version $Z_{n,t(n)}\left(\psi^\dagger,F_O\right)$ of the limit random variable $Z_n\left(\psi^\dagger,F_O\right)$, where $Z_n(\psi,F_O)=Z(\psi)=$

$$\begin{aligned}&Z\left(\psi,s^*_{eff},c^{s^*_{eff}},\alpha^*,\varsigma^*,\varkappa^*,\zeta^*\right)\\&=\widehat{\Sigma}^{-1/2}\left(\psi,s^*_{eff},c^{s^*_{eff}},\alpha^*,\varsigma^*,\varkappa^*,\zeta^*\right)n^{+1/2}\times\\&P_n\left[U^\dagger_{\mathrm{mod}\,adj}\left(\psi,s^*_{eff},c^{s^*_{eff}},\alpha^*,\varsigma^*,\varkappa^*,\zeta^*\right)\right]\end{aligned}$$

with the * limits evaluated under F_O;that is, for any $\epsilon>0$, there exists $N^*(\epsilon)$ such that for $n>N^*(\epsilon)$

$$\sup_{F_O} pr_{F_O}\left\{\left|\widehat{Z}_{n,t(n)}\left(\psi^\dagger\right)-Z_{n,t(n)}\left(\psi^\dagger(F_O),F_O\right)\right|>\epsilon\right\}<\epsilon$$

and (iii) the limit variable $Z_{n,t(n)}\left(\psi^\dagger(F_O),F_O\right)$ is uniformly asymptotically t-truncated normal *i.e.*, given u, for any $\epsilon>0$, there exists $N^{**}(\epsilon)$ such that for $n>N^{**}(\epsilon)$

$$\sup_{F_O}\left|pr_{F_O}\left[Z_{n,t(n)}\left(\psi^\dagger(F_O),F_O\right)>u\right]-\Phi_{t(n),normal}(u)\right|<\epsilon$$

where $\phi_{t-Normal}(u)$ is the distribution function of a $t-$truncated MVN$(0,I)$ random variable. In each case the sup is over the F_O in the union model of Theorem 3.4. Thus we have a uniform large sample approximation to the likelihood function based on observing $Z_t\left(\cdot,\widehat{s}_{eff},c^{\widehat{s}_{eff}},\widehat{\alpha},\widehat{\varsigma},\widehat{\varkappa},\widehat{\zeta}\right)$.

We give a rough sketch of the proof. The key idea is to show that Theorem 3.4 implies that under the $\overline{d}_{op},\overline{0}$ union model, for $\left(\psi-\psi^\dagger\right)$ of $O\left(n^{-1/2}\right)$,

$$n^{+1/2}P_n\left[U^{\dagger}_{\text{mod }adj}\left(\psi,\widehat{s}_{eff},c^{\widehat{s}_{eff}},\widehat{\alpha},\widehat{\varsigma}\ ,\widehat{\varkappa},\widehat{\zeta}\right)\right] \tag{5.2}$$
$$=n^{+1/2}P_n\left[U^{\dagger}_{\text{mod }adj}\left(\psi^{\dagger},s^{*}_{eff},c^{s^{*}_{eff}},\alpha^{*},\varsigma^{*}\ ,\varkappa^{*},\zeta^{*}\right)\right]+\kappa\left(\psi,\psi^{\dagger},F_O\right)+o_P\left(1\right)$$

where $\psi^{\dagger}=\psi^{\dagger}\left(F_O\right)$, the $o_P\left(1\right)$ is uniform over the model, and $\kappa\left(\cdot,\cdot,\cdot\right)$ is a non-random function that equals 0 at $\psi=\psi^{\dagger}$, is everywhere continuous in ψ, and differentiable when F_O is a non-exceptional law. Note if it were not for exceptional laws, where $E\left[U^{\dagger}_{\text{mod }adj}\left(\psi,s^{*}_{eff},c^{s^{*}_{eff}},\alpha^{*},\varsigma^{*}\ ,\varkappa^{*},\zeta^{*}\right)\right]$ may be non-differentiable in ψ (although continuous), standard arguments could be used to show that Eq.(5.2) was true with $\kappa\left(\psi,\psi^{\dagger},F_O\right)$ equal to

$$\partial E\left[U^{\dagger}_{\text{mod }adj}\left(\psi,s^{*}_{eff},c^{s^{*}_{eff}},\alpha^{*},\varsigma^{*}\ ,\varkappa^{*},\zeta^{*}\right)\right]/\partial\psi^{T}_{|\psi=\psi^{\dagger}}\ n^{1/2}\left(\psi-\psi^{\dagger}\right)$$

Because it would be notationally complex and thereby obscure the essential idea, we do not provide a general proof of (5.2); rather in appendix $A1.2$, we provide, with essentially no loss of generality, an explicit proof for a simple paradigmatic example.

Remark 5.2: Under an exceptional law F_O, for all ψ (except $\psi^{\dagger}$) in a neighborhood of $\psi^{\dagger}$ $DER\left(\psi\right)=DER\left(\psi,F_O\right)$
$=\partial E\left[U^{\dagger}_{\text{mod }adj}\left(\psi,s^{*}_{eff},c^{s^{*}_{eff}},\alpha^{*},\varsigma^{*}\ ,\varkappa^{*},\zeta^{*}\right)\right]/\partial\psi^{T}$ exists and
$\kappa\left(\psi,\psi^{\dagger}\left(F_O\right),F_O\right)=DER\left(\psi\right)n^{1/2}\left(\psi-\psi^{\dagger}\right)$; however, $DER\left(\psi^{\dagger}\right)$ is undefined because for different sequences $\text{seq}_m=\{\psi_{mj};j=1,2,...\}$ all converging to
$\psi^{\dagger}$, $\{DER\left(\psi_{mj}\right);j=1,2,...\}$ may converge to different limits depending on the particular seq_m. Nonetheless since the set of limits is bounded, $\kappa\left(\psi,\psi^{\dagger}\left(F_O\right),F_O\right)$ converges to 0 as $\psi\rightarrow\psi^{\dagger}$. See Appendix A1.2 for an example. One can view the lack of existence of the derivative of $DER\left(\psi^{\dagger}\right)$ at $\psi^{\dagger}$ under an exceptional law as the reason that $\widehat{\psi}$ solving
$n^{-1/2}P_n\left[U^{\dagger}_{\text{mod }adj}\left(\psi,\widehat{s}_{eff},c^{\widehat{s}_{eff}},\widehat{\alpha},\widehat{\varsigma}\ ,\widehat{\varkappa},\widehat{\zeta}\right)\right]=0$ is asymptotically biased since, plugging $\widehat{\psi}$ in place of ψ in (5.2) with $\kappa\left(\psi,\psi^{\dagger},F_O\right)=DER\left(\psi\right)n^{1/2}\left(\psi-\psi^{\dagger}\right)$, we have

$$\begin{aligned}&n^{+1/2}P_n\left[U^{\dagger}_{\text{mod }adj}\left(\widehat{\psi},\widehat{s}_{eff},c^{\widehat{s}_{eff}},\widehat{\alpha},\widehat{\varsigma}\ ,\widehat{\varkappa},\widehat{\zeta}\right)\right] \\ &=n^{+1/2}P_n\left[U^{\dagger}_{\text{mod }adj}\left(\psi^{\dagger},s^{*}_{eff},c^{s^{*}_{eff}},\alpha^{*},\varsigma^{*}\ ,\varkappa^{*},\zeta^{*}\right)\right]+ \\ &DER\left(\widehat{\psi},F_O\right)n^{1/2}\left(\widehat{\psi}-\psi^{\dagger}\right)+o_P\left(1\right)\end{aligned} \tag{5.3}$$

so $n^{1/2}\left(\widehat{\psi}-\psi^{\dagger}\right)=$

$$-\left\{DER\left(\widehat{\psi}\right)\right\}^{-1}n^{-1/2}P_n\left[U^{\dagger}_{\text{mod }adj}\left(\psi^{\dagger},s^{*}_{eff},c^{s^{*}_{eff}},\alpha^{*},\varsigma^{*}\ ,\varkappa^{*},\zeta^{*}\right)\right]+o_P\left(1\right).$$

The bias results from the fact that $DER\left(\widehat{\psi}\right)$is not converging in probability but rather, due to the lack of continuity of the derivative $DER\left(\psi\right)$ at $\psi^{\dagger}$,

$DER\left(\widehat{\psi}\right)$ has variance $O\left(1\right)$. Eq (5.2) implies that, uniformly, for $\left(\psi-\psi^{\dagger}\right)$ in a ball of radius $O\left(n^{-1/2}\right)$,

$$\widehat{Z}\left(\psi\right)=Z\left(\psi^{\dagger}\right)+\kappa^{*}\left(\psi,\psi^{\dagger},F_{O}\right)+o_{P}\left(1\right) \tag{5.4}$$

where $\kappa^{*}\left(\psi,\psi^{\dagger},F_{O}\right)$ is a non-random function that is continuous in ψ, differentiable at non-exceptional laws, equal to 0 at $\psi=\psi^{\dagger}$, and, for $\psi\neq\psi^{\dagger}$, equal to

$$\kappa^{*}\left(\psi,\psi^{\dagger},F_{O}\right)=\left[\partial E_{F_{O}}\left[Z\left(\psi\right)\right]/\partial\psi\right]n^{1/2}\left(\psi-\psi^{\dagger}\right) \tag{5.5}$$

in a neighborhood of $\psi^{\dagger}$.

It follows from Eqs. (5.4) and (5.5) that if, as we assume, $n^{1/2}E\left[U_{\mathrm{mod}\,adj}^{\dagger}\left(\psi,s_{eff}^{*},c^{s_{eff}^{*}},\alpha^{*},\varsigma^{*}\ ,\varkappa^{*},\zeta^{*}\right)\right]$ is greater than $O\left(1\right)$ whenever $\left(\psi-\psi^{\dagger}\right)$ is greater than $O\left(n^{-1/2}\right)$, the stochastic process $\widehat{Z}_{t}\left(\cdot\right)$ (with t going to infinity slowly enough with n) is asymptotically degenerate under F_{O}, because (i) the local process $\left\{\widehat{Z}\left(\psi\right);\left(\psi-\psi^{\dagger}\right)\text{ in a ball of radius }O\left(n^{-1/2}\right)\right\}$ is, to order $o_{p}\left(1\right)$, a deterministic function of $Z\left(\psi^{\dagger}\right)$ and (ii) for $\left(\psi-\psi^{\dagger}\right)$ greater than $O\left(n^{-1/2}\right)$, $\widehat{Z}\left(\psi\right)$ is greater than $O\left(1\right)$ so $\widehat{Z}_{t}\left(\psi\right)=\mathbf{t}$ with probability going to one. Finally, the limiting distribution of $Z\left(\psi^{\dagger}\right)$ under F_{O} is $MVN(0,I)$.

We can use the approximation $\left(2\pi\right)^{-p/2}\exp\left(-\widehat{Z}_{t}\left(\psi^{\dagger}\right)^{T}\widehat{Z}_{t}\left(\psi^{\dagger}\right)/2\right)$ to the likelihood $f\left(Z_{t}\left(\cdot,\widehat{s}_{eff},c^{\widehat{s}_{eff}},\widehat{\alpha},\widehat{\varsigma}\ ,\widehat{\varkappa},\widehat{\zeta}\right)|F_{O}\right)$ to compute an asymptotic approximation

$$\pi_{post}\left(\psi^{\dagger}\right)=\frac{I\left\{\psi^{\dagger};\left|\left|\widehat{Z}_{t}\left(\psi^{\dagger}\right)\right|\right|<t\dim\left(\psi^{\dagger}\right)^{1/2}\right\}\exp\left(-\widehat{Z}_{t}\left(\psi^{\dagger}\right)^{T}\widehat{Z}_{t}\left(\psi^{\dagger}\right)/2\right)\pi\left(\psi^{\dagger}\right)}{\int_{\left\{\psi^{\dagger};\left|\left|\widehat{Z}_{t}\left(\psi^{\dagger}\right)\right|\right|<t\dim\left(\psi^{\dagger}\right)^{1/2}\right\}}\exp\left(-\widehat{Z}_{t}\left(\psi^{\dagger}\right)^{T}\widehat{Z}_{t}\left(\psi^{\dagger}\right)/2\right)\pi\left(\psi^{\dagger}\right)d\psi^{\dagger}}$$

to the posterior $\pi\left(\psi^{\dagger}|Z_{t}\left(\cdot,\widehat{s}_{eff},c^{\widehat{s}_{eff}},\widehat{\alpha},\widehat{\varsigma}\ ,\widehat{\varkappa},\widehat{\zeta}\right)\right)$ given $Z_{t}\left(\cdot,\widehat{s}_{eff},c^{\widehat{s}_{eff}},\widehat{\alpha},\widehat{\varsigma}\ ,\widehat{\varkappa},\widehat{\zeta}\right)$ for $t=t\left(n\right)$, n sufficiently large, with $t\left(n\right)$ going slowly to ∞ with n, so we can ignore the set where $\left|\left|\widehat{Z}_{t}\left(\psi^{\dagger}\right)\right|\right|>t\dim\left(\psi\right)^{1/2}$. Consider a prior $\pi\left(\psi^{\dagger}\right)$ that is absolutely continuous wrt Lesbegue measure, and charges $\psi^{\dagger}$ in a volume of radius $O\left(1\right)$ (that includes the true parameter value). Such a prior is effectively uniform on a volume with radius $O\left(n^{-1/2}\right)$ around the truth. Since the likelihood is highly peaked on a volume of $O\left(n^{-1/2}\right)$, it follows that the approximate posterior based on a prior $\pi\left(\psi^{\dagger}\right)$ is just the rescaled likelihood i.e. set the prior $\pi\left(\psi^{\dagger}\right)$ to be 1 in the previous display.

Thus for sufficiently large t and n, an approximate highest posterior $(1-\alpha)$ credible regions $C_{cred}(1-\alpha)$ for $\psi^\dagger$ is
$C_{cred}(1-\alpha) = \left\{\psi; \widehat{Z}_t(\psi)^T \widehat{Z}_t(\psi) < c_\alpha^2 \right\}$ where c_α^2 satisfies

$$\begin{aligned} & 1-\alpha \\ &= \frac{\int_{\left\{\psi^\dagger; \widehat{Z}_t(\psi^\dagger)^T \widehat{Z}_t(\psi^\dagger) < c_\alpha^2\right\}} \exp\left(-\widehat{Z}_t\left(\psi^\dagger\right)^T \widehat{Z}_t\left(\psi^\dagger\right)/2\right) d\psi^\dagger}{\int_{\left\{\psi^\dagger; \left|\left|\widehat{Z}_t(\psi^\dagger)\right|\right| < t \dim(\psi^\dagger)^{1/2}\right\}} \exp\left(-\widehat{Z}_t\left(\psi^\dagger\right)^T \widehat{Z}_t\left(\psi^\dagger\right)/2\right) d\psi^\dagger} \\ &\approx \frac{\int_0^{c_\alpha} \exp\left(-u^2/2\right) area(u)\, du}{\int_0^{\infty} \exp\left(-u^2/2\right) area(u)\, du} \end{aligned}$$

and

$$area(u)\, du = \int_{\left\{\psi^\dagger; \widehat{Z}_t(\psi^\dagger)^T \widehat{Z}_t(\psi^\dagger) = (u+du)^2\right\}} d\psi^\dagger - \int_{\left\{\psi^\dagger; \widehat{Z}_t(\psi^\dagger)^T \widehat{Z}_t(\psi^\dagger) = u^2\right\}} d\psi^\dagger$$

is the volume of the infinitesmal annulus between
$\left\{\psi^\dagger; \widehat{Z}_t\left(\psi^\dagger\right)^T \widehat{Z}_t\left(\psi^\dagger\right) = (u+du)^2\right\}$ and $\left\{\psi^\dagger; \widehat{Z}_t\left(\psi^\dagger\right)^T \widehat{Z}_t\left(\psi^\dagger\right) = u^2\right\}$. Thus $area(u)$ is the (surface) area of the set $\left\{\psi^\dagger; \widehat{Z}_t\left(\psi^\dagger\right)^T \widehat{Z}_t\left(\psi^\dagger\right) = u^2\right\}$ of dimesion $R^{\dim(\psi^\dagger)-1}$ imbedded in $R^{\dim(\psi^\dagger)}$. The surface area of this set is well-defined with probability one since, under any law including an exceptional law, $\widehat{Z}_t\left(\psi^\dagger\right)$ is almost surely a continuous function of $\psi^\dagger$ and smooth except on a set of $\psi^\dagger$ of Lesbegue measure 0. We have used the fact that $\int_t^\infty \exp\left(-u^2/2\right) area(u)\, du$ goes to zero as $t = t(n) \to \infty$

Henceforth, it will be convenient to index the likelihood approximation $\exp\left(-\widehat{Z}_t(\psi)^T \widehat{Z}_t(\psi)/2\right)$ by ψ and again only use $\psi^\dagger$ to denote the true value of ψ. Upon inserting the expansion (5.4) – (5.5) into the approximate likelihood $(2\pi)^{-p/2} \exp\left(-\widehat{Z}_t(\psi)^T \widehat{Z}_t(\psi)/2\right)$, we see that with probability going to one, the posterior for ψ is asymptotically quadratic in $E\left[\partial Z(\psi)/\partial\psi^T\right]\left(\psi-\psi^\dagger\right)$ for $\left(\psi-\psi^\dagger\right)$ of $O\left(n^{-1/2}\right)$. Thus the posterior will be asymptotically quadratic in $\left(\psi-\psi^\dagger\right)$ (and indeed asymptotically normal with mean $\widehat{\psi}$) if $E\left[\partial U^\dagger_{\text{mod}\, adj}(\psi)/\partial\psi^T\right] = E\left[\partial U^\dagger_{\text{mod}\, adj}\left(\psi^\dagger\right)/\partial\psi^T\right] + o(1)$ and thus $E\left[\partial Z(\psi)/\partial\psi^T\right] = E\left[\partial Z\left(\psi^\dagger\right)/\partial\psi^T\right] + o(1)$ does not depend on ψ to this order. Here $\widehat{\psi}$ is our locally efficient doubly robust estimator solving $\widehat{Z}_t(\psi) = \widehat{Z}(\psi) = 0$. This lack of dependence on ψ will be true, when, as we are assuming, $r\left(m, \overline{l}_m, \overline{a}_{m-1}, \psi\right)$ is smooth in ψ, except for the case where the data were generated under an exceptional law. Thus, in large samples, under non-exceptional laws, the posterior distribution of ψ will be normal. It follows that under non-exceptional laws we can substitute for $area(u)$ the suface area of the $\dim(\psi)$ ball of radius u in $R^{\dim(\psi)}$. Thus, for example, if $\dim(\psi) = 3$, then $area(u) = 4\pi u^2$. In that case c_α is equal to the upper

α quantile $\chi_{\alpha,\dim(\psi)}$ of a χ^2 random variable on $\dim(\psi)$-d.o.f. and thus our highest posterior credible interval $C_{cred}(1-\alpha)$ for ψ is exactly the same as the frequentist confidence interval $C_{op}(1-\alpha)$ based on the pivotal statistic $\widehat{Z}(\psi)$. Thus a result which is well known for regular parametric models continues to hold for our semiparametric model at non-exceptional laws.

However, under the exceptional laws, the posterior distribution for $\psi^\dagger$ will be non-quadratic and the surface $\left\{\psi; \widehat{Z}_t(\psi)^T \widehat{Z}_t(\psi) = u^2\right\}$ will not be a sphere of radius u. Thus area(u) will not be proportional to the suface area of the $\dim(\psi)$ ball of radius u in $R^{\dim(\psi)}$. Nonetheless, the frequentist interval $C_{op}(1-\alpha)$ will remain a highest posterior credible interval $C_{cred}(1-\alpha^*)$ with the 'frequency-calibrated' α^* given by

$$\alpha^* = \alpha^*(\alpha) \approx \frac{\int_0^{\chi_{\alpha,\dim(\psi)}} \exp\left(-u^2/2\right) area(u)\, du}{\int_0^{\infty} \exp\left(-u^2/2\right) area(u)\, du}$$

where $\chi_{\alpha,\dim(\psi)}$ is again the upper α quantile $\chi_{\alpha,\dim(\psi)}$ of a χ^2 random variable on $\dim(\psi)$ d.o.f.

Indeed even when the data are not generated under an exceptional law, for $\left(\psi - \psi^\dagger\right)$ of $O\left(n^{-1/2}\right)$, the $o(1)$ in the equation $E\left[\partial U^\dagger_{\text{mod } adj}(\psi)/\partial\psi^T\right] - E\left[\partial U^\dagger_{\text{mod } adj}\left(\psi^\dagger\right)/\partial\psi^T\right] = o(1)$ is non-uniform in $\psi^\dagger$ because $\psi^\dagger$ may still be close to an exceptional law (see Appendix A1.2 for additional discussion). However the Lebesgue measure of the set of $\psi^\dagger$ for which this $o(1)$ difference will not be small will decrease as n increases so that posterior associated with a prior absolutely continuous wrt Lesbegue measure that charges $\psi^\dagger$ in a volume of radius $O(1)$ will eventually be quadratic.

Nonetheless, in practice, if frequentist interval $C_{op}(1-\alpha)$ includes exceptional laws (or laws very close to exceptional laws) and thus the set where the likelihood is relatively large contains an exceptional law, it is best not to use a normal approximation, but rather to use either Markov chain Monte Carlo or rejection sampling techniques to generate a sample $\psi^{(v)}$, $v = 1, ..., V$ from a density proportional to $\exp\left(-\widehat{Z}(\psi)^T \widehat{Z}(\psi)/2\right)$ to construct a highest posterior credible intervals, even if one had a prior mass of zero on the exceptional laws. Informally, this recommendation represents the possibility that one may be in asymptopia in regards to using the uniform approximation $\exp\left(-\widehat{Z}(\psi)^T \widehat{Z}(\psi)/2\right)$ to the likelihood of $Z\left(\cdot, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right)$, but may not be in asymptopia as far as the non-uniform approximation $E\left[\partial U^\dagger_{\text{mod } adj}(\psi)/\partial\psi^T\right] - E\left[\partial U^\dagger_{\text{mod } adj}\left(\psi^\dagger\right)/\partial\psi^T\right] = o(1)$ is concerned.

A way to formalize this approach to approximate Bayes inference when the set where the likelihood is relatively large contains an exceptional law is to assume that our prior $\pi(\psi)$, although still absolutely continuous, only charges a volume with radius $O\left(n^{-1/2}\right)$ but that volume includes the set where the likelihood is relatively large. [Robins, Scheines, Spirtes, and Wasserman (2003)

also discuss the fact that Bayesian inference can be made to qualitatively agree with uniform asymptotic frequentist inference by letting the support of the prior vary with the sample size.] When the support of the prior only charges a volume with radius $O\left(n^{-1/2}\right)$, the posterior may be non-normal even under non-exceptional laws and even if the prior $\pi(\psi)$ is uniform on its support. In finite samples, one's substantive priors in high dimensional semiparametric problems are often not 'washed out' by the data and thus an asymptotics that assumes the prior charges a volume with radius $O\left(n^{-1/2}\right)$ may be more relevant to finite sample inference than an asymptotics that takes priors as being $O(1)$ since the latter priors are not sensitive to the fact that, due to non-uniform convergence, even at a reasonably large sample size n, the uniform asymptotic frequentist interval $C_{op}(1-\alpha)$ may contain several subsets with nearly equal measure such that the subset-specific typical values of $E\left[\partial U^{\dagger}_{\mathrm{mod}\, adj}(\psi)/\partial\psi^{T}\right]$ and thus $\overline{d}_{op}(\psi)$, differ between subsets by $O(1)$. Indeed, as discussed in the introduction, in many biomedical studies, one may wish not to use an absolutely continuous prior but rather to give a positive prior mass to $\psi = 0$ representing laws (essentially always exceptional) satisfying the g-null hypothesis. In that case, the posterior can often be far from normal.

Failure of the Bernstein–Von Mises Theorem Based on all the Data: It is natural to inquire how a highest posterior $(1-\alpha)$ credible region based on a posterior $\pi(\psi|O_i, i = 1, ..., n)$ that conditions on all the data compares to that based on observing only $Z_t\left(\cdot, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right)$. In regular parametric models, the standardized estimated efficient score process for a subset of the model parameters is asymptotically sufficient so the intervals should be asymptotically equivalent. A formal proof relies on the Bernstein–Von Mises theorem that shows that $\pi(\psi|O_i, i = 1, ..., n)$ is asymptotically normal centered on an efficient estimator of $\psi^{\dagger}$ with variance converging to the Cramer Rao variance bound for $\psi^{\dagger}$. In particular it implies that the posterior median is $n^{1/2}-$*consistent* under the true data generating process F_O. In semiparametric models, a general proof of the Bernstein–Von Mises theorem does not yet exist. However we can prove that, if L_m has continuous components, then in model (a.1) with the $p_m\left(a_m|\overline{L}_m, \overline{A}_{m-1}\right)$ known (and thus in the larger $\overline{d}_{op}, \overline{0}$ union model of Theorem 3.4 as well), the Bernstein Von Mises theorem cannot hold for any Bayesian who would use the same prior $\pi(f_{res})$ for the response densities f_{res}, whatever be the known treatment probabilities $p_m\left(a_m|\overline{L}_m, \overline{A}_{m-1}\right)$. This follows from the fact that, under such an 'independence' prior, the posterior $\pi\left(\psi^{\dagger}|O_i, i = 1, ..., n\right)$ and thus the posterior median is not a function of $p_m\left(a_m|\overline{L}_m, \overline{A}_{m-1}\right)$. But Robins and Ritov (1997) showed that any estimator of $\psi^{\dagger}$ that does not depend on the $p_m\left(a_m|\overline{L}_m, \overline{A}_{m-1}\right)$ must converge at rate no greater than $\log(\log n)$ under some treatment process $p_m\left(a_m|\overline{L}_m, \overline{A}_{m-1}\right)$ and response law f_{res}, contradicting the conclusion of the Bernstein–Von Mises theorem. Indeed Robins

and Ritov (1997) prove that under an 'independence' prior, the highest posterior $(1-\alpha)$ credible regions based on posterior $\pi(\psi|O_i, i=1,...,n)$ cannot both have asymptotic frequentist coverage of $1-\alpha$ and volume converging to zero as $n \to \infty$. Since both these desirable properties are true of the intervals based on $\pi\left(\psi|Z_t\left(\cdot, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right)\right)$, it follows that intervals based on $\pi(\psi|O_i, i=1,...,n)$ and $\pi\left(\psi|Z_t\left(\cdot, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right)\right)$ are not asymptotically equivalent and from a frequentist point of view those based on $\pi\left(\psi|Z_t\left(\cdot, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right)\right)$ are preferable. The good performance of intervals based on $\pi\left(\psi|Z_t\left(\cdot, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right)\right)$ in our union model depends critically on the fact that $Z_t\left(\cdot, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right)$ is a function of either the true $p_m\left(a_m|\overline{L}_m, \overline{A}_{m-1}\right)$ (when the latter is known) or of a model-based estimate of $p_m\left(a_m|\overline{L}_m, \overline{A}_{m-1}\right)$ (when the true function is unknown). It can be argued from general decision theoretic results that there exists a particular prior for $f_O = f_{res} f_{tr,obs}$ with f_{res} and the $f_{tr,obs} = \prod_m p_m\left(a_m|\overline{L}_m, \overline{A}_{m-1}\right)$ apriori dependent such that $\pi(\psi|O_i, i=1,...,n)$ is an accurate large sample approximation to $\pi\left(\psi|Z_t\left(\cdot, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right)\right)$ when the treatment probabilities are known. However the prior may be hard to derive and will essentially never represent the analyst's subjective beliefs.

Marginal Bayesian inference: Bayesian inference concerning the marginal ψ_{int} is based on integration of $\pi\left(\psi|Z_t\left(\cdot, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right)\right) \propto (2\pi)^{-p/2} \exp\left(-\widehat{Z}_t(\psi)^T \widehat{Z}_t(\psi)/2\right) \pi(\psi)$ over ψ_{-int}. Thus, if $\pi(\psi)$ is smooth and charges $\psi^\dagger$ in a volume of $O(1)$, then, in large samples, under non-exceptional laws, the posterior distribution of ψ_{int} will be normal if, as we have assumed, $r\left(m, \overline{l}_m, \overline{a}_{m-1}, \psi\right)$ is smooth in ψ and and thus (a) the posterior median and posterior mean of ψ_{int} will agree in sign and (b) furthermore the highest $(1-\alpha)$ posterior credible region for ψ_{int} will be asymptotically equivalent to the univariate (non-simultaneous) $(1-\alpha)$ confidence interval for ψ_{int} considered in the last subsection. However because of the possibility that the data were generated under an exceptional law or because we are not in asymptopia in regards to the non-uniform approximation $E\left[\partial U^\dagger_{\mathrm{mod}\, adj}(\psi)/\partial\psi^T\right] - E\left[\partial U^\dagger_{\mathrm{mod}\, adj}\left(\psi^\dagger\right)/\partial\psi^T\right] = o(1)$, it is again best not to use a normal approximation but rather to use either Markov chain Monte Carlo or rejection sampling techniques to generate a sample $\psi^{(v)}$, $v = 1,...,V$ from a density proportional to $\exp\left(-\widehat{Z}_t(\psi)^T \widehat{Z}_t(\psi)/2\right)$ or $\exp\left(-\widehat{Z}_t(\psi)^T \widehat{Z}_t(\psi)/2\right)\pi(\psi)$ and then report whether the sample mean or median of the $\psi^{(v)}_{int}$ exceeds zero as an estimate of the posterior mean $E\left(\psi_{int}|Z_t\left(\cdot, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right)\right)$ or median of ψ_{int}. In this setting, the

posterior mean and median may fail to agree in sign, leading to different decisions under our two different decision rules. Further, the highest $(1-\alpha^*)$ posterior credible region for ψ_{int}, in contrast to the credible interval $C_{cred}(1-\alpha^*)$ for the entire vector ψ, (i) may have frequentist coverage less than $(1-\alpha)$ when we choose α^* equal to $\alpha^*(\alpha)$ and (ii) may not be asymptotically equivalent (in terms of shape and volume) to the univariate (non-simultaneous) $(1-\alpha)$ confidence interval for ψ_{int} considered in the last subsection for any value of α^*.

Optimal Bayes Decison Rules: The actual fact is that the optimal Bayes decision rule (i.e. the rule that maximizes posterior expected utility) that a Bayesian should use in deciding whether to treat a patient with past data $\left(\overline{l}_m, \overline{a}_{m-1}\right)$ on day t_m may be captured by neither of our decision rules. Thus, in principle, we should not choose between them but calculate the optimal Bayes rule. First suppose we have a correctly specified model $\gamma^{\overline{0}}\left(\overline{l}_m, \overline{a}_m, \psi\right) = a_m r\left(m, \overline{l}_m, \overline{a}_{m-1}, \psi\right)$ for $\gamma^{\overline{d}_{op},\overline{0}}\left(\overline{l}_m, \overline{a}_m\right)$ so the true optimal regime $d_{op} = d_{op}\left(\psi^\dagger\right)$ is a deterministic function of the true value $\psi^\dagger$. Let Ψ be a set that contains ψ with posterior probability one so $D_{op}(\Psi) = \{d_{op}(\psi); \psi \in \Psi\}$ contains the optimal regime with posterior probability one. Further we can assume the true $\psi^\dagger$ that actually generated the data is in Ψ. Nonetheless the optimal Bayes decision rule (i.e. treatment regime) need not be an element of $D_{op}(\Psi)$. Here is a specific simple example. Suppose the data is $A_0, A_1,$ and $Y = L_2$ with no L_0 or L_1. Suppose we have a saturated model $\gamma^{\overline{0}}\left(\overline{l}_m, \overline{a}_m, \psi\right) = \gamma^{\overline{0}}\left(\overline{a}_m, \psi\right)$, $\psi = (\psi_{m=0}, \psi_{m=1,a_0=1}, \psi_{m=1,a_0=0})$ for $\gamma^{\overline{d}_{op},\overline{0}}\left(\overline{l}_m, \overline{a}_m\right)$ and the posterior distribution is discrete with 2 support points: $\psi_{m=0} = -100, \psi_{m=1,a_0=1} = 0, \psi_{m=1,a_0=0} = 10$ with posterior probability .4 and $\psi_{m=0} = 50$, $\psi_{m=1,a_0=1} = 1, \psi_{m=1,a_0=0} = -20$ with posterior probability .6. Then, with posterior probability .4, the optimal treatment regime is $d_0^{op} = 0$, $d_1^{op} = 1$ and with posterior probability .6, $d_0^{op} = 1$, $d_1^{op} = 1$. So the set Ψ consists of two vectors ψ and $D_{op}(\Psi)$ contains only the two regimes $d_0^{op} = 0$, $d_1^{op} = 1$ and $d_0^{op} = 1$, $d_1^{op} = 1$. But the optimal Bayes decision rule is $d_0^{op} = 0, d_1^{op} = 0$ as it has posterior expected utility of $.4 \times (-10 + c_1) + .6 \times (-50 + c_2)$ while the posterior expected utility of $d_0^{op} = 0$, $d_1^{op} = 1$ is $.4 \times c_1 + .6 \times (-50 - 20 + c_2)$, of $d_0^{op} = 1$, $d_1^{op} = 1$ is $.4 \times (-100 + c_1) + .6 \times (c_2)$ and of $d_0^{op} = 1$, $d_1^{op} = 0$ is $.4 \times (-100 + c_1) + .6 \times (-1 + c_2)$, where c_1 is the posterior utility of $d_0^{op} = 0$, $d_1^{op} = 1$ conditional on the first support point and c_2 is the posterior utility of $d_0^{op} = 1$, $d_1^{op} = 1$ conditional on the second support point.

Note in this simple example, the optimal Bayes decision rule was a function of the marginal posterior distribution of ψ. In the next paragraph we will see that, when the sample size is large, this remains true even when data on covariates L_m are available, provided the data is generated under a nonexceptional law and the prior for ψ charges a volume with radius $O(1)$. However, we show in the next paragraph but one that, even in large samples, if either the prior for ψ only charges a volume of radius $O\left(n^{-1/2}\right)$ or the data were gener-

ated under an exceptional law, the optimal Bayes decision rule will generally be a complex function of the posterior distribution of the infinite-dimensional nuisance parameters.

Suppose the data were generated under an unexceptional law and the prior $\pi(\psi)$ is absolutely continuous and charges a volume with radius $O(1)$. We therefore approximate the posterior $\pi\left(\psi|Z_t\left(\cdot,\widehat{s}_{eff},c^{\widehat{s}_{eff}},\widehat{\alpha},\widehat{\varsigma},\widehat{\varkappa},\widehat{\zeta}\right)\right)$ by the approximate posterior

$$\pi_{post}(\psi)=\frac{I\left\{\psi;\left|\left|\widehat{Z}_t(\psi)\right|\right|<t\dim(\psi)^{1/2}\right\}\exp\left(-\widehat{Z}_t(\psi)^T\widehat{Z}_t(\psi)/2\right)}{\int_{\left\{\psi;||\widehat{Z}_t(\psi)||<t\dim(\psi)^{1/2}\right\}}\exp\left(-\widehat{Z}_t(\psi)^T\widehat{Z}_t(\psi)/2\right)d\psi}.$$

Then, for sufficiently large n and t, the posterior mean $\int a_K r\left(K,\overline{l}_K,\overline{a}_{K-1},\psi\right)\pi_{post}(\psi)\,d\psi$ of $\gamma^{\overline{d}_{op},\overline{0}}\left(\overline{l}_K,\overline{a}_K\right)$ based on a correct smooth model $\gamma^{\overline{0}}\left(\overline{l}_m,\overline{a}_m,\psi\right)=a_m r\left(m,\overline{l}_m,\overline{a}_{m-1},\psi\right)$ will to $o(1)$ be $a_K r\left(K,\overline{l}_K,\overline{a}_{K-1},\widehat{\psi}\right)$ because $\pi_{post}(\psi)$ is normal with mean equal to the locally efficient doubly robust estimator $\widehat{\psi}$ solving $\widehat{Z}_t(\psi)=\widehat{Z}(\psi)=0$, the variance of $\widehat{\psi}$ is $O(1/n)$, and $\widehat{\psi}$ is greater than $O\left(n^{-1/2}\right)$ away from any ψ for which $\arg\max_{a_m\in\mathcal{A}_m}\gamma^{\overline{0}}\left(\overline{L}_m,\overline{A}_{m-1},a_m,\psi\right)$ is not unique with positive probability (by our assumption the data were generated under a nonexceptional law.) The optimal Bayes decison $d_{bayes,K}\left(\overline{l}_K,\overline{a}_{K-1}\right)$ for a subject known to have history $\overline{l}_K,\overline{a}_{K-1}$ is $\arg\max_{a_k\in\mathcal{A}_k}$ of the posterior mean, which under our assumptions is equal to $\arg\max_{a_k\in\mathcal{A}_k}\ a_K r\left(K,\overline{l}_K,\overline{a}_{K-1},\widehat{\psi}\right)$ with probability going to 1. Further because $\widehat{\psi}$ is $n^{1/2}-$*consistent* for $\psi^\dagger$, results described in the following paragraph imply that, with probability going to one, the optimal Bayes decison $d_{bayes,K-1}\left(\overline{l}_{K-1},\overline{a}_{K-2}\right)$ for a subject known to have history $\overline{l}_{K-1},\overline{a}_{K-2}$ is $\arg\max_{a_{k-1}\in\mathcal{A}_{k-1}}$ of the posterior mean $\int a_{K-1}r\left(K-1,\overline{l}_{K-1},\overline{a}_{K-2},\psi\right)\pi_{post}(\psi)\,d\psi$ of $\gamma^{\overline{d}_{op},\overline{0}}\left(\overline{l}_{K-1},\overline{a}_{K-1}\right)=a_{K-1}r\left(K-1,\overline{l}_{K-1},\overline{a}_{K-2},\psi^\dagger\right)$ which, under our assumptions is equal, with probability approaching 1, to $\arg\max_{a_{k-1}\in\mathcal{A}_{k-1}}a_{K-1}r\left(K-1,\overline{l}_{K-1},\overline{a}_{K-2},\widehat{\psi}\right)$. Continuing in this manner we see that the optimal bayes decision rule $d_{bayes,m}\left(\overline{l}_m,\overline{a}_{m-1}\right)$ is $d_{op,m}\left(\overline{l}_m,\overline{a}_{m-1},\widehat{\psi}\right)$ for each m.

Suppose next the data were generated under an exceptional law and/or the prior $\pi(\psi)$ only charges a volume with radius $O\left(n^{-1/2}\right)$. Although we still approximate the posterior by $\pi_{post}(\psi)$, now $\pi_{post}(\psi)$ may neither have mean $\widehat{\psi}$ nor be normal, and thus

$$d_{bayes,K}\left(\overline{l}_K,\overline{a}_{K-1}\right)=\arg\max_{a_k\in\mathcal{A}_k}\int a_K r\left(K,\overline{l}_K,\overline{a}_{K-1},\psi\right)\pi_{post}(\psi)\,d\psi$$

may differ from $d_{op,K}\left(\overline{l}_K,\overline{a}_{K-1},\widehat{\psi}\right)$. Now, by definition, $d_{bayes,K-1}\left(\overline{l}_{K-1},\overline{a}_{K-2}\right)$ is the optimal Bayes choice for a_{K-1} given that a_K will equal $d_{bayes,K}\left(\overline{l}_K,\overline{a}_{K-1}\right)$.

That is $d_{bayes,K-1}\left(\overline{l}_{K-1}, \overline{a}_{K-2}\right)$ equals $\arg\max_{a_{k-1}\in\mathcal{A}_{k-1}}$ of the posterior mean of $E\left[Y_{\overline{a}_{K-2},a_{K-1},d_{bayes,K}}|\overline{A}_{K-2}=\overline{a}_{K-2},\overline{L}_{K-1}=\overline{l}_{K-1}\right]$ and, thus, of

$$E\left[Y_{\overline{a}_{K-2},a_{K-1},d_{bayes,K}}-Y_{\overline{a}_{K-2},0,d_{op,K-1}}|\overline{A}_{K-2}=\overline{a}_{K-2},\overline{L}_{K-1}=\overline{l}_{K-1}\right].$$

At $\psi=\psi^{\dagger}$, this contrast can be written as

$$\begin{aligned}
&E\left[Y_{\overline{a}_{K-2},a_{K-1},d_{bayes,K}}-Y_{\overline{a}_{K-2},a_{K-1},d_{op,K-1}}|\overline{A}_{K-2}=\overline{a}_{K-2},\overline{L}_{K-1}=\overline{l}_{K-1}\right]+\\
&\gamma^{\overline{d}_{op},\overline{0}}\left(\overline{l}_{K-1},\overline{a}_{K-1},\psi\right)\\
&=E\left\{\begin{array}{c}E\left[Y_{\overline{a}_{K-2},a_{K-1},d_{bayes,K}}-Y_{\overline{a}_{K-2},a_{K-1},d_{op,K-1}}|\overline{A}_{K-1}=\overline{a}_{K-1},\overline{L}_{K}\right]\\ |\overline{A}_{K-1}=\overline{a}_{K-1},\overline{L}_{K-1}=\overline{l}_{K-1}\end{array}\right\}+\\
&+\gamma^{\overline{d}_{op},\overline{0}}\left(\overline{l}_{K-1},\overline{a}_{K-1},\psi\right)\\
&=E\left[j_{Bayes}\left(\overline{L}_{K},\overline{A}_{K-1}\right)|\overline{A}_{K-1}=\overline{a}_{K-1},\overline{L}_{K-1}=\overline{l}_{K-1}\right]+\gamma^{\overline{d}_{op},\overline{0}}\left(\overline{l}_{K-1},\overline{a}_{K-1},\psi\right)
\end{aligned}$$

where

$$\begin{aligned}
&j_{Bayes}\left(\overline{l}_{K},\overline{a}_{K-1}\right)\\
&=\gamma^{\overline{d}_{op},\overline{0}}\left(\overline{l}_{K},\overline{a}_{K-1},d_{bayes,K}\left(\overline{l}_{K},\overline{a}_{K-1}\right),\psi\right)-\\
&\gamma^{\overline{d}_{op},\overline{0}}\left(\overline{l}_{K},\overline{a}_{K-1},d_{op,K}\left(\overline{l}_{K},\overline{a}_{K-1},\psi\right),\psi\right)\\
&=\left\{d_{bayes,K}\left(\overline{l}_{K},\overline{a}_{K-1}\right)-d_{op,K}\left(\overline{l}_{K},\overline{a}_{K-1},\psi\right)\right\}r\left(K,\overline{l}_{K},\overline{a}_{K-1},\psi\right)
\end{aligned}$$

Hence

$$\begin{aligned}
&d_{bayes,K-1}\left(\overline{l}_{K-1},\overline{a}_{K-2}\right)\\
&=\arg\max_{a_{k-1}\in\mathcal{A}_{k-1}}\\
&[\int\left\{d_{bayes,K}\left(\overline{l}_{K},\overline{a}_{K-1}\right)-d_{op,K}\left(\overline{l}_{K},\overline{a}_{K-1},\psi\right)\right\}\times\\
&r\left(K,\overline{l}_{K},\overline{a}_{K-1},\psi\right)dF\left(l_{K}|\overline{l}_{K-1},\overline{a}_{K-1};\eta\right)\pi_{post}\left(\eta|\psi\right)\pi_{post}\left(\psi\right)d\psi d\mu\left(\eta\right)+\\
&\int a_{K-1}r\left(K-1,\overline{l}_{K-1},\overline{a}_{K-2},\psi\right)\pi_{post}\left(\psi\right)d\psi]
\end{aligned}$$

where η denotes the parameter governing the density $f\left(l_{K}|\overline{l}_{K-1},\overline{a}_{K-1}\right)$ and $\pi_{post}\left(\eta|\psi\right)$ is the conditonal posterior of η with respect to the measure $\mu\left(\cdot\right)$. One possible approach to obtaining $d_{bayes,K-1}\left(\overline{l}_{K-1},\overline{a}_{K-2}\right)$ would be to specify a parametric model $f\left(l_{m}|\overline{l}_{m-1},\overline{a}_{m-1};\eta\right)$, estimate η by the MLE $\widehat{\eta}$ and take $\pi_{post}\left(\eta|\psi\right)=\pi_{post}\left(\eta\right)$ to be normal with mean $\widehat{\eta}$ and variance given by the inverse Hessian matrix for $\widehat{\eta}$. A second approach, analogous to that taken in Section 6.2 and (more specifically) Section 7.2 below, is to assume the law of $j\left(\overline{L}_{K},\overline{A}_{K-1}\right)|\overline{A}_{K-1}=\overline{a}_{K-1},\overline{L}_{K-1}=\overline{l}_{K-1}$ is normal with mean $\nu\left(\overline{a}_{K-1},\overline{l}_{K-1};\beta_{K}\right)$ and, say, variance σ_{K}^{2} where $\nu\left(\overline{a}_{K-1},\overline{l}_{K-1};\beta_{K}\right)$ is

a known function and β_K and σ_K^2 are unknown parameters. We might take the posterior of β_K given ψ to be normal with mean equal to the (possibly nonlinear) least squares regession estimator $\widehat{\beta}_K(\psi)$ from the regression of $j\left(\overline{L}_K, \overline{A}_{K-1}\right) = j\left(\overline{L}_K, \overline{A}_{K-1}, \psi\right)$ on $\left(\overline{L}_{K-1}, \overline{A}_{K-1}\right)$ with regression function $\nu\left(\overline{a}_{K-1}, \overline{l}_{K-1}; \beta_K\right)$. Then take $d_{bayes,K-1}\left(\overline{l}_{K-1}, \overline{a}_{K-2}\right) =$
$\arg\max_{a_{k-1}\in\mathcal{A}_{k-1}} \int\int \left\{\nu\left(\overline{a}_{K-1}, \overline{l}_{K-1}; \beta_K\right) + a_{K-1} r\left(K-1, \overline{l}_{K-1}, \overline{a}_{K-2}, \psi\right)\right\} \times$
$\pi_{post}\left(\beta_K|\psi\right)\pi_{post}\left(\psi\right) d\psi d\mu\left(\beta_K\right)$.. More generally, under this second approach, it follows from results in Section 6.2 and 7.2 that $d_{bayes,m}\left(\overline{l}_m, \overline{a}_{m-1}\right)$ is

$$\begin{aligned}
&\arg\max_{a_m\in\mathcal{A}_m} [\int \left\{\nu\left(\overline{a}_m, \overline{l}_m; \beta_{m+1}\right) + a_m r\left(m, \overline{l}_m, \overline{a}_{m-1}, \psi\right)\right\} \times \\
&\pi_{post}\left(\beta_{m+1}|\psi, \underline{\beta}_{m+2}\right) \times \\
&\pi_{post}\left(\psi, \underline{\beta}_{m+2}\right) d\psi d\mu\left(\psi, \underline{\beta}_{m+2}, \beta_{m+1}\right)]
\end{aligned}$$

where $\nu\left(\overline{a}_m, \overline{l}_m; \beta_{m+1}\right)$ is a parametric model for
$E\left[j\left(\overline{L}_{m+1}, \overline{A}_m\right)|\overline{L}_m = \overline{l}_m, \overline{A}_m = \overline{a}_m\right]$ with
$j\left(\overline{L}_m, \overline{A}_{m-1}\right) = E\left[Y_{\overline{A}_{m-1}, \underline{d}_{bayes,m}} - Y_{\overline{A}_{m-1}, \underline{d}_{op,m}}|\overline{L}_m, \overline{A}_{m-1}\right]$.

Under either of the 2 approaches, the principal benefit of specifying a optimal drSNMM $\gamma^{\overline{d}_{op},\overline{0}}\left(\overline{l}_m, \overline{a}_m, \psi\right)$ is lost when computing the optimal Bayes decision, in the sense that, even when the treatment probabilities are known, we must model, in addition to $\gamma^{\overline{d}_{op},\overline{0}}\left(\overline{l}_m, \overline{a}_m\right)$, other aspects of the joint distribution of the observed data. Further the second approach may result in incompatible models in the sense that there is no joint distribution for the observed data satisfying all the functional form restrictions imposed by the models $\nu\left(\overline{a}_m, \overline{l}_m; \beta_{m+1}\right)$ and $\gamma^{\overline{d}_{op},\overline{0}}\left(\overline{l}_m, \overline{a}_m, \psi\right)$. Indeed, Robins (1994) shows that even the first approach may suffer from model incompatibility. At the cost of a complex reparametrization of the joint distribution of the observed data described in the Appendix of Robins (1994), the possibility of model incompatibility when using the first approach can be resolved.

To overcome the need to model other aspects of the joint distribution of the observed data we might specify a drSNMM $\gamma^{\overline{d}_{Bayes},\overline{0}}\left(\overline{l}_m, \overline{a}_m, \psi\right)$ for
$E\left[Y_{\overline{A}_{m-1}, a_m, \underline{d}_{bayes,m+1}} - Y_{\overline{A}_{m-1}, 0, \underline{d}_{bayes,m+1}}|\overline{L}_m = \overline{l}_m, \overline{A}_{m-1} = \overline{a}_{m-1}\right]$
$= \gamma^{\overline{d}_{Bayes},\overline{0}}\left(\overline{l}_m, \overline{a}_m\right)$ so that $d_{bayes,m}\left(\overline{l}_m, \overline{a}_{m-1}, \psi\right)$
$= \arg\max_{a_m\in\mathcal{A}_m} \int \gamma^{\overline{d}_{Bayes},\overline{0}}\left(\overline{l}_m, \overline{a}_m, \psi\right)\pi_{post}\left(\psi\right) d\psi$. This idea raises all sorts of interesting and unresolved philosophical and statistical questions because of course $d_{bayes,m}\left(\overline{l}_m, \overline{a}_{m-1}, \psi\right)$ is a function of the data through the posterior $\pi_{post}\left(\psi\right)$. [However the above expectation is to be computed treating $\underline{d}_{bayes,m+1}$ as a given fixed regime rather than as a random regime that depends on the data.] This idea will be further pursued elsewhere.

6 Comparison of Optimal drSNMMs with Alternative Approaches

6.1 Susan Murphy's semiparametric regret model are SNMMS

I now show that, under sequential randomization Murphy's semiparametric regret model is a particular parametrization of an additive $\overline{d}-regime$ srSNMM with $\overline{d}$ the optimal regime $\overline{d}_{op}$. Specifically, Murphy*'s* semiparametric regret model specifies that the regret $E\left[Y_{\overline{a}_{m-1},\underline{d}_{op,m}} - Y_{\overline{a}_m,\underline{d}_{op,m+1}} | \overline{L}_{\overline{a}_{m-1},m} = \overline{l}_m\right]$ equals $u_m\left(\overline{l}_m, \overline{a}_m, \beta^\dagger\right)$ where $u_m\left(\overline{l}_m, \overline{a}_m, \beta\right) = \eta_m\left(\overline{l}_m, \overline{a}_{m-1}, \beta_{scale}\right) \times$

$f\left(a_m - d_{op,m}\left(\overline{l}_m, \overline{a}_{m-1}, \beta_{regime}\right)\right)$, $\beta = (\beta_{scale}, \beta_{regime})$ is a finite dimensional parameter vector, $f(\cdot)$ is a known non-negative function satisfying $f(0) = 0, \eta_m\left(\overline{l}_m, \overline{a}_{m-1}, \beta_{scale}\right)$ is a known non-negative scale function, and $d_{op,m}\left(\overline{l}_m, \overline{a}_{m-1}, \beta^\dagger_{regime}\right)$ is the optimal regime $d_{op,m}\left(\overline{l}_m, \overline{a}_{m-1}\right)$. In all her examples $\beta = (\beta_{scale}, \beta_{regime})$ had (i) β_{scale} and β_{regime} as variation independent and (ii) $\eta_m\left(\overline{l}_m, \overline{a}_{m-1}, \beta_{scale}\right) = 0$ if and only if $\beta_{scale} = 0$. [Murphy also allows the possibility that $u_m\left(\overline{l}_m, \overline{a}_m, \beta\right)$ is a sum of J_m terms indexed by j of the form $\eta_{j,m}\left(\overline{l}_m, \overline{a}_{m-1}, \beta_{j,scale}\right) \times f_j\left(a_m - d_{j,op,m}\left(\overline{l}_m, \overline{a}_{m-1}, \beta_{j,regime}\right)\right)$]. It follows that Murphy's model is a $\overline{d}_{op}-regime$ srSNMM with $-u_m\left(\overline{l}_m, \overline{a}_m, \beta\right)$ equal to $\gamma^{\overline{d}_{op}}\left(\overline{l}_m, \overline{a}_m; \beta\right)$ in my notation. Note that Murphy uses a particular parametrization, similar to one suggested by Robins (1999, p.125), under which the "scale" components β_{scale} of β being zero implies both the g-null hypothesis and that the other components of β_{regime} are undefined. That is her parametrization is such that the parameters β_{regime} are only defined under the alternative $\beta_{scale} \neq 0$. (This parametrization can result in certain additional inferential difficulties that have been discussed frequently in the statistical literature; however, as discussed by Robins (1999) the parametrization has a certain conceptual justification.)

Comparisons

Limitations of Murphy's regret model include a) estimation of $\beta^\dagger$ based on smooth function optimization methods requires (differentiable) approximations (e.g., with sigmoid functions) of indicator functions and b) regrets are not effect measures about which scientists have clear substantive opinions amenable to easy modelling. Optimal drSNMMs do not suffer as severely from these limitations. We will consider these two limitations within the context of one specific example offered by Murphy. She considers the model for dichotomous A_m *and* univariate positive L_m given by $u_m\left(\overline{l}_m, \overline{a}_m, \beta\right) = \beta_{scale}\left\{a_m - d_{op,m}\left(\overline{l}_m, \overline{a}_{m-1}, \beta_{regime}\right)\right\}^2$ with $d_{op,m}\left(\overline{l}_m, \overline{a}_{m-1}, \beta_{regime}\right)$
$= I\left(l_m > \beta_{regime}\right)$, so that treatment is preferred whenever L_m exceeds $\beta^\dagger_{regime}$. Because β_{regime} lies inside an indicator function, Murphy's criterion function is not differentiable with respect to β_{regime}. In order to use

smooth function optimization methods, she uses a differentiable approximation $e^{30(l_m-\beta_{regime})}/\left\{1+e^{30(l_m-\beta_{regime})}\right\}$ for $I\left(l_m>\beta_{regime}\right)$. The performance of this approximation in her simulations was rather poor (in the sense that estimates of the model parameters were biased).

Murphy's model also has substantive limitations: it assumes apriori that the dose response as a function of L_m is positive in the sense that if treatment is preferred at a given level of L_m, it is preferred at all higher levels; if the data may imply a negative dose response this cannot be detected using this model. Further the model implies that the regret β_{scale} of taking treatment when one should not $\left(L_m<\beta^{\dagger}_{regime}\right)$ is exactly equal to the regret of not taking treatment when one should $\left(L_m>\beta^{\dagger}_{regime}\right)$. Further it assumes that the advantage β_{scale} of taking treatment once $L_m>\beta^{\dagger}_{regime}$ is independent of the value of L_m and the advantage β_{scale} of not taking treatment once $L_m<\beta^{\dagger}_{regime}$ also does not depend on L_m. It is substantively hard to imagine such a sharp cut-point $\beta^{\dagger}_{regime}$. Now Murphy clearly did not intend this regret model to be substantively realistic and one could clearly elaborate the model , perhaps by including additional parameters, in order to make it substantively realistic. I will argue that precisely how to do so requires quite a bit of thought, tedious calculation, and some mathematical ability.

The substantive difficulty in specifying Murphy's regret model is already evident in the simplest of all settings : the setting in which we have a single time-independent treatment $(K=0)$. Thus assume the data are $L=L_0, A=A_0, Y=L_1$ with A dichotomous, L univariate and positive and Y continuous. Then under sequential randomization any $\left(\overline{d},\overline{0}\right)-drSNMM\ \gamma^{\overline{0}}\left(\overline{l}_0,\overline{a}_0;\psi\right)$ for $\gamma^{\overline{d},\overline{0}}\left(\overline{l}_0,\overline{a}_0\right)=\gamma^{\overline{0}}\left(l,a\right)$ does not depend on $\overline{d}$ as our only treatment decision is the final one; in particular $\gamma^{\overline{d}_{op},\overline{0}}\left(\overline{l}_m,\overline{a}_m\right)$ equals $\gamma^{\overline{d},\overline{0}}\left(\overline{l}_m,\overline{a}_m\right)=\gamma^0\left(l,a\right)=E\left[Y|L=l,A=a\right]-E\left[Y|L=l,A=0\right]$ under sequential randomization. Thus an optimal drSNMM model is just a model $\gamma^0\left(l,a,\psi\right)$ for $E\left[Y|L=l,A=a\right]-E\left[Y|L=l,A=0\right]$. The simplest such model that includes the possibility that optimal treatment may change at some value of L is the simple (semiparametric) linear regression model $\gamma^0\left(l,a,\psi\right)=a\left(\psi_0+\psi_1 l\right)$ [i.e. the model $E\left[Y|L=l,A=a\right]=a\left(\psi_0+\psi_1 l\right)+b(l)$ with $b(l)$ unrestricted.] It is trivial to derive the optimal regime and the regret from this linear regression model. Specifically, $d_{op}\left(l,\psi\right)=\arg\max_{a\in\{0,1\}}\gamma^0\left(l,a,\psi\right)=I\left(\psi_0+\psi_1 l>0\right)$ and the regret $E\left[Y_{d_{op}}-Y_a|L=l,\psi\right]=\left\{I\left(\psi_0+\psi_1 l>0\right)-a\right\}\left(\psi_0+\psi_1 l\right)$. Thus in our optimal drSNMM approach, we take our beliefs as to the functional form of the dose response $E\left[Y|L=l,A=1\right]-E\left[Y|L=l,A=0\right]$ as primary. We encode these beliefs in a regression model $\gamma^0\left(l,a,\psi\right)$ for $E\left[Y|L=l,A=a\right]-E\left[Y|L=l,A=0\right]$ and then derive the optimal regime and the regret function from our regression model. I believe that most scientists would use this same approach in working out their beliefs about the likely functional form of the regret.

Murphy's approach is just the converse. She specifies a model $d_{op}(l, \beta_{regime}) = I(l > \beta_{regime})$ for the optimal regime and for the regret $E\left[Y_{d_{op}} - Y_a | L = l, \psi\right] = \beta_{scale}\{a - d_{op}(l, \beta_{regime})\}^2$. This of course induces an optimal dr SNMM $\gamma^0_{Murphy}(l, a, \beta) = u(l, 0, \beta) - u(l, 1, \beta)$
$= a\{\beta_{scale} I(l > \beta_{regime}) - \beta_{scale} I(l \leq \beta_{regime})\} = a\beta_{scale}[2I(l > \beta_{regime}) - 1]$ for $E[Y|L = l, A = a] - E[Y|L = l, A = 0] = \gamma^0(l, a)$. Written in terms of the regression function $\gamma^0(l, a)$, one can immediately see how substantively unusual Murphy's regret model is with a jump discontinuity at β_{regime}, and why smooth optimization methods cannot be applied without approximation. In contrast, in section 4.1.1, we have seen how to obtain closed form g-estimates of the parameters of the semiparametric linear model $\gamma^0(l, a, \psi) = a(\psi_0 + \psi_1 l)$ when the treatment probabilities $p(a|l)$ are known or can be modelled.

We have seen it is easy to derive the optimal drSNMM model $\gamma^0_{Murphy}(l, a, \beta)$ implied by Murphy's regret model. We now show the converse is not true. That is the functional form of Murphy's regret model implied by the simple optimal drSNMM model $\gamma^0(l, a, \psi) = a(\psi_0 + \psi_1 l)$ for $\gamma^0(l, a) = E[Y|L = l, A = a] - E[Y|L = l, A = 0]$ is tedious to derive. Above we saw that the regret is

$\{I(\psi_0 + \psi_1 l > 0) - a\}(\psi_0 + \psi_1 l)$. But this is not in the Murphy form of a sum over j of functions $u_j(l, a, \beta) = \eta_j(l, \beta_{j,scale}) f_j\{a - d_{op,j}(l, \beta_{j,regime})\}$ with $\eta_j(l, \beta_{scale})$ nonnegative and $f_j(u)$ minimized at $u = 0$. To put it in the Murphy form we define $\beta_1 = |\psi_1|$, $\beta_2 = -\psi_0/\psi_1, \beta_3 = I(\psi_1 > 0)$. Then some tedious algebra shows that the regret $\{I(\psi_0 + \psi_1 l > 0) - a\}(\psi_0 + \psi_1 l)$

$= \sum_{j=1}^4 \eta_j(l, \beta_{j,scale}) f_j\{a - d_{j,op}(l, \beta_{j,regime})\}$ where $f_j(u) = u^2$ for all j,

$\eta_1(l, \beta_{1,scale}) = (1 - \beta_3) I(\beta_2 > 0) \beta_1 |1 - \beta_2 l|$,

$d_{op,1}(l, \beta_{1,regime}) = I\{l < \beta_2\}; \eta_2(l, \beta_{2,scale}) = \beta_3 I(\beta_2 > 0) \beta_1 |1 - \beta_2 l|$,

$d_{op,2}(l, \beta_{2,regime}) = I\{l \geq \beta_2\}$;

$\eta_3(l, \beta_{3,scale}) = \beta_3 I(\beta_2 \leq 0) \beta_1 |1 - \beta_2 l|$,

$d_{op,3}(l, \beta_{3,regime}) = 1; \eta_4(l, \beta_{4,scale}) = (1 - \beta_3) I(\beta_2 \leq 0) \beta_1 |1 - \beta_2 l|$,

$d_{op,4}(l, \beta_{4,regime}) = 0$. Note in particular that $\beta_{1,scale}$ and $\beta_{1,regime}$ are not variation independent and that $\beta_{3,regime}$ and $\beta_{4,regime}$ do not exist. Given we choose $f_j(u) = u^2$, this is the unique expression for the regret in the Murphy parametrization. Further if we were given the model only in its Murphy parametrization with parameters buried within indicator functions, it would not be immediately obvious without some calculations that we could obtain closed form estimates of the parameter vector β by reexpressing the model in its alternative form $E[Y|L = l, A = a] - E[Y|L = l, A = 0] = a(\psi_0 + \psi_1 l)$ with ψ and β related as described above and then fitting using g-estimation. Indeed, we can obtain closed form dr-lse estimates as described in section 4.

To summarize we note that any optimal drSNMM model $\gamma^{\overline{0}}(\overline{l}_m, \overline{a}_m, \psi)$ for

$E\left[Y_{\overline{a}_m, \underline{d}_{op,m+1}} - Y_{\overline{a}_{m-1}, 0_m, \underline{d}_{op,m+1}} | \overline{L}_{\overline{a}_{m-1}, m} = \overline{l}_m\right]$ induces a $\overline{d}_{op} - srSNNM$

$$\varrho^{\overline{d}_{op}}\left(\overline{l}_m, \overline{a}_m, \psi\right)$$
$$= \gamma^{\overline{0}}\left(\overline{l}_m, \overline{a}_m, \psi\right) - \gamma^{\overline{0}}\left(\overline{l}_m, \overline{a}_{m-1}, d_{op,m}\left(\overline{l}_m, \overline{a}_{m-1}\right), \psi\right)$$

for $E\left[Y_{\overline{a}_m, \underline{d}_{op,m+1}} - Y_{\overline{a}_{m-1}, \underline{d}_{op,m}} | \overline{L}_{\overline{a}_{m-1}, m} = \overline{l}_m\right]$, which is the negative regret. Further the induced regret model $-\varrho^{\overline{d}_{op}}\left(\overline{L}_m, \overline{A}_m, \psi\right)$ can, after some tedious calculation, always be reparametrized $u_m\left(\overline{l}_m, \overline{a}_m, \beta\right)$ where $u_m\left(\overline{l}_m, \overline{a}_m, \beta\right)$ satisfies Murphy's parametrization $u_m\left(\overline{l}_m, \overline{a}_m, \beta\right)$

$= \eta_m\left(\overline{l}_m, \overline{a}_{m-1}, \beta_{scale}\right) f\left(a_m - d_{op,m}\left(\overline{l}_m, \overline{a}_{m-1}, \beta_{regime}\right)\right)$. Conversely any Murphy regret model $u_m\left(\overline{l}_m, \overline{a}_m, \beta\right)$ induces a optimal drSNMM model $\gamma^{\overline{0}}\left(\overline{l}_m, \overline{a}_m, \beta\right)$ via $\gamma^{\overline{0}}\left(\overline{l}_m, \overline{a}_m, \beta\right) = u_m\left(\overline{l}_m, \overline{a}_{m-1}, 0_m, \beta\right) - u_m\left(\overline{l}_m, \overline{a}_m, \beta\right)$. It follows that there is a clear sense in which optimal drSNMM models and Murphy regret models are mathematically equivalent. In my opinion, however, the advantages of optimal drSNMM models are that (i) it is easier to directly specify scientifically meaningful models for a) the mean effect $\gamma^{\overline{d}_{op}, \overline{0}}\left(\overline{l}_m, \overline{a}_m\right)$ of treatment level a_m (versus level "zero") at m before following $\overline{d}_{op}$ from $m+1$ onwards than for b1) the scale component $\eta_m\left(\overline{l}_m, \overline{a}_{m-1}\right)$ and b2) the optimal treatment regime $d_{op,m}\left(\overline{l}_m, \overline{a}_{m-1}\right)$ of the Murphy parametrized regret $u_m\left(\overline{l}_m, \overline{a}_m\right)$ for given a function $f(\cdot)$, (ii) it is straightforward to compute both the optimal regime $d_{op,m}\left(\overline{l}_m, \overline{a}_{m-1}\right) = argmax_{a_m} \gamma^{\overline{d}_{op}, \overline{0}}\left(\overline{l}_m, \overline{a}_m\right)$ and the regret $\gamma^{\overline{d}_{op}, \overline{0}}\left(\overline{l}_m, \overline{a}_{m-1}, d_{op,m}\left(\overline{l}_m, \overline{a}_{m-1}\right)\right) - \gamma^{\overline{d}_{op}, \overline{0}}\left(\overline{l}_m, \overline{a}_m\right)$ from $\gamma^{\overline{d}_{op}, \overline{0}}\left(\overline{l}_m, \overline{a}_m\right)$, (iii) the map from $\gamma^{\overline{d}_{op}, \overline{0}}\left(\overline{l}_m, \overline{a}_m\right)$ to Murphy's $u_m\left(\overline{l}_m, \overline{a}_m\right)$ is tedious to compute and of no additional utility, and (iv) for a dr SNMM it is usually immediately obvious when it is possible to obtain sequential closed form estimates of the model parameters by noting whether $\gamma^{\overline{0}}\left(\overline{l}_m, \overline{a}_m, \psi\right)$ can be embedded in a model $\gamma^{\overline{0}}\left(\overline{l}_m, \overline{a}_m, \psi^*_m, \underline{\psi^*}_{m+1}\right)$ linear in ψ^*_m with ψ a function of $\underline{\psi^*}_0$.

6.2 Comparison of Optimal drSNMMS with DP-regression SNMMs

Estimation of an Optimal Regime with DP-regression SNMMs

In this subsection we describe how to use a DP-like regression model applied to an estimated srSNMM or drSNMM to estimate the optimal treatment regime. In the following subsection we compare and contrast the optimal regime drSNMM methodology with this DP-regression SNMM methodology.

To avoid complex notation we will study the case where our SNMM is a srSNMM and the single regime is the regime that is always 0. It may be a standard or non-standard $\overline{0}$ regime. Generalization to other srSNMMs and drSNMMs is straightforward and is given explicitly in Section 7.3 . Recall the srSNMM $\gamma^{\overline{0}}\left(\overline{l}_m, \overline{a}_m, \psi\right)$ is a model for for $\gamma^{\overline{0}, \overline{0}}\left(\overline{l}_m, \overline{a}_m\right)$. If $\gamma^{\overline{0}}\left(\overline{l}_m, \overline{a}_m, \psi\right)$ is smooth in ψ we can obtain lse-dr estimators of $\psi^\dagger$ as described in Section

3. We now provide a DP-like algorithm for computing the optimal treatment regime from knowledge of $\gamma^{\overline{0}}\left(\overline{l}_m, \overline{a}_m, \psi\right)$. It is a special case of Theorem 7.6 below. It can also be seen as the consequence of part (iii) of Theorem 3.3.

Theorem 6.1: Under sequential randomization (2.5), the following recursive DP-like algorithm computes $\overline{d}_{op}$. Define

$$q\left(\overline{L}_K, \overline{A}_{K-1}, a_K\right) = \gamma^{\overline{0},\overline{0}}\left(\overline{L}_K, \overline{A}_{K-1}, a_K\right)$$

For $m = K, ..., 0$ set

$$\begin{gathered} d_{op,m}\left(\overline{L}_m, \overline{A}_{m-1}\right) = \arg\max_{a_m} q\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right), \\ j\left(\overline{L}_m, \overline{A}_{m-1}\right) = q\left(\overline{L}_m, \overline{A}_{m-1}, d_{op,m}\left(\overline{L}_m, \overline{A}_{m-1}\right)\right), \\ q\left(\overline{L}_{m-1}, \overline{A}_{m-2}, a_{m-1}\right) = E\left[j\left(\overline{L}_m, \overline{A}_{m-1}\right) | \overline{L}_{m-1}, \overline{A}_{m-2}, A_{m-1} = a_{m-1}\right] + \\ \gamma^{\overline{0},\overline{0}}\left(\overline{L}_{m-1}, \overline{A}_{m-2}, a_{m-1}\right) \end{gathered}$$

Further $j\left(\overline{L}_m, \overline{A}_{m-1}\right) = E\left[Y_{\overline{A}_{m-1},\underline{d}_{op,m}} - Y_{\overline{A}_{m-1},\underline{0}_m} | \overline{L}_m, \overline{A}_{m-1}\right]$ and thus $E\left[Y_{\overline{d}_{op}}\right] = E\left[j\left(\overline{L}_0, \overline{A}_{-1}\right)\right] + E\left[Y_{\overline{0}}\right]$.

Note $j\left(\overline{L}_m, \overline{A}_{m-1}\right)$ measures the difference in average utility of subjects with observed history $\left(\overline{L}_m, \overline{A}_{m-1}\right)$ were they were to follow the optimal regime from time t_m onward rather than the 0 regime. The above theorem motivates the following.

DP-regression srSNMM Fitting Algorithm: Let $\widehat{\psi}$ be a dr-lse efficient estimator of the parameter $\psi^\dagger$ of a srSNMM calculated under the union model of Theorem 3.4. For $m = K, ..., 1$, we specify regression models

$$E\left[j\left(\overline{L}_m, \overline{A}_{m-1}\right) | \overline{L}_{m-1}, \overline{A}_{m-2}, A_{m-1} = a_{m-1}\right] = r\left(\overline{L}_{m-1}, \overline{A}_{m-2}, a_{m-1}; \beta_m\right)$$

and compute $Q_K\left(\widehat{\psi}, a_K\right) = \gamma^{\overline{0}}\left(\overline{L}_K, \overline{A}_{K-1}, a_K, \widehat{\psi}\right)$. Then recursively, for $m = K, ..., 0$, (with β_{K+1} the null set),

$$\begin{gathered} d_{op,m}\left(\overline{L}_m, \overline{A}_{m-1}, \widehat{\psi}, \underline{\widehat{\beta}}_{m+1}\right) = \arg\max_{a_m} Q_m\left(\widehat{\psi}, a_m, \underline{\widehat{\beta}}_{m+1}\right), \\ J_m\left(\widehat{\psi}, \underline{\widehat{\beta}}_{m+1}\right) = Q_m\left(\widehat{\psi}, d_{op,m}\left(\overline{L}_m, \overline{A}_{m-1}, \widehat{\psi}, \underline{\widehat{\beta}}_{m+1}\right), \underline{\widehat{\beta}}_{m+1}\right), \\ Q_{m-1}\left(\widehat{\psi}, a_{m-1}; \underline{\widehat{\beta}}_m\right) = r\left(\overline{L}_{m-1}, \overline{A}_{m-2}, a_{m-1}; \widehat{\beta}_m, \widehat{\psi}, \underline{\widehat{\beta}}_{m+1}\right) + \\ \gamma^{\overline{0}}\left(\overline{L}_{m-1}, \overline{A}_{m-2}, a_{m-1}, \widehat{\psi}\right), \end{gathered}$$

where $\widehat{\beta}_m$ is the possibly non-linear least squares estimate of β_m from the regression of $J_m\left(\widehat{\psi}, \underline{\widehat{\beta}}_{m+1}\right)$ on $\overline{L}_{m-1}, \overline{A}_{m-2}, A_{m-1} = a_{m-1}$ based on the regression function $r\left(\overline{L}_{m-1}, \overline{A}_{m-2}, a_{m-1}; \beta_m\right)$.

Finally calculate

$$\widehat{E}_{\widehat{\psi},\underline{\widehat{\beta}}_0}\left[Y_{\overline{d}_{op,m}}\right] = P_n\left[J_0\left(\widehat{\psi},\underline{\widehat{\beta}}_1\right)\right] + P_n\left[H_0^{\underline{0},\underline{0}}\left(\widehat{\psi}\right)\right]$$

Note that if the $H_m^{\underline{0},\underline{0}}(\psi)$ are linear in ψ as for the model $\gamma^{\overline{0}}\left(\overline{l}_m,\overline{a}_m,\psi\right) = a_m\left(1,l_m,a_{m-1}\right)\psi$, then $\widehat{\psi}$ will exist in closed form. If the $q\left(\overline{L}_{m-1},\overline{A}_{m-2},a_{m-1};\beta_m\right)$ are linear in β_m and $\widehat{\beta}_m$ is the OLS estimator, then $d_{op,m}\left(\overline{L}_m,\overline{A}_{m-1},\widehat{\psi},\underline{\widehat{\beta}}_{m+1}\right)$ will exist in closed form (provided the argmax function can be evaluated in closed form). However the model will share the inferential difficulties we noted in our study of the closed-form inefficient estimator of an optimal drSNMM. Specifically although the parameter ψ of our srSNMM, in contrast to that of an optimal drSNMM, is a regular parameter, the parameters $\beta_m = \beta_m\left(\psi,\underline{\beta}_{m+1}\right)$ for $m < K$ are not, as $\beta_m\left(\psi,\underline{\beta}_{m+1}\right)$ is not an everywhere differentiable function of $\left(\psi,\underline{\beta}_{m+1}\right)$. Specifically, $J_m\left(\widehat{\psi},\underline{\widehat{\beta}}_{m+1}\right) = Q_m\left(\widehat{\psi},d_{op,m}\left(\overline{L}_m,\overline{A}_{m-1},\widehat{\psi},\underline{\widehat{\beta}}_{m+1}\right),\underline{\widehat{\beta}}_{m+1}\right)$ and, for dichotomous A_m, $d_{op,m}\left(\overline{L}_m,\overline{A}_{m-1},\widehat{\psi},\underline{\widehat{\beta}}_{m+1}\right)$ will jump from 1 to 0 or vice-versa as $\widehat{\psi}$ or $\underline{\widehat{\beta}}_{m+1}$ are continuously varied. However a large sample confidence interval for $\overline{d}_{op,m}$ can be obtained because (i) $d_{op,m}$ is a function of $\left(\psi,\underline{\beta}_{m+1}\right)$ and (ii) a joint large sample confidence interval for $(\psi,\beta) = \left(\psi,\underline{\beta}_1\right)$ can be obtained based on inverting a χ^2 statistic for the joint estimating functions for (ψ,β) (which for β are the least squares normal equations). Because the estimating functions for ψ do not depend on β, the confidence interval is guaranteed to be valid under the g-null hypothesis when the union model of theorem 3.4 is correct.

It follows that the misspecification of the regression model $E\left[j\left(\overline{L}_m,\overline{A}_{m-1}\right)|\overline{L}_{m-1},\overline{A}_{m-2},A_{m-1}\right] = r\left(\overline{L}_{m-1},\overline{A}_{m-2},A_{m-1};\beta_m\right)$ does not lead to bias in estimating $d_{op,m}\left(\overline{L}_m,\overline{A}_{m-1}\right)$ or $E\left[Y_{\overline{d}_{op,m}}\right]$ under the g-null hypothesis that $\psi^\dagger$ is zero as $q\left(\overline{L}_m,\overline{A}_{m-1},a_m\right)$ and $j\left(\overline{L}_m,\overline{A}_{m-1}\right)$ are identically zero. To obtain a consistent estimate of $\overline{d}_{op}$ under the alternative $\psi^\dagger \neq 0$, correct specification of this regression model will be necessary, which is not feasible due to the high dimension of the vector $\left(\overline{L}_{m-1},\overline{A}_{m-1}\right)$. However, I do not consider this to be a major shortcoming of the DP-regression srSNMM methodology compared to the optimal drSNMM methodology for the following reason. To obtain a consistent estimate of $\overline{d}_{op}$ under an optimal drSNMM, the optimal drSNMM $\gamma^{\overline{0}}\left(\overline{l}_m,\overline{a}_m,\psi\right)$ for $\gamma^{\overline{d_{op}},\overline{0}}\left(\overline{l}_m,\overline{a}_m\right)$ must be correct, but this is also not feasible due to the high dimension of $\left(\overline{L}_m,\overline{A}_m\right)$. That is, because of the high dimensionality of the problem, no method can provide a consistent estimator for $\overline{d}_{op}$ under the alternative, even when the treatment probabilities are known. The question is then do we expect to obtain less biased estimates of $\overline{d}_{op}$ with the DP-regression srSNMM methodology that requires us to specify models for

both $\gamma^{\overline{0},\overline{0}}\left(\overline{l}_m,\overline{a}_m\right)$ and $E\left[j\left(\overline{L}_m,\overline{A}_{m-1}\right)|\overline{L}_{m-1},\overline{A}_{m-2},A_{m-1}=a_{m-1}\right]$ or with the optimal drSNMM that requires a model for $\gamma^{\overline{d_{op}},\overline{0}}\left(\overline{l}_m,\overline{a}_m\right)$. In general that will depend on whether it is easier to use our substantive subject-matter knowledge to model $\gamma^{\overline{d_{op}},\overline{0}}\left(\overline{l}_m,\overline{a}_m\right)$ or to model both $\gamma^{\overline{0},\overline{0}}\left(\overline{l}_m,\overline{a}_m\right)$ and $E\left[j\left(\overline{L}_m,\overline{A}_{m-1}\right)|\overline{L}_{m-1},\overline{A}_{m-2},A_{m-1}=a_{m-1}\right]$. There is no general rule as to which is easier even when we use a standard (i.e. substantively meaningful) zero regime. To understand why it is only an issue about the ease of applying substantiative knowledge, I will now show there is a precise sense in which fitting optimal drSNMM models and DP- srSNMM models can be made algebraically equivalent. To do so I shall use the following Lemma.

Lemma 6.1: Under sequential randomization (2.5),

$$\begin{aligned}&\gamma^{\overline{d}_{op},\overline{0}}\left(\overline{L}_m,\overline{A}_m\right)-\gamma^{\overline{d}_{op},\overline{0}}\left(\overline{L}_m,\overline{A}_{m-1},d_{op,m}\left(\overline{L}_m,\overline{A}_{m-1}\right)\right)\\&=\gamma^{\overline{0},\overline{0}}\left(\overline{L}_m,\overline{A}_m\right)+j\left(\overline{L}_m,\overline{A}_{m-1}\right)-E\left[j\left(\overline{L}_{m+1},\overline{A}_m\right)|\overline{L}_m,\overline{A}_m\right]\\&=j\left(\overline{L}_m,\overline{A}_{m-1}\right)+q\left(\overline{L}_m,\overline{A}_m\right).\end{aligned}$$

Further

$$H_m^{\overline{d}_{op},\overline{0}}-E\left[H_m^{\overline{d}_{op},\overline{0}}|\overline{L}_m,\overline{A}_{m-1}\right]=H_m^{\overline{0},\overline{0}}-E\left[H_m^{\overline{0},\overline{0}}|\overline{L}_m,\overline{A}_{m-1}\right]+Z_m,$$

with

$$Z_m=\sum_{j=m+1}^{K}j\left(\overline{L}_j,\overline{A}_{j-1}\right)-E\left[j\left(\overline{L}_j,\overline{A}_{j-1}\right)|L_{j-1},\overline{A}_{j-1}\right]-$$

$$E\left[\sum_{j=m+1}^{K}j\left(\overline{L}_j,\overline{A}_{j-1}\right)-E\left[j\left(\overline{L}_j,\overline{A}_{j-1}\right)|L_{j-1},\overline{A}_{j-1}\right]|\overline{L}_m,\overline{A}_{m-1}\right]$$

.

Proof: By definition $\gamma^{\overline{d}_{op},\overline{0}}\left(\overline{L}_m,\overline{A}_m\right)-\gamma^{\overline{d}_{op},\overline{0}}\left(\overline{L}_m,\overline{A}_{m-1},d_{op,m}\left(\overline{L}_m,\overline{A}_{m-1}\right)\right)-\gamma^{\overline{0},\overline{0}}\left(\overline{L}_m,\overline{A}_m\right)$

$=-E\left[Y_{\overline{A}_{m-1},\underline{d}_{op,m}}-Y_{\overline{A}_m,\underline{d}_{op,m+1}}+Y_{\overline{A}_m,\underline{0}_{m+1}}-Y_{\overline{A}_{m-1},\underline{0}_m}|\overline{L}_m,\overline{A}_m\right]$

$=-E\left[Y_{\overline{A}_{m-1},\underline{d}_{op,m}}-Y_{\overline{A}_{m-1},\underline{0}_m}|\overline{L}_m,\overline{A}_m\right]+E\left[-Y_{\overline{A}_m,\underline{d}_{op,m+1}}+Y_{\overline{A}_m,\underline{0}_{m+1}}|\overline{L}_m,\overline{A}_m\right]$

$=-\left\{j\left(\overline{L}_m,\overline{A}_{m-1}\right)-E\left[j\left(\overline{L}_{m+1},\overline{A}_m\right)|\overline{L}_m,\overline{A}_m\right]\right\}$, where we have used sequential randomization in the final step. Thus by Lemma 6.1, a model $\gamma^{\overline{0},\overline{0}}\left(\overline{L}_m,\overline{A}_m,\psi\right)$ for $\gamma^{\overline{0},\overline{0}}\left(\overline{L}_m,\overline{A}_m\right)$ plus regression models

$$E\left[j\left(\overline{L}_m,\overline{A}_{m-1}\right)|\overline{L}_{m-1},\overline{A}_{m-1}\right]=r\left(\overline{L}_{m-1},\overline{A}_{m-1};\beta_m\right),$$

$m=K,...,1$, induce a model $\gamma^{\overline{d}_{op},\overline{0}}\left(\overline{L}_m,\overline{A}_m,\psi,\underline{\beta}_m\right)$ for $\gamma^{\overline{d}_{op},\overline{0}}\left(\overline{L}_m,\overline{A}_m\right)$. Conversely a model $\gamma^{\overline{d}_{op},\overline{0}}\left(\overline{L}_m,\overline{A}_m,\psi\right)$ for $\gamma^{\overline{d}_{op},\overline{0}}\left(\overline{L}_m,\overline{A}_m\right)$ plus regression models $E\left[j\left(\overline{L}_m,\overline{A}_{m-1}\right)|\overline{L}_{m-1},\overline{A}_{m-1}\right]=r\left(\overline{L}_{m-1},\overline{A}_{m-1};\beta_m\right)$ induces a model $\gamma^{\overline{0},\overline{0}}\left(\overline{L}_m,\overline{A}_m,\psi,\underline{\beta}_m\right)$ for $\gamma^{\overline{0},\overline{0}}\left(\overline{L}_m,\overline{A}_m\right)$.

Given models $\gamma^{\overline{0},\overline{0}}\left(\overline{L}_m, \overline{A}_m, \psi\right)$ for $\gamma^{\overline{0},\overline{0}}\left(\overline{L}_m, \overline{A}_m\right)$ plus regression models $E\left[j\left(\overline{L}_m, \overline{A}_{m-1}\right)|\overline{L}_{m-1}, \overline{A}_{m-1}\right] = r\left(\overline{L}_{m-1}, \overline{A}_{m-1}; \beta_m\right)$ an alternative way to estimate $\psi^\dagger$ would be to solve an estimating equation based on the induced optimal drSNMM model. For example we can solve $0 = P_n\left[U^{\dagger \overline{d}_{op},\overline{0}}\left(\psi, \underline{\widehat{\beta}}_0\left(\psi\right), s, c^{s,\overline{d}_{op},\overline{0}}, \widehat{\varsigma}\right)\right]$ where $\underline{\widehat{\beta}}_0\left(\psi\right)$ is obtained as in the DP srSNMM fitting algorithm. But from Lemma (6.1), this will equal

$$P_n\left[U^{\dagger \overline{0},\overline{0}}\left(\psi, s, c^{s,\overline{0},\overline{0}}, \widehat{\varsigma}\right)\right] +$$
$$P_n\left[\sum_{m=0}^{K} Z_m\left(\psi, \underline{\widehat{\beta}}_{m+1}\left(\psi\right), \widehat{\varsigma}\right)\left\{S_m\left(A_m\right) - E\left[S_m\left(A_m\right) \mid \overline{A}_{m-1}, \overline{L}_m\right]\right\}\right].$$

We can always choose our regression models $E\left[j\left(\overline{L}_m, \overline{A}_{m-1}\right)|\overline{L}_{m-1}, \overline{A}_{m-1}\right] = r\left(\overline{L}_{m-1}, \overline{A}_{m-1}; \beta_m\right)$ such that $P_n\left[\sum_{m=0}^{K} Z_m\left(\psi, \underline{\widehat{\beta}}_{m+1}\left(\psi\right), \widehat{\varsigma}\right)\left\{S_m\left(A_m\right) - E\left[S_m\left(A_m\right) \mid \overline{A}_{m-1}, \overline{L}_m\right]\right\}\right]$ is zero with probability one. Specifically we choose $r\left(\overline{L}_{m-1}, \overline{A}_{m-1}; \beta_m\right) = \beta_m^T W_m$ with W_m including each $\left\{S_j\left(A_j\right) - E\left[S_j\left(A_j\right) \mid \overline{A}_{j-1}, \overline{L}_j\right]\right\}$ as a covariate for every $j < m$.

This will guarantee that we obtain the exact same estimates of ψ and $\overline{d}_{op}$ by directly solving the induced optimal drSNMM estimating equation $0 = P_n\left[U^{\dagger \overline{d}_{op},\overline{0}}\left(\psi, \underline{\widehat{\beta}}_0\left(\psi\right), s, c^{s,\overline{d}_{op},\overline{0}}, \widehat{\varsigma}\right)\right]$ as by first solving $P_n\left[U^{\dagger \overline{0},\overline{0}}\left(\psi, s, c^{s,\overline{0},\overline{0}}, \widehat{\varsigma}\right)\right] = 0$ and then implementing the DP srSNMM fitting algorithm. Thus the only issue is whether it is an easier substantive task to model $\gamma^{\overline{d_{op}},\overline{0}}\left(\overline{l}_m, \overline{a}_m\right)$ or to model both $\gamma^{\overline{0},\overline{0}}\left(\overline{l}_m, \overline{a}_m\right)$ and $E\left[j\left(\overline{L}_m, \overline{A}_{m-1}\right)|\overline{L}_{m-1}, \overline{A}_{m-2}, A_{m-1} = a_{m-1}\right]$, as either model can then be fit using the methods described in section 4 for fitting optimal drSNMM models.

An example of a setting in which it might be easier to model $\gamma^{\overline{d_{op}},\overline{0}}\left(\overline{l}_m, \overline{a}_m\right)$ is one in which one believes that current medical practice fluctuates around the optimal regime, because the direct experience of clinicians will then be with subjects who followed treatment plans close to the optimal, providing a basis for developing a good intuition for the functional form of $\gamma^{\overline{d_{op}},\overline{0}}\left(\overline{l}_m, \overline{a}_m\right)$.

7 Sensitivity Analysis and Decisions with Information Loss:

In subsections 7.1 and 7.2 we no longer assume that sequential randomization holds and develops a sensitivity analysis methodology. In the next subsection, we develop two simple methods for estimating a $\left(\overline{d}, \overline{d}^*\right)$ drSNMM model $\gamma^{\overline{d}^*}\left(\overline{l}_m, \overline{a}_m; \psi\right)$ and the corresponding regime -specific mean $E\left[Y^{\overline{d}}\right]$ in

the absence of sequential randomization. The first method requires us to treat as known two different nonidentifiable functions unless $\overline{d} = \overline{d}^*$. These functions are then varied in a sensitivity analysis. The second, more parsimonious, method only requires us to treat as known a single nonidentifiable function. However we shall see in section 7.2 that two nonidentifiable functions will always be required to estimate $E\left[Y^{\overline{d}_{op}}\right]$ precisely because $\overline{d}_{op}$ is not known. In Section 7.2 and 7.3, we consider settings in which the decison maker can only use a subset of the past information to make a current decision.

7.1 Regime Specific SNMMs

Method 1:

We turn to our first method. Under the assumption of sequential randomization, the following function is identically zero for each m. Define

$$r^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right) = r^{\underline{d}_m, d_m^*}\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right) =$$
$$E\left[Y_{\overline{A}_{m-1}, a_m, \underline{d}_{m+1}} - Y_{\overline{A}_{m-1}, d_m^*, \underline{d}_{m+1}} | \overline{L}_m, \overline{A}_{m-1}, A_m = a_m\right] - \tag{7.1}$$
$$E\left[Y_{\overline{A}_{m-1}, a_m, \underline{d}_{m+1}} - Y_{\overline{A}_{m-1}, d_m^*, \underline{d}_{m+1}} | \overline{L}_m, \overline{A}_{m-1}, A_m \neq a_m\right]$$

We refer to $r^{\underline{d}_m, d_m^*}\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right)$ as a regime $d - specific$ current treatment interaction function since, among subjects with history $\left(\overline{L}_m, \overline{A}_{m-1}\right)$, it compares the magnitude of the effect of a last blip of treatment of dose a_m compared to dose $d_m^*\left(\overline{L}_m, \overline{A}_{m-1}\right)$ before following regime d among those who received treatment a_m to the same effect among those who did not receive the treatment a_m. If, as in this section, we do not assume sequential randomization but do assume the support of *each* counterfactual response may be the whole real line, then the function $r^{\underline{d}_m, d_m^*}\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right)$ is completely nonidentified in the sense that the distribution of the observed data O places no restrictions on $r^{\underline{d}_m, d_m^*}\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right)$ except for the definitional restriction that $r^{\underline{d}_m, d_m^*}\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right) = 0$ if $a_m = d_m^*\left(\overline{L}_m, \overline{A}_{m-1}\right)$. This is immediately evident because there can be no data evidence restricting the mean of $Y_{\overline{A}_{m-1}, a_m, \underline{d}_{m+1}}$ among subjects with $A_m \neq a_m$. Thus we will regard the unidentified function $r^{\underline{d}_m, d_m^*}\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right)$ as known and vary it in a sensitivity analysis. Note that under the non-identifiable assumption of additive local rank preservation $r^{\underline{d}_m, d_m^*}\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right) \equiv 0$. Further $r^{\underline{d}_m, d_m^*}\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right) \equiv 0$ under the sharp null hypothesis that $Y_{\overline{a}} = Y$ for all $\overline{a} \in \overline{\mathcal{A}}$ $w.p.1$ of no treatment effect. Thus if one wishes to test the sharp null hypothesis one must do so assuming $r^{\underline{d}_m, d_m^*}\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right) \equiv 0$. Define the function $r^{\underline{d}_m, d_m^*}\left(\overline{L}_m, \overline{A}_{m-1}\right)$ of $\left(\overline{L}_m, \overline{A}_{m-1}\right)$ by

$$r^{\underline{d}_m, d_m^*}\left(\overline{L}_m, \overline{A}_{m-1}\right) \equiv r^{\underline{d}_m, d_m^*}\left(\overline{L}_m, \overline{A}_{m-1}, d_m\left(\overline{L}_m, \overline{A}_{m-1}\right)\right).$$

The following theorem states that knowledge of $r^{\underline{d}_m,d_m^*}\left(\overline{L}_m,\overline{A}_{m-1}\right)$ plus identification of $\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_{m-1},a_m\right)$ suffices to identify $E\left[Y^{\overline{d}}\right]$ and more generally $E\left[Y_{\overline{A}_{m-1},\underline{d}_m}|\overline{L}_m,\overline{A}_{m-1}\right]$. Note by the definition of $\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_{m-1},a_m\right)$,

$$r^{\underline{d}_m,d_m^*}\left(\overline{L}_m,\overline{A}_{m-1}\right)=$$
$$\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_{m-1},d_m\left(\overline{L}_m,\overline{A}_{m-1}\right)\right)-$$
$$E\left[Y_{\overline{A}_{m-1},d_m,\underline{d}_{m+1}}-Y_{\overline{A}_{m-1},d_m^*,\underline{d}_{m+1}}|\overline{L}_m,\overline{A}_{m-1},A_m\neq d_m\left(\overline{L}_m,\overline{A}_{m-1}\right)\right]$$

We shall need the following definitions which reduce to our previous definitions (3.5) under sequential randomization. Let $\gamma^{\overline{d}*}*\left(\overline{L}_m,\overline{A}_m\right)$ satisfy (3.6) and define $a_m=d_m\left(\overline{L}_m,\overline{A}_{m-1}\right)$ for the remainder of this paragraph. Define

$H_{K+1}\left(\gamma^{\overline{d}^**}\right)=Y,$

$$H_K\left(\gamma^{\overline{d}^**}\right)=Y-\gamma^{\overline{d}^**}\left(\overline{L}_K,\overline{A}_K\right)+$$
$$\gamma^{\overline{d}^**}\left(\overline{L}_K,\overline{A}_{K-1},d_K\left(\overline{L}_K,\overline{A}_{K-1}\right)\right)-\ r^{\underline{d}_K,d_K^*}\left(\overline{L}_K,\overline{A}_{K-1}\right)\left\{1-f\left(a_K|\overline{L}_K,\overline{A}_{K-1}\right)\right\}$$
,

$$H_m^{\underline{d}_m}\left(\gamma^{\overline{d}^**}\right)=$$
$$H_{m+1}^{\underline{d}_{m+1}}\left(\gamma^{\overline{d}^**}\right)+\gamma^{\overline{d}^**}\left(\overline{L}_m,\overline{A}_{m-1},d_m\left(\overline{L}_m,\overline{A}_{m-1}\right)\right)-$$
$$r^{\underline{d}_m,d_m^*}\left(\overline{L}_m,\overline{A}_{m-1}\right)\left\{1-f\left(a_m|\overline{L}_m,\overline{A}_{m-1}\right)\right\}-\gamma^{\overline{d}^**}\left(\overline{L}_m,\overline{A}_m\right)$$
$$=Y-\sum_{j=m}^{K}\gamma^{\overline{d}^**}\left(\overline{L}_j,\overline{A}_j\right)+$$
$$\sum_{j=m}^{K}\gamma^{\overline{d}^**}\left(\overline{L}_j,\overline{A}_{j-1},d_j\left(\overline{L}_j,\overline{A}_{j-1}\right)\right)-\ r^{\underline{d}_j,d_j^*}\left(\overline{L}_j,\overline{A}_{j-1}\right)\left\{1-f\left(a_j|\overline{L}_j,\overline{A}_{j-1}\right)\right\}$$
.

Note $H_m^{\underline{d}_m}\left(\gamma^{\overline{d},\overline{d}^*}\right)$ depends only on the data O and the functions $r^{\underline{d}_m,d_m^*}$ and $\gamma^{\overline{d},\overline{d}^*}$. In appendix 3 we prove the following.

Theorem 7.1: With $H_m^{\underline{d}_m}\left(\gamma^{\overline{d}^**}\right)$ as defined in the previous paragraph, $E\left[H_m^{\underline{d}_m}\left(\gamma^{\overline{d}^**}\right)|\overline{L}_m,\overline{A}_m\right]=E\left[Y_{\overline{A}_{m-1},\underline{d}_m}|\overline{L}_m,\overline{A}_m\right]$ for all m if and only if

$$\gamma^{\overline{d}^**}\left(\overline{L}_m,\overline{A}_m\right)=\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)\ w.p.1 \text{ for all } m \tag{7.2}$$

In particular $E\left[Y_{\overline{d}}\right]=E\left[H_0\left(\gamma^{\overline{d}}\right)\right]$.

To see why we require $r^{\underline{d}_K,d_K^*}\left(\overline{L}_K,\overline{A}_{K-1}\right)$ in addition to $\gamma^{\overline{d},\overline{d}^*}\left(\overline{l}_m,\overline{a}_m\right)$ to identify $E\left[Y_{\overline{A}_{K-1},d_K}|\overline{L}_K,\overline{A}_{K-1}\right]$ first note that knowledge of $\gamma^{\overline{d},\overline{d}^*}\left(\overline{l}_K,\overline{a}_K\right)$ allows us to identify $E\left[Y_{\overline{A}_{K-1},d_K^*}|\overline{L}_K,\overline{A}_{K-1},A_K=a_K\right]$ for each a_K and thus to identify $E\left[Y_{\overline{A}_{K-1},d_K^*}|\overline{L}_K,\overline{A}_{K-1}\right]$. Since we know $E\left[Y_{\overline{A}_{K-1},d_K}-Y_{\overline{A}_{K-1},d_K^*}|\overline{L}_K,\overline{A}_{K-1},A_K=d_K\left(\overline{L}_K,\overline{A}_{K-1}\right)\right]$, knowledge of $r^{\underline{d}_K,d_K^*}\left(\overline{L}_K,\overline{A}_{K-1}\right)$ allows us to calculate $E\left[Y_{\overline{A}_{K-1},d_K}-Y_{\overline{A}_{K-1},d_K^*}|\overline{L}_K,\overline{A}_{K-1},A_K\neq d_K\left(\overline{L}_K,\overline{A}_{K-1}\right)\right]$ and, using the law of O, $E\left[Y_{\overline{A}_{K-1},d_K}-Y_{\overline{A}_{K-1},d_K^*}|\overline{L}_K,\overline{A}_{K-1}\right]$ as well. Thus we can compute $E\left[Y_{\overline{A}_{K-1},d_K}|\overline{L}_K,\overline{A}_{K-1}\right]$.

In the absence of sequential randomization, knowledge of $r^{\underline{d}_m,d_m^*}\left(\overline{L}_m,\overline{A}_{m-1}\right)$ is not sufficient to identify $\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$ from the law of O. Now under sequential randomization the function

$$v^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)=E\left[Y_{\overline{A}_{m-1},d_m^*,\underline{d}_{m+1}}|\overline{L}_m,\overline{A}_m\right]-E\left[Y_{\overline{A}_{m-1},d_m^*,\underline{d}_{m+1}}|\overline{L}_m,\overline{A}_{m-1},A_m=d_m^*\left(\overline{L}_m,\overline{A}_{m-1}\right)\right]$$

takes the value zero. Hence $v^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$ is a measure of the magnitude of confounding due to unmeasured factors among subjects with history $\left(\overline{L}_m,\overline{A}_{m-1}\right)$, as it compares the mean of the same counterfactual in those who received treatment A_m to that in those who received treatment $d_m^*\left(\overline{L}_m,\overline{A}_{m-1}\right)$at t_m. Further knowledge of the law of O and of $r^{\underline{d}_m,d_m^*}\left(\overline{L}_m,\overline{A}_{m-1}\right)$ together place no restrictions on $v^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$ beyond the definitional restriction that $v^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)=0$ if $A_m=d_m^*\left(\overline{L}_m,\overline{A}_{m-1}\right)$. Thus we will regard the unidentified functions $r^{\underline{d}_m,d_m^*}\left(\overline{L}_m,\overline{A}_{m-1}\right)$ and $v^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$ both as known and vary both in a sensitivity analysis. The following theorem, proved in Appendix 3, states that knowledge of $v^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$ and $r^{\underline{d}_m,d_m^*}\left(\overline{L}_m,\overline{A}_{m-1}\right)$ identifies $\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$ and thus by Theorem 7.1 $E\left[Y^{\overline{d}}\right]$ as well.

Theorem 7.2: Let $H_{m+1}^{\underline{d}_m}\left(\gamma^{\overline{d}^*}*\right)$ be defined as in Theorem 7.1. Then

$$E\left[H_{m+1}^{\underline{d}_m}\left(\gamma^{\overline{d}^*}*\right)-\gamma^{\overline{d}^*}*\left(\overline{L}_m,\overline{A}_m\right)-v^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)|\overline{L}_m,\overline{A}_m\right]$$

is not a function of A_m for all m if and only if Eq (7.2) of Theorem 7.1 holds if and only if $E[H_m^{\underline{d}_m}\left(\gamma^{\overline{d}^*}*\right)-v^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)|\overline{L}_m,\overline{A}_m]$ is not a function of A_m for all m.

Inference under Method 1:

The following two corollaries are special cases of Theorem 4.3 in Robins and Rotnizky (2003)

Corollary 7.2a: Consider again the semiparametric models (a.1) - (a.3) of Theorem 3.3, except now the assumption of sequential randomization is replaced by the assumption that the functions $v^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$ and $r^{\underline{d}_m,d_m^*}\left(\overline{L}_m,\overline{A}_{m-1}\right)$ are known (but may be non-zero). Then,
a) part *(i)* of Theorem 3.3 remains true if we replace $H_m^{\overline{d},\overline{d}^*}(\psi)$ in Eq. (3.9) with $H_m^{\overline{d},\overline{d}^*}(\psi) - v^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$ where (a) $H_m^{\overline{d},\overline{d}^*}(\psi)$ is now $H_{m+1}^{\underline{d}_m}\left(\gamma^{\overline{d}^**}\right)$ as defined before Theorem 7.1 with $\gamma^{\overline{d}^**}\left(\overline{L}_m,\overline{A}_m\right) = \gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m,\psi\right)$ being the dr SNMM model $\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m,\psi\right)$ for $\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$. However, part *(ii)* and (iii) must be modified as follows.
b) When $v^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$ and $r^{\underline{d}_m,d_m^*}\left(\overline{L}_m,\overline{A}_{m-1}\right)$ are identically zero, (ii) of Theorem 3.3 holds. However, if $v^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$ and $r^{\underline{d}_m,d_m^*}\left(\overline{L}_m,\overline{A}_{m-1}\right)$ are not identically zero, then $\widehat{\psi}(s,c^s) \equiv \widehat{\psi}\left(s,c^s,\alpha^\dagger\right)$ has asymptotic variance less than or equal to that of $\widehat{\psi}(s,c^s,\widehat{\alpha})$ which has an asymptotic variance less than or equal to that of $\widehat{\psi}(s,c^s,\widehat{\alpha}_{smooth})$. Further

$$C_m^s - E\left[C_m^s|\overline{A}_{m-1},\overline{L}_m\right] \neq E\left[U_m\left(\psi^\dagger,s\right) \mid \overline{A}_m,\overline{L}_m\right] - E\left[U_m\left(\psi^\dagger,s\right)|\overline{A}_{m-1},\overline{L}_m\right]$$

and thus

$$C^s \neq \widetilde{C}^s \equiv \sum_m \widetilde{C}_m^s - E\left[\widetilde{C}_m^s|\overline{A}_{m-1},\overline{L}_m\right] \text{ with } \widetilde{C}_m^s = E\left[U_m\left(\psi^\dagger,s\right) \mid \overline{A}_m,\overline{L}_m\right]$$

Explicit expressions for the influence functions are as follows. $\widehat{\psi}(s,c)$ has influence function $I^{-1}U^\dagger(s,c)$ with variance
$I^{-1}\left\{E\left[U^\dagger(s,c)^{\otimes 2}\right]\right\}I^{-1,T}$ where
$E\left[U^\dagger(s,c)^{\otimes 2}\right] = E\left[U^\dagger(s,c^s)^{\otimes 2}\right] + E\left[(C_s - C)^{\otimes 2}\right]$, $\widehat{\psi}(s,c,\widehat{\alpha})$ has influence function $I^{-1}\left\{U^\dagger(s,c) + E[\partial U^\dagger(s,c,\alpha)/\partial\alpha]\left\{E\left(S_{part}^{\otimes 2}\right]\right\}^{-1}S_{part}\right\}$ with variance $I^{-1}\times$

$$\left\{E\left[U^\dagger(s,c^s)^{\otimes 2}\right] + E\left[\left\{C_s - C + E\left[\partial U^\dagger(s,c,\alpha)/\partial\alpha\right]\left[S_{part}^{\otimes 2}\right]^{-1}S_{part}\right\}^{\otimes 2}\right]\right\}I^{-1,T}.$$

Finally, $\widehat{\psi}(s,c,\widehat{\alpha}_{smooth})$ has has influence function $I^{-1}U^\dagger(s,\widetilde{c}^s)$
with variance $I^{-1}\left\{E\left[U^\dagger(s,\widetilde{c}^s)^{\otimes 2}\right]\right\}I^{-1,T}$.
c) When $v^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$ and $r^{\underline{d}_m,d_m^*}\left(\overline{L}_m,\overline{A}_{m-1}\right)$ are identically zero, (iii) of Theorem 3.3 holds. However, if $v^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$ and $r^{\underline{d}_m,d_m^*}\left(\overline{L}_m,\overline{A}_{m-1}\right)$ are not identically zero, the semiparametric variance bound is smallest in model (a.1) and greatest in model (a.3). The bound in model (a.3) is equal to the

asymptotic variance of $\widehat{\psi}\left(s_{eff}, \widetilde{c}^{s_{eff}}\right)$. However a simple explicit expression for s_{eff} is no longer available since even when

$$var\left(H_m\left(\psi^\dagger\right) - v^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_m\right)|\overline{L}_m, \overline{A}_m\right)$$
$$= var\left(H_m\left(\psi^\dagger\right) - v^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_m\right)|\overline{L}_m, \overline{A}_{m-1}\right)$$

does not depend on A_m for all m, it is no longer the case that, when $v^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_m\right)$ is non-zero, $E\left[U_m^\dagger\left(s, \widetilde{c}\right) U_j^\dagger\left(s, \widetilde{c}\right)|\overline{L}_m, \overline{A}_m\right] = 0$ for $m \neq j$ where $U_m^\dagger\left(s, \widetilde{c}\right) = U_m\left(s\right) - \left\{\widetilde{C}_m^s - E\left[\widetilde{C}_m^s|\overline{A}_{m-1}, \overline{L}_m\right]\right\}$.

Next consider a model $\varsigma^T W_m$ for $E\left[H_m\left(\psi^\dagger\right) - v^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_m\right)|\overline{L}_m, \overline{A}_{m-1}\right]$ where $W_m = w_m\left(\overline{A}_{m-1}, \overline{L}_m\right)$ is a known vector function of $\left(\overline{A}_{m-1}, \overline{L}_m\right)$ and ς is an unknown parameter. Define $\widehat{\psi}\left(s, \widetilde{c}^s; \widehat{\alpha}, \widehat{\varsigma}\right)$ to be a solution to $P_n\left[U^\dagger\left(\psi^\dagger, s, \widetilde{c}^s; \widehat{\alpha}, \widehat{\varsigma}\right)\right] = 0$ where

$$U^\dagger\left(\psi^\dagger, s, \widetilde{c}^s; \widehat{\alpha}, \widehat{\varsigma}\right) = \sum_{m=0}^{K}\left\{H_m\left(\psi\right) - v^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_m\right) - \widehat{\varsigma}^T\left(\psi\right) W_m\right\} \times$$
$$\left\{S_m\left(A_m\right) - E\left[S_m\left(A_m\right) \mid \overline{A}_{m-1}, \overline{L}_m\right]\right\}$$

and $\widehat{\varsigma}^T\left(\psi\right)$ solves the OLS estimating equation
$P_n\left[\sum_{m=0}^{K}\left(\left\{H_m\left(\psi\right) - v^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_m\right)\right\} - \varsigma^T W_m\right) W_m\right] = 0$. The following Corollary describes the so called "double-robustness" properties of $\widehat{\psi}\left(s, \widetilde{c}^s; \widehat{\alpha}, \widehat{\varsigma}\right)$ and $U^\dagger\left(\psi^\dagger, s, \widetilde{c}^s; \widehat{\alpha}, \widehat{\varsigma}\right)$.

Corollary 7.2b : Consider the $\overline{d}, \overline{d}^*$"union" model characterized by (a) $v^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_m\right)$ and $r^{\underline{d}_m, d_m^*}\left(\overline{L}_m, \overline{A}_{m-1}\right)$ known , (b) a correctly specified $(\bar{d}, \bar{d}^*)$-double-regime-specific SNMM $\gamma\left(\overline{l}_m, \overline{a}_m, \psi\right)$ and (c) that either (but not necessarily both) the parametric model $p_m\left[A_m \mid \overline{L}_m, \overline{A}_{m-1}; \alpha\right]$ for $p_m\left[A_m \mid \overline{L}_m, \overline{A}_{m-1}\right]$ is correct or the regression model
$\varsigma^T W_m$ for $E\left[H_m\left(\psi^\dagger\right) - v^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_m\right)|\overline{L}_m, \overline{A}_{m-1}\right]$ is correct. Then, under standard regularity conditions, the conclusions (i) and (ii) of Theorem 3.4 hold when we replace c by $\widetilde{c}$.

It follows that the algorithm following Theorem 3.4 can be used if we replace c by $\widetilde{c}$ and $H_m\left(\psi\right)$ by $H_m\left(\psi\right) - v^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_m\right)$ as defined above. Of course the resulting estimator will only be locally efficient when $v^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_m\right)$ is zero. Further we can also choose to replace $H_m\left(\psi\right)$ by $H_{\mathrm{mod},m}\left(\psi\right)$.

Method 2:

In the special case that $\overline{d} = \overline{d}^*$ we, of course, do not need to vary
$r^{\underline{d}_m, d_m^*}\left(\overline{L}_m, \overline{A}_{m-1}\right)$ as it is 0 by definition. This raises the question if we

might circumvent the need for $r^{\underline{d}_m,d_m^*}\left(\overline{L}_m,\overline{A}_{m-1}\right)$ to identify $E\left[Y_{\overline{d}}\right]$ even when $\overline{d}\neq\overline{d}^*$. We shall now show this is possible if we define a modified version $\gamma_{\text{mod}}^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$ of $\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$, which equals $\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$ under sequential randomization or when $\overline{d}=\overline{d}^*$. In Section 7.2 we will see that, in the absence of sequential randomization, it is $\gamma_{\text{mod}}^{\overline{d}_{op},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$, rather than $\gamma^{\overline{d}_{op}\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$, that is the essential function we shall need to estimate in order to find the optimal treatment regime d_{op}.

Define

$$\gamma_{\text{mod}}^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)=$$
$$\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)+E\left[Y_{\overline{A}_{m-1},d_m,\underline{d}_{m+1}}-Y_{\overline{A}_{m-1},d_m^*,\underline{d}_{m+1}}|\overline{L}_m,\overline{A}_{m-1},A_m\right]-$$
$$E\left[Y_{\overline{A}_{m-1},d_m,\underline{d}_{m+1}}-Y_{\overline{A}_{m-1},d_m^*,\underline{d}_{m+1}}|\overline{L}_m,\overline{A}_{m-1},A_m=d_m^*\left(\overline{L}_m,\overline{A}_{m-1}\right)\right]$$
$$=E\left[Y_{\overline{A}_m,\underline{d}_{m+1}}|\overline{L}_m,\overline{A}_m\right]-$$
$$E\left[Y_{\overline{A}_{m-1},d_m^*,\underline{d}_{m+1}}|\overline{L}_m,\overline{A}_{m-1},A_m=d_m^*\left(\overline{L}_m,\overline{A}_{m-1}\right)\right]-$$
$$m^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right),$$
$$\text{where } m^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)=E\left[Y_{\overline{A}_{m-1},\underline{d}_m}|\overline{L}_m,\overline{A}_m\right]-$$
$$E\left[Y_{\overline{A}_{m-1},\underline{d}_m}|\overline{L}_m,\overline{A}_{m-1},A_m=d_m^*\left(\overline{L}_m,\overline{A}_{m-1}\right)\right].$$

Under sequential randomization the function $m^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$ takes the value zero. Hence $m^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$, like $v^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$, is a measure of the magnitude of confounding due to unmeasured factors among subjects with history $\left(\overline{L}_m,\overline{A}_{m-1}\right)$, as it compares the mean of the same counterfactual in those who received treatment A_m to that in those who received treatment $d_m^*\left(\overline{L}_m,\overline{A}_{m-1}\right)$ at t_m. In one case the counterfactual being compared is $Y_{\overline{A}_{m-1},d_m^*,\underline{d}_{m+1}}$ and in the other $Y_{\overline{A}_{m-1},\underline{d}_m}$. We have the following.

Theorem 7.3: Let $H_m^{\underline{d}_m}\left(\gamma^{\overline{d}^*}*\right)$ again be defined as in equation (3.5). Then Theorem 7.1 and 7.2 hold when we replace $\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$ by $\gamma_{\text{mod}}^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$ and $v^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$ by $m^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$.

Proof: We only prove $\Leftarrow$ since the proof of $\Rightarrow$ mimics the proof of $\Rightarrow$ in Theorems 7.1 and 7.2 given in the Appendix.

¿From the definition of $\gamma_{\text{mod}}^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$, we have

$$\gamma_{\text{mod}}^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_{m-1},d_m^*\left(\overline{L}_m,\overline{A}_{m-1}\right)\right)-\gamma_{\text{mod}}^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$$
$$=E\left[Y_{\overline{A}_{m-1},\underline{d}_m}|\overline{L}_m,\overline{A}_m\right]-E\left[Y_{\overline{A}_m,\underline{d}_{m+1}}|\overline{L}_m,\overline{A}_m\right].$$

We proceed by induction in reverse time order.

Case 1:

$m = K : H_K\left(\gamma_{\text{mod}}^{\overline{d},\overline{d}^*}\right) = Y_{\overline{A}_K} - E\left[Y_{\overline{A}_K}|\overline{L}_K, \overline{A}_K\right] + E\left[Y_{\overline{A}_{K-1},\underline{d}_K}|\overline{L}_K, \overline{A}_K\right]$,

so $E\left[H_K\left(\gamma_{\text{mod}}^{\overline{d},\overline{d}^*}\right)|\overline{L}_K, \overline{A}_K\right] = E\left[Y_{\overline{A}_{K-1},\underline{d}_K}|\overline{L}_K, \overline{A}_K\right]$.

Case 2: $m < K$: Assume $E\left[H_{m+1}\left(\gamma_{\text{mod}}^{\overline{d},\overline{d}^*}\right)|\overline{L}_{m+1}, \overline{A}_{m+1}\right]$

$= E\left[Y_{\overline{A}_{m},\underline{d}_{m+1}}|\overline{L}_{m+1}, \overline{A}_{m+1}\right]$. By definition $H_m\left(\gamma_{\text{mod}}^{\overline{d},\overline{d}^*}\right) = H_{m+1}\left(\gamma_{\text{mod}}^{\overline{d},\overline{d}^*}\right) + \gamma_{\text{mod}}^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_{m-1}, d_m^*\left(\overline{L}_m, \overline{A}_{m-1}\right)\right) - \gamma_{\text{mod}}^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_m\right)$

$= \left\{H_{m+1}\left(\gamma_{\text{mod}}^{\overline{d},\overline{d}^*}\right) - E\left[Y_{\overline{A}_{m},\underline{d}_{m+1}}|\overline{L}_m, \overline{A}_m\right]\right\} + E\left[Y_{\overline{A}_{m-1},\underline{d}_m}|\overline{L}_m, \overline{A}_m\right]$.

So

$$
\begin{aligned}
&E\left[H_m\left(\gamma_{\text{mod}}^{\overline{d},\overline{d}^*}\right)|\overline{L}_m, \overline{A}_m\right] \\
&= E\left\{E\left[Y_{\overline{A}_{m},\underline{d}_{m+1}}|\overline{L}_{m+1}, \overline{A}_{m+1}\right]|\overline{L}_m, \overline{A}_m\right\} - \\
&E\left[Y_{\overline{A}_{m},\underline{d}_{m+1}}|\overline{L}_m, \overline{A}_m\right] + E\left[Y_{\overline{A}_{m-1},\underline{d}_m}|\overline{L}_m, \overline{A}_m\right] \\
&= E\left[Y_{\overline{A}_{m-1},\underline{d}_m}|\overline{L}_m, \overline{A}_m\right],
\end{aligned}
$$

proving the analogue of Theorem 7.1.

Thus

$$
\begin{aligned}
&E\left[H_m\left(\gamma_{\text{mod}}^{\overline{d},\overline{d}^*}\right)|\overline{L}_m, \overline{A}_m\right] - m^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_m\right) \\
&= E\left[Y_{\overline{A}_{m-1},\underline{d}_m}|\overline{L}_m, \overline{A}_m\right] - m^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_m\right) \\
&= E\left[Y_{\overline{A}_{m-1},\underline{d}_m}|\overline{L}_m, \overline{A}_{m-1}, A_m = d_m^*\left(\overline{L}_m, \overline{A}_{m-1}\right)\right]
\end{aligned}
$$

which is not a function of A_m, proving the analogue of Theorem 7.2.

Corollary 7.3: Given $m^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_m\right)$, both $\gamma_{\text{mod}}^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_m\right)$ and $E\left[Y_{\overline{d}}\right]$ are identified.

Note knowledge of $r^{\underline{d}_m, d_m^*}\left(\overline{L}_m, \overline{A}_{m-1}\right)$ is no longer required for identification as, in contrast with Theorems 7.1 and 7.2, $H_m^{\underline{d}_m}\left(\gamma^{\overline{d}^*}\right)$ in Theorem 7.3 is not a function of $r^{\underline{d}_m, d_m^*}\left(\overline{L}_m, \overline{A}_{m-1}\right)$.

Inference under Method 2:

Under method 2, we can use a drSNMM model $\gamma^{\overline{d}^*}\left(\overline{l}_m, \overline{a}_m; \psi\right)$ to model and estimate $\gamma_{\text{mod}}^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_m\right)$ and $E\left[Y^{\overline{d}}\right]$ since $\gamma_{\text{mod}}^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_m\right)$, like $\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_m\right)$, is only restricted by the definitional constraint $\gamma_{\text{mod}}^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_{m-1}, d_m^*\left(\overline{L}_m, \overline{A}_{m-1}\right)\right) = 0$. Specifically it follows from Theorem 4.3 in Robins and Rotnitzky (2003) that Corollary 7.2a and 7.2b continue to hold when we replace $v^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_m\right)$ by $m^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_m\right)$, references to $r^{\underline{d}_m, d_m^*}\left(\overline{L}_m, \overline{A}_{m-1}\right)$ are deleted, $H_m^{\underline{d}_m}\left(\gamma^{\overline{d}^*}\right)$

is again defined as in equation (3.5), and $\gamma^{\overline{d}^*}\left(\overline{l}_m, \overline{a}_m; \psi\right)$ is interpreted as a model for $\gamma^{\overline{d},\overline{d}^*}_{\text{mod}}\left(\overline{l}_m, \overline{a}_m\right)$.

Instrumental Variable Estimation under Method 2:

Suppose $A_m = (A_{pm}, A_{dm})$ where A_{pm} is the prescribed dose of a medicine and A_{dm} is the actual consumed dose. Since we often have good measures of why doctors prescribe a given dose but poor measures of why patients comply, it would often be reasonable to assume Eq (2.5) was false but the *partial sequential randomization* assumption

$$\{\overline{L}_{\overline{a}}; \overline{a} \in \overline{\mathcal{A}}\} \coprod A_{Pk} \mid \overline{L}_k, \overline{A}_{k-1} \text{ w.p.1, for } k = 0, 1, \ldots, K \tag{7.3}$$

was true. Under this assumption we can estimate $\gamma^{\overline{d},\overline{d}^*}_{\text{mod}}\left(\overline{l}_m, \overline{a}_m\right)$ and, by Theorem 7.3, $E\left[Y_{\overline{d}}\right]$ as well using Theorems 3.3 and 3.4 plus the algorithm following 3.4. Specifically given a correctly specified drSNMM model $\gamma^{\overline{d}^*}\left(\overline{l}_m, \overline{a}_m; \psi\right)$ for $\gamma^{\overline{d},\overline{d}^*}_{\text{mod}}\left(\overline{l}_m, \overline{a}_m\right)$, it follows from Theorem 7.3 that $H_m\left(\gamma^{\overline{d},\overline{d}^*}_{\text{mod}}\right) \equiv H^{\overline{d},\overline{d}^*}_m\left(\psi^\dagger\right)$ satisfies $E\left[H^{\overline{d},\overline{d}^*}_m\left(\psi^\dagger\right)|\overline{L}_m, \overline{A}_m\right] = E\left[Y_{\overline{A}_{m-1},\underline{d}_m}|\overline{L}_m, \overline{A}_m\right]$. Hence $E\left[H^{\overline{d},\overline{d}^*}_m\left(\psi^\dagger\right)|\overline{L}_m, \overline{A}_{m-1}, A_{Pm}\right]$ does not depend on A_{Pm} under partial sequential randomization. Thus, when $\psi^\dagger$ is identified, we can estimate the parameter $\psi^\dagger$ of our model for $\gamma^{\overline{d},\overline{d}^*}_{\text{mod}}\left(\overline{l}_m, \overline{a}_m\right)$ using Theorems 3.3 and 3.4 plus the algorithm following 3.4, provided we use functions $S_m\left(A_{Pm}\right) \equiv s_m\left(A_{Pm}, \overline{A}_{m-1}, \overline{L}_m\right)$ that only depend on A_m only through A_{pm} and our models for treatment indexed by parameter α are models for $p_m\left(A_{Pm}|\overline{A}_{m-1}, \overline{L}_m\right)$ rather than for $p_m\left(A_m|\overline{A}_{m-1}, \overline{L}_m\right)$. If the exclusion restriction that $\gamma^{\overline{d},\overline{d}^*}_{\text{mod}}\left(\overline{l}_m, \overline{a}_m\right)$ is a function of a_m only through actual dose a_{dm} holds, we say the A_{pm} are "instrumental variables" for the A_{dm}.

Remark: Even though under assumption (7.3) the function $\gamma^{\overline{d},\overline{d}^*}_{\text{mod}}\left(\overline{l}_m, \overline{a}_m\right)$ is not non-parametrically identified from the law of the observered data, nonetheless the parameter $\psi^\dagger$ of most models for $\gamma^{\overline{d},\overline{d}^*}_{\text{mod}}\left(\overline{l}_m, \overline{a}_m\right)$will be identified provided the dimension of $\psi^\dagger$ is not too great. Robins(1994) gives precise conditions for identification.

7.2 Optimal Regime SNMM

Recall that, by definition,

$$d_{op,K} = d_{op,K}\left(\overline{L}_K, \overline{A}_{K-1}\right) = \arg\max_{a_K \in \mathcal{A}_K} E\left[Y_{\overline{A}_{K-1},a_K}|\overline{L}_K, \overline{A}_{K-1}\right]$$

and for $m = K-1, ...0$, $d_{op,m} = d_{op,m}\left(\overline{L}_m, \overline{A}_{m-1}\right) =$ $\arg\max_{a_m \in \mathcal{A}_m} E\left[Y_{\overline{A}_{m-1},a_m,\underline{d}_{op,m+1}}|\overline{L}_m, \overline{A}_{m-1}\right]$ where

$\underline{d}_{op,m+1} = (d_{op,m+1}, ..., d_{op,K})$. In Theorem 7.6 and 7.7, we will show that $\overline{d}_{op}$ is identified if, for a pair of regimes $\overline{d}$ and $\overline{d}^*$, the nonidentifiable functions $m^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_m\right)$ and
$g^{\overline{d}_{op},\overline{d}^*}\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right) = g^{\underline{d}_{op,m},\underline{d}^*_m}\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right)$ are known where
$g^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right)$ is a generalization of $r^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right)$ of Eq. 7.1 that is identically 0 under sequential randomization. Specifically, we define $g^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right) =$

$$
\begin{aligned}
&E\left[Y_{\overline{A}_{m-1},a_m,\underline{d}_{m+1}} - Y_{\overline{A}_{m-1},\underline{d}^*_m} | \overline{L}_m, \overline{A}_{m-1}, A_m = a_m\right] - \\
&E\left[Y_{\overline{A}_{m-1},a_m,\underline{d}_{m+1}} - Y_{\overline{A}_{m-1},\underline{d}^*_m} | \overline{L}_m, \overline{A}_{m-1}, A_m \neq a_m\right] \\
&= r^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right) + \\
&E\left[Y_{\overline{A}_{m-1},d^*_m,\underline{d}_{m+1}} - Y_{\overline{A}_{m-1},\underline{d}^*_m} | \overline{L}_m, \overline{A}_{m-1}, A_m = a_m\right] - \\
&E\left[Y_{\overline{A}_{m-1},d^*_m,\underline{d}_{m+1}} - Y_{\overline{A}_{m-1},\underline{d}^*_m} | \overline{L}_m, \overline{A}_{m-1}, A_m \neq a_m\right]
\end{aligned}
$$

so that $g^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_m - 1, a_m\right) = r^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_m - 1, a_m\right)$ if $\overline{d} = \overline{d}^*$. Note that $g^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_m - 1, a_m\right)$ is identically zero under the sharp null hypothesis of no effect of treatment. In this sense $g^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_m - 1, a_m\right)$, like $r^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_m - 1, a_m\right)$, is a measure of treatment interaction.

However, as discussed in remark 4.1, in the absence of sequential randomization, $\overline{d}_{op}$ may not be the optimal regime in the sense that $\overline{d}_{op}$ may not be the maximizer of $E\left[Y^{\overline{d}}\right]$ over $\overline{d} \in \overline{\mathcal{D}}$ and thus methods based on backward induction may not be available. We now characterize those settings in which $\overline{d}_{op}$ is the maximizer of $E\left[Y^{\overline{d}}\right]$

Definition: We say the backward induction feasibility assumption holds if for all m, all a_m, and any $\underline{d}_{m+1} \in \underline{\mathcal{D}}_{m+1}$,

$$
\begin{aligned}
&E\left[Y_{\overline{A}_{m-1},a_m,\underline{d}_{op,m+1}} - Y_{\overline{A}_{m-1},a_m,\underline{d}_{m+1}} | \overline{L}_m, \overline{A}_{m-1}, A_m = a_m\right] \times \\
&E\left[Y_{\overline{A}_{m-1},a_m,\underline{d}_{op,m+1}} - Y_{\overline{A}_{m-1},a_m,\underline{d}_{m+1}} | \overline{L}_m, \overline{A}_{m-1}\right] \geq 0
\end{aligned}
\tag{7.4}
$$

so that the two factors never have opposite signs.

This nonidentifiable assumption implies that if $\underline{d}_{op,m+1}$ is the optimal regime among subjects with observed history $\left(\overline{L}_m, \overline{A}_{m-1}, A_m = a_m\right)$ then it is the optimal regime from time t_{m+1} onwards for subjects with observed history $\left(\overline{L}_m, \overline{A}_{m-1}\right)$ when they are forced to take treatment a_m at t_m. In Appendix 3, we prove the following.

Theorem 7.4: The regime $\underline{d}_{op,m}$ maximizes $E\left[Y_{\overline{A}_{m-1},\underline{d}_m} | \overline{L}_m, \overline{A}_{m-1}\right]$ over all regimes $\underline{d}_m \in \underline{\mathcal{D}}_m$ for each m and $\overline{d}_{op} =$

$\underline{d}_{op,0}$ maximizes $E\left[Y_{\overline{d}}\right]$ over all $\overline{d} \in \overline{\mathcal{D}}$ if and only if the backward induction feasibility assumption holds.

Note that under sequential randomization the backward induction feasibility assumption holds and so $\overline{d}_{op}$ is the optimal regime. Further knowledge of $m^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_m\right)$, $r^{\underline{d}_{op,m+1},d_m^*}\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right)$ and the law of O do not determine whether the backward induction feasibility assumption holds. We will not consider the problem of estimating the true optimal regime when this assumption does not hold and thus $\overline{d}_{op}$ is not the optimal regime, because the computational problem of estimating the true optimal regime is then intractable without strong additional assumptions.

When sequential randomization does not hold, additional problematic issues arise even in the case of single stage decisions (i.e. the special case of $K = 0$). Specifically $E\left[Y_{d_{op}}|L_0\right]$ may be less than $E\left[Y|L_0\right]$ and thus $E\left[Y_{d_{op}}\right]$ may be less than $E\left[Y\right]$. This possibility reflects the fact that by definition any deterministic regime assigns the same treatment to all subjects with a given value of L_0. Thus it might be preferable to allow each subject to choose their own treatment than to impose the optimal treatment $d_{op}\left(L_0\right)$. For example, this will be the true when randomization does not hold because for a subgroup with a common L_0, the best treatment varies among individuals, and each individual is aware of some unmeasured factor that allows her to self-select the best treatment. Under the non-identifiable assumption of additive local rank preservation, the backward feasibility assumption always holds and $E\left[Y_{d_{op}}|L_0\right]$ can never be less than $E\left[Y|L_0\right]$.

Optimal w-compatible treatment regimes: We will actually prove identifiability of $\overline{d}_{op}$ given $m^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_m\right)$ and $g^{\overline{d}_{op},\overline{d}}\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right)$ in a broader setting. Specifically we no longer assume that decision rule at t_m can necessarily depend on the entire past history $\left(\overline{L}_m, \overline{A}_{m-1}\right)$. Rather we assume there exists a known vector-valued function $W_m = w_m\left(\overline{L}_m, \overline{A}_{m-1}\right)$ representing the data available at time t_m on which a decision is to be based. This would be the case if either (i) rules that can based on all of the data $\left(\overline{L}_m, \overline{A}_{m-1}\right)$ are viewed as too complex or (ii) because in the future, after the study is finished, the subject-specific data available to the decision makers (i.e. physicians) at t_m months from, say, onset of HIV infection will only be a subset $w_m\left(\overline{L}_m, \overline{A}_{m-1}\right)$ of the data $\left(\overline{L}_m, \overline{A}_{m-1}\right)$ available to the investigators who analyzed the study data. Let w_m denote a realization of W_m.

Let $\overline{d}_{op}^{w} = \left(\overline{d}_{op,1}^{w}, ..., \overline{d}_{op,K}^{w}\right)$ be the regime defined recursively as follows.

$$\overline{d}_{op,K}^{w}\left(W_K\right) = \arg\max_{a_K} E\left[Y_{\overline{A}_{K-1},a_K}|W_K\right],$$
$$\overline{d}_{op,m}^{w}\left(W_m\right) = \arg\max_{a_m} E\left[Y_{\overline{A}_{m-1},a_m,\underline{d}_{op,m+1}}|W_m\right]$$

We say a regime $\overline{d} = \overline{d}^{w} \in \overline{\mathcal{D}}$ is $w-$*compatible* if $d_m\left(\overline{L}_m, \overline{A}_{m-1}\right) = d_m\left(W_m\right)$ for each m and we let $\overline{\mathcal{D}}^{w}$ denote the set of $w-$*compatible* regimes. We say the regime $\overline{d}_{op}^{w}$ is $w-$*optimal* if $\underline{d}_{op,m}^{w}$ maximizes

$E\left[Y_{\overline{A}_{m-1},\underline{d}_m}|W_m\right]$ over all regimes $\underline{d}^w{}_m \in \underline{\mathcal{D}}^w{}_m$ for each m and $\overline{d}^w_{op} = \underline{d}^w_{op,0}$ maximizes $E\left[Y_{\overline{d}^w}\right]$ over all $\overline{d}^w \in \overline{\mathcal{D}}^w$. We say w is increasing if (W_m, A_m) is a function of W_{m+1} with probability one, i.e., there is no 'forgetting'. We recover our previous set-up when W_m is equal to $\left(\overline{L}_m, \overline{A}_{m-1}\right)$ so no information has been lost. We then have:

Theorem 7.5: The regime $\overline{d}^w_{op}$ is $w-optimal$ if w is increasing and for all m, all a_m, and any $\underline{d}^w{}_{m+1} \in \underline{\mathcal{D}}^w{}_{m+1}$,

$$E\left[Y_{\overline{A}_{m-1},a_m,\underline{d}^w{}_{op,m+1}} - Y_{\overline{A}_{m-1},a_m,\underline{d}^w{}_{m+1}}|W_m, A_m = a_m\right] \geq 0 \tag{7.5}$$
$$\Rightarrow E\left[Y_{\overline{A}_{m-1},a_m,\underline{d}^w{}_{op,m+1}} - Y_{\overline{A}_{m-1},a_m,\underline{d}^w{}_{m+1}}|W_m\right] \geq 0.$$

The proof of theorem 7.5 is analogous to that of 7.4. It is well known that, even under sequential randomization, backward induction cannot be used to compute the true $w-optimal$ regime unless w is increasing.

Now let $\overline{d}, \overline{d}^*$ be two other regimes, either or both of which may themselves equal $\overline{d}^w_{op}$ or one another.

Theorem 7.6 : Given the non-identifiable functions $g^{\overline{d}^w_{op},\overline{d}}\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right)$ and $\gamma^{\overline{d},\overline{d}^*}_{\text{mod}}\left(\overline{L}_m, \overline{A}_m\right)$, the quantities

$$\overline{d}^w_{op}, E\left[Y_{\overline{A}_{m-1},a_m,\underline{d}^w_{op,m+1}}|\overline{L}_m, \overline{A}_{m-1}\right], E\left[Y_{\overline{A}_{m-1},a_m,\underline{d}^w_{op,m+1}}|W_m\right], E\left[Y_{\overline{d}^w_{op}}\right]$$

are identified from the law of $O = \left(\overline{L}, \overline{A}\right)$.

Proof: Define

$$\begin{aligned} q\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right) &= E\left[Y_{\overline{A}_{m-1},a_m,\underline{d}^w_{op,m+1}} - Y_{\overline{A}_{m-1},\underline{d}_m}|\overline{L}_m, \overline{A}_{m-1}\right] \\ &= E\left[Y_{\overline{A}_{m-1},a_m,\underline{d}^w_{op,m+1}} - Y_{\overline{A}_{m-1},\underline{d}_m}|\overline{L}_m, \overline{A}_{m-1}, A_m = a_m\right] - \\ &\quad \left\{1 - f\left(a_m|\overline{L}_m, \overline{A}_{m-1}\right)\right\} g^{\overline{d}^w_{op},\overline{d}}\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right), \\ &\text{and } j\left(\overline{L}_m, \overline{A}_{m-1}\right) = q\left(\overline{L}_m, \overline{A}_{m-1}, d^w_{op,m}\left(\overline{L}_m, \overline{A}_{m-1}\right)\right) \\ &= E\left[Y_{\overline{A}_{m-1},\underline{d}^w_{op,m}} - Y_{\overline{A}_{m-1},\underline{d}_m}|\overline{L}_m, \overline{A}_{m-1}\right]. \end{aligned} \tag{7.6}$$

Note , from their definitions, $E\left[Y|\overline{L}_K, \overline{A}_K\right] = E\left[Y_{\overline{A}_{K-1},A_K}|\overline{L}_K, \overline{A}_K\right]$ equals
$\gamma^{\overline{d},\overline{d}^*}_{\text{mod}}\left(\overline{L}_K, \overline{A}_K\right) - \gamma^{\overline{d},\overline{d}^*}_{\text{mod}}\left(\overline{L}_K, \overline{A}_{K-1}, d_K\left(\overline{L}_K, \overline{A}_{K-1}\right)\right) + E\left[Y_{\overline{A}_{K-1},d_K}|\overline{L}_K, \overline{A}_K\right]$.
Hence $E\left[Y_{\overline{A}_{K-1},d_K}|\overline{L}_K, \overline{A}_K\right]$ is identified. Further

$$\begin{aligned} d^w_{op,K}\left(W_K\right) &= \arg\max_{a_K} E\left[Y_{\overline{A}_{K-1},a_K} - Y_{\overline{A}_{K-1},d_K}|W_K\right] \\ &= \arg\max_{a_K} E\left[q\left(\overline{L}_K, \overline{A}_{K-1}, a_K\right)|W_K\right] \end{aligned}$$

But $q\left(\overline{L}_K, \overline{A}_{K-1}, a_K\right)$ is identified by

$$
\begin{aligned}
& q\left(\overline{L}_K, \overline{A}_{K-1}, a_K\right) \qquad (7.7)\\
& = \gamma^{\overline{d},\overline{d}^*}_{\text{mod}}\left(\overline{L}_K, \overline{A}_{K-1}, a_K\right) - \gamma^{\overline{d},\overline{d}^*}_{\text{mod}}\left(\overline{L}_K, \overline{A}_{K-1}, d_K\left(\overline{L}_K, \overline{A}_{K-1}\right)\right) - \\
& \left\{1 - f\left(a_K | \overline{L}_K, \overline{A}_{K-1}\right)\right\} g^{\overline{d}^w_{op},\overline{d}}\left(\overline{L}_K, \overline{A}_{K-1}, a_K\right).
\end{aligned}
$$

Thus $d^w_{op,K}\left(W_K\right)$, $j\left(\overline{L}_K, \overline{A}_{K-1}\right)$, and $E\left[Y_{\overline{A}_{K-1}, d^w_{op,K}} | \overline{L}_K, \overline{A}_{K-1}\right] = E\left[Y_{\overline{A}_{K-1}, d_K} | \overline{L}_K, \overline{A}_K\right] + j\left(\overline{L}_K, \overline{A}_{K-1}\right)$ are identified.

We now proceed by reverse induction. Specifically we show that if $j\left(\overline{L}_{m+1}, \overline{A}_m\right)$ is identified then $E\left[Y_{\overline{A}_{m-1}, a_m, \underline{\underline{d}}^w_{op,m+1}} - Y_{\overline{A}_{m-1}, \underline{d}_m} | \overline{L}_m, \overline{A}_{m-1}, A_m = a_m\right]$ is identified and hence $q\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right)$,

$$
d^w_{op,m}\left(W_m\right) = \arg\max_{a_m} E\left[q\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right) | W_m\right] \qquad (7.8)
$$

and $j\left(\overline{L}_m, \overline{A}_{m-1}\right) = q\left(\overline{L}_m, \overline{A}_{m-1}, d^w_{op,m}\left(\overline{L}_m, \overline{A}_{m-1}\right)\right)$ are identified. Write

$$
\begin{aligned}
& E\left[Y_{\overline{A}_{m-1}, a_m, \underline{\underline{d}}^w_{op,m+1}} - Y_{\overline{A}_{m-1}, \underline{d}_m} | \overline{L}_m, \overline{A}_{m-1}, A_m = a_m\right] \\
& = E\left[Y_{\overline{A}_{m-1}, a_m, \underline{\underline{d}}^w_{op,m+1}} - Y_{\overline{A}_{m-1}, a_m, \underline{d}_{m+1}} | \overline{L}_m, \overline{A}_{m-1}, A_m = a_m\right] + \\
& E\left[Y_{\overline{A}_{m-1}, a_m, \underline{d}_{m+1}} - Y_{\overline{A}_{m-1}, \underline{d}_m} | \overline{L}_m, \overline{A}_{m-1}, A_m = a_m\right] \\
& = E\left\{ E\left[Y_{\overline{A}_{m-1}, a_m, \underline{\underline{d}}^w_{op,m+1}} - Y_{\overline{A}_{m-1}, a_m, \underline{d}_{m+1}} | \overline{L}_{m+1}, \overline{A}_{m-1}, A_m = a_m\right] | \right. \\
& \qquad \left. \overline{L}_m, \overline{A}_{m-1}, A_m = a_m \right\} + \\
& \gamma^{\overline{d},\overline{d}^*}_{\text{mod}}\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right) - \gamma^{\overline{d},\overline{d}^*}_{\text{mod}}\left(\overline{L}_m, \overline{A}_{m-1}, d_m\left(\overline{L}_m, \overline{A}_{m-1}\right)\right) \\
& = E\left\{ j\left(\overline{L}_{m+1}, \overline{A}_{m-1}, a_m\right) | \overline{L}_m, \overline{A}_{m-1}, A_m = a_m \right\} + \\
& \gamma^{\overline{d},\overline{d}^*}_{\text{mod}}\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right) - \gamma^{\overline{d},\overline{d}^*}_{\text{mod}}\left(\overline{L}_m, \overline{A}_{m-1}, d_m\left(\overline{L}_m, \overline{A}_{m-1}\right)\right)
\end{aligned}
$$

Thus, in particular, by (7.6),

$$
\begin{aligned}
& q\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right) = E\left\{ j\left(\overline{L}_{m+1}, \overline{A}_{m-1}, a_m\right) | \overline{L}_m, \overline{A}_{m-1}, A_m = a_m \right\} + \\
& \gamma^{\overline{d},\overline{d}^*}_{\text{mod}}\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right) - \gamma^{\overline{d},\overline{d}^*}_{\text{mod}}\left(\overline{L}_m, \overline{A}_{m-1}, d_m\left(\overline{L}_m, \overline{A}_{m-1}\right)\right) - \qquad (7.9a)\\
& \left\{1 - f\left(a_m | \overline{L}_m, \overline{A}_{m-1}\right)\right\} g^{\overline{d}^w_{op},\overline{d}}\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right) = \\
& E\left[q\left(\overline{L}_{m+1}, \left(\overline{A}_{m-1}, a_m\right), d^w_{op,m+1}\left(\overline{L}_{m+1}, \left\{\overline{A}_{m-1}, a_m\right\}\right)\right) | \overline{L}_m, \overline{A}_{m-1}, A_m = a_m\right] + \\
& \gamma^{\overline{d},\overline{d}^*}_{\text{mod}}\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right) - \gamma^{\overline{d},\overline{d}^*}_{\text{mod}}\left(\overline{L}_m, \overline{A}_{m-1}, d_m\left(\overline{L}_m, \overline{A}_{m-1}\right)\right) - \qquad (7.9b)\\
& \left\{1 - f\left(a_m | \overline{L}_m, \overline{A}_{m-1}\right)\right\} g^{\overline{d}^w_{op},\overline{d}}\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right)
\end{aligned}
$$

Finally since identification of $q\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right)$ and $E\left[Y_{\overline{A}_{m-1}, \underline{d}_m} | \overline{L}_m, \overline{A}_{m-1}\right]$ implies identification of $E\left[Y_{\overline{A}_{m-1}, a_m, \underline{\underline{d}}^w_{op,m+1}} | \overline{L}_m, \overline{A}_{m-1}\right]$, to complete the

proof it suffices to prove that if $E\left[Y_{\overline{A}_m,\underline{d}_{m+1}}|\overline{L}_{m+1},\overline{A}_{m+1}\right]$ is identified, then $E\left[Y_{\overline{A}_{m-1},\underline{d}_m}|\overline{L}_m,\overline{A}_m\right]$ is identified. This implication follows from the fact that, by definition,

$$E\left[Y_{\overline{A}_{m-1},\underline{d}_m}|\overline{L}_m,\overline{A}_m\right] = E\left[Y_{\overline{A}_m,\underline{d}_{m+1}}|\overline{L}_m,\overline{A}_m\right] - \gamma^{\overline{d},\overline{d}^*}_{\text{mod}}\left(\overline{L}_{m+1},\overline{A}_{m+1}\right) + \gamma^{\overline{d},\overline{d}^*}_{\text{mod}}\left(\overline{L}_{m+1},\overline{A}_m,d_{m+1}\left(\overline{L}_{m+1},\overline{A}_m\right)\right).$$

The following corollary is an immediate consequence of corollary 7.3.

Corollary 7.6 : Given the non-identifiable functions $g^{\overline{d}^w_{op},\overline{d}}\left(\overline{L}_m,\overline{A}_{m-1},a_m\right)$ and $m^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$, we have that $\gamma^{\overline{d},\overline{d}^*}_{\text{mod}}\left(\overline{L}_m,\overline{A}_m\right)$ is identified and thus the conclusions of Theorem 7.6 hold.

Using Theorem 7.6 and Corollary 7.6 to Estimate $\overline{d}^w_{op}$ in Various Settings:

Our basic approach will be to estimate the unknown quantities in Eq (7.7) and then for $m = K, ..., 0$, to alternate estimating the unknown quantities in Eqs (7.8) and Eq (7.9a) (or equivalently Eq (7.9b)). Here are some specific examples.

(1):To generalize estimation of optimal drSNMMs to the setting where we do not assume sequential randomization but $W_m = \left(\overline{L}_m,\overline{A}_{m-1}\right)$ so there is no information loss, we would typically choose (i) $\overline{d} = \overline{d}_{op}$ which implies that a) from its definition, $g^{\overline{d}^w_{op},\overline{d}^w_{op}}\left(\overline{L}_m,\overline{A}_{m-1},a_m\right)$ equals $r^{\overline{d}^w_{op},\overline{d}^w_{op}}\left(\overline{L}_m,\overline{A}_{m-1},a_m\right)$), b) by Eq. (7.6), for all m, $j\left(\overline{L}_{m+1},\overline{A}_{m-1},a_m\right) \equiv 0$, and c), by Eq. (7.9a),

$$q\left(\overline{L}_m,\overline{A}_{m-1},a_m\right) = \gamma^{\overline{d},\overline{d}^*}_{\text{mod}}\left(\overline{L}_m,\overline{A}_{m-1},a_m\right) - \gamma^{\overline{d},\overline{d}^*}_{\text{mod}}\left(\overline{L}_m,\overline{A}_{m-1},d_m\left(\overline{L}_m,\overline{A}_{m-1}\right)\right) - \left\{1 - f\left(a_m|\overline{L}_m,\overline{A}_{m-1}\right)\right\} r^{\overline{d}^w_{op},\overline{d}}\left(\overline{L}_m,\overline{A}_{m-1},a_m\right), \quad (7.10)$$

and (ii) choose $\overline{d}^*$ to be the regime that is identically zero. Thus to estimate the optimal treatment regime $\overline{d}^w_{op}$, we treat $g^{\overline{d}^w_{op},\overline{d}}\left(\overline{L}_m,\overline{A}_{m-1},a_m\right)$ as known (but vary it in a sensitivity analysis), specify and fit a model for $f\left(a_m|\overline{L}_m,\overline{A}_{m-1}\right)$, and finally carry out doubly robust estimation, as described in the last subsection, of a drSNMM model $\gamma^{\overline{d}^*}\left(\overline{l}_m,\overline{a}_m;\psi\right)$ for $\gamma^{\overline{d},\overline{d}^*}_{\text{mod}}\left(\overline{L}_m,\overline{A}_m\right)$ either by assuming $m^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$ is a known function (to be varied in a sensitivity analysis) or, if $\overline{A}_m$ has two components $\left(\overline{A}_{pm},\overline{A}_{dm}\right)$, possibly, by assuming partial sequential randomization. Note that although based on Theorem 3.4 and its extensions discussed in the last subsection we can obtain doubly robust estimators of $\psi^\dagger$ and thus of $\gamma^{\overline{d},\overline{d}^*}_{\text{mod}}\left(\overline{L}_m,\overline{A}_m\right)$, we cannot obtain doubly robust estimators of $\overline{d}^w_{op}$ when $g^{\overline{d}^w_{op},\overline{d}}\left(\overline{L}_m,\overline{A}_{m-1},a_m\right)$ is not assumed

to be identically zero, since consistent estimation of $f\left(a_m|\overline{L}_m,\overline{A}_{m-1}\right)$ is then required.

Remark : Because $j\left(\overline{L}_{m+1},\overline{A}_{m-1},a_m\right)\equiv 0$ when $\overline{d}=\overline{d}_{op}$, we did not have to specify and fit $DP{-}like$ regression models for $E\left\{j\left(\overline{L}_{m+1},\overline{A}_m\right)|\overline{L}_m,\overline{A}_m\right\}$ in Eq (7.9a).

(2): To generalize estimation of DP-regression srSNMMs with the single regime being the identically zero treatment regime to the setting where we do not assume sequential randomization but assume $W_m=\left(\overline{L}_m,\overline{A}_{m-1}\right)$, we would typically choose $\overline{d}^*=\overline{d}$ to be the zero regime $\overline{0}$ and proceed as in (1) just above except now we would have to fit $DP{-}$ regression models for $E\left\{j\left(\overline{L}_{m+1},\overline{A}_m\right)|\overline{L}_m,\overline{A}_m\right\}$ in Eq (7.9a). Inference can proceed as described under the DP- srSNMM fitting algorithm of Section 6 except as modified below under (3).

Remark: The only simplifications due to choosing $\overline{d}^*=\overline{d}$ are that $\gamma^{\overline{d},\overline{d}^*}_{\text{mod}}\left(\overline{L}_m,\overline{A}_{m-1},d_m\left(\overline{L}_m,\overline{A}_{m-1}\right)\right)=0$ in Eq (7.9a) and that $\gamma^{\overline{d},\overline{d}^*}_{\text{mod}}\left(\overline{L}_m,\overline{A}_m\right)=\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$

(3):For estimation of DP-regression drSNMMs with $W_m=\left(\overline{L}_m,\overline{A}_{m-1}\right)$ we proceed as under (2), except now $\gamma^{\overline{d},\overline{d}^*}_{\text{mod}}\left(\overline{L}_m,\overline{A}_{m-1},d_m\left(\overline{L}_m,\overline{A}_{m-1}\right)\right)\neq 0$ and we make the appropriate substitutions of regimes $\overline{d},\overline{d}^*$ for the regimes $\overline{0},\overline{0}$ in the DP- srSNMM fitting algorithm of Section 6. Specifically now

$$Q_K\left(\widehat{\psi},a_K\right)= \tag{7.11}$$
$$\gamma^{\overline{d},\overline{d}^*}_{\text{mod}}\left(\overline{L}_K,\overline{A}_{K-1},a_K,\widehat{\psi}\right)-\gamma^{\overline{d},\overline{d}^*}_{\text{mod}}\left(\overline{L}_K,\overline{A}_{K-1},d_K\left(\overline{L}_K,\overline{A}_{K-1}\right),\widehat{\psi}\right)-$$
$$\left\{1-f\left(a_K|\overline{L}_K,\overline{A}_{K-1},\widehat{\alpha}\right)\right\}g^{\overline{d}^w_{op},\overline{d}}\left(\overline{L}_K,\overline{A}_{K-1},a_K\right),$$
$$Q_{m-1}\left(\widehat{\psi},a_{m-1},\underline{\widehat{\beta}}_m\right)= \tag{7.12}$$
$$r\left\{\overline{L}_{m-1},\overline{A}_{m-2},a_{m-1};\widehat{\beta}_m,\widehat{\psi},\underline{\widehat{\beta}}_{m+1}\right\}+\gamma^{\overline{d},\overline{d}^*}_{\text{mod}}\left(\overline{L}_{m-1},\overline{A}_{m-2},a_{m-1};\widehat{\psi}\right)-$$
$$\gamma^{\overline{d},\overline{d}^*}_{\text{mod}}\left(\overline{L}_{m-1},\overline{A}_{m-2},d_m\left(\overline{L}_{m-1},\overline{A}_{m-2}\right);\widehat{\psi}\right)-$$
$$\left\{1-f\left(a_{m-1}|\overline{L}_{m-1},\overline{A}_{m-2}\right)\right\}g^{\overline{d}^w_{op},\overline{d}}\left(\overline{L}_{m-1},\overline{A}_{m-2},a_{m-1}\right)$$

(4): Consider again Murphy's regret model $u_m\left(\overline{l}_m,\overline{a}_m,\beta^\dagger\right)$
$=E\left[Y_{\overline{a}_{m-1},\underline{d}_{op,m}}-Y_{\overline{a}_m,\underline{d}_{op,m+1}}|\overline{L}_{\overline{a}_{m-1},m}=\overline{l}_m\right]$ where
$u_m\left(\overline{l}_m,\overline{a}_m,\beta\right)=\eta_m\left(\overline{l}_m,\overline{a}_{m-1},\beta_{scale}\right)f\left(a_m-d_{op,m}\left(\overline{l}_m,\overline{a}_{m-1},\beta_{regime}\right)\right)$,
$\beta=(\beta_{scale},\beta_{regime})$ is a finite dimensional parameter vector, $f(\cdot)$ is a known non-negative function satisfying $f(0)=0$, and $\eta_m\left(\overline{l}_m,\overline{a}_{m-1},\beta_{scale}\right)$ is a known non-negative scale function. Suppose $W_m=\left(\overline{L}_m,\overline{A}_{m-1}\right)$ so there is no information loss. Murphy only considered her regret model in the case of sequential randomization. Since under sequential randomization

$$
\begin{aligned}
&E\left[Y_{\overline{a}_{m-1},\underline{d}_{op,m}} - Y_{\overline{a}_m,\underline{d}_{op,m+1}} | \overline{L}_{\overline{a}_{m-1},m} = \overline{l}_m\right] \qquad (7.13)\\
&= E\left[Y_{\overline{a}_{m-1},\underline{d}_{op,m}} - Y_{\overline{a}_m,\underline{d}_{op,m+1}} | \overline{L}_m = \overline{l}_m, \overline{A}_{m-1} = \overline{a}_{m-1}\right]\\
&= E\left[Y_{\overline{a}_{m-1},\underline{d}_{op,m}} - Y_{\overline{a}_m,\underline{d}_{op,m+1}} | \overline{L}_m = \overline{l}_m, \overline{A}_{m-1} = \overline{a}_{m-1}, A_m = a_m\right],
\end{aligned}
$$

Murphy's regret model is also a model for $E\left[Y_{\overline{a}_{m-1},\underline{d}_{op,m}} - Y_{\overline{a}_m,\underline{d}_{op,m+1}} | \overline{L}_m = \overline{l}_m, \overline{A}_{m-1} = \overline{a}_{m-1}\right]$ and for $E\left[Y_{\overline{a}_{m-1},\underline{d}_{op,m}} - Y_{\overline{a}_m,\underline{d}_{op,m+1}} | \overline{L}_m = \overline{l}_m, \overline{A}_{m-1} = \overline{a}_{m-1}, A_m = a_m\right]$. In the absence of sequential randomization, none of the equalities in Eq.(7.13) necessarily hold and we must choose which quantity $u_m\left(\overline{l}_m, \overline{a}_m, \beta\right)$ to model.

We now argue that the appropriate generalization of Murphy's regret model to the setting without sequential randomization is as a model $u_m\left(\overline{l}_m, \overline{a}_m, \beta\right)$ for $E\left[Y_{\overline{a}_{m-1},\underline{d}_{op,m}} - Y_{\overline{a}_m,\underline{d}_{op,m+1}} | \overline{L}_m = \overline{l}_m, \overline{A}_{m-1} = \overline{a}_{m-1}\right]$. The reason is that we would like to be able to determine the optimal treatment strategy beginning at t_m for a new subject who has data on $\overline{L}_m, \overline{A}_{m-1}$ available and who is exchangeable with the subjects in our study, but on whom we are unable to intervene with an optimal strategy prior to time t_m (say, because we did not have the ability to apply our optimal strategy prior to time t_m). It follows from Theorem 7.4 that given the the backward induction feasibility assumption holds, we can succeed in achieving this goal if we can successfully model the quantity $E\left[Y_{\overline{a}_{m-1},\underline{d}_{op,m}} - Y_{\overline{a}_m,\underline{d}_{op,m+1}} | \overline{L}_m = \overline{l}_m, \overline{A}_{m-1} = \overline{a}_{m-1}\right]$.

Consider then the model

$$
u_m\left(\overline{l}_m, \overline{a}_m, \beta^\dagger\right) = E\left[Y_{\overline{a}_{m-1},\underline{d}_{op,m}} - Y_{\overline{a}_m,\underline{d}_{op,m+1}} | \overline{L}_m = \overline{l}_m, \overline{A}_{m-1} = \overline{a}_{m-1}\right]. \qquad (7.14)
$$

Recall $u_m\left(\overline{l}_m, \overline{a}_m, \beta\right)$ attains its minimum of 0 at $a_m = d_{op,m}\left(\overline{l}_m, \overline{a}_{m-1}, \beta_{regime}\right)$. Further, from our definitions, we have $\gamma^{\overline{d}^w_{op},\overline{d}^w_{op}}_{\text{mod}}\left(\overline{L}_m, \overline{A}_m\right) = \gamma^{\overline{d}^w_{op},\overline{d}^w_{op}}\left(\overline{L}_m, \overline{A}_m\right)$ is given by

$$
\begin{aligned}
&\gamma^{\overline{d}^w_{op},\overline{d}^w_{op}}\left(\overline{L}_m, \overline{A}_m\right)\\
&= -u_m\left(\overline{L}_m, \overline{A}_m, \beta^\dagger\right) - r^{\overline{d},\overline{d}^*}\left(\overline{L}_m, \overline{A}_{m-1}, A_m\right)\left\{1 - f\left(A_m | \overline{L}_m, \overline{A}_{m-1}\right)\right\}.
\end{aligned}
$$

Thus $\gamma^{\overline{d}^w_{op},\overline{d}^w_{op}}\left(\overline{L}_m, \overline{A}_m\right)$ may not be either maximized or minimized at $A_m = d_{op,m}\left(\overline{L}_m, \overline{A}_{m-1}\right)$. To estimate the model $u_m\left(\overline{l}_m, \overline{a}_m, \beta\right)$, we put $\overline{d}^* = \overline{d} = \overline{d}^w_{op}$, and note our model $\gamma^{\overline{d}^w_{op},\overline{d}^w_{op}}\left(\overline{L}_m, \overline{A}_m, \psi\right)$ is given by $-u_m\left(\overline{L}_m, \overline{A}_m, \beta\right) - r^{\overline{d}^w_{op},\overline{d}^w_{op}}\left(\overline{L}_m, \overline{A}_{m-1}, A_m\right)\left\{1 - f\left(A_m | \overline{L}_m, \overline{A}_{m-1}; \alpha\right)\right\}$ for $\psi = (\beta, \alpha)$. We then regard $r^{\overline{d}^w_{op},\overline{d}^w_{op}}\left(\overline{L}_m, \overline{A}_{m-1}, A_m\right)$ as known (but vary it in a sensitivity analysis), estimate α with the maximum partial likelihood estimator $\widehat{\alpha}$, and finally estimate the remaining component β of ψ as in (1) above. The remarks found under both (1) and (2) apply in this setting. Furthermore the maximization in Eq (7.8) need not be explicitly carried out, as it will always return $d_{op,m}\left(\overline{L}_m, \overline{A}_{m-1}, \beta_{regime}\right)$ for the current value of β_{regime}.

(5) When $W_m \neq \left(\overline{L}_m, \overline{A}_{m-1}\right)$ so there is information loss, we proceed as described in (1)-(4), except now, for each m, we estimate the unknown random function $B\left(a_m\right) = E\left[q\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right)|W_m\right]$ of a_m in Eq.(7.8) by specifying a multivariate parametric regression model for the mean of the vector $\left\{B\left(a_m\right); a_m \in \mathcal{A}_m\right\}$ given W_m. We then regress the vector $\left\{B\left(a_m\right); a_m \in \mathcal{A}_m\right\}$ (or an estimate thereof) on W_m under this model. Note that a_m is fixed rather than random in the regression so a regression model for $B\left(a_m\right)$ given W_m would be a function of a_m, W_m, and a regression parameter.

7.3 Optimal Marginal drSNMMs vs. Optimal drSNMMs for Estimation of Optimal w-compatible treatment regimes:

Suppose the sequential randomization assumption (2.5) holds and we want to estimate $\overline{d}^w_{op}$ for an increasing w with $W_m \neq \left(\overline{L}_m, \overline{A}_{m-1}\right)$, so there is information loss. In this section we will compare the approach described in the last subsection of specifying a optimal drSNMM model $\gamma^{\overline{d}^w_{op},\overline{0}}\left(\overline{L}_m, \overline{A}_m, \psi\right)$ for $\gamma^{\overline{d}^w_{op},\overline{0}}\left(\overline{L}_m, \overline{A}_m\right) = E\left[Y_{\overline{A}_{m-1},A_m,\underline{d}^w_{op,m+1}} - Y_{\overline{A}_{m-1},0_m,\underline{d}^w_{op,m+1}}|\overline{L}_m, \overline{A}_{m-1}\right]$ versus modelling $\gamma^{\overline{d}^w_{op},\overline{0}}\left(W_m\right) = E\left[Y_{\overline{A}_{m-1},A_m,\underline{d}^w_{op,m+1}} - Y_{\overline{A}_{m-1},0_m,\underline{d}^w_{op,m+1}}|W_m\right]$ directly with a drSNMM model $\gamma^{\overline{d}^w_{op},\overline{0}}\left(W_m, A_m, \omega\right)$ with parameter ω, say, and thus avoiding the need to specify a multivariate regression model for $B\left(a_m\right) = E\left[q\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right)|W_m\right] = E\left[\gamma^{\overline{d}^w_{op},\overline{0}}\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right)|W_m\right] - \gamma^{\overline{d}^w_{op},\overline{0}}\left(\overline{L}_m, \overline{A}_{m-1}, d_m\left(\overline{L}_m, \overline{A}_{m-1}\right)\right)$ in order to evaluate $d^w_{op,m}\left(W_m\right) = \arg\max_{a_m} E\left[q\left(\overline{L}_m, \overline{A}_{m-1}, a_m\right)|W_m\right]$ in (7.8). Note, however, that if $\gamma^{\overline{d}^w_{op},\overline{0}}\left(\overline{L}_m, \overline{A}_m, \psi\right)$ has the functional form $A_m r\left(\overline{L}_m, \overline{A}_{m-1}, \psi\right)$ for some given function $r\left(\overline{L}_m, \overline{A}_{m-1}, \psi\right)$, then $\gamma^{\overline{d}^w_{op},\overline{0}}\left(\overline{L}_m, \overline{A}_m, \psi\right)$ is linear in A_m, and

$$d^w_{op,m}\left(W_m\right) = \max\left\{a_m \in \mathcal{A}_m\right\} I\left\{E\left[r\left(\overline{L}_m, \overline{A}_{m-1}, \psi^\dagger\right)|W_m\right] > 0\right\} + \min\left\{a_m \in \mathcal{A}_m\right\} I\left\{E\left[r\left(\overline{L}_m, \overline{A}_{m-1}, \psi^\dagger\right)|W_m\right] \leq 0\right\}.$$

It then follows that we only need to specify a univariate regression model for $E\left[r\left(\overline{L}_m, \overline{A}_{m-1}, \psi^\dagger\right)|W_m\right]$ in order to estimate $d^w_{op,m}\left(W_m\right)$ even though $\mathcal{A}_m$ might be an uncountable set.

Under (2.5), using ideas in van der Laan, Murphy, and Robins (2003) summarized in van der Laan and Robins (2002, Chapter 6), we could estimate the model $\gamma^{\overline{d}^w_{op},\overline{0}}\left(W_m, A_m, \omega\right)$ (which, by an extension of their nomenclature should be referred to a optimal marginal drSNMM) directly by (i) creating a pseudo-data set with WT_i copies of each subject i, where $WT = \prod_{m=0}^{K} 1/f\left(A_m|\overline{L}_m, \overline{A}_{m-1}\right)$ if $f\left(A_m|\overline{L}_m, \overline{A}_{m-1}\right)$ is known or with $\widehat{WT}_i$

copies where $\widehat{WT} = \prod_{m=0}^{K} 1/f\left(A_m|\overline{L}_m, \overline{A}_{m-1}; \widehat{\alpha}\right)$ if $f\left(A_m|\overline{L}_m, \overline{A}_{m-1}\right)$ is unknown and must be modelled, (ii) retaining in the pseudo -data set only the data $(W_K, A_K, Y) = \left(\overline{V}_K, \overline{A}_K, Y\right)$ where $V_m = W_m \setminus (A_{m-1} \cup W_{m-1})$ are the non-treatment components of W_m that were not in W_{m-1} (since by w increasing, $\overline{A}_{m-1}$ is a function of W_m and (W_m, A_m) can be written as $\left(\overline{V}_m, \overline{A}_m\right)$), and (iii) fitting the optimal marginal drSNMM model $\gamma^{\overline{d}^w_{op},\overline{0}}\left(W_m, A_m, \omega\right) = \gamma^{\overline{d}^w_{op},\overline{0}}\left(\overline{V}_m, \overline{A}_m, \omega\right)$ to the retained part of the pseudo -data set as in Section 4 except with $\overline{V}_m$ replacing $\overline{L}_m$ and ω replacing ψ. This approach succeeds because (2.5) holding in the actual population implies that, in the pseudo-poulation based on weighting by WT_i, (2.5) holds with L replaced by V and $V_{K+1} \equiv Y$.

However when $f\left(A_m|\overline{L}_m, \overline{A}_{m-1}\right)$ is unknown and must be modelled, the approach based on modelling $\gamma^{\overline{d}^w_{op},\overline{0}}\left(\overline{L}_m, \overline{A}_m\right)$ may be preferred to that based on modelling $\gamma^{\overline{d}^w_{op},\overline{0}}\left(W_m\right)$ directly, because the former approach has robustness properties under the g-null mean hypothesis not shared by the latter. To see this sharply, imagine the L_m are discrete with only a moderate number of levels and the A_m are continuous. Then, even with $f\left(A_m|\overline{L}_m, \overline{A}_{m-1}\right)$ totally unrestricted , an asymptotically distribution-free test of $\psi^\dagger = 0$ (and thus of g-null mean hypothesis) exists based on the model $\gamma^{\overline{d}^w_{op},\overline{0}}\left(\overline{L}_m, \overline{A}_m, \psi\right)$, but, because of the curse of dimensionality, not based on the model $\gamma^{\overline{d}^w_{op},\overline{0}}\left(\overline{V}_m, \overline{A}_m, \omega\right)$. Specifically, suppose in the interest of robustness to misspecification, we use model (a.3) that regards $f\left(A_m|\overline{L}_m, \overline{A}_{m-1}\right)$ as completely unknown and also use the empirical conditional distribution of $f\left(A_m|\overline{L}_m, \overline{A}_{m-1}\right)$ as its estimator. Then our test of the hypothesis $\psi^\dagger = 0$ reduces to a nonparametric test of independence of Y and A_0 within strata (where each $l_0 \in \mathcal{L}_0$ defines a separate stratum), because at times other than t_0 no two subjects have the same history $\overline{L}_m, \overline{A}_{m-1}$ so $S\left(A_m\right) - \widehat{E}\left[S\left(A_m\right)|\overline{L}_m, \overline{A}_{m-1}\right]$ will be zero for $m > 0$. This test will have some, although limited, power to detect alternatives with $\psi^\dagger \neq 0$. In contrast, as discussed in Robins (1999), a test of $\omega^\dagger = 0$ will have the correct $\alpha - level$ even asymptotically only if WT can be uniformly consistently estimated which requires we can uniformly consistently estimate $f\left(A_m|\overline{L}_m, \overline{A}_{m-1}\right)$ for all m which is not possible under model (a.3). Quite generally when estimating $\psi^\dagger$ we can trade off efficiency in return for protection against bias caused by possible model misspecification in a way that is not available when we are estimating $\omega^\dagger$.

On the other hand, the approach based on modelling $\gamma^{\overline{d}^w_{op},\overline{0}}\left(W_m\right)$ may be preferred to that based on modelling $\gamma^{\overline{d}^w_{op},\overline{0}}\left(\overline{L}_m, \overline{A}_m\right)$ if it is important that we succeed in specifying a correct or nearly correct blip model because the dimension of $\left(\overline{L}_m, \overline{A}_m\right)$ vastly exceeds that of W_m. For example if based on prior beliefs we are essentially certain $\gamma^{\overline{d}^w_{op},\overline{0}}\left(W_m\right)$ only depends on on $W_m = w_m\left(\overline{L}_m, \overline{A}_{m-1}\right)$ through the few components of W_m that are functions

of the data (L_j, A_{j-1}) obtained at times for $j = m, m-1, m-2$ and W_m is discrete, it is possible we could fit a model saturated in those few components of W_m, thus avoiding major worries about model misspecification. See the Remark on Marginal drSNMMs in Section 9.1 below.

7.4 Bayesian Decision Making in the Prescence of Unmeasured Confounding:

In this section we go beyond sensitivity analysis and consider optimal Bayes decision making in the presence of unmeasured confounding, although we only consider the single occasion setting.

The Single Decision Problem:

Consider again the setting in which we have a single time-independent treatment $(K = 0)$. Thus assume the data are $L = L_0, A = A_0$, $Y = L_1$ with A dichotomous, L multivariate with continuous components and Y continuous. Suppose we have a known correct linear model $\gamma(l, a, \psi) = ar(l, \psi) = a\psi^T w(L) = a\psi^T W$ with true value $\psi^\dagger$ for $\gamma(l, a) = E[Y_a - Y_0 | L = l, A = a]$ and for simplicity assume that it is known that $E[Y_a - Y_0 | L = l, A = a] = E[Y_a - Y_0 | L = l, A \neq a]$ so $r^{\overline{d},\overline{0}}(l, a)$ of (7.1) is identically 0. Then $d_{op}(l) = \arg\max_a \gamma(l, a, \psi) = I\{r(l, \psi) > 0\}$. Further supppose that there may be unmeasured confounding with $v^{\overline{d},\overline{0}}(l, a) = E[Y_0 | L = l, A = a] - E[Y_0 | L = l, A = 0] = a\delta^{\dagger,T} W$ with $\delta^\dagger$unknown. Then if we define $\theta = \psi + \delta$, $H(\theta) = Y - A\theta^T W$, we have that

$$E\left[H\left(\theta^\dagger\right) | L, A\right] = E\left[H\left(\theta^\dagger\right) | L\right], \ \theta^\dagger = \psi^\dagger + \delta^\dagger \tag{7.15}$$

It follows from (7.15) if we define $\widehat{Z}_t^\circ\left(\theta, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right)$ exactly as we did $\widehat{Z}_t\left(\psi, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right)$ except with θ everywhere replacing ψ, we obtain an asymptotic approximation $\pi_{post}\left(\theta^\dagger\right)$ to the posterior for $\theta^\dagger$ given the stochastic process $\widehat{Z}_t^\circ\left(\cdot, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right)$ indexed by θ given by

$$\pi_{post}\left(\theta^\dagger\right) \tag{7.16}$$
$$= \frac{I\left\{\theta^\dagger; \left|\left|\widehat{Z}_t^\circ\left(\theta^\dagger\right)\right|\right| < t \dim\left(\theta^\dagger\right)^{1/2}\right\} \exp\left(-\widehat{Z}_t^\circ\left(\theta^\dagger\right)^T \widehat{Z}_t^\circ\left(\theta^\dagger\right)/2\right) \pi\left(\theta^\dagger\right)}{\int_{\left\{\theta^\dagger; \left|\left|\widehat{Z}_t^\circ(\theta^\dagger)\right|\right| < t \dim(\theta^\dagger)^{1/2}\right\}} \exp\left(-\widehat{Z}_t^\circ\left(\theta^\dagger\right)^T \widehat{Z}_t^\circ\left(\theta^\dagger\right)/2\right) \pi\left(\theta^\dagger\right) d\theta^\dagger}.$$

Thus since $\psi^\dagger$ is independent of $\widehat{Z}_t^\circ\left(\cdot, \widehat{s}_{eff}, c^{\widehat{s}_{eff}}, \widehat{\alpha}, \widehat{\varsigma}, \widehat{\varkappa}, \widehat{\zeta}\right)$ given $\theta^\dagger$, we have that

$$\pi_{post}\left(\psi^\dagger\right) = \int \pi_{post}\left(\theta^\dagger\right) \pi\left(\psi^\dagger | \theta^\dagger\right) d\theta^\dagger$$

where $\pi\left(\psi^{\dagger}|\theta^{\dagger}\right)$ is the conditonal prior of $\psi^{\dagger}$ given $\theta^{\dagger}=\psi^{\dagger}+\delta^{\dagger}$. The optimal Bayes decison rule is, of course,

$$d_{bayes}\left(l\right)=\arg\max_{a\in\mathcal{A}=\{0,1\}}\int ar\left(l,\psi^{\dagger}\right)\pi_{post}\left(\psi^{\dagger}\right)d\psi^{\dagger}$$

Assume a correct model for either $E\left[H\left(\theta^{\dagger}\right)|L\right]$ or f$(A|L)$, so that $\theta^{\dagger}$ is estimated at rate $n^{1/2}$. Then $\pi_{post}\left(\theta^{\dagger}\right)$ charges a volume with radius $O\left(n^{-1/2}\right)$. It follows that if $\pi\left(\psi^{\dagger}|\theta^{\dagger}\right)$ charges a volume of radius $O\left(1\right)$ [as would be the case if $\psi^{\dagger}$ and $\delta^{\dagger}$ were apriori independent with priors charging a volume of radius $O\left(1\right)$] $\pi_{post}\left(\psi^{\dagger}\right)/\pi\left(\psi^{\dagger}|\theta^{\dagger}\right)$ is approximately 1 . If $\psi^{\dagger}$ and $\delta^{\dagger}$ were apriori independent with $\pi\left(\psi^{\dagger}\right)$ charging a volume of radius $O\left(n^{-1/2}\right)$ and $\pi\left(\delta^{\dagger}\right)$ charging a volume of radius $O\left(1\right)$, then $\pi_{post}\left(\psi^{\dagger}\right)/\pi\left(\psi^{\dagger}\right)$ is approximately 1. A limitation of the above analysis is that we assumed $v^{\overline{d},\overline{0}}\left(l,a\right)=a\delta^{\dagger,T}W$ and $\gamma\left(l,a,\psi\right)=a\psi^{\dagger,T}W$ had the same functional form so that the approximate distribution of $\widehat{Z}_{t}^{\circ}\left(\cdot,\widehat{s}_{eff},c^{\widehat{s}_{eff}},\widehat{\alpha},\widehat{\varsigma}\,,\widehat{\varkappa},\widehat{\zeta}\right)$ depended on $\left(\delta^{\dagger},\psi^{\dagger}\right)$ only through their sum $\theta^{\dagger}$. This limitation is more apparent than real as it can be modified by appropriate specification of the joint prior $\pi\left(\delta^{\dagger},\psi^{\dagger}\right)$. For example suppose suppose that W was 20 dimensional and one believed that $\gamma\left(l,a\right)$ only depended on the first 2 components of W but $v^{\overline{d},\overline{0}}\left(l,a\right)$ depended on all 20. Then one could take $\pi\left(\psi^{\dagger}\right)$ to place all its mass on the last 18 components of $\psi^{\dagger}$ being 0. Note that in such a case, since $\theta^{\dagger}$ is estimated at rate $n^{1/2}$, we obtain $n^{1/2}-$*consistent* estimates of the last 18 components of $\delta^{\dagger}$.

8 Continuous Time Optimal drSNMMs

Continuous Time drSNMMs: To extend some of our results to continuous time, let $A\left(u\right)$ and $L\left(u\right)$ be recorded treatment and covariates at time u, where in this section we use parentheses rather than subscripts to denote the time of an event. We shall assume that a subject's observed data $O=\left(\overline{L}\left(K\right),\overline{A}\left(K\right)\right)$ with $\overline{L}\left(t\right)=\left\{L\left(u\right);0\leq u\leq t\right\}$ and $\overline{A}\left(t\right)=\left\{A\left(u\right);0\leq u\leq t\right\}$ are generated by a marked point process such that (i) $L\left(t\right)$ and $A\left(t\right)$ have sample paths that are CADLAG step functions, i.e. they are right-continuous with left-hand limits and we write $\left(\overline{A}\left(t^{-}\right),\overline{L}\left(t^{-}\right)\right)$ for a subject's history up to but not including t ; (ii) the $L\left(t\right)$ and $A\left(t\right)$ process do not jump simultaneously or both in an interval of time $o\left(1\right)$; and (iii) both processes have continuous time bounded intensities so the total number of jumps K^{*} of the joint $\left(\overline{A}\left(t\right),\overline{L}\left(t\right)\right)$ process in $[0,K]$ is random and finite, occurring at random times $T_{1},\ldots,T_{K^{*}}$. That is we assume $\lambda_{A}\left(t\mid\overline{A}\left(t^{-}\right),\overline{L}\left(t^{-}\right)\right)=\lim_{\delta\,t\rightarrow0}pr[A\left(t+\delta t\right)\neq A\left(t^{-}\right)\mid\overline{A}\left(t^{-}\right),\overline{L}\left(t^{-}\right)]/\delta t$ and $\lambda_{L}\left(t\mid\overline{A}\left(t^{-}\right),\overline{L}\left(t^{-}\right)\right)=\lim_{\delta\,t\rightarrow0}pr[L\left(t+\delta t\right)\neq L\left(t^{-}\right)\mid\overline{A}\left(t^{-}\right),\overline{L}\left(t^{-}\right)]/\delta t$ are bounded and measurable on $[0,K]$ where, for example, $A\left(t^{-}\right)=\lim_{u\uparrow t}A\left(u\right)$ is well defined because the sample path have left hand

limits. We choose this restricted class of sample paths because their statistical properties are well understood.

We only consider the case of sequential randomization. Generalizations similar to those in Section 7 are straightforward. The assumption (2.5) of sequential randomization becomes

$$\lambda_A\left[t \mid \overline{L}\left(t^-\right), \overline{A}\left(t^-\right), \{\overline{L}_{\overline{a}}; \overline{a} \in \overline{\mathcal{A}}\}\right] = \lambda_A\left[t \mid \overline{L}\left(t^-\right), \overline{A}\left(t^-\right)\right] \tag{8.1}$$

and

$$\begin{aligned} f[A(t) \mid \overline{L}\left(t^-\right), \overline{A}\left(t^-\right), A(t) \neq A\left(t^-\right), \{\overline{L}_{\overline{a}}; \overline{a} \in \overline{\mathcal{A}}\}] = \\ f\left[A(t) \mid \overline{L}\left(t^-\right), \overline{A}\left(t^-\right), A(t) \neq A\left(t^-\right)\right]. \end{aligned} \tag{8.2}$$

Eq. (8.1) says that given past treatment and confounder history, the probability that the A process jumps at t does not depend on the joint counterfactual histories $\overline{L}_{\overline{a}} = \{L_{\overline{a}}(u); 0 \leq u \leq K\}$ corresponding to the non-dynamic treatment histories $\overline{a} = \{a(u); 0 \leq u \leq K\} \in \overline{\mathcal{A}}$. Eq. (8.2) says that given that the treatment process did jump at t, the probability it jumped to a particular value of $A(t)$ does not depend on the counterfactual histories.

A regime $\overline{d} = \left\{d\left(t, \overline{l}\left(t^-\right), \overline{a}\left(t^-\right)\right); t \in [0, K]\right\}$ is a collection of functions $d\left(t, \overline{l}\left(t^-\right), \overline{a}\left(t^-\right)\right)$ indexed by t mapping $\left(\overline{l}\left(t^-\right), \overline{a}\left(t^-\right)\right) \in \overline{\mathcal{L}}\left(t^-\right) \times \overline{\mathcal{A}}\left(t^-\right)$ into the support $\mathcal{A}(t)$ of $A(t)$ such that the law $f_{res}(o) f_{tr,\overline{d}}(o)$ that replaces the observed treatment process $f_{tr,\overline{p}_{obs}}$ with the deterministic process $f_{tr,\overline{d}}(o)$ has sample paths satisfying (i)-(iii) above with probability one. This is a limitation on the collection of functions that can constitute a valid regime $\overline{d}$. Let $\overline{\mathcal{D}}$ denote the collection of all such regimes. We assume all regimes are feasible in the sense that $f\left[a(t) \mid \overline{L}\left(t^-\right), \overline{A}\left(t^-\right), A(t) \neq A\left(t^-\right)\right]$ is non-zero w.p.1 for all $a(t) \in \mathcal{A}(t) \setminus \{A(t^-)\}$.

Example 8.1: Suppose $A(t)$ is Bernoulli and consider $d\left(t, \overline{L}\left(t^-\right), \overline{A}\left(t^-\right)\right) = I[A(t-\delta) < 1/2] = 1 - A(t-\delta)$ for a given known $\delta > 0$ where $A(t) = 0$ w.p.1 for $t < 0$. Then a subject starts treatment at time 0 and is to take treatment until time δ, to take no treatment from δ to 2δ, to again take treatment from 2δ to 3δ and continue alternating thereafter. Suppose we tried to replace $A(t-\delta)$ by $A(t-) = \lim_{\delta \uparrow t} A(\delta)$ but kept the same parameter values. We will call such a replacement the case $\delta = 0$. Then we have a contradiction as the regime for a subject starting from time zero is not defined, because treatment cannot alternate infinitely quickly at all t. So the regime $\delta = 0$ is not in $\overline{\mathcal{D}}$.

Given a treatment regime $\overline{d}$, let $\underline{d}(u) = \left\{d\left(t, \overline{l}\left(t^-\right), \overline{a}\left(t^-\right)\right); t \in [u, K]\right\}$. Then, given $h \geq 0$, $\overline{d}$ and $\overline{a}\left(t^-\right)$, and a define $\left(\overline{a}\left(t^-\right), a, \underline{d}(t+h)\right)$ to be the regime in which the nondynamic regime $\overline{a}\left(t^-\right)$ is followed on $[0, t)$, the constant dose a is given on $[t, t+h)$and $\overline{d}$ is followed from $t+h$.

Let $Y = y\left(\overline{L}(K), \overline{A}(K)\right)$ be a utility and $Y_{\overline{d}}$ be its counterfactual version under a regime $\overline{d}$. Let $V^{\overline{d},\overline{0}}(t, h, a)$
$= E\left[Y_{\left(\overline{A}(t^-), a, \underline{d}(t+h)\right)} - Y_{\left(\overline{A}(t^-), 0, \underline{d}(t+h)\right)} \mid \overline{L}(t), \overline{A}\left(t^-\right)\right]$ be the mean causal

effect on subjects with observed history $\left(\overline{L}(t), \overline{A}(t^-)\right)$ of a final blip of constant treatment a in the interval $[t, t+h)$ compared to no treatment before changing to the regime $\overline{d}$ at $t+h$. Note $V^{\overline{d},\overline{0}}(t,h,0) = 0$. We restrict attention to treatments for which an instantaneously brief bit of treatment has a negligible effect on Y. We formalize this as follows.

Assumption Of Negligible Instantaneous Effects: We assume that, for all $\overline{d} \in \overline{\mathcal{D}}$, (i) $V^{\overline{d},\overline{0}}(t,h,a)$ is continuous in h and $M^{\overline{d},\overline{0}}(t,a) \equiv \lim_{h \downarrow 0} V^{\overline{d},\overline{0}}(t,h,a)/h$ exists for all $t \in [0,K)$ and (ii) with probability one, $M^{\overline{d},\overline{0}}(t,a) = \partial V^{\overline{d},\overline{0}}(t,0,a)/\partial h$ is continuous in t (uniformly over $\overline{d} \in \overline{\mathcal{D}}, a \in \mathcal{A}(t)$ and $t \in [0,K)$) for all $t \in [T_m, T_{m+1}), m = 0, \ldots, K^* + 1$ where $T_0 \equiv 0, T_{K^*+1} \equiv K$.

$V^{\overline{d},\overline{0}}(t,h,a)$ and $M^{\overline{d},\overline{0}}(t,a)$ may be discontinuous in t at the jump times T_m because of the abrupt change in the conditioning event defining $V^{\overline{d},\overline{0}}(t,h,a)$ at $t = T_m$. Note $M^{\overline{d},\overline{0}}(t,a)\,dt$ is the effect on the mean of Y of a last blip of treatment a compared to no treatment both sustained for "instantaneous" time dt before resorting to regime $\overline{d}$. Hence, $M^{\overline{d},\overline{0}}(t, A(t)) \equiv m^{\overline{d},\overline{0}}\left(t, \overline{L}(t), \overline{A}(t)\right)$ is the instantaneous version of the function $\gamma^{\overline{d},\overline{0}}\left(\overline{L}_m, \overline{A}_m\right)$ of Sec. 3.

Define $H^{\overline{d},\overline{0}}(u) = Y + \int_u^K \left[M^{\overline{d},\overline{0}}\left\{t, d\left(t, \overline{L}(t^-), \overline{A}(t^-)\right)\right\}) - M^{\overline{d},\overline{0}}\{t, A(t)\}\right] dt$ and define $H^{\overline{d},\overline{0}}$ to be $H^{\overline{d},\overline{0}}(0)$. Then our main result is the following which is proved exactly like it analogue in the appendix of Robins (1998).

Theorem 8.1: Under sequential randomization (8.1)-(8.2) and the Assumption of negligible instantaneous effects, if $\overline{d} \in \overline{\mathcal{D}}$, then
$E\left[H^{\overline{d},\overline{0}}(t) \mid \overline{L}(t), \overline{A}(t)\right] = E\left[Y_{\left(\overline{A}(t^-), \underline{d}(t)\right)} \mid \overline{L}(t), \overline{A}(t)\right]$ and
$E\left[Y_{\left(\overline{A}(t^-), \underline{d}(t)\right)} \mid \overline{L}(t), \overline{A}(t)\right]$ is not a function of $A(t)$. In particular,
$E\left(H^{\overline{d},\overline{0}}(0)\right) = E\left[Y_{\overline{d}}\right]$.

We say the data follows a continuous-time drSNMM $M^{\overline{d},\overline{0}}(t,a,\psi)$ if $M^{\overline{d},\overline{0}}(t,a) \equiv m^{\overline{d},\overline{0}}\left(t, \overline{L}(t), \overline{A}(t^-), a\right)$ equals
$M^{\overline{d},\overline{0}}\left(t,a,\psi^\dagger\right) \equiv m^{\overline{d},\overline{0}}\left(t, \overline{L}(t), \overline{A}(t^-), a, \psi^\dagger\right)$ where $\psi^\dagger$ is an unknown parameter to be estimated and $M^{\overline{d},\overline{0}}(t,a,\psi)$ is a known random function continuous in t on $[T_m, T_{m+1})$ satisfying $M^{\overline{d},\overline{0}}(t,a,\psi) = 0$ if $\psi = 0$ or $a = 0$.

First suppose that $A(t)$ is Bernoulli and let $H^{\overline{d},\overline{0}}(t,\psi)$ be $H^{\overline{d},\overline{0}}(t)$ with $M^{\overline{d},\overline{0}}(t,a,\psi)$ replacing $M^{\overline{d},\overline{0}}(t,a)$. Given a Cox model for jumps in the treatment process

$$\lambda_A\left(t \mid \overline{L}(t^-), \overline{A}(t^-)\right) = \lambda_0(t) \exp\left[\alpha' W(t)\right] \tag{8.3}$$

where $W(t)$ is a vector function of $\left\{\overline{L}(t^-), \overline{A}(t^-)\right\}$, α is an unknown vector parameter, and $\lambda_0(t)$ is an unrestricted baseline hazard function, we obtain a G-estimate of the parameter ψ of the continuous-time drSNMM $M^{\overline{d},\overline{0}}(t,a,\psi)$ by adding the term $\theta' H^{\overline{d},\overline{0}}(t,\psi)\, b\left(t, \overline{L}(t^-), \overline{A}(t^-)\right)$ to model (8.3) where

$b\left(t,\overline{l}\left(t^{-}\right),\overline{a}\left(t^{-}\right)\right)$ is a known vector function of the dimension of ψ chosen by the investigator. Specifically, the G-estimate $\widehat{\psi}_{ge}$ is the value of ψ for which the Cox partial likelihood estimator $\widehat{\theta}=\widehat{\theta}\left(\psi\right)$ of θ in the expanded model

$$\begin{aligned}&\lambda_{A}\left(t\mid\overline{L}\left(t^{-}\right),\overline{A}\left(t^{-}\right)\right) \\ &=\lambda_{0}\left(t\right)\left\{\exp\left[\alpha'W\left(t\right)\right]+\theta'H^{\overline{d},\overline{0}}\left(t,\psi\right)b\left(t,\overline{L}\left(t^{-}\right),\overline{A}\left(t^{-}\right)\right)\right\}\end{aligned} \tag{8.4}$$

is zero. Then, under sequential randomization, Theorem 8.1 can be used to show that $\widehat{\psi}_{ge}$ and $n^{-1}\sum_{i}H_{i}\left(0,\widehat{\psi}_{ge}\right)$ will be $n^{\frac{1}{2}}$-consistent for $\psi^{\dagger}$ and $E\left[Y_{\overline{d}}\right]$ provided the drSNMM $M^{\overline{d},\overline{0}}\left(t,a,\psi\right)$ and the Cox model (8.3) are correctly specified. It is also possible to construct doubly robust estimators of $\psi^{\dagger}$. In addition, confidence intervals for $\psi^{\dagger}$ and $E\left[Y_{\overline{d}}\right]$ can also be obtained . Futhermore, if $A\left(t\right)$ is not a dichotomous random variable, then we can obtain more efficient estimators of $\psi^{\dagger}$ and $E\left[Y_{\overline{d}}\right]$ by exploiting (8.2). Technical details will be presented elsewhere.

Example 8.1 (cont): Suppose again $d\left(t,\overline{L}\left(t^{-}\right),\overline{A}\left(t^{-}\right)\right)=1-A\left(t-\delta\right)$ **.** Further $M^{\overline{d},\overline{0}}\left(t,a,\psi\right)=ar\left(t,\overline{L}\left(t\right),\overline{A}\left(t^{-}\right),\psi\right)$ with

$$r\left(t,\overline{L}\left(t\right),\overline{A}\left(t^{-}\right),\psi\right)=\left(1,A\left(t-\delta\right)\right)\psi=\psi_{1}+\psi_{3}A\left(t-\delta\right)$$

where δ is the same known non-negative number. Then $H^{\overline{d},\overline{0}}\left(t,\psi\right)=Y+\psi_{1}cum_{1,\delta}\left(\underline{A}\left(t\right)\right)-\left(\psi_{1}+\psi_{3}\right)cum_{2,\delta}\left(\underline{A}\left(t-\delta\right)\right)$ where $cum_{1,\delta}\left(\underline{A}\left(t\right)\right)$ is the measure of the set $\left\{u;u\in\left[t,K\right]\text{ and }A\left(u\right)=A\left(u-\delta\right)=0\right\}$ and $cum_{2,\delta}\left(\underline{A}\left(t-\delta\right)\right)$ is the measure of the set $\left\{u;u\in\left[t,K\right]\text{ and }A\left(u\right)=A\left(u-\delta\right)=1\right\}$.

Extension to Continuous Time Optimal drSNMMs: The proof of the existence of an optimal regime under our assumptions is subtle and will be given elsewhere. Here we will simply suppose there exists an optimal regime $\overline{d}_{op}$ i.e. a regime $\overline{d}_{op}\in\overline{\mathcal{D}}$ satisfying $E\left[Y_{\overline{A}(t^{-}),\underline{d}_{op}(t)}\mid\overline{L}\left(t^{-}\right),\overline{A}\left(t^{-}\right)\right]\geq E\left[Y_{\overline{A}(t^{-}),\underline{d}(t)}\mid\overline{L}\left(t^{-}\right),\overline{A}\left(t^{-}\right)\right]$ for all $\overline{d}\in\overline{\mathcal{D}}$.

Remark: To see why it is not completely trivial to prove the existence of an optimal regime, given a regime $\overline{d}$, define $z_{op}\left(t,\overline{L}\left(t\right),\overline{A}\left(t^{-}\right),\overline{d}^{*}\right)=\arg\max_{a}m^{\overline{d}^{*},\overline{0}}\left(t,\overline{L}\left(t\right),\overline{A}\left(t^{-}\right),a\right)$. Define

$$\begin{aligned}&d_{op}\left(t,\overline{L}\left(t^{-}\right),\overline{A}\left(t^{-}\right),\overline{d}^{*}\right) \\ &=\arg\max_{a}E\left[m^{\overline{d}^{*},\overline{0}}\left(t,a,\overline{L}\left(t\right),\overline{A}\left(t^{-}\right)\right)|\overline{L}\left(t^{-}\right),\overline{A}\left(t^{-}\right)\right] \\ &=z_{op}\left(t,\overline{L}\left(t^{-}\right),L\left(t\right)=L\left(t^{-}\right),\overline{d}^{*}\right)\end{aligned}$$

where the last equality follows from the fact that the event $L\left(t\right)=L\left(t^{-}\right)$ has probability one (because the probability the covariate process jumps at

any given t is zero). Informally $d_{op}\left(t, \overline{L}\left(t^{-}\right), \overline{A}\left(t^{-}\right), \overline{d}^{*}\right)$ is the optimal treatment decision at t given the information in $\overline{L}\left(t^{-}\right), \overline{A}\left(t^{-}\right)$ if one is to follow $\overline{d}^{*}$ from $t+h$ onwards for h sufficiently small. But since the number of times t are now uncountable we cannot use simple backward induction to define $d_{op}\left(t, \overline{L}\left(t^{-}\right), \overline{A}\left(t^{-}\right)\right)$ in terms of the $d_{op}\left(t, \overline{L}\left(t^{-}\right), \overline{A}\left(t^{-}\right), \overline{d}^{*}\right)$. The smoothness assumptions embedded in the Assumption of negligible instantaneous effects will be needed.

We now consider an optimal drSNMM model $M^{\overline{d}_{op}, \overline{0}}\left(t, a, \psi\right) = m^{\overline{d}_{op}, \overline{0}}\left(t, \overline{L}\left(t\right), \overline{A}\left(t^{-}\right), a, \psi\right)$ for $M^{\overline{d}_{op}, \overline{0}}\left(t, a\right)$. Then, we define

$$H^{\overline{d}_{op}, \overline{0}}\left(u, \psi\right) = Y + \int_{u}^{K} \left[M^{\overline{d}_{op}, \overline{0}}\left\{t, d_{op}\left(t, \overline{L}\left(t^{-}\right), \overline{A}\left(t^{-}\right), \psi\right)\right\} - M^{\overline{d}_{op}, \overline{0}}\left\{t, A(t), \psi\right\}\right]$$

where

$$\begin{aligned} d_{op}\left(t, \overline{L}\left(t^{-}\right), \overline{A}\left(t^{-}\right), \psi\right) &= \arg\max_{a} E\left[m^{\overline{d}_{op}, \overline{0}}\left(t, \overline{L}\left(t\right), \overline{A}\left(t^{-}\right), a, \psi\right) | \overline{L}\left(t^{-}\right), \overline{A}\left(t^{-}\right)\right] \\ &= z_{op}(t, \overline{L}\left(t^{-}\right), L\left(t\right) = L\left(t^{-}\right), \psi) \end{aligned}$$

and $z_{op}(t, \overline{L}\left(t^{-}\right), L\left(t\right), \psi) = \arg\max_{a} m^{\overline{d}_{op}, \overline{0}}\left(t, a, \overline{L}\left(t\right), \overline{A}\left(t^{-}\right), \psi\right)$. We estimate $\psi^{\dagger}$ and $E\left[Y_{\overline{d}_{op}}\right]$ by g-estimation as in the paragraph above.

Example 8.1 (cont): Consider the model $M^{\overline{d}_{op}, \overline{0}}\left(t, a, \psi\right) = ar\left(t, \overline{L}\left(t\right), \overline{A}\left(t^{-}\right), \psi\right)$ with $r\left(t, \overline{L}\left(t\right), \overline{A}\left(t^{-}\right), \psi\right) = \left(1, L\left(t\right), A\left(t-\delta\right)\right)\psi$ where δ is the same non-negative number. Then

$$\begin{aligned} d_{op}\left(t, \overline{L}\left(t^{-}\right), \overline{A}\left(t^{-}\right), \psi\right) &= I\left(r\left(t, \overline{L}\left(t^{-}\right), L\left(t\right) = L\left(t^{-}\right), \overline{A}\left(t^{-}\right), \psi\right) > 0\right) \\ &= I\left(\left(1, L\left(t^{-}\right), A\left(t-\delta\right)\right)\psi > 0\right). \end{aligned}$$

Thus $d_{op}\left(t, \overline{L}\left(t^{-}\right), \overline{A}\left(t^{-}\right), \psi\right) = I\left[A\left(t-\delta\right) < \left\{\psi_1 + \psi_2 L\left(t\right)\right\}/\psi_3\right]$ if $\psi_3 < 0$. Suppose that $\psi_3^{\dagger} = 2\ \psi_1^{\dagger} < 0$ and $\psi_2^{\dagger} = 0$ so $d_{op}\left(t, \overline{L}\left(t^{-}\right), \overline{A}\left(t^{-}\right), \psi^{\dagger}\right) = I\left[A\left(t-\delta\right) < 1/2\right]$. Then the optimal regime for a subject starting at time 0 is to take treatment until time δ, to take no treatment from δ to 2δ, to again take treatment from 2δ to 3δ and continue alternating thereafter as we saw before. Suppose we tried to replace $A\left(t-\delta\right)$ by $A\left(t-\right) = \lim_{\delta \uparrow t} A\left(\delta\right)$ but kept the same parameter values. Then as above, we have a contradiction as the optimal regime for a subject starting from time zero is not defined, because treatment cannot alternate infinitely quickly.

9 Some Important Alternative Approaches:

Results in Sections 1-8 rely on two assumptions that will never be strictly correct: the first that our optimal drSNMM is correct and the second that

either (but not necessarily both) a low dimensional model for the conditional law of treatment or a low dimensional model for the mean of the counterfactual utility given the past is correct. In this section, we relax both assumptions, although not simultaneously. To accomplish this, we use recent results of van der Laan and Dudoit (2003) on model selection via cross-validation and of Robins and van der Vaart (2003,2004) on adaptive non-parametric confidence intervals and inference based on higher order influence functions. Robins and van der Vaart (2004) consider simultaneously relaxing both assumptions.

Selection of a Candidate Optimal Treatment Regime by Cross-Validation

In this section we study a frequentist approach to selecting a single best candidate for the optimal treatment regime using cross-validation regime. The motivation for the method relies heavily on recent work by Wegkamp (2003) and, more particularly and crucially, van der Laan and Dudoit (2003) on model selection by cross-validation. For the remainder of this subsection we assume sequential randomization (i.e. no unmeasured confounders). A major limitation of the methods we proposed in sections 1-5 is the assumption that we have a correctly specified parametric dr SNMM model $\gamma^{\overline{d}^*}\left(\overline{l}_m, \overline{a}_m; \psi\right)$ with true parameter $\psi^\dagger$ for the very high dimensional function $\gamma^{\overline{d},\overline{d}^*}\left(\overline{l}_m, \overline{a}_m\right)$ determining the optimal regime. Here rather than assuming a single correct model we will assume a large list of candidate models $\gamma^j\left(\overline{l}_m, \overline{a}_m; \psi^j\right),\ j = 1, ..., J$ for the optimal regime where we have dropped the $\overline{d}^*$ from the notation and ψ^j denotes the parameter vector corresponding to the j^{th} model. Further, we will no longer assume that any of the J models are true. Rather our approach will be as follows. We will randomly split the n study subjects into two subsamples - the estimation subsample and the validation subsample. We will obtain estimates $\widehat{\psi^j}$by fitting each of the J models to the estimation subsample. We will use the validation subsample to select among the candidates.

The Single Decision Problem:

We begin with the single occassion problem. Thus suppose we have n iid copies of data $O = (A, L, Y)$. Then $d_{op}(l) = \arg\max_{a \in \mathcal{A}} \gamma(l, a)$ where $\gamma(l, a) = E(Y_a | L = l) - E(Y_{a=0} | L = l)$. We assume sequential randomization but now leave $\gamma(l, a)$ unspecified. In terms of the distribution F_O of the observables randomization implies that

$$E(Y|A = a, L = l) - E(Y|A = 0, L = l) = \gamma(l, a) \tag{9.1}$$

and thus that $E[Y - \gamma(L, A)|A, L] = E\{Y - \gamma(L, A)|L\} = E(Y|A = 0, L = l)$ $= E(Y_{a=0}|L = l)$, i.e., $E[H|A, L] = E[H|L] =$ with $H = Y - \gamma(L, A)$.

A First Risk Function:

The key to the first approach employed in this section is the following characterization theorem for $\gamma(l, a)$ of (9.1). Let $E[\cdot]$ denote expectation with

respect to the distribution F_O generating the data and let $E^*[\cdot]$ denote an expectation with respect to an arbitary distribution F^* . Then we have the following.

Theorem 9.1:(i): If $E^*[Y-\gamma(L,A)|L] = E[Y-\gamma(L,A)|L]$, then, for any $g(l)$ such that $g(L) \neq 0$ w.p.1, $\gamma(l,a)$ of (9.1) is the unique function $c(L,A)$ minimizing

$$E\left[g^2(L)\{Y-c(L,A)-E^*[Y-c(L,A)|L]\}^2\right] \tag{9.2}$$

subject to $c(L,0)=0$;

(ii): For all functions $b(l)$ and $g(l)$ such that $g(L) \neq 0$ $w.p.1, \gamma(l,a)$ of (9.1) is the unique function $c(L,A)$ minimizing

$$E\left[g^2(L)\{Y-c(L,A)+E[c(L,A)|L]-b(L)\}^2\right] \tag{9.3}$$

subject to $c(L,0)=0$

Proof: To prove (ii), write $\{Y-c(L,A)\}g(L)-g(L)b(L)+E[g(L)c(L,A)|L]$ as $R+S(c)$, where $R=\{Y-\gamma(L,A)\}g(L)-g(L)b(L)+E[g(L)\gamma(L,A)|L]$ and $S(c)=\{\gamma(L,A)-c(L,A)\}g(L)-E[g(L)\gamma(L,A)|L]+E[g(L)c(L,A)|L]$. Then (9.3) is $E\left[R^2\right]+2E[RS(c)]+E\left[S^2(c)\right]$. Thus $c(L,A)$ minimizing (9.3) is the minimizer of $2E[RS(c)]+E\left[S^2(c)\right]$. But $E[RS(c)]=0$ because $E[S(c)|L]=0$ so $E[RS(c)]=E[\{Y-\gamma(L,A)\}g(L)S(c)]=E[E\{Y-\gamma(L,A)|L,A\}g(L)S(c)]$

$=E[E(Y|A=0,L)g(L)E[S(c)|L]]=0$. Finally $E\left[S^2(c)\right]$ takes it minimum at 0 when $\gamma(L,A)=c(L,A)$.

To prove (i), write $\{Y-c(L,A)\}g(L)-g(L)E^*[Y-c(L,A)|L]$ as $R^*+S^*(c)$ where $R^*=\{Y-\gamma(L,A)\}g(L)-g(L)E^*[Y-\gamma(L,A)|L]$ and $S^*(c)=\{\gamma(L,A)-c(L,A)\}g(L)-g(L)E^*[\gamma(L,A)-c(L,A)|L]$. Note $R^*=\{Y-\gamma(L,A)\}g(L)-g(L)E[Y-\gamma(L,A)|L]$ under the supposition of the theorem. Thus $c(L,A)$ minimizing (9.2) is the minimizer of $2E[R^*S^*(c)]+E\left[S^{*2}(c)\right]$. But $E[R^*S^*(c)]=0$ because $E[R^*|A,L]=0$. Finally $E\left[S^{*2}(c)\right]$ takes it minimum at 0 when $\gamma(L,A)=c(L,A)$.

Corollary 9.1: Doubly Robust Minimizer of Risk:
If $E^*[Y-\gamma(L,A)|L]=E[Y-\gamma(L,A)|L]$ or $E^*[c(L,A)|L]=E[c(L,A)|L]$ for all $c(L,A)$, then for all functions $g(l)$ such that $g(L) \neq 0$ $w.p.1$, $\gamma(l,a)$ of (9.1) is the unique function $c(L,A)$ minimizing $risk(c,F^*,g) = E[Loss(c,F^*,g)]$ with $Loss(c,F^*,g)=loss(O,c,F^*,g)=$

$$g^2(L)\{[Y-c(L,A)]-E^*[Y-c(L,A)|L]\}^2 \tag{9.4}$$

subject to $c(L,0)=0$.

Proof: If $E^*[Y-\gamma(L,A)|L]=E[Y-\gamma(L,A)|L]$ this follows immediately from Theorem 9.1(i). If $E^*[c(L,A)|L]=E[c(L,A)|L]$, the corollary follows from Theorem 9.1(ii) upon writing $risk(c,F^*,g)$ as $E\left[\{[Y-c(L,A)]g(L)+g(L)E^*[c(L,A)|L]-g(L)b(L)\}^2\right]$ with $b(L)=E^*[Y|L]$.

Corollary 9.1 provides a characterization of $\gamma(L,A)$ as the minimizer over functions $c(L,A)$ of the particular risk function $risk(c,F^*,g)$. Suppose we have data from a randomized trial with known randomization probabilities $p(a|l)$, say $p(a|l)=1/2$, and choose F^* such that $F^*_{A|L}$ is $F_{A|L}$ generating the data and $E^*(Y|L)$ is set to a fixed function $b(L)$. We use $F_{A|L},b$ as shorthand for this F^* and so write $risk(c,F^*,g)$ as $risk\left(c,F_{A|L},b,g\right)$.

Suppose one is given J candidates models $\gamma^j\left(L,A,\psi^j\right), j=1,...,J$, for $\gamma(L,A)$ where the dimension of ψ^j and the function $\gamma^j(.,.,.)$ can vary with j, and, based only on the estimation (i.e training) sample data, locally efficient estimators $\widehat{\psi}^j$of ψ^j and thus $\widehat{\gamma}^j(L,A)=\gamma^j\left(L,A,\widehat{\psi}^j\right)$ of $\gamma^j\left(L,A,\psi^j\right)$ are obtained as in Sections 3 and 4 . Then given user-supplied functions $b(l)$ and $g(l)$ we select the index $\widehat{j}$ minimizing $\widehat{risk}\left(\widehat{\gamma}^j,F_{A|L},b,g\right)=P^{val}_{n^{val}}\left[Loss\left(\widehat{\gamma}^j,F_{A|L},b,g\right)\right]=$

$P^{val}_{n^{val}}\left[g^2(L)\left\{Y-\widehat{\gamma}^j(L,A)-b(L)+E\left[\widehat{\gamma}^j(L,A)|L\right]\right\}^2\right]$ over the J candidates functions $\widehat{\gamma}^j$ where $P^{val}_{n^{val}}[\cdot]$ is the sample average over the validation sample. Let j_{oracle} be the j minimizing $risk\left(\widehat{\gamma}^j,F_{A|L},b,g\right)$
$=E\left[g^2(L)\left\{Y-\widehat{\gamma}^j(L,A)-b(L)+E\left[\widehat{\gamma}^j(L,A)|L\right]\right\}^2\right]$. If our goal were to minimize $risk\left(\widehat{\gamma}^j,F_{A|L},b,g\right)$ over our J candidiates, $\widehat{\gamma}^{j_{oracle}}(L,A)$ is the optimal but unobtainable solution. However van der Laan and Dudoit (1993) show that, with high probability, provided the number of models J is not too large compared to n, $risk\left(\widehat{\gamma}^{\widehat{j}},F_{A|L},b,g\right)$ is very close to $risk\left(\widehat{\gamma}^{j_{oracle}},F_{A|L},b,g\right)$ even though $\widehat{j}$ only minimized $\widehat{risk}\left(\widehat{\gamma}^j,F_{A|L},b,g\right)$. Indeed, the number of candidates J can increase as $e^{(n^\alpha)}$ with $\alpha<1$ and yet, under regularity conditions, $risk\left(\widehat{\gamma}^{\widehat{j}},F_{A|L},b,g\right)/risk\left(\widehat{\gamma}^{j_{oracle}},F_{A|L},b,g\right)$ will still approach 1 as $n\to\infty$.

One might reasonably wonder why, if we are nearly certain that model j is misspecified, we use the locally efficient estimator $\widehat{\psi}^j$to estimate the parameter ψ^j, since the desirable properties of $\widehat{\psi}^j$ described above only hold if model j is correctly specified. Our justification is (i) $\widehat{\psi}^j$ should perform well if model j is correct or nearly correct and (ii) if model j is far wrong our cross validation procedure will appropriately eliminate model j from consideration.

Remark on Marginal drSNMMs: The usefulness of having $\widehat{\psi}^j$ correct or nearly correct suggests one might use as candidates at least some optimal marginal drSNMMs with discrete W_m for the reasons described in the final paragraph of Section 7.3.

Minus Expected Utility As A Better Risk Function:

Given these encouraging results for our cross-validated selected model, the question becomes: Is minimizing $risk\left(\widehat{\gamma}^j,F_{A|L},b,g\right)$ over our J candidates really the optimization criteria we wish to use when the true optimal function of interest $\gamma=\gamma(A,L)$ is unknown. Now, of course, our goal is not to mini-

mize $risk\left(\widehat{\gamma}^{j}, F_{A|L}, b, g\right)$ over our J available candidates but rather to maximize expected utility $E\left[Y_{\widehat{d}^{j}_{op}}\right]$ (i.e to minimize the risk $E\left[-Y_{\widehat{d}^{j}_{op}}\right]$ of the loss $-Y_{\widehat{d}^{j}_{op}}$) over the J candidates $\widehat{\gamma}^{j}$ where, we write $\widehat{d}^{j}_{op}$ (or sometimes even d^{j}_{op}) as short-hand for $d^{\widehat{\gamma}^{j}}_{op}$ and, as usual, $d^{\widehat{\gamma}^{j}}_{op} = d^{\widehat{\gamma}^{j}}_{op}(l) = \arg\max_{a} \widehat{\gamma}^{j}(a,l)$. To be explict about the distinction, suppose that we have a dichotomous treatment only taking the values 1 or 0 so $\gamma(A,L)$ can be written $A\gamma(L)$for some function $\gamma(L)$. Then any candidate function $c(a,l) = ac(l)$ for $\gamma(a,l)$ is associated with a candidate optimal regime $d^{c}_{op}(l) = \arg\max_{a} c(a,l) = I\{c(l) > 0\}$ while the true optimal regime $d_{op}(l)$=$d^{\gamma}_{op}(l)$ is $I\{\gamma(l) > 0\}$. Now the expected utility of c is $E\left[Y_{d^{c}_{op}}\right] = E\left[I\{c(L) > 0\} Y_{a=1}\right] + E\left[I\{c(L) \leq 0\} Y_{a=0}\right] = E\left[I\{c(L) > 0\}\gamma(L)\right] + E\left[Y_{a=0}\right]$ while $E\left[Y_{d_{op}}\right] = E\left[I\{\gamma(L) > 0\}\gamma(L)\right] + E\left[Y_{a=0}\right]$. Thus it is clear that $\sup_{c} E\left[Y_{d^{c}_{op}}\right] = E\left[Y_{d_{op}}\right]$ and that the supremum is attained not only at γ but at any c^{*} for which the sets $\{l; c^{*}(l) > 0\} = \{l; \gamma(l) > 0\}$ on which c^{*} and γ are positive are equal. Thus one also can characterize the optimal regime(s) as the maximizer over c of
$E\left[I\{c(L) > 0\} c^{*}(L)\right] + E\left[Y_{a=0}\right]$ as well as the oracle maximizer $\widehat{\gamma}^{j_{util-orac}}$ of the expected utility $E\left[I\{c(L) > 0\}\gamma(L)\right] + E\left[Y_{a=0}\right]$. However this last result does not imply that, if one has only J candidates $\widehat{\gamma}^{j}(a,l)$ available (none of which typically includes any optimal regime), the oracle $\widehat{\gamma}^{j^{*}}$ that maximizes $E\left[I\left\{\widehat{\gamma}^{j}(L) > 0\right\} c^{*}(L)\right]$ will have expected utility $E\left[Y_{\widehat{d}^{j^{*}}_{op}}\right]$ close to the expected utility $E\left[Y_{\widehat{d}^{j_{util-orac}}_{op}}\right]$ of the oracle maximizer $\widehat{\gamma}^{j_{util-orac}}$ of expected utility $E\left[I\left\{\widehat{\gamma}^{j}(L) > 0\right\}\gamma(L)\right] + E\left[Y_{a=0}\right]$ over $j = 1, ..., J$.

Likewise, the maximizer $\widehat{\gamma}^{j_{oracle}}$ of $risk\left(\widehat{\gamma}^{j}, F_{A|L}, b, g\right)$ over the J candidates may have expected utility $E\left[Y_{\widehat{d}^{j_{oracle}}_{op}}\right]$ much less than $E\left[Y_{\widehat{d}^{j_{util-orac}}_{op}}\right]$, even though $risk\left(c, F_{A|L}, b, g\right)$, $E\left[I\{c(L) > 0\} c^{*}(L)\right]$ and
$E\left[I\{c(L) > 0\}\gamma(L)\right]$ are all maximized over all c by γ. The result in the preceding clause is only useful in an asymptopia which, with realistic sized samples and L high dimesional, we can never reach.

Remark A: The point being made here is different from the equally valid point that even if we are fortunate and one of the J models $\gamma^{j}\left(L, A, \psi^{j}\right)$ happens to be correctly specifed, i.e., $\gamma^{j_{correct}}\left(\cdot, \cdot, \psi^{\dagger j_{correct}}\right) = \gamma(\cdot, \cdot)$ for some $j_{correct}$ and some parameter value $\psi^{\dagger j_{correct}}$, if $\psi^{\dagger j_{correct}}$ is sufficiently high dimensional, the huge variability of $\widehat{\psi}^{j_{correct}}$ compared to the smaller variablity of the $\widehat{\psi}^{j}$ of incorrectly specified lower dimensional models may mean that, with high probability, $\widehat{\gamma}^{j_{correct}}$ does much worse with respect to $risk\left(c, F_{A|L}, b, g\right)$, $E\left[I\{c(L) > 0\} c^{*}(L)\right]$ and $E\left[I\{c(L) > 0\}\gamma(L)\right]$ than the corresponding aforementioned oracles.

If we had an unbiased estimator of expected utility
$E\left[Y_{d^{c}_{op}}\right] = E\left[I\{c(L) > 0\}\gamma(L)\right] + E\left[Y_{a=0}\right]$ or $E\left[I\{c(L) > 0\}\gamma(L)\right]$ we could use cross validation, as we did with $risk\left(c, F_{A|L}, b, g\right)$, to obtain an

estimator whose expected utility was close to that of the oracle maximizer $\widehat{\gamma}^{j_{util-orac}}$ and all would seem well.

Now for any $d = d(l)$ an unbiased estimator of $E[Y_d]$ under sequential randomization with $p(a|l)$ known and A discrete is the Horvitz-Thompson-like estimator $P_n[YI\{A = d(L)\}/p(A|L)]$. Thus for any candidate $c(a,l)$ for $\gamma(a,l)$ with associated candidate optimal regime $d^c_{op}(l) = \arg\max_a c(a,l)$, $P_n\left[YI\left\{A = d^c_{op}(L)\right\}/p(A|L)\right]$ is an unbiased estimator of $E\left[Y_{d^c_{op}}\right]$. Thus, under van der Laan and Dudoit's regularity conditions, the $\widehat{\gamma}^{\widehat{j}_{util}}$ maximizing $P^{val}_{n^{val}}\left[YI\left\{A = d^{\widehat{\gamma}^j}_{op}(L)\right\}/p(A|L)\right]$ has expected utility $E\left[Y_{d^{\widehat{j}_{util}}_{op}}\right]$ close to expected utility $E\left[Y_{d^{j_{util-orac}}_{op}}\right]$ of the utility oracle, provided J is not too large. Thus it appears, at least in the one occassion problem, we have developed a quite reasonable approach that selects $d^{\widehat{j}_{util}}_{op}(l) = \arg\max_a \gamma^{\widehat{j}_{util}}\left(a, l, \widehat{\psi}^{\widehat{j}_{util}}\right)$ as the estimated optimal regime with which to treat new patients.

Remaining philosophical and practical difficulties:

But there are remaining philosophical and practical difficulties. For example consider using the data from a study of a population with distribution F to determine the optimal regime for a new and different population in which Y_a has the same conditional distribution given L as the study poulation but $f_{new}(l) \neq f(l)$. The expected utility of a candidate optimal regime $c(l,a)$ in the new population is then $E\left[\left\{YI\left\{A = d^c_{op}(L)\right\}/p(A|L)\right\}w(L)\right]$ with weight function $w(L) = f_{new}(L)/f(L)$ which is still maximized over all c at $c = \gamma$ but is no longer necessarily maximixed over the J candidates $\widehat{\gamma}^j$ at $\widehat{\gamma}^{j_{util-orac}}$ because of the weight function $w(L) = f_{new}(L)/f(L)$. Thus, all would agree that cross validation of the $\widehat{\gamma}^j$should be done by maximizing $P^{val}_{n^{val}}\left[\left\{YI\left\{A = d^{\widehat{\gamma}^j}_{op}(L)\right\}/p(A|L)\right\}w(L)\right]$ if reasonable smooth and reliable estimates of the densities $f_{new}(L)$ and $f(L)$ can be obtained.

The Single Patient Problem : But now consider a single new patient with $L_{new} = l$ for whom a physician needs to select a treatment.The patient certainly constitutes a population with a point mass at $L = l$. Now even if there were a validation sample member with $L = l$, the above maximization would be based on just that one validation sample member, and thus is too variable to be useful. But why should the physician be interested in the loss $-E\left[Y_{d^{\widehat{\gamma}^j}_{op}}\right]$ for a candidate regime $\widehat{\gamma}^j$ rather than, for example, $-E\left[Y_{d^{\widehat{\gamma}^j}_{op}}|L \in rel\right]$ where rel is a subset of covariate values that includes his patient's l, excludes values of l the physician believes irrelevant for determining treatment for his patient, and that contains a sufficiently large fraction of the validation sample so that the expected utility of the $\widehat{\gamma}^j$ selected by cross-validation restricted to validations member with $L \in rel$ will

be close to the oracle maximizer over the $\widehat{\gamma}^j$of $E\left[Y_{d_{op}^{\widehat{j}^j}}|L\in rel\right]$. For example in an AIDS study suppose the CD4 count of the patient was 100. Then the physician might include in rel only the 60%, say, of validation subjects with CD4 counts less than 250. The use of rel is a special case of a more general strategy wherein the physician would use $d_{op}^{\widehat{j}_w}$ with $\widehat{j}_w$ the j that maximizes $P_{n^{val}}^{val}\left[\left\{YI\left\{A=d_{op}^{\widehat{\gamma}^j}(L)\right\}/p(A|L)\right\}w(L)\right]$ for a weight function $w(l)$ supplied by the physician that attains its maximum height at his patient's l, and has relatively less height at those $l's$ the physician believes less relevant to his patient. Determining how quickly $w(l)$ plummets from its maximum is a classic variance bias trade off since the more peaked is $w(l)$ (i) the greater is the probability (owing to sampling variability) that the w-risk $E\left[\left\{YI\left\{A=d_{op}^{\widehat{j}_w}(L)\right\}/p(A|L)\right\}w(L)\right]$ of the selected model $\widehat{\gamma}^{\widehat{j}_w}$ differs greatly from the $w-risk$ $E\left[\left\{YI\left\{A=d_{op}^{j_{w,util-orac}}(L)\right\}/p(A|L)\right\}w(L)\right]$ of the $w-$oracle regime $\widehat{\gamma}^{j_{w,util-orac}}$ that maximizes $E\left[\left\{YI\left\{A=d_{op}^{\widehat{\gamma}^j}(L)\right\}/p(A|L)\right\}w(L)\right]$, but (ii) the less the (doctor's subjective) probability of large bias where we measure bias as the absolute difference between the oracle $w-$risk and the oracle *patient*-risk $\max_j E\left[\left\{YI\left\{A=d_{op}^{\widehat{\gamma}^j}(L)\right\}/p(A|L)\right\}w_{patient}(L)\right]$ and the patient-risk $E\left[\left\{YI\left\{A=d_{op}^{j_{w,util-orac}}(L)\right\}/p(A|L)\right\}w_{patient}(L)\right]$ of the $w-$oracle regime $\widehat{\gamma}^{j_{w,util-orac}}$, where $w_{patient}(l)=0$ for all l other than the patient's. To help understand the doctor's subjective probability of bias suppose, after defining the subset rel based on his own knowledge base, the doctor was persuaded by others that the (i) some of the proposers of the optimal regime SNNMs $\gamma^j\left(a,l,\psi^j\right)$ had an understanding of the relevant biology superior to his and, therefore, (ii) to the extent their models $\gamma^j(a,l,\psi)$ borrow information from subjects with $l\notin rel$ to estimate the effect of treatment at his patient's $l\in rel$ (say, by assuming $\gamma^j\left(a,l,\psi^j\right)=a\gamma^j\left(l,\psi^j\right)$ had a quadratic dependence $\psi_1^jCD4+\psi_2^jCD4^2$ over the entire CD4 range), this decison to borrow is based on sound biological knowledge. In that case the physician might assume that the oracle $w-$risk even for $w(l)$ constant would not differ greatly from the oracle *patient*-risk so to decrease variability the doctor would choose treatment simply by maximizing $P_{n^{val}}^{val}\left[\left\{YI\left\{A=d_{op}^{\widehat{\gamma}^j}(L)\right\}/p(A|L)\right\}\right]$. But if the doctor could not be so persuaded he would use a non-constant weight function. Software to allow the doctor to input a his preferred weight function and to select among the J offered treatment by cross-validation could be implemented. Clearly it would be important to derive the distribution of and confidence intervals for the difference between the w-risk $E\left[\left\{YI\left\{A=d_{op}^{\widehat{j}_w}(L)\right\}/p(A|L)\right\}w(L)\right]$ of the randomly selected model $\widehat{\gamma}^{j_w}$ and the oracle $w-risk$

$E\left[\left\{YI\left\{A=d_{op}^{j_{w,util-orac}}(L)\right\}/p(A|L)\right\}w(L)\right]$, as a function of $J, n, w(\cdot)$, and various other parameters. This is an open statistical problem.

Adding Regimes by Voting : When one is given a collection of J candidate regimes $d_{op}^{\widehat{\gamma}^j}(l)$ it is natural to add one or more regimes to the collection before choosing among them using the above methods. Specifically we add the "vote regime" $d_{op}^{vote}(l)$ that selects $\arg\max_a\left(\sum_j I\left(d_{op}^{\widehat{\gamma}^j}(l)=a\right)\right)$ that selects the most recommended regime. If several values of a tie at a given l then add both regimes, unless the combinatorics from ties at different values of l would add a prohibitively large numer of *vote* regimes; in that case, one can select randomly among ties.

Continuous Treatment: Suppose now that the treatment A is continuous with conditional density given L absolutely continuous wrt Lebesgue measure rather than binary. For example A may represent the number of milligrams of a drug that one takes and all take some, but differing amounts of the drug. In that case even when $p(A|L)$ is known, there exists no unbiased estimator of $E[Y_d] = E[E\{Y|A=d(L),L\}]$. For example $P_{n^{val}}^{val}[YI\{A=d(L)\}/p(A|L)]$ is undefined since the event $I\{A=d(L)\}$ is 0 with probability one. Nonetheless $P_{n^{val}}^{val}\left[Loss\left(c,F_{A|L},b,g\right)\right]$ remains an unbiased estimate of $risk\left(c,F_{A|L},b,g\right) = E\left[Loss\left(c,F_{A|L},b,g\right)\right]$ so we can continue to estimate the less desirable risk function $risk\left(c,F_{A|L},b,g\right)$. An alternative, perhaps preferred, approach based on ideas in Murphy, van der Laan, and Robins (1998) and Gill and Robins (2001) is given a candidate $c(a,l)$ to convert d_{op}^c to a random regime p^c in which when $L=l$ we treat with A drawn from $p^c(a|l)$, where, say, $p^c(a|l)$ could be a uniform distribution with support on $\left(d_{op}^c(l)-\sigma, d_{op}^c(l)+\sigma\right)$ or more precisely on $\left(min\left\{d_{op}^c(l)-\sigma, a_{\min}\right\}, min\left\{d_{op}^c(l)+\sigma, a_{\max}\right\}\right)$ where $a_{\min}$ and $a_{\max}$ are the extremes of ethically allowed doses and σ is a positive constant. Letting Y_{p^c} represent the counterfactual response under the random regime, the expected utility $E[Y_{p^c}]$ under $p^c(a|l)$ is $\int E\{Y|A=a,L\}p^c(a|L)\,dadF(l) = E[p^c(A|L)Y/p(A|L)]$ which admits the unbiased estimator $P_{n^{val}}^{val}[p^c(A|L)Y/p(A|L)]$ since $p^c(a|l)$ is absoutely continuous with respect to $p(a|l)$. Now we would wish to choose σ very small so that the $E\left[Y_{p^{\widehat{\gamma}^j}}\right]$ approximate the $E\left[Y_{d^{\widehat{\gamma}^j}}\right]$, our utilities of interest. However this is not generally possible since σ must be chosen large enough for the expected utility of the random regime $p^{\widehat{\widehat{\gamma}^j}}$ maximizing $P_{n^{val}}^{val}\left[Yp^{\widehat{\gamma}^j}(A|L)/p(A|L)\right]$ over J to be close to oracle utility $max_j E\left[Yp^{\widehat{\gamma}^j}(A|L)/p(A|L)\right]$. Having selected $\widehat{\widehat{\gamma}^j}$ the question then remains whether to treat randomly with $p^{\widehat{\widehat{\gamma}^j}}$ or deterministically with $d_{op}^{\widehat{\widehat{\gamma}^j}}$ I would opt for the latter but with no strong justification.

Attempts to Acheive Double Robustness : Assume again A is binary and $w(L)=1$, but now suppose that, as in an observational study, $p(A|L)$ is unknown. Then, following van der Laan and Dudoit, we might try to construct

doubly robust estimators of the risk $E\left[Y_{d^c_{op}}\right]$ of a candidate regime $c(a,l) = ac(l)$ based on the following double robust identity (9.5).

Theorem (9.2): Let F^* be an arbitrary distribution. Then, under the assumption of no unmeasured confounders, if $p^*(A|L) = p(A|L)$ or $E^*\left[Y|A = d^c_{op}(L), L\right] = E\left[Y|A = d^c_{op}(L), L\right]$ then $E\left[Y_{d^c_{op}}\right]$ is given by

$$E\left[Y_{d^c_{op}}\right] \tag{9.5a}$$

$$= E\left[\frac{YI\left(A = d^c_{op}(L)\right)}{p^*(A|L)} - E^*\left[Y|A = d^c_{op}(L), L\right]\left(\frac{I\left(A = d^c_{op}(L)\right)}{p^*(A|L)} - 1\right)\right]$$

$$= E\left[\frac{I\left(A = d^c_{op}(L)\right)\left\{Y - E^*\left[Y|A = d^c_{op}(L), L\right]\right\}}{p^*(A|L)} + E^*\left[Y|A = d^c_{op}(L), L\right]\right] \tag{9.5b}$$

It follows that the cross validated estimate of expected utility

$$P^{val}_{n^{val}}$$
$$\left[\frac{YI\left\{A = d^c_{op}(L)\right\}}{p(A|L;\widehat{\alpha})} - E\left[Y|A = d^c_{op}(L), L;;\widehat{\eta}\right]\left(\frac{I\left(A = d^c_{op}(L)\right)}{p(A|L;\widehat{\alpha})} - 1\right)\right]$$

will be a $n^{1/2}$ CAN estimator of $E\left[Y_{d^c_{op}}\right]$ if we have estimated from estimation sample or validation sample data parametric models $E[Y|A,L;\eta]$ and $p(A|L;\alpha)$ for $E[Y|A,L]$ and $p(A|L)$ and either (but not necessarily both) are correct. But note that $E[Y|A,L] = \{\gamma(A,L) + E[Y|A=0,L]\}$. Now given a correct model $\gamma(A,L;\psi)$ for $\gamma(A,L)$, we are familiar from Sections 3 and 4 of having to model $p(A|L)$ or $E[Y|A=0,L]$ correctly to obtain CAN estimates of the parameter ψ and thus of $\gamma(A,L)$ and $d_{op}(l)$. But now we find, that if our model for $E[Y|A=0,L]$ is correct and our model for $p(A|L)$ is mispecified, we must still model $\gamma(A,L)$ correctly to obtain CAN estimates of the expected utilities $E\left[Y_{d^c_{op}}\right]$ of candidate regimes $d^c_{op}(l) = \arg\max_a c(a,l)$. One might suppose this is of no help for if we could specify a correct model $\gamma^{j_{correct}}\left(A,L,\psi^{j_{correct}}\right)$ for $\gamma(A,L)$ and a correct model for $E[Y|A=0,L]$, we could immediately obtain a CAN estimate $d_{op}(L)$, namely $\arg\max_a \gamma\left(A,L;\widehat{\psi}\right)$, where $\widehat{\psi}$ is the DR estimator of sections 3 and 4 without needing to resort to cross-validation. However, as discussed in Remark A just above, since our goal is to minimize expected utility based on a sample of size n if $\psi^{j_{correct}}$ is high dimesional so that $\widehat{\psi}^{j_{correct}}$ is highly variable, the model $j_{util-orac}$ that maximizes expected utility might not be the model $j_{correct}$ and thus our cross validation procedure would correctly and usefully fail to select model $j_{correct}$. Thus there is a meaningful, sense in

which we can obtain useful cross-validated DR estimators of $E\left[Y_{d^c_{op}}\right]$ that are robust to misspecification of the model for $p(A|L)$.

The situation appears to be different if we decide to use $risk(c, F^*, g) = E[Loss(c, F^*, g)]$ as a criterion, where we recall that
$Loss(c, F^*, g) = loss(O, c, F^*, g) = g^2(L)\{[Y - c(L, A)] - E^*[Y - c(L, A)|L]\}^2$
and that, by Corollary (9.1), $\gamma(L, A)$ is the minimizer of $risk(c, F^*, g)$ over all c if (i) either $E^*[Y - \gamma(L, A)|L] = E[Y - \gamma(L, A)|L]$ or (ii) $E^*[c(L, A)|L] = E[c(L, A)|L]$ for all $c(L, A)$. Thus if we have estimated, from the estimation sample or validation sample data, a correct parametric model $p(A|L; \alpha)$ for $p(A|L)$, then
$\widehat{r}isk\left(c, \widehat{F}_{A|L}, b, g\right) = P^{val}_{n^{val}}\left[g^2(L)\{Y - c(L, A) - b(L) + E[c(L, A)|L; \widehat{\alpha}]\}^2\right]$
is a CAN estimator of
$risk(c, F^*, g)$ where $F^* = \left(F_{A|L}, b\right)$ satisifies (ii). If, separately for each candidate $c(L, A)$,we obtain, from estimation sample data or validation sample data, an estimate $\widehat{\varsigma}(c)$ of the fit of a correct parametric model $b(L; \varsigma)$ for $E[Y|A = 0, L]$ based on regressing $Y - c(L, A)$ on L wiith regression function $b(L; \varsigma)$, then $\widehat{r}isk\left(c, \widehat{F}_{Y|A=0,L}, g\right) = P^{val}_{n^{val}}\left[g^2(L)\{Y - c(L, A) - b(L; \widehat{\varsigma}(c))\}^2\right]$ is a CAN estimator of $risk(c, F^*(c), g)$ where $F^*(c) = \left(F^*_{Y|A=0,L}(c)\right)$ satisifies (i) (since $F^*_{Y|A=0,L}(\gamma) = F_{Y|A=0,L}$). Thus, we see that, unlike when we used expected utility as a criterion, we can obtain a CAN estimator of a risk function $risk(c, F^*(c), g)$ that is minimized over c at $\gamma(L, A)$ if we have a correct model for $E[Y|A = 0, L]$. However in contrast to the spirit of the doubly robust estimators of $\gamma(L, A)$ studied in Sections 3 and 4, we cannot obtain a CAN estimator of a risk function $risk(c, F^*(c), g)$ that is minimized over c at $\gamma(L, A)$ when either (but not necessarily both) of the parametric models $b(L; \varsigma)$ and $p(A|L; \alpha)$ for $E[Y|A = 0, L]$ and $p(A|L)$ are correct. Thus a true double robustness property is lacking.

Beyond Double Robustness:

Heretofore we have assumed that we have been able to correctly specify either (but not necessarily both) models for the law of treatment A_m given the past or a model for the mean of $H_m\left(\psi^\dagger\right)$. With high dimensional data such an assumption will never be exactly true and so the question arises as to whether we can obtain additional robustness to misspecification beyond double robustness and if so what will be the cost in terms of variance. We investigate that question in this section using a new theory of higher dimensional influence functions based on $U-$ statistics due to Robins and van der Vaart (2004). We consider the simplest example in order to make the ideas clear. For more general examples see Robins and van der Vaart (2004). We do not consider the estimation of nonparametric $\gamma(A, L)$ as above but rather assume $\gamma(A, L) = \psi A$ and focus on estimating ψ.

Consider the analysis of the normal semiparametric regression model based on n iid observations $O_i = (Y_i, A_i, X_i)$

$$Y = \psi^{\dagger} A + b\left(X; \eta^{\dagger}\right) + e \tag{9.6}$$

where $e \sim N(0,1)$, $b\left(X; \eta^{\dagger}\right)$ is an unknown function, and A is dichotomous. For simplicity we will assume the law of e is known. The law $F\left(a|X; \alpha^{\dagger}\right)$ of $A|X$ is unknown as is the law $F\left(x; \omega^{\dagger}\right)$ of X. We consider first the case with X discrete with either finite or countable support. All quantities will be allowed to depend on the sample size n, including the support of X and the true parameters generating the data. We suppress the dependence on n in the notation except for emphasis.

The likelihood with respect to a dominating measure for one observation is

$$\begin{aligned} &f(O; \theta = (\psi, \eta, \alpha, \omega)) \\ &= \phi\left\{Y - \psi A - b(X; \eta)\right\} f(A|X; \alpha) f(X; \omega); \theta \in \Theta = \Psi \times \mathcal{N} \times \mathring{A} \times \Omega \end{aligned} \tag{9.7}$$

The following argument suggests that there should exist estimators that are superior to our doubly robust estimators. Our doubly robust estimators of $\psi^{\dagger}$ are (i) $n^{1/2}-$ consistent estimators if we succeed in specifiying a correct lower dimensional model for either $b\left(X; \eta^{\dagger}\right)$ or $f\left(A|X; \alpha^{\dagger}\right)$ but (ii) our estimators are inconsistent if both models are incorrect. It seems logical to suppose that by specifying larger models for $b\left(X; \eta^{\dagger}\right)$ and/or $f\left(A|X; \alpha^{\dagger}\right)$ we should be able to obtain doubly robust confidence intervals and point estimators whose length and standard deviation are $n^{-\alpha}$ for $\alpha < 1/2$, thus allowing us to give up efficiency for further protection against bias. We shall see that this is indeed possible. Indeed this approach can result in triply robust or even infinitely robust (i.e. exactly unbiased) estimating functions in certain settings.

We will analyze this model using a new theory of higher order influence functions due to Robins and van der Vaart (2004) that extends the first order semiparametric efficiency theory of Bickel et al. (1993) and van der Vaart (1991) by incorporating the theory of higher order scores and Bhattacharyya bases due to McLeish and Small (1994) and Lindsay and Waterman (1996). The following follows the development in McLeish and Small (1994) in many aspects.

A Theory of Higher Order Influence Functions :

Suppose we observe n iid observations $O_i, i = 1, ..., n$, from a model $M(\Theta) = \{F(o; \theta), \theta \in \Theta\}$ and we wish to make inference on a particular functional $\widetilde{\psi}(F) \in R^{p^*}$ or equivalently $\psi(\theta) = \widetilde{\psi}(F(\theta))$. In general the functional $\widetilde{\psi}(F)$ can infinite dimensional but here for simplicity we only consider consider the finite dimensional case.

Given a possibly vector valued function $b(\varsigma)$, $\varsigma = \{\varsigma_1, ..., \varsigma_p\}^T$, define for $m = 0, 1, 2$, $b_{\setminus i_1 ... i_m}(\varsigma) = \partial^m b(\varsigma) / \partial \varsigma_{i_1} ... \partial \varsigma_{i_m}$ with $i_s \in \{1, ..., p\}$, for $s = 1, 2, ..., m$ where the $\setminus$ symbol denotes differentiation by the variables occurring to its right. Given a sufficiently smooth $p-$ dimensional parametric submodel $\widetilde{\theta}(\varsigma)$ mapping $\varsigma \in R^p$ injectively into Θ, define $\psi_{\setminus i_1 ... i_m}(\theta)$ to be

$\left(\psi \circ \widetilde{\theta}\right)_{\backslash i_1 \ldots i_m}(\varsigma)\,|_{\varsigma=\widetilde{\theta}^{-1}\{\theta\}}$ and $f_{\backslash i_1 \ldots i_m}(\mathbf{O};\theta)$ to be $\left(f \circ \widetilde{\theta}\right)_{\backslash i_1 \ldots i_m}(\varsigma)\,|_{\varsigma=\widetilde{\theta}^{-1}\{\theta\}}$ where $f(\mathbf{O};\theta) \triangleq \prod_{i=1}^{n} f(O_i;\theta)$ and each $i_s \in \{1,...,p\}$.

Definition of a kth order estimation influence function: A vector U-statistic $U_k(\theta) = u_k(\mathbf{O};\theta)$ of order k, dimension of p^* of $\psi(\theta)$ and finite variance is said to be an kth order estimation influence function for $\psi(\theta)$ if (i) $E_\theta[U_k(\theta)] = 0$, $\theta \in \Theta$ and (ii) for $m = 1, 2, \ldots, k$, and every suitably smooth p dimensional parametric submodel $\widetilde{\theta}(\varsigma)$, $p = 1, 2, ...,$

$$\psi_{\backslash i_1 \ldots i_m}(\theta) = E_\theta\left[U_k(\theta)\, S_{i_1 \ldots i_m}(\theta)\right]$$

where $S_{i_1 \ldots i_m}(\theta) \triangleq f_{\backslash i_1 \ldots i_m}(\mathbf{O};\theta)/f(\mathbf{O};\theta)$. We refer to $S_{i_1 \ldots i_m}(\theta)$ as an mth order score associated with the model $\widetilde{\theta}(\varsigma)$. If $\psi_{\backslash i_1 \ldots i_m}(\theta) = 0$, we refer to $S_{i_1 \ldots i_m}(\theta)$ as an estimation nuisance score.

Remark: The scores $S_{i_1 \ldots i_m}(\theta)$ are U statistics of order m. For later use it will be useful here to collect formula for the an arbitrary score $S_{i_1 \ldots i_s}(\theta)$ of order s in terms of the subject specific scores $S_{i_1 \ldots i_m, j}(\theta) = f_{/i_1 \ldots i_m, j}(O_j;\theta)/f_j(O_j;\theta), j = 1, ..., n$ for $s = 1, 2, 3$. Resullts in Waterman and Lindsay (1996) imply

$$S_{i_1} = \sum_j S_{i_1, j} \tag{9.8a}$$

$$S_{i_1 i_2} = \sum_j S_{i_1 i_2, j} + \sum_{l \neq j} S_{i_1, j} S_{i_2, l} \tag{9.8b}$$

$$S_{i_1 i_2 i_3} = \sum_j S_{i_1 i_2 i_3, j} + \sum_{l \neq j} S_{i_1 i_2, j} S_{i_3, l} + S_{i_3 i_2, j} S_{i_1, l} + S_{i_1 i_3, j} S_{i_2, l} + \sum_{l \neq j \neq t} S_{i_1, j} S_{i_2, l} S_{i_3, t} \tag{9.8c}$$

Note these formulae are examples of the following canonical representation of an arbitrary s^{th} order U statistic.

$$U_m = \sum_{m=1}^{m=s} D_m \tag{9.9}$$

$$D_m(\theta) = \sum_{\{i_1 \neq i_2 \neq \ldots \neq i_m; i_l \in \{1,2,\ldots,n\}, l \in =1,\ldots,m\}} d_m(O_{i_1}, O_{i_2}, ..., O_{i_m}),$$

For all m and $l, 1 \leq l \leq m$, with $O_{-i_l} = \left(O_{i_1}, ...O_{i_{l-2}}, O_{i_{l-1}}, O_{i_{l+1}}, O_{i_{l+2}}..., O_{i_m}\right)$

$$E\left[d_m(O_{i_1}, O_{i_2}, ..., O_{i_m})\,|O_{-i_l}\right] = 0,$$

$d_m(O_{i_1}, O_{i_2}, ..., O_{i_m})$ need not be symmetric in $O_{i_1}, O_{i_2}, ..., O_{i_m}$

We also consider a U statistic of order $m < s$ to also be a U statistic of order s with $d_j\left(O_{i_1}, O_{i_2}, ..., O_{i_j}\right) = 0$ for $s \geq j > m$.

Estimation influence functions will be useful for deriving point estimators of ψ with small bias and for deriving interval estimators centered on an estimate of ψ. We also define testing influence functions both to test hypotheses

about ψ and to form confidence intervals for ψ whose expected length may be less than that of intervals based on an estimation influence function.

Definition of a kth order testing influence function: A U-statistic $U_k(\theta) = u_k(\mathbf{O};\theta)$ of order k, dimension p^*, and finite variance is said to be an kth order testing influence function for testing $\psi(\theta) = \psi^\dagger$ if in the restricted model $M\left(\Theta\left(\psi^\dagger\right)\right) = M(\Theta) \cap \left\{F; \widetilde{\psi}(F) = \psi^\dagger\right\}$ (i.e the submodel with parameter space $\Theta\left(\psi^\dagger\right) = \Theta \cap \left\{\theta; \psi(\theta) = \psi^\dagger\right\}$) (i) $E_\theta[U_k(\theta)] = 0$, $\theta \in \Theta\left(\psi^\dagger\right)$ and (ii) for $m = 1, 2, ..., k$, and every suitably smooth p dimensional parametric submodel $\widetilde{\theta}(\varsigma)$ with range $\Theta\left(\psi^\dagger\right)$, $p = 1, 2, ...,$ $\psi_{\setminus i_1...i_m}(\theta) = E_\theta[U_k(\theta) S_{i_1...i_m}(\theta)]$ where $S_{i_1...i_m}(\theta) \triangleq f_{\setminus i_1...i_m}(\mathbf{O};\theta)/f(\mathbf{O};\theta)$. Since in model $M\left(\Theta\left(\psi^\dagger\right)\right)$, $\psi_{\setminus i_1...i_m}(\theta) = 0$ for all $S_{i_1...i_m}(\theta)$, all scores are nuisance scores.

Remark: Suppose that $\psi(\theta) = \psi^\dagger$ and $U_k(\theta)$ is a kth order estimation influence function, then it is a kth order testing influence function, since every smooth submodel through θ in model $M\left(\Theta\left(\psi^\dagger\right)\right)$ is a smooth submodel through θ in model $M(\Theta)$. Further the set of estimation nuisance scores includes the set of testing scores. The converses need not be true.

Remark: Henceforth, in any statement in which we do not mention whether the parameter space under consideration is $M\left(\Theta\left(\psi^\dagger\right)\right)$ or $M(\Theta)$, our results hold for for both. When we wish to distinguish the 2 cases we use 'est' and 'test' to discriminate.

Definition of the Bias Function of a kth order influence function: We call $B_k\left[\theta^\dagger, \theta\right] = E_{\theta^\dagger}[U_k(\theta)]$ the bias function of $U_k(\theta)$.

Given a parametric submodel $\widetilde{\theta}(\varsigma)$, define $B_{k,i_1^*...i_m^* i_{m+1},...i_s}[\theta,\theta]$ $= \partial^s B_k\left[\widetilde{\theta}(\varsigma^*), \widetilde{\theta}(\varsigma)\right] / \partial\varsigma_{i_1...}^* \partial\varsigma_{i_m}^* \partial\varsigma_{i_{m+1}...} \partial\varsigma_{i_s}|_{\varsigma^* = \widetilde{\theta}^{-1}\{\theta\}, \varsigma = \widetilde{\theta}^{-1}\{\theta\}}$ where we reserve $*$ for differentiation with respect to the first argument of $B_k[\cdot,\cdot]$. Thus, under regularity conditions, by the definition of a kth order influence function $U_k(\theta)$, $B_{k,i_1^*...i_s^*}[\theta,\theta] = \psi_{\setminus i_1...i_s}(\theta)$.

The following Theorem is closely analagous to related results in McLeish and Small (1994).

Extended Information Equality Theorem: Given a kth order influence function $U_k(\theta)$, for all smooth submodels $\widetilde{\theta}(\varsigma)$ and all $i_1^*...i_m^* i_{m+1}...i_s$, $s \leq k$, (i) $B_{k,i_1^*...i_m^* i_{m+1},...i_s}[\theta,\theta] = 0$ if $s > m > 0$, but (ii) $B_{k,i_1^*...i_m^* i_{m+1},...i_s}[\theta,\theta]$ $\equiv B_{k,i_1...i_s}[\theta,\theta] = -\psi_{\setminus i_1...i_s}(\theta)$ if $m = 0$

Proof: See Robins and van der Vaart (2004).

Let $V_m(\theta) = S_{i_1...i_m}(\theta)$ denote a generic m^{th} order score at θ in model . Let $\{V_m(\theta)\}$ be the set of m^{th} order scores at θ as we vary over both the parametric submodels $\widetilde{\theta}(\varsigma)$ of our model and the indices $i_1...i_m$. Let $\cup_{l=1}^{l=m}\{V_l(\theta)\}$ *be* the collection of scores of order m or less and $\overline{\mathcal{B}}_m(\theta)$ be the closed linear span of $\cup_{l=1}^{l=m}\{V_l(\theta)\}$ in the Hilbert space $\mathcal{U}_m$ composed of all $U-statistics$ of order m with mean zero and finite variance and dimension of p^* of ψ with inner product defined by covariances with respect to the product

measure $F^n(\cdot;\theta)$.. We refer to $\overline{\mathcal{B}}_m(\theta)$ as the mth order tangent space for the model. $\overline{\mathcal{B}}_m(\theta)$ is parametrization invariant and thus a "geometric quantity."

Repeating the above for the estimation nuisance scores $V_m^{est,nuis}(\theta)$ in the 'estimation' model with parameter space $M(\Theta)$, we refer to the closed linear span $\overline{\Lambda}_m(\theta)$ of $\cup_{l=1}^{l=m}\left\{V_m^{est,nuis}(\theta)\right\}$ to be the mth order estimation nuisance tangent space. We write $\overline{\mathcal{B}}_m^{est}(\theta)$ and $\overline{\Lambda}_m^{est}(\theta)$ for the tangent space and nuisance tangent space in model $M(\Theta)$ ..We write $\overline{\mathcal{B}}_m^{test}(\theta)$ for the tangent space in model $M\left(\Theta\left(\psi^{\dagger}\right)\right)$. Note $\overline{\mathcal{B}}_m^{test}(\theta) \subseteq \overline{\Lambda}_m^{est}(\theta) \subseteq \overline{\mathcal{B}}_m^{est}(\theta)$.

Given any kth order estimation influence function $U_k^{est}(\theta)$, define $IF_k^{est}(\theta) = \Pi_\theta\left[U_k(\theta)\,|\overline{\mathcal{B}}_k^{est}(\theta)\right]$ where the projection operator $\Pi_\theta[\cdot|\cdot]$ is the projection operator in the Hilbert space $\mathcal{U}_k(\theta)$. .

Efficient Influence Function Theorem : (i) $IF_k^{est}(\theta)$ is unique in the sense that for any two kth order influence functions $\Pi_\theta\left[U_k^{est(1)}(\theta)\,|\overline{\mathcal{B}}_k^{est}(\theta)\right]$ **and** $\Pi_\theta\left[U_k^{est(2)}(\theta)\,|\overline{\mathcal{B}}_k^{est}(\theta)\right]$ are equal almost surely.

(ii) $IF_k^{est}(\theta)$ is a kth order estimation influence function and has variance less than or equal to any other kth order estimation influence function.

(iii)$U_k(\theta)$ is a kth order estimation influence function if and only if $U_k(\theta) \in \left\{IF_k^{est}(\theta) + \overline{B}_k^{est,\perp}(\theta); \overline{B}_k^{est,\perp}(\theta) \in \overline{\mathcal{B}}_k^{est,\perp}(\theta)\right\}$ where $\overline{\mathcal{B}}_k^{est,\perp}(\theta)$ is the orthocomplement of $\overline{\mathcal{B}}_k^{est}(\theta)$

(iv) For $m < k$, $\Pi\left[IF_k^{est}(\theta)\,|\overline{\mathcal{B}}_m^{est}(\theta)\right] = IF_m^{est}(\theta)$

Proof: See Robins and van der Vaart (2004):

Definition of the kth order efficient influence function and variance: $IF_k^{est}(\theta)$ is referred to as the the kth order efficient estimation influence function and its variance as the kth *order* efficient estimation variance.

Consider again model $M(\Theta)$ with parameter space Θ. When the parameter space $\Theta = \prod_{r=1}^{r=R}\Theta_r$ is a cartesian product of sets Θ_r, so $\theta = (\theta_1, ...\theta_R), \theta_r \in \Theta_r$,we say a parametric submodel $\widetilde{\theta}(\varsigma), \varsigma \in \mathcal{Z}$, is variation independent wrt the Θ_r if $\varsigma = (\varsigma_1, ..., \varsigma_R), \varsigma_r \in \mathcal{Z}_r, \mathcal{Z} = \prod_{r=1}^{r=R}\mathcal{Z}_r$, and $\widetilde{\theta}(\varsigma) = \left(\widetilde{\theta}_1(\varsigma), ..., \widetilde{\theta}_R(\varsigma)\right) = \left(\widetilde{\theta}_1(\varsigma_1), ..., \widetilde{\theta}_R(\varsigma_R)\right)$ with $\widetilde{\theta}_r(\varsigma_r) \in \Theta_r$. An m dimensional score $S_{i_1...i_m}(\theta)$ of a variation independent submodel is a member of a particular set of scores B_{mt_m} with generic member $V_{\theta_1^{t_{m1}},...,\theta_R^{t_{mR}}}$ where the t_{mr} are components of an $R-$vector $t_m = (t_{m1}, ..., t_{mR})$ satisfying $\sum_{r=1}^{R} t_{mr} = m$ with t_{mr} determined by $S_{i_1...i_m}(\theta)$ via $t_{mr} = \sum_{j=1}^{m} I\left(\varsigma_{i_j} \in \mathcal{Z}_r\right)$. The components t_{mr} tell how many of the m derivatives in $S_{i_1...i_m}(\theta)$ were with respect to components of ς that lay in $\mathcal{Z}_r$. We let $B_{mt_m} = \left\{V_{\theta_1^{t_{m1}},...,\theta_R^{t_{mR}}}\right\}$ denote the set of all order m scores of variation independent parametric submodels with a given value of t_m. Then in general $\overline{\mathcal{B}}_k$ is the closed linear span of the union $\cup_{m=1}^{k}\cup_{\{t_m\}} B_{mt_m}$ of variation independent scores where $\cup_{\{t_m\}} refers$

to the *union* over all vectors of length R with nonnegative integer components whose components sum to m.

Example: If $m = 2$ and $R = 4$, the number of sets B_{2t_2} is 10 as there are 10 vectors of length 4 with nonnegative integer components whose components sum to 2.

Suppose the model can be parametrized as $\theta = (\psi, \gamma)$, $\psi \in \Psi, \gamma \in \Gamma, \Theta = \Psi \times \Gamma$ (at least locally). That is, in the above notation, with $\Theta = \prod_{r=1}^{r=2} \Theta_r$, we can then take $\Theta_1 = \Psi$ and $\Theta_2 = \Gamma$. We refer to the 1st order ($m = 1$) scores $V_\gamma = V_{\gamma^1}$ as (pure) nuisance scores and $V_\psi = V_{\psi^1}$ as the score for ψ. For $m > 1$, we refer to generic scores (i) V_{γ^m}as mth order pure nuisance scores ; (ii) V_{ψ^m} as mth order scores for ψ and (iii) $V_{\psi^c \gamma^{m-c}}, m > c > 0$ as mth order mixed scores. The closed linear span of $\cup_{m=1}^{k} \{V_{\gamma^m}\}$ of all pure nuisance scores of order k or less is $\overline{\mathcal{B}}_k^{test}$. The closed linear span $\left[\cup_{m=1}^{k} \{V_{\gamma^m}\}\right] \cup \left[\cup_{m=2}^{k} V_{\psi^m}\right] \cup \left[\cup_{m=1}^{k} \cup_{c=1}^{c=m-1} \left\{V_{\psi^c \gamma^{m-c}}\right\}\right]$ of all scores excepting the 1st order ($m = 1$) score V_ψ for ψ is the estimation nuisance tangent space $\overline{\Lambda}_k^{est}(\theta)$. Finally the estimation tangent space $\overline{\mathcal{B}}_k^{est}(\theta)$ is the closed linear span of all the scores of order k or less. Note if $\Theta = \prod_{r=1}^{r=R} \Theta_r$ and the likelihood for one observation factors as $f(O;\theta) = \prod_{r=1}^{r=R} L_r(\theta_r)$ and $S_{i_1 \ldots i_R, j}(\theta)$ is a mixed score $V_{\theta_1 \theta_2 \cdots \theta_R}$, then $S_{i_1 \ldots i_R, j} = \prod_{r=1}^{r=R} S_{i_r, j}$.

Definition of the kth order efficient testing and estimation scores and information: Suppose $\theta = (\psi, \gamma)$, $\psi \in \Psi, \gamma \in \Gamma, \Theta = \Psi \times \Gamma$ (at least locally). We define the kth order efficient testing score $ES_k^{test}(\theta) = \Pi_\theta \left[V_\psi(\theta) | \overline{\mathcal{B}}_k^{test,\perp}\right]$ to be the projection of the first order score $V_\psi(\theta)$ on the orthogonal complement in $\mathcal{U}_k(\theta)$ of the kth order testing tangent space. We define the kth order efficient estimation score $ES_k^{est}(\theta) = \Pi_\theta \left[V_\psi(\theta) | \overline{\Lambda}_k^{est\perp}(\theta)\right]$ to be the projection of the first order score $V_\psi(\theta)$ on the orthogonal complement in $\mathcal{U}_k(\theta)$ of the kth order estimation nuisance tangent space. We call the variances $E_\theta \left[ES_k^{test}(\theta) ES_k^{test}(\theta)^T\right]$ and $E_\theta \left[ES_k^{est}(\theta) ES_k^{est}(\theta)^T\right]$ of the kth order efficient scores the kth order testing and estimation efficient informations.

Remark: Note that $E_\theta \left[ES_k^{test}(\theta) ES_k^{test}(\theta)^T\right] \geq E_\theta \left[ES_k^{est}(\theta) ES_k^{est}(\theta)^T\right]$ since $\overline{\Lambda}_k^{est\perp} \subseteq \overline{\mathcal{B}}_k^{test,\perp}$ so the efficient testing information is greater than or equal to the efficient estimation information. For $k = 1$, $\overline{\Lambda}_k^{est\perp} = \overline{\mathcal{B}}_k^{test,\perp}$, so $ES_1^{test}(\theta) = ES_1^{est}(\theta)$

Efficient Score Lemma: Suppose the model can be parametrized as $\theta = (\psi, \gamma)$, $\psi \in \Psi, \gamma \in \Gamma, \Theta = \Psi \times \Gamma$ (at least locally), then $IF_k(\theta) = \left\{E_\theta \left[ES_k^{est}(\theta) ES_k^{est}(\theta)^T\right]\right\}^{-1} ES_k^{est}(\theta)$, so the kth order estimation efficient variance is the inverse $\left\{E_\theta \left[ES_k^{est}(\theta) ES_k^{est}(\theta)^T\right]\right\}^{-1}$ of the kth order estimation information.

Proof: see Robins and van der Vaart (2004).

The main ideas: Here are the main ideas behind using higher order influence functions in models in which they exist. [In models which are so large that higher order influence functions do not exist, we will consider a lower dimensional working model that admits higher order influence functions and allow the dimension of the working model (the sieve) to increase with sample size. A worked example is given later.] Consider the estimator $\widehat{\psi}_k = \psi\left(\widehat{\theta}\right) + U_k\left(\widehat{\theta}\right)$ based on a sample size n where $\widehat{\theta}$ is an initial estimator of θ from a separate sample (based on random sample splitting) that perhaps obtains the optimal rate of convergence for θ and $U_k(\theta)$ is a kth order estimation influence function. It would be optimal to choose $U_k(\theta)$ equal to $IF_k^{est}(\theta)$. Expanding and evaluating conditonally on $\widehat{\theta}$, we have

$$\begin{aligned}\widehat{\psi}_k - \psi(\theta) &= \left\{\psi\left(\widehat{\theta}\right) - \psi(\theta) + U_k\left(\widehat{\theta}\right) - U_k(\theta)\right\} + U_k(\theta) \\ &= U_k(\theta) + \left\{U_k\left(\widehat{\theta}\right) - U_k(\theta) - E_\theta\left[U_k\left(\widehat{\theta}\right) - U_k(\theta)\,|\widehat{\theta}\right]\right\} + \\ &\qquad \left\{\psi\left(\widehat{\theta}\right) - \psi(\theta) + E_\theta\left[U_k\left(\widehat{\theta}\right)|\widehat{\theta}\right]\right\}\end{aligned}$$

Now under weak conditions $var\left\{U_k\left(\widehat{\theta}\right) - U_k(\theta)\,|\widehat{\theta}\right\}/var\left[U_k(\theta)\right] = o_p(1)$ unconditonally since we assume, with unconditional probability approaching one, $\left|\left|\widehat{\theta} - \theta\right|\right| \to 0$ as $n \to \infty$ for a norm on Θ for which $var\left\{U_k\left(\widehat{\theta}\right) - U_k(\theta)\,|\widehat{\theta}\right\}/var\left[U_k(\theta)\right]$ is continuous in $\widehat{\theta}$ at θ. Thus given $\widehat{\theta}$, the distance $d\left(\widehat{\psi}_k - \psi(\theta), U_k(\theta) + \left\{\psi\left(\widehat{\theta}\right) - \psi(\theta) + E_\theta\left[U_k\left(\widehat{\theta}\right)\right]\right\}\right)$ is converging to 0 where $d(\cdot,\cdot)$ is a distance that metrizes weak convergence. If $\psi\left(\widehat{\theta}\right) - \psi(\theta) + E_\theta\left[U_k\left(\widehat{\theta}\right)\right]$ has $k+1$ Frechet derivatives in $\widehat{\theta}$ in a neighborhood of $\widehat{\theta} = \theta$, then by part (ii) of the extended information equality theorem, we expect that the multilinear operator of m arguments corresponding to the mth Frechet derivative at $\widehat{\theta} = \theta$ would be zero for $m = 1, ..., k$. Hence $\psi\left(\widehat{\theta}\right) - \psi(\theta) + E_\theta\left[U_k\left(\widehat{\theta}\right)\right]$ [and thus the bias of $\widehat{\psi}_k - \psi(\theta)$] will be $O_p\left(\left|\left|\widehat{\theta} - \theta\right|\right|^{k+1}\right)$, which decreases with k for fixed n. On the other hand, in view of part (iv) of the efficient influence function theorem, we know the variance $var\left[U_k(\theta)\right]$ increases with k. Thus the rate of convergence of $\widehat{\psi}_k$ to $\psi(\theta)$ is minimized at $k(n)$ equal to $k_{balance} = k_{balance}(n)$ at which the squared bias and the variance are of the same order. Further the shortest conservative uniform asymptotic confidence intervals will be based on $\widehat{\psi}_{k_{conf}} \pm var\left[U_{k_{conf}}\left(\widehat{\theta}\right)|\widehat{\theta}\right]^{1/2} z_\alpha c_{est}$ where $k_{conf} = k_{conf}(n)$ is the smallest value of k such that $var\left[U_k(\theta)\right]$ is of higher

order than the squared bias and $c_{est}, c_{est} \geq 1$, is an appropriate constant chosen to guarantee coverage $1-\alpha$, as $U_k\left(\widehat{\theta}\right)$ may not be normal and we might use a tail bound based on, say, Markov's inequality.

In this same setting if the model can be parametrized as $\theta = (\psi, \gamma)$, then, for an appropriate choice of k, conservative uniform asymptotic confidence intervals can be constructed as

$$\left\{\psi; \left|ES_k^{test}\left(\psi, \widehat{\gamma}\left(\psi\right)\right) / \widehat{var}\left[ES_k^{test}\left(\psi, \widehat{\gamma}\left(\psi\right)\right)|\widehat{\gamma}\left(\psi\right)\right]^{1/2}\right| < z_\alpha c_{test}\right\}$$

where $\widehat{var}\left[ES_k^{test}\left(\psi, \widehat{\gamma}\left(\psi\right)\right)|\widehat{\gamma}\left(\psi\right)\right]$ is an appropriate variance estimator and c_{test} is similiar to c_{est}. Because $ES_k^{test}(\psi, \gamma)$ is not orthogonal to the higher order scores for ψ and the mixed $\psi-\gamma$ scores, in an expansion of $ES_k^{test}\left(\psi^*, \widehat{\gamma}\left(\psi^*\right)\right)$ around $\psi^\dagger$, the $\psi-$derivative $ES_{k,\psi}^{test}\left(\dagger, \widehat{\gamma}\left(\psi^\dagger\right)\right)$ will typically be of the same order as the mixed derivatives $ES_{k,\psi^c\gamma^{m-c}}^{test}\left(\psi^\dagger, \widehat{\gamma}\left(\psi^\dagger\right)\right)$ and the higher order $\psi-$derivatives $ES_{k,\psi^m}^{test}\left(\psi^\dagger, \widehat{\gamma}\left(\psi^\dagger\right)\right)$. Nonetheless, because in the expansion $ES_{k,\psi^c\gamma^{m-c}}^{test}\left(\psi^\dagger, \widehat{\gamma}\left(\psi^\dagger\right)\right)$ is multiplied by $\left(\psi^* - \psi^\dagger\right)^c\left(\widehat{\gamma}\left(\psi^*\right) - \widehat{\gamma}\left(\psi^\dagger\right)\right)^{m-c}$ and $ES_{k,\psi^m}^{test}\left(\psi^\dagger, \widehat{\gamma}\left(\psi^\dagger\right)\right)$ is multiplied by $\left(\psi^* - \psi^\dagger\right)^m, m > 1$, but $ES_{k,\psi}^{test}\left(\psi^\dagger, \widehat{\gamma}\left(\psi^\dagger\right)\right)$ is only multiplied by $\left(\psi^* - \psi^\dagger\right)$, the asymptotic distribution of the solution $\widehat{\psi}_{eff,k}^{test}$to $ES_k^{test}(\psi, \gamma) = 0$ will often be unaffected by the fact that $ES_k^{test}(\psi, \gamma)$ is not orthogonal to the higher order scores for ψ and the mixed $\psi - \gamma$ scores. Thus $\widehat{\psi}_{eff,k_{conf}}^{test}$ may have smaller limiting variance than the solution $\widehat{\psi}_{eff,k_{conf}}^{est}$ to $ES_k^{est}\left(\psi, \widehat{\gamma}\left(\psi\right)\right) = 0$ without incurring greater bias. Under further regularity conditions, for an appropriate choice of k, $\widehat{\psi}_{eff,k}^{test}$ will typically have 'limiting' variance $\left\{var_{(\psi,\gamma)}\left[ES_k^{test}(\psi, \gamma)\right]\right\}^{-1}$. Note, for $m \leq k$,

$$\left\{var_{(\psi,\gamma)}\left[ES_k^{test}(\psi, \gamma)\right]\right\}^{-1} = \tau_{km}^{-1} var_{(\psi,\gamma)}\left[ES_k^{test}(\psi, \gamma)\right]\tau_{km}^{-1,T}$$

$$\tau_{km} = E_{(\psi,\gamma)}\left[ES_k^{test}(\psi, \gamma)\, ES_m^{test}(\psi, \gamma)^T\right]$$

.

The lesson here is that for a given functional $\psi(\theta)$ the optimal procedure is not necessarily based on $IF_k^{est}(\theta)$ for any k as one only needs to consider those components of θ that can make the bias exceed the variance. As an example we have just seen that when $\theta = (\psi, \gamma)$ it is often not important to be orthogonal to the nuisance scores corresponding to the higher order scores for ψ and the mixed $\psi - \gamma$ scores, even though it may be important to be orthogonal to higher order scores for some or all components of γ (depending on the rate at which particular components are estimable).

We now apply this methodology to our semiparametric regression model. In Robins and van der Vaart (2003, 2004) we used this methodology (i) to obtain an alternative derivation of some results due to Ritov and Bickel (1988), Laurent and Massart (2000), and Laurent (1996) concerning the estimation of $\int f^2(x)\, dx$,(ii) to construct conditonal interval estimates for the functionals

$\int \left[\widehat{f}(x) - f(x)\right]^2 dx$ and $\int \left[\widehat{E}[Y|X=x] - E[Y|X=x]\right]^2 dx$, obtaining as a by-product improved adaptive confidence intervals for $f(x)$ and $E[Y|X=x]$ compared to those of Lepski and Hoffmann (2002), and (iii) to construct point and interval estimators for finite dimensional parameters in complex missing and censored data models that improve on the doubly robust estimators of Robins, Rotnitzky and van der Laan (2000) and van der Laan and Robins (2002).

Application To Semiparametric Regression: We are now ready to return to the semiparametric regression example. $\Theta = \prod_{r=1}^{r=4} \Theta_r = \Psi \times \mathcal{N} \times \mathring{A} \times \Omega$. Let $\Delta(\alpha) = A - E_\alpha(A|X)$, $e(\psi,\eta) = Y - \psi A - b(X;\eta)$.Then the generic subject -specific first order scores for $\Psi, \mathcal{N}, \mathring{A}$, and Ω are $V_{\psi i}(\psi,\eta) = Ae_i(\psi,\eta)$, $V_{\eta i}(\psi,\eta,g_\eta) = g_\eta(X_i)\, e_i(\psi,\eta)$, $V_{\alpha i}(\alpha,g_\alpha)$
$= g_\alpha(X_i)\,\Delta_i(\alpha)$, $V_{\omega i}(\omega,g_\omega) = g_\omega(X_i)$,where $E_\omega[g_\omega(X_i)] = 0$. Thus the set of composite first order scores evaluated at the truth $\theta^\dagger$ are

$$\left\{\sum_i V_{\psi i}\right\} = \left\{\sum_i Ae_i\right\} \tag{9.10a}$$

$$\left\{\sum_i V_{\eta i}(g_\eta)\right\} = \left\{\sum_i g_\eta(X_i)\, e_i; g(.)\ \ unrestricted\right\} \tag{9.10b}$$

$$\left\{\sum_i V_{\alpha i}(g_\alpha)\right\} = \left\{\sum_i g_\alpha(X_i)\,\Delta_i; \Delta = A - E_{\alpha^\dagger}(A|X), g_a(.)\, unrestricted\right\} \tag{9.10c}$$

$$\left\{\sum_i V_{\omega i}(g_\omega)\right\} = \{g_\omega(X_i); E_{\omega^\dagger}[g_\omega(X_i)] = 0\} \tag{9.10d}$$

The first order estimation and testing tangent space $\overline{\mathcal{B}}_1^{est}$ is the closed linear span (cls) of the union of the sets (9.10a)-(9.10d). The first order estimation nuisance tangent space $\overline{\mathcal{B}}_1^{test}$and the first order testing nuisance tangent space $\overline{\Lambda}_1^{est}$ equal the cls of (9.10b)-(9.10d) The second order estimation tangent $\overline{\mathcal{B}}_2^{est}$ space is the cls of the union of $\overline{\mathcal{B}}_1^{est}$ and the 10 sets of second order scores.

$\{V_{\psi\psi}\} = \left\{\sum_i V_{\psi i}V_{\psi i} - A_i^2 + \sum\sum_{i\neq j} V_{\psi i}V_{\psi j}\right\}$

$\{V_{\psi\eta}(g_\eta)\} = \left\{\sum_i V_{\psi i}V_{\eta i}(g_\eta) - g_\eta(X_i)A_i + \sum\sum_{i\neq j} V_{\psi i}V_{\eta j}(g_\eta)\right\}$

$\{V_{\psi\omega}(g_\omega)\} = \left\{\sum_i V_{\psi i}V_{\omega i}(g_\omega) + \sum\sum_{i\neq j} V_{\psi i}V_{\omega j}(g_\omega)\right\}$

$\{V_{\alpha\psi}(g_\alpha)\} = \left\{\sum_i V_{\alpha i}(g_\alpha)V_{\psi i} + \sum\sum_{i\neq j} V_{\alpha i}(g_\alpha)V_{\psi j}\right\}$

$= \left\{\sum_i g_\alpha(X_i)\,\Delta_i e_i A_i + \sum\sum_{i\neq j} g_\alpha(X_i)\,\Delta_i e_j A_j\right\}$

$\{V_{\eta\omega}(g_\eta,g_\omega)\} = \left\{\sum_i V_{\eta i}(g_\eta)V_{\omega i}(g_\omega) + \sum\sum_{i\neq j} V_{\eta i}(g_\eta)V_{\omega j}(g_\omega)\right\}$

$$\{V_{\alpha\omega}(g_\alpha, g_\omega)\} = \left\{\sum_i V_{\alpha i}(g_\alpha) V_{\omega i}(g_\omega) + \sum\sum_{i\neq j} V_{\alpha i}(g_\alpha) V_{\omega j}(g_\omega)\right\}$$
$$\{V_{\alpha\eta}((g_\eta, g_\alpha))\} = \left\{\sum_i V_{\alpha i}(g_\alpha) V_{\eta i}(g_\eta) + \sum\sum_{i\neq j} V_{\alpha i}(g_\alpha) V_{\eta j}(g_\eta)\right\}$$
$$\left\{V_{\eta\eta}\left(g_\eta, g_\eta^*, g_\eta^{**}\right)\right\}$$
$$= \left\{\sum_i \left[e_i^2 - 1\right] g_\eta(X_i) g_\eta^*(X_i) + e_i g_\eta^{**}(X_i)\right] + \sum\sum_{i\neq j} g_\eta(X_i) e_i e_j g_\eta^*(X_j)\right\}$$
$$\{V_{\alpha\alpha}(g_\alpha, g_\alpha^*, g_\alpha^{**})\} = \left\{\sum_i \Delta_i g_\alpha^{**}(X_i) + \sum\sum_{i\neq j} g_\alpha(X_i) g_\alpha^*(X_i) \Delta_i \Delta_j\right\}$$
$$\{V_{\omega\omega}(g_\omega, g_\omega^*, g_\omega^{**})\} = \left\{\sum_i g_\omega^{**}(X_i) + \sum\sum_{i\neq j} g_\omega(X_i) g_\omega^*(X_i)\right\},$$

with $g_\omega^{**}(X_i), g_\omega(X_i)$, and $g_\omega^*(X_i)$ having mean zero. The second estimation nuisance tangent $\overline{\Lambda}_2^{est}$ space is the cls of the union of $\overline{\Lambda}_1^{est}$ and the 10 sets of second order scores. The second order testing tangent space is the cls of the union of $\overline{\mathcal{B}}_1^{test}$ and the 6 sets of second order scores that do not involve ψ.

The first order efficient (testing and estimation) score ES_1 for ψ is easily seen to be $\Pi\left[V_\psi | \Lambda_1^\perp\right] = \sum_i \Delta_i e_i$. The following Lemma proved in Robins and van der Vaart (2004) gives ES_2^{test}. Recall X is discrete.

Lemma: Let $c^*(X_i) = \{1 + (n-1) f(X_i)\}^{-1}, v(X_i) = var(A_i | X_i)$

$$ES_2^{test} = ES_1 - \Pi\left[ES_1 | \overline{\{V_{\alpha\eta}\}}\right] \tag{9.11}$$
$$= \sum_i \Delta_i e_i \{(n-1) f(X_i)\} c^*(X_i) - \sum\sum_{i\neq j} \Delta_i e_j I(X_i = X_j) c^*(X_i)$$

Remark: ES_2^{test} differs from ES_1 because ES_1 is not orthogonal to the scores $\{V_{\alpha\eta}\}$. Robins and van der Vaart also show that ES_2^{test} differs from ES_2^{est} because ES_2^{test} is not orthogonal to the mixed scores $\{V_{\alpha\psi}\}$ and the use of ES_2^{est} may result in a loss of efficiency. ES_1 and ES_2^{test} are the same whether the marginal distribution X is known, known to lie in a low dimensional model, or completely unknown. Recall $ES_1(\psi, \alpha, \eta)$ is doubly robust in the sense $E_{(\psi^\dagger, \alpha^\dagger, \eta^\dagger, \omega^\dagger)}\left[ES_1\left(\psi^\dagger, \alpha^\dagger, \eta\right)\right] = E_{(\psi^\dagger, \alpha^\dagger, \eta^\dagger, \omega^\dagger)}\left[ES_1\left(\psi^\dagger, \alpha, \eta^\dagger\right)\right]$. Strikingly ES_2^{test} is triply robust in that it has mean 0 if one of the three nuisance parameters $\left(\alpha^\dagger, \eta^\dagger, \omega^\dagger\right)$ are correct. That is,

$$E_{(\psi^\dagger, \alpha^\dagger, \eta^\dagger, \omega^\dagger)}\left[ES_2^{test}\left(\psi^\dagger, \alpha^\dagger, \eta, \omega\right)\right] = E_{(\psi^\dagger, \alpha^\dagger, \eta^\dagger, \omega^\dagger)}\left[ES_2^{test}\left(\psi^\dagger, \alpha, \eta^\dagger, \omega\right)\right] \tag{9.12}$$
$$= E_{(\psi^\dagger, \alpha^\dagger, \eta^\dagger, \omega^\dagger)}\left[ES_2^{test}\left(\psi^\dagger, \alpha, \eta, \omega^\dagger\right)\right] = 0$$

ES_2^{test} has variance

$$E\left[v(X) f^2(X)\left\{b^*(X_i)\right\}^2 (n-1)^2 n + f(X_i) v(X)\left\{b^*(X_i)\right\}^2 (n-1) n\right]$$
$$= E\left[\frac{(n-1) f(X_i)}{\{1 + (n-1) f(X_i)\}} n [v(X)]\right]$$

while ES_1 has the greater variance $E[v(X) n]$. It is interesting to consider the case where X_i has n^ρ levels and $f(X_i) = n^{-\rho}$. Then $\{var[ES_1]\}^{-1}$ is always $O\left(n^{-1}\right)$ while $var[ES_2^{test}]$ is $O\left(\min\left\{n, n^2/n^\rho\right\}\right)$ and $\{var[ES_2^{test}]\}^{-1}$ is

$O\left(\max\left\{n^{-1}, n^{\rho}/n^{2}\right\}\right)$. On the other hand the conditional bias of $ES_1(\psi, \widehat{\eta}, \widehat{\alpha})$ (i.e., $E[ES_1(\psi, \widehat{\eta}, \widehat{\alpha})|\widehat{\eta}, \widehat{\alpha}]$), where $(\widehat{\eta}, \widehat{\alpha})$ is obtained from an independent sample, is $n\sum_l \left\{p\left(l, \alpha_l^{\dagger}\right) - p(l, \widehat{\alpha}_l)\right\}\left\{b\left(l, \eta_l^{\dagger}\right) - b(l, \widehat{\eta}_l)\right\}$. If $\rho < 1$ and $b(l, \widehat{\eta}_l)$ is the empirical mean of $Y - \psi A$ in statum l and $p(l, \widehat{\alpha}_l)$ is the empirical mean of A in stratum l, the conditional bias is approximated by $n\left\{E\left[var\{p(L, \widehat{\alpha}_L)|L\}^{1/2} var\{b(L, \widehat{\eta}_L)|L\}^{1/2}\right]\right\} = nO\left(\left[\left\{n^{-(1-\rho)}\right\}^{1/2}\right]^2\right) = O(n^{\rho})$. Now to be able to set conditional confidence intervals given $(\widehat{\alpha}, \widehat{\eta})$ based on $ES_1(\psi, \widehat{\eta}, \widehat{\alpha})$ we need the bias squared of $O\left(n^{2\rho}\right)$ to be less than the variance of $O(n)$,thus requiring $\rho < 1/2$. In contrast the bias of $ES_2^{test}\left(\psi, \widehat{\eta}, \widehat{\alpha}, \omega^{\dagger}\right)$ is always zero when $\omega^{\dagger}$ is known and has variance $O\left(\min\left\{n, n^2/n^{\rho}\right\}\right)$. In fact $ES_2^{test}\left(\psi, \widehat{\eta}, \widehat{\alpha}, \omega^{\dagger}\right)$ has variance converging to the first order efficient information $E[v(X)n]$ for $\rho < 1$. Let $\widehat{\psi}_1(\widehat{\eta}, \widehat{\alpha})$ be the solution to $ES_1(\psi, \widehat{\eta}, \widehat{\alpha}) = 0$. Conditional on $(\widehat{\eta}, \widehat{\alpha})$, the estimator $\widehat{\psi}_2^{test}(\widehat{\eta}, \widehat{\alpha})$ solving $ES_2^{test}\left(\psi, \widehat{\eta}, \widehat{\alpha}, \omega^{\dagger}\right) = 0$, with $\omega^{\dagger}$ assumed known, is semiparametric efficent for $\rho < 1$, $n^{1/2} - consistent$ for $\rho = 1$, converges at rate $n^{1-\rho/2}$ for $2 > \rho \geq 1$, and is inconsistent for $2 > \rho$.

In summary, conditional on $(\widehat{\eta}, \widehat{\alpha})$, our 2nd order $U - statistic$ is necessary to obtain first order semiparametric efficiency for $1/2 \leq \rho < 1$, because the bias of our usual doubly robust estimator $\widehat{\psi}_1(\widehat{\eta}, \widehat{\alpha})$ is too large. Our second order estimator corrects the bias without adding to the limiting variability of $\widehat{\psi}_1(\widehat{\eta}, \widehat{\alpha})$ for $1/2 \leq \rho < 1$. Results exactly analogous to ours have been obtained for other quadratic functions by a number of other investigators such as Bickel and Ritov (1988) and Laurent and Massart (2000).

The reason that it is not possible to obtain a consistent estimator if $\rho > 2$ is that to control bias, it was necessary that in the "degenerate" part of the U-statistic, the pair i and j only contribute if they have the same value of X. When we toss n subjects randomly onto a grid with $n^{\rho}, \rho > 1$, compartments, one can show that the number of compartments containing more than one subject goes as $n^{(2-\rho)}$ and thus we will obtain an infinite amount of information as $n \to \infty$ only if $\rho \leq 2$.

Suppose now we do not know the law of X apriori. When $f(X)$ (i.e., $\psi^{\dagger}$) is known ES_2^{test} is uncorrelated with $\overline{\mathcal{B}}_k^{test}$ for $k \geq 2$, since $E_{(\psi^{\dagger}, \alpha^{\dagger}, \eta^{\dagger}, \omega^{\dagger})}\left[ES_2^{test}\left(\psi^{\dagger}, \alpha, \eta, \omega^{\dagger}\right)\right] = 0$ for all (α, η). However when $\omega^{\dagger}$ is unknown, the third order scores

$$
\begin{aligned}
&V_{\alpha\eta\omega}\left(g_\omega, g_\alpha, g_\eta\right) \\
&= \sum_i g_\omega\left(X_i\right) g_\alpha\left(X_i\right) g_\eta\left(X_i\right) \Delta_i e_i + \sum\sum_{i\neq j} g_\omega\left(X_i\right) g_\alpha\left(X_i\right) g_\eta\left(X_j\right) \Delta_i e_j + \\
&\sum\sum_{i\neq j} g_\omega\left(X_i\right) g_\alpha\left(X_j\right) g_\eta\left(X_j\right) \Delta_j e_j + g_\omega\left(X_j\right) g_\alpha\left(X_i\right) g_\eta\left(X_j\right) \Delta_i e_j + \\
&\sum\sum\sum_{i\neq j\neq s} g_\omega\left(X_s\right) g_\alpha\left(X_i\right) \Delta_i e_j g_\eta\left(X_j\right)
\end{aligned}
$$

are correlated with ES_2^{test}, *i.e.*, $E\left[ES_2^{test} V_{\alpha\eta\omega}\left(g_\omega, g_\alpha, g_\eta\right)\right] \neq 0$, implying third order bias. We could elimate the third order bias by calculating $ES_3^{test} = \Pi\left[ES_2^{test} | \overline{\mathcal{B}}_3^{test,\perp}\right]$. However rather than take this approach define $ES_2^{test}\left(c\right) = \sum_i \Delta_i e_i \left\{\left(n-1\right) f\left(X_i\right)\right\} c\left(X_i\right) - \sum\sum_{i\neq j} \Delta_i e_j I\left(X_i = X_j\right) c\left(X_i\right)$ for any $c\left(X_i\right)$. Note $f\left(X_i\right)$ appears only at one place in $ES_2^{test}\left(c\right)$. Let $ES_2^{test}\left(\psi, \widehat{\eta}, \widehat{\alpha}, \widehat{\omega}_-; c\right)$ be $ES_2^{test}\left(\psi, \widehat{\eta}, \widehat{\alpha}, \omega^\dagger\right)$ with c replacing c^* and with $f\left(X_i\right)$ replaced by

$$
\widehat{f}_{-i}\left(X_i\right) = \sum_{\{j: j\neq i\}} I\left(X_j = X_i\right) / \left(n-1\right).
$$

Define $H\left(\psi\right) = Y - \psi A$, $e_i\left(\psi, \widehat{\eta}\right) = H\left(\psi\right) - b\left(X, \widehat{\eta}\right)$, and $\Delta_i\left(\widehat{\alpha}\right) = A_i - E_{\widehat{\alpha}}\left[A | X_i\right]$. Then

$$
\begin{aligned}
&ES_2^{test}\left(\psi, \widehat{\eta}, \widehat{\alpha}, \widehat{\omega}_-; c\right) \\
&= \sum_i \Delta_i\left(\widehat{\alpha}\right) e_i\left(\psi, \widehat{\eta}\right) \widehat{f}_{-i}\left(X_i\right) c\left(X_i\right)\left(n-1\right) - \\
&\sum\sum_{i\neq j} c\left(X_i\right) I\left(X_i = X_j\right) \Delta_i\left(\widehat{\alpha}\right) e_j\left(\psi, \widehat{\eta}\right) \qquad (9.13a) \\
&= ES_2^{test}\left(\psi, \widehat{\eta}, \widehat{\alpha}, \omega^\dagger; c\right) + \sum\sum_{i\neq j} \left\{I\left(X_i = X_j\right) - f\left(X_i\right)\right\} c\left(X_i\right) \Delta_i\left(\widehat{\alpha}\right) e_i\left(\psi, \widehat{\eta}\right) \qquad (9.13b) \\
&= \sum\sum_{i\neq j} I\left(X_i = X_j\right) c\left(X_i\right) H_i\left(\psi\right)\left(A_i - A_j\right) \qquad (9.13c)
\end{aligned}
$$

where $|\mathcal{X}|$ is the cardinality of the support $\mathcal{X}$ of X. Note that (9.13b) shows that $ES_2^{test}\left(\psi, \widehat{\eta}, \widehat{\alpha}, \widehat{\omega}_-; c\right)$ has mean zero and (9.13c) shows that $ES_2^{test}\left(\psi, \widehat{\eta}, \widehat{\alpha}, \widehat{\omega}_-; c\right) = ES_2^{test}\left(\psi, \widehat{\omega}_-; c\right)$ does not depend on the nuisance parameters (η, ω, α) and thus is orthogonal to $\overline{\mathcal{B}}_k^{test}$ regardless of the order k. That is, $ES_2^{test}\left(\psi, \widehat{\omega}_-; c\right)$ is exactly unbiased for 0. [This would not be the case had we estimated $f\left(X_i\right)$ by $\widehat{f}\left(X_i\right) = \sum_{j=1}^{n} I\left(X_j = X_i\right)/n$]. However even though $ES_2^{test}\left(\psi, \widehat{\omega}_-; c\right)$ is orthogonal to the 3rd order testing nuisance tangent space $\overline{\mathcal{B}}_3^{test}$, nonetheless $ES_2^{test}\left(\psi, \widehat{\omega}_-; c\right)$ with $c\left(X_i\right) = c^*\left(X_i\right)$ is not the

residual from the projection of ES_2^{test} on $\overline{\mathcal{B}}_3^{test}$, so it is not fully third order efficient.

Remark: Define $\overline{Z}_x = \sum_i Z_i I(X_i = x) / \sum_i I(X_i = x)$ for any Z_i and $\overline{Z}_{x,-j} = \sum_{\{i;i\neq j\}} Z_i I(X_i = x) / \sum_{\{i;i\neq j\}} I(X_i = x)$. Let $\widetilde{\eta}(x) = \overline{H}_x(\psi), \widetilde{\alpha}(x) = \overline{A}_x$. Note that the statistic
$ES_2^{test}(\psi, \widehat{\omega}_-; c) = \sum_{x=1}^{|\mathcal{X}|} \sum_i I(X_i = x) c(x) H_i(\psi)(n-1)\left(A_i - \overline{A}_{x,-i}\right)$. Hence, it is closely related to the statisitic $ES_1(\psi, \widetilde{\eta}, \widetilde{\alpha}) =$

$$\sum_{x=1}^{|\mathcal{X}|} \sum_i I(X_i = x)\left(H_i(\psi) - \overline{H}_x(\psi)\right)\left(A_i - \overline{A}_x\right)$$

$$= \left\{\frac{n-1}{n}\right\}^2 \sum_{x=1}^{|\mathcal{X}|} \sum_i I(X_i = x)\left(H_i(\psi) - \overline{H}_{x,-i}(\psi)\right)\left(A_i - \overline{A}_{x,-i}\right)$$

$$= \left\{\frac{n-1}{n^2}\right\} ES_2^{test}(\psi, \widehat{\omega}_-; c) - \left\{\frac{n-1}{n}\right\}^2 \sum_{x=1}^{|\mathcal{X}|} \sum_i I(X_i = x) \overline{H}_{x,-i}(\psi)\left(A_i - \overline{A}_{x,-i}\right)$$

that uses the same sample rather than a different sample to estimate (η, α). $ES_1(\psi, \widetilde{\eta}, \widetilde{\alpha})$ is thus also orthogonal to all $\overline{\mathcal{B}}_k^{test}$ and is unbiased for 0 (Donald and Newey, 1994). $ES_2^{test}(\psi, \widehat{\omega}_-; c)$ and $ES_2^{test}(\psi, \widehat{\omega}_-; c)$ are asymptotically normal under weak conditions. To see this let
$R(x) = \sum_i I(X_i = x)\left(H_i - \overline{H}_x\right)\left(A_i - \overline{A}_x\right)$ and $N(x) = \sum_i I(X_i = x)$. Then, given
$\mathbf{X} = \{X_1, ..., X_n\}$, $ES_1(\psi, \widetilde{\eta}, \widetilde{\alpha}) = \sum_{x=1}^{|\mathcal{X}|} R(x) = \sum_{x=1}^{|\mathcal{X}|} I\{N(x) \geq 2\} R(x)$ is a sum of $IS = \sum_{x=1}^{|\mathcal{X}|} I\{N(x) \geq 2\}$ independent mean zero random variables $I\{N(x) \geq 2\} R(x)$ with $IS \to \infty$ as $n \to \infty$ provided $|\mathcal{X}| = O(n^\rho), 0 < \rho < 2$. The ability to substitute an 'own' sample estimate of η and α into $ES_1(\psi, \eta, \alpha)$ without incurring bias results from our assuming that $H(\psi^\dagger)$ and A are (mean) independent given X and will not happen in most models. For example, suppose we had defined $\psi^\dagger = E\{Var[A|X]\}^{-1} E\{Cov[Y, A|X]\}$ to be the unique solution to

$$E[\{Y - A\psi - E[Y - A\psi|X]\}\{A - E[A|X]\}] = 0$$

in the nonparametric model that does not assume $E\left[Y - A\psi^\dagger|A, X\right] = E\left[Y - A\psi^\dagger|X\right]$. In this nonparametric model, $ES_1(\psi, \eta, \alpha)$ is still the first order efficient score for ψ and $ES_2^{test}(\psi, \widehat{\omega}_-; c)$ is still an unbiased estimating function. However, if $E\left[Y - A\psi^\dagger|A, X\right] = E\left[Y - A\psi^\dagger|X\right]$ is false, then, owing to the term $\sum_i I(X_i = x) \overline{H}_{x,-i}(\psi)\left(A_i - \overline{A}_{x,-i}\right)$, $ES_1(\psi, \widetilde{\eta}, \widetilde{\alpha})$ will not be unbiased for 0 and will have second order bias if the number of levels of X exceed $n^{1/2}$. An analogous remark applies to the continuous covariate case discussed below.

The ability to obtain an unbiased estimator in our semiparametric regression model by estimating the nuisance parameters in ES_1 using "own"sample

estimates is not only unusual but obscures the fundamental connection of the inference problem to the need for orthogonality with higher order nuisance scores. By using "independent" sample estimates of nuisance parameters and considering inference conditional on these, the underlying general structure of the inference problem is revealed.

Continuous Covariates: Consider the analysis of the semiparametric regression model

$$Y = \psi^{\dagger} A + b^{\dagger}(X) + e$$

with $e \tilde{} N(0,1)$ having a known distribution, X absolutely continuous wrt Lesbegue measure on $[0,1]$, and A dichotomous based on n iid observations $O_i = (Y_i, A_i, L_i)$. Let $F_X\left(X;\omega^{\dagger}\right)$ be the marginal of X

Consider the model with likelihood with respect to a dominating measure

$$\begin{aligned} & f(O;\theta = (\psi,\eta,\alpha,\omega)) \\ & = \phi_{std-n}\left\{e(\psi,\eta,\omega)\right\} p(X;\alpha,\omega)^{A} \left\{1 - p(X;\alpha,\omega)\right\}^{1-A} f(X;\omega) \\ & = \theta \in \Theta = \Psi \times \mathcal{N} \times \mathring{A} \times \Omega \end{aligned}$$

$$\begin{aligned} & with\ e(\psi,\eta,\omega) = Y - \psi A - b(X,\eta,\omega), \\ & p(X;\alpha,\omega) = E[A|X;\alpha,\omega] \end{aligned}$$

where we use the model

$$b(x,\eta,\omega) = b^{*}(x,\eta)/f(x;\omega), b^{*}(x,\eta) \in \left\{ \sum_{r=1}^{k_\eta} \eta_r \varphi_r(x) \right\}, \quad (9.14a)$$

$$p(x;\alpha,\omega) = p^{*}(x;\alpha)/f(x;\omega), p^{*}(x;\alpha) \in \left\{ \sum_{r=1}^{k_\alpha} \alpha_r \varphi_r(x) \right\}, \quad (9.14b)$$

$$\int \varphi_r(x)\varphi_j(x)dx = I(r=j), \ \{\varphi_r(x), \text{ a complete orthonormal basis for } L_2(\mu)\}$$

We assume k_η and k_α are known functions of n that may be infinite for all n. When k_η and k_α are infinite we consider 2 kinds of models

$$\text{mod}\, el\ \ (i): \sum_{r=1}^{\infty} \eta_r^2 < C_\eta, \sum_{r=1}^{\infty} \alpha_r^2 < C_\alpha \quad (9.14c)$$

$$\text{mod}\, el\ (ii):\ p = 1, \sum_{r=1}^{\infty} \eta_r^2 r^{2\beta_\eta} < C_\eta, \sum_{r=1}^{\infty} \alpha_r^2 r^{2\beta_\alpha} < C_\alpha \quad (9.14d)$$

If we let $k_\eta = k_\alpha = \infty$, and chose model (i) so the only restriction is that $h^{*}\left(x;\eta^{\dagger}\right)$ and $p^{*}(x;\alpha)$ are in L_2 balls, $\psi^{\dagger}$ is not a kth order pathwise differentiable parameter except for the case $k = 1$ and this case is misleading

in the sense that the first order asymptotic bound (the Cramer Rao bound) is not attainable. In fact no uniformly consistent estimators of $\psi^\dagger$ exist (Ritov and Bickel,1990). Thus we use model (ii) for some apriori choice of (β_η, C_η) and (β_α, C_α). In model (ii), $\psi^\dagger$ still does not have a kth order influence function with finite variance for $k > 1$. But we can make progress by considering a sequence of (misspecified) working models (sieves) changing with n with finite $k_\eta(n)$ and $k_\alpha(n)$ in such a way that the bias due to setting coefficients past $k_\eta(n)$ and $k_\alpha(n)$ to 0 is properly controlled. Thus we first study models with a finite $k_\eta(n)$ and $k_\alpha(n)$. These models have kth order influence functions with finite variance for $k > 1$.

Here we only give detailed results for the case where the marginal of X (i.e., $\omega^\dagger$ is known). leaving the general case to Robins and van der Vaart (2004). Robins and van der Vaart show that $U_1^{test}(\theta) = \sum_i e_i \Delta_i$ is a 1st order testing influence function and that it equals ES_1^{test} when $k_\alpha \leq k_\eta$. Further with $k^* = \min(k_\alpha, k_\eta)$,

$$U_2^{test}(\psi, \eta, \alpha, \omega) = \sum_i \Delta_i(\alpha, \omega) e_i(\psi, \eta, \omega) f(X_i; \omega)(n-1) - \sum_{l=1}^{k^*} \sum \sum_{i \neq j} e_i(\psi, \eta, \omega) \varphi_l(X_i) \Delta_j(\alpha, \omega) \varphi_l(X_j)$$

is a 2nd order testing influence function under law $(\psi, \eta, \alpha, \omega)$ with ω known. Further $U_2^{test}(\psi, \eta, \alpha, \omega)$ is strongly doubly robust in the sense that

$$E_{(\psi^\dagger, \alpha^\dagger, \eta^\dagger, \omega^\dagger)}\left[U_2^{test}\left(\psi^\dagger, \alpha, \eta, \omega^\dagger\right)\right] = 0, \tag{9.15}$$

but only for $p^*(x; \alpha)$ and $h^*(x; \eta)$ in model $(9.14a - 9.14b)$. That is, $U_2^{test}\left(\psi^\dagger, \alpha, \eta, \omega^\dagger\right)$ has mean zero even if both α and η are incorrect.

Sketch of proof: Note that, with $\omega^\dagger$ known, double robustness in the sense of (9.15) implies U_2^{test} is a 2nd order testing influence function as it implies orthogonality to $\{V_{\alpha\eta}\}, \{V_\alpha\}$, and $\{V_\eta\}$. Now suppressing dependence on the true parameter values and setting $H = Y - \psi^\dagger A$, $E\left[U_2^{test}\left(\psi^\dagger, \alpha, \eta, \omega^\dagger\right)\right] =$

$$nE\left[(n-1)\left\{H - b\left(X, \eta, \omega^\dagger\right)\right\} f\left(X; \omega^\dagger\right)\left\{A - p\left(X, \alpha, \omega^\dagger\right)\right\}\right] -$$
$$n(n-1) \sum_{l=1}^{k^*} E\left[\left\{H - b\left(X, \eta, \omega^\dagger\right)\right\} \varphi_l(X)\right] E\left[\left\{A - p\left(X, \alpha, \omega^\dagger\right)\right\} \varphi_l(X)\right]$$

Now

$$
\begin{aligned}
&E\left[\left\{H-b\left(X,\eta,\omega^{\dagger}\right)\right\}f\left(X;\omega^{\dagger}\right)\left\{A-p\left(X,\alpha,\omega^{\dagger}\right)\right\}\right]\\
&=E\left[\left\{b\left(X,\eta^{\dagger},\omega^{\dagger}\right)-b\left(X,\eta,\omega^{\dagger}\right)\right\}f\left(X;\omega^{\dagger}\right)\left\{p\left(X,\alpha^{\dagger},\omega^{\dagger}\right)-p\left(X,\alpha,\omega^{\dagger}\right)\right\}\right]\\
&=E\left[\begin{array}{l}\left\{\sum_{r=1}^{\infty}\left(\eta_r^{\dagger}-\eta_r I\left(r<k_{\eta}\right)\right)\varphi_r\left(X\right)/f\left(X;\omega^{\dagger}\right)\right\}f\left(X;\omega^{\dagger}\right)\times\\ \left\{\sum_{r=1}^{\infty}I\left(r<k_{\alpha}\right)\left(\alpha_r^{\dagger}-\alpha_r I\left(r<k_{\alpha}\right)\right)\varphi_r\left(X\right)/f\left(X;\omega^{\dagger}\right)\right\}\end{array}\right]\\
&=\sum_{r=1}^{\infty}\left(\eta_r^{\dagger}-\eta_r I\left(r<k_{\eta}\right)\right)\left(\alpha_r^{\dagger}-\alpha_r I\left(r<k_{\alpha}\right)\right)=\\
&\sum_{r=1}^{\min(k_{\eta},k_{\alpha})}\left(\eta_r^{\dagger}-\eta_r\right)\left(\alpha_r^{\dagger}-\alpha_r\right)+\sum_{\min(k_{\eta},k_{\alpha})+1}^{\infty}\eta_r^{\dagger}\alpha_r^{\dagger}-\qquad(9.16)\\
&\sum_{\min(k_{\eta},k_{\alpha})+1}^{k_{\alpha}}\eta_r^{\dagger}\alpha_r-\sum_{\min(k_{\eta},k_{\alpha})+1}^{k_{\eta}}\eta_r\alpha_r^{\dagger}
\end{aligned}
$$

When $\eta_r^{\dagger}=\eta_r=0$ for $r>k_{\eta}$ and $\alpha_r^{\dagger}=\alpha_r=0$ for $r>k_{\alpha}$, (9.16) equals $\sum_{r=1}^{\min(k_{\eta},k_{\alpha})}\left(\eta_r^{\dagger}-\eta_r\right)\left(\alpha_r^{\dagger}-\alpha_r\right)$. The following calculation completes the proof.

$$
\begin{aligned}
&\sum_{l=1}^{k^*}E\left[\left\{H-b\left(X,\eta,\omega^{\dagger}\right)\right\}\varphi_l\left(X\right)\right]E\left[\left\{A-p\left(X,\alpha,\omega^{\dagger}\right)\right\}\varphi_l\left(X\right)\right]\\
&=\sum_{l=1}^{k^*}E\left[\left\{\sum_{r=1}^{\infty}\left(\eta_r^{\dagger}-\eta_r I\left(r<k_{\eta}\right)\right)\varphi_r\left(X\right)/f\left(X;\omega^{\dagger}\right)\right\}\varphi_l\left(X\right)\right]\times\\
&E\left[\left\{\sum_{r=1}^{\infty}\left(\alpha_r^{\dagger}-\alpha_r I\left(r<k_{\alpha}\right)\right)\varphi_r\left(X\right)/f\left(X;\omega^{\dagger}\right)\right\}\varphi_l\left(X\right)\right]\\
&=\sum_{l=1}^{k^*}\left(\eta_l^{\dagger}-\eta_l I\left(l<k_{\eta}\right)\right)\left(\alpha_l^{\dagger}-\alpha_l I\left(l<k_{\alpha}\right)\right)\qquad(9.17)
\end{aligned}
$$

When $\eta_r^{\dagger}=\eta_r=0$ for $r>k_{\eta}$ and $\alpha_r^{\dagger}=\alpha_r=0$ for $r>k_{\alpha}$, (9.17) equals $\sum_{r=1}^{\min(k_{\eta},k_{\alpha})}\left(\eta_r^{\dagger}-\eta_r\right)\left(\alpha_r^{\dagger}-\alpha_r\right)$ for $k^*=\min\left(k_{\alpha},k_{\eta}\right)$.

Thus since the variance of the estimator solving $0=U_2^{test}\left(\psi^{\dagger},\alpha,\eta,\omega^{\dagger}\right)$ is $O\left(n^{-1}+k^*/n^2\right)$ it has rate of convergence $n^{1/2}$ if $\min\left(k_{\alpha},k_{\eta}\right)$ is $O\left(n\right)$ and convergence rate $n^{1-\rho/2}$ if $\min\left(k_{\alpha},k_{\eta}\right)=n^{\rho},1<\rho<2$.

We now turn our attention to model (9.14d) which allows $\eta_r^{\dagger}$ and $\alpha_r^{\dagger}$ to exceed 0 for all r. Suppose $\beta_{\eta}\leq\beta_{\alpha}$. We shall still choose a "working " model of the form $(9.14a-9.14b)$ in which $\eta_r=0$ and $\alpha_r=0$ for $r>k^*=k_{\eta}=k_{\alpha}$. Then (9.17) equals $\sum_{r=1}^{k^*}\left(\eta_r^{\dagger}-\eta_r\right)\left(\alpha_r^{\dagger}-\alpha_r\right)$ and (9.16) equals $\sum_{r=1}^{\infty}\left(\eta_r^{\dagger}-\eta_r I\left(r<k_{\eta}\right)\right)\left(\alpha_r^{\dagger}-\alpha_r I\left(r<k_{\alpha}\right)\right)$ so the squared bias is $\left(\sum_{k^*+1}^{\infty}\eta_r^{\dagger}\alpha_r^{\dagger}\right)^2$ which is less than or equal to $\sum_{k^*+1}^{\infty}\left(\eta_r^{\dagger}\right)^2\sum_{k^*+1}^{\infty}\left(\alpha_r^{\dagger}\right)^2<\left(k^*\right)^{-2(\beta_{\eta}+\beta_{\alpha})}$. Thus the maximum squared bias of $U_2^{test}\left(\psi^{\dagger},\alpha,\eta,\omega^{\dagger}\right)$ is $\left\{n\left(n-1\right)\right\}^2\left(k^*\right)^{-2(\beta_{\eta}+\beta_{\alpha})}$. Since the variance of $U_2^{test}\left(\psi^{\dagger},\alpha,\eta,\omega^{\dagger}\right)$ is $O\left(n+k^*\right)$,

the maximum bias squared and variance are balanced for $k^*_{balance} = O\left[n^{1/\left(\frac{1}{2}+\beta_\eta+\beta_\alpha\right)\cdot}\right]$. To construct confidence intervals we would choose k^* slightly larger. The order of the length of those confidence intervals would just exceed $O\left[\left\{\kappa^{-2} var\left(U_2^{test}\left(\psi^\dagger,\alpha,\eta,\omega^\dagger\right)\right)\right\}^{1/2}\right] = O\left[\min\left(n^{-1/2},\left\{k^*_{balance}/n^2\right\}^{1/2}\right)\right] = O\left[\min\left(n^{-1/2},n^{-2(\beta_\eta+\beta_\alpha)/1+2(\beta_\eta+\beta_\alpha)}\right)\right]$ where $\kappa =$

$$
\begin{aligned}
&E\left[\partial U_2^{test}\left(\psi^\dagger,\alpha,\eta,\omega^\dagger\right)/\partial\psi\right] \\
&= E\left[U_2^{test}\left(\psi^\dagger,\alpha,\eta,\omega^\dagger\right) ES_1\right] \\
&= E\left[U_2^{test}\left(\psi^\dagger,\alpha,\eta,\omega^\dagger\right) ES_2^{test}\right]
\end{aligned}
$$

Remark: If we do not wish to assume β_η and β_α are known apriori we can use the data itself to choose these parameters and thereby obtain adaptive estimators of $\psi^\dagger$. See Robins and van der Vaart (2004). However, although one can obtain adaptive point estimators, it is not possible to obtain adaptive confidence intervals for $\psi^\dagger$.

The theory of higher order influence functions and their associated U-statistics can, in certain cases, be used to improve upon the cross-validation results of van der Laan and Dudoit (2003) by replacing their loss functions, which are first order U-statistics, by loss functions that are higher order U-statistics, thereby decreasing sensitivity to the bias in estimators of the unknown parameters in the loss function. This will be true even when van der Laan and Dudoit's loss functions are doubly robust. Indeed van der Laan and Dudoit's results for first order U-statistic loss functions can be fairly straightforwardly generalized to higher order U-statistics by replacing the maximal inequalities used in their proofs by maximal inequalities for higher order U-statistics.

In addition our theory of inference based on higher order U statistics can be applied to data generated under exceptional laws at which the parameter $\psi(\theta)$ is not differentiable by using a higher order generalization of Theorem 5.1.

Unknown marginal for X : When the marginal of X is unknown, Robins and van der Vaart (2004) show that when $\beta_\eta + \beta_\alpha > 1/2$, root-n estimation of $\psi^\dagger$ is possible with no assumptions on the rate at which rate the density $f(x)$ is estimable (a result previously obtained by Donald and Newey (1994)). However if $\beta_\eta + \beta_\alpha < 1/2$, the optimal rate of estimation of $\psi^\dagger$ depends not only on $\beta_\eta + \beta_\alpha$, but also on the rate at which the density $f(x)$ can be estimated. Thus for valid confidence interval construction, we require that, in addition to β_η and β_α, we be given β_ω specifying the maximum apriori complexity of $f(x)$.

Confidence Intervals After Selection:

Intervals based on Negative Utility as A Loss Function and $\widehat{d^{j}_{op}}^{util}(l)$ **as Input :** Suppose again that A is binary, sequential randomization holds,

and $p(a|l)$ is known. Recall that of the J candidate regimes $\widehat{\gamma}^j$, $\widehat{\widehat{\gamma}^j}_{util}$ maximizes $P_{n^{val}}^{val}\left[YI\left\{A=\widehat{d_{op}^{\gamma^j}}(L)\right\}/p(A|L)\right]$ and will have expected utility $E\left[Y_{\widehat{d_{op}^j}_{util}}\right]$ close to expected utility $E\left[Y_{d_{op}^{j_{util-orac}}}\right]$ of the utility oracle, provided J is not too large. Suppose rather than using $\widehat{d_{op}^j}_{util}(l)$ to determine our treatment choice, we wish to use $\widehat{d_{op}^j}_{util}(l)$ to 'center' in some sense a $(1-\alpha)$ confidence set for the unknown $d_{op}=d_{op}^{\gamma}=\arg\max_a \gamma(l,a)$, under the theory that (i) one should report confidence intervals to show the regimes compatible with the data and (ii) one's 'best' estimate should be used to center a confidence interval.

We shall see go in this direction and shall fail. But the failure will point us in a better direction. We assume that we have available a third group of subjects the confidence group, for which we assume no unmeasured confounders. We will construct confidence intervals for this group's optimal regime $d_{op}^{\gamma}(l)$ which will be a confidence interval for the estimation and validation sample's $d_{op}^{\gamma}(l)$ as well if $\gamma(l,a)$ for the confidence group equals the common $\gamma(l,a)$ in the other groups. [Most often the confidence group will have arisen by a random division of the original study into the three groups - estimation, validation, and confidence.]

To begin we shall be satisfied with constructing a confidence interval for $d_{op}^{\gamma}(l)$ only at the n_{conf} values of l in the set $OLC=\{l_i; i\in \text{confidence group}\}$ corresponding to the observed values of l in the confidence group, where we have assumed no ties. [Thus, for the present, we view L as having support OLC and a regime $d(l)$ is a n_{conf} vector of $0's$ and $1's$ corresponding to whether on not to treat at each $l\in OLC$.] To do so we will use a new construction due to Robins and van der Vaart (2003) to form conservative uniform $(1-\alpha)$ large sample confidence region for $d_{op}^{\gamma}(l)$ for $l\in OLC$. In the following expectations and probabilities are wrt the distribtion of the confidence population. First note that since $E\left[YI\left\{A=\widehat{d_{op}^j}_{util}(L)\right\}/p(A|L)|L\right]=\gamma(L)\widehat{d_{op}^j}_{util}(L)+E[Y|L,A=0]$ where $\gamma(L)=\gamma(L,1)$, we have that

$$n_{conf}^{1/2}P_{n_{conf}}^{conf}\left[Y\frac{\left[I\{A=0\}-I\left\{A=\widehat{d_{op}^j}_{util}(L)\right\}\right]}{p(A|L)}+\gamma(L)\widehat{d_{op}^j}_{util}(L)\}\right]$$
$$=n_{conf}^{1/2}P_{n_{conf}}^{conf}\left[Y\frac{\{1-2A\}\widehat{d_{op}^j}_{util}(L)}{p(A|L)}+\gamma(L)\widehat{d_{op}^j}_{util}(L)\right]$$

is conditionally asymptotically normal with mean 0 conditional on the estimation, and validation sample and the the set OLC of confidence sample $L's$ (so the support of L is OLC.) The conditional variance $\tau^2=P_{n_{conf}}^{conf}\left\{Var\left[Y\frac{\{1-2A\}\widehat{d_{op}^j}_{util}(L)}{p(A|L)}|L\right]\right\}$ is $O_p(1)$. Thus

$$C_\gamma(1-\alpha) = \begin{array}{c} \left\{\gamma^*(L); P^{conf}_{n_{conf}}\left[-\gamma^*(L)\, d^{\widehat{j}_{util}}_{op}(L)\right]\right. \\ \left. < n^{-1/2}\tau z_\alpha + P^{conf}_{n_{conf}}\left[Y\frac{\{1-2A\}d^{\widehat{j}_{util}}_{op}(L)}{p(A|L)}\right]\right\} \end{array}$$

is a uniform $1-\alpha$ asymptotic confidence for the true $\gamma(L)$, where z_α is the upper α quantile of a standard normal so $z_{.05} = 1.64$. Thus

$$C_{d_{op}}(1-\alpha) = \{d^*(L) = I\{\gamma^*(L) > 0\}; \gamma^*(L) \in C_\gamma(1-\alpha)\}$$

is a uniform $1-\alpha$ asymptotic confidence for the true $d_{op}(L) = I\{\gamma(L) > 0\}$.

We shall see this confidence set has terrible power properties, failing to exclude $d^*(L)$ that are far from d_{op}. To understand what the confidence set $C_{d_{op}}(1-\alpha)$ looks like note that it follows from above that $P^{conf}_{n_{conf}}\left[Y\frac{\{1-2A\}d^{\widehat{j}_{util}}_{op}(L)}{p^*(A|L)}\right] = P^{conf}_{n_{conf}}\left[-\gamma(L)\, d^{\widehat{\gamma}^{\widehat{j}}_{util}}_{op}(L)\right] + \tau n^{-1/2}Z + o_p\left(n^{-1/2}\right)$ where Z is independent standard normal.

Thus

$$C_\gamma(1-\alpha)$$
$$= \left\{\gamma^*(L); P^{conf}_{n_{conf}}\left[(\gamma(L) - \gamma^*(L))\, d^{\widehat{j}_{util}}_{op}(L)\right\}\right] < n^{-1/2}\tau(z_\alpha + Z) + o_p\left(n^{-1/2}\right)\Big\}$$

Now consider the extreme "perfect " case in which $d^{\widehat{j}_{util}}_{op}(L) = d_{op}(L)$. Suppose the subset $POS = \{l_1, ..., l_{100}\}$ of OLC on which $\gamma(L) > 0$ contains 100 points and at each such point $\gamma(L) = 100$. Consider now $\gamma^*(L)$ such that $\gamma^*(l_1) = 10^5$ and $\gamma^*(l_k) = -101$ for the remaining 99 l_k in POS. Then $P^{conf}_{n_{conf}}\left[(\gamma(L) - \gamma^*(L))\, d^{\widehat{j}_{util}}_{op}(L)\right\}\right]$ will be less than -7×10^4 so $\gamma^*(L)$ will be in $C_\gamma(1-\alpha)$ and thus $d^*(L) = I(\gamma^*(L) > 0)$ in $C_{d_{op}}(1-\alpha)$ even though $d^*(L)$ incorrectly withholds treatment at all but one l_k in POS. Part of the problem is the input $d^{\widehat{j}_{util}}_{op}(L)$, even when optimal, does not contain enough information to sufficiently constrain the possible $\gamma^*(L)$ that could have generated the data. That is although $d_{op}(L) = I(\gamma(L) > 0)$ solves (i) $\text{argmax}_{c(L)} E[I(c(L) > 0)\gamma(L)]$, $\gamma(L)$ does not solve (ii) $\text{argmax}_{c(L)} E[I(\gamma(L) > 0)c(L)]$, and indeed a $c(L)$ that, with high probability, differs in sign from $\gamma(L)$ can have $E[I(\gamma(L) > 0)c(L)] >> E[I(\gamma(L) > 0)\gamma(L)]$. The good properties of $\widehat{\gamma}^{\widehat{j}}_{util}$ derives from the fact that it approximately solves (i). The bad properties of the above confidence interval derives from the fact that it is based on 'approximately solving' (ii).

Intervals based on $Loss(c, F_{A|L}, b, g)$ as A Loss Function and an Estimate of γ as Input :

The Fixed L Case: One lesson of the last subsection is that if we wish our confidence regions to be sufficiently small that they have the power to exclude implausible candidates $d^*(l)$ for $d_{op}(l)$, we at least need a confidence procedure that uses as input the full information contained in a cross-validated

selected approximation, such as $\widehat{\widehat{\gamma^j}}_{util}(l)$, to $\gamma(l)$ rather than the lesser information contained in, say, $\widehat{d^j_{op}}^{util}(l) = I\left[\widehat{\widehat{\gamma^j}}_{util}(l) > 0\right]$. To do so we consider intervals based based on

$$Loss\left(\widehat{\gamma}, F_{A|L}, b, g\right) = \left[g^2(L)\left\{Y - \widehat{\gamma}(L,A) - b(L) + E\left[\widehat{\gamma}(L,A)|L\right]\right\}^2\right]$$

where $\widehat{\gamma}$ is some cross-validated selected estimate such as $\widehat{\widehat{\gamma^j}}_{util}(l)$ or $\widehat{\widehat{\gamma^j}}$ that we hope might be a good approximation to the true γ. We will use the construction due to Robins and van der Vaart (2003) to form conservative uniform $(1-\alpha)$ large sample confidence region for γ and thus $d^{\gamma}_{op}(l)$ for $l \in OLC$. We will see that the size is adaptive in the sense that the set of regimes contained in the confidence set for d^{γ}_{op}will be few if $\widehat{\gamma}$ is a good estimate of γ but will be many if the estimate $\widehat{\gamma}$ is poor. Recall that

$$\begin{aligned} R &= g(L)\left[\{Y - \gamma(L,A)\} - b(L) + E\left[\gamma(L,A)|L\right]\right] \\ &= g(L)\left[\{Y - \gamma(L,A)\} - E\left[\{Y - \gamma(L,A)\}|L\right]\right] + g(L)\left\{E\left[Y|L\right] - b(L)\right\} \end{aligned}$$

and

$$\begin{aligned} S(c) &= g(L)\left[\{\gamma(L,A) - c(L,A)\} - E\left[\gamma(L,A) - c(L,A)|L\right]\right] \\ &= g(L)\left[\{\gamma(L) - c(L)\}\{A - E\left[A|L\right]\}\right]. \end{aligned}$$

Therefore $E\left[S^2(c)|L\right] = g^2(L)\left\{\gamma(L) - c(L)\right\}^2 var\left[A|L\right]$ and

$$E\left[R^2|L\right] = g^2(L)\left[\left\{var\left[Y - \gamma(L,A)|L\right]\right\} + \left\{E\left[Y|L\right] - b(L)\right\}^2\right].$$

We know from our prior results that

$$n^{1/2}_{conf} P^{conf}_{n_{conf}}\left[Loss\left(\widehat{\gamma}, F_{A|L}, b, g\right) - \left\{E\left[R^2|L\right] + E\left[S^2(\widehat{\gamma})|L\right]\right\}\right]$$

is conditionally asymptotically normal with mean 0 conditional on the estimation, and validation sample and the the set OLC of confidence sample $L's$ (so the support of L is OLC) because $Loss\left(\widehat{\gamma}, F_{A|L}, b, g\right)$ has conditional mean $E\left[R^2|L\right] + E\left[S^2(c)|L\right]$. The conditional variance $\tau^2 = \tau^2(\gamma^*) = P^{conf}_{n_{conf}}\left\{Var\left[Loss\left(\widehat{\gamma}, F_{A|L}, b, g\right)|L\right]\right\}$ is $O_p(1)$. Thus

$$\begin{aligned} C_{\gamma}(1-\alpha) = \Big\{&\gamma^*(L); P^{conf}_{n_{conf}}\left[g^2(L)\left[\left\{\widehat{\gamma}(L) - \gamma^*(L)\right\}^2 var\left[A|L\right]\right]\right] \\ &< n^{-1/2}\tau z_{\alpha} + P^{conf}_{n_{conf}}\left[Loss\left(\widehat{\gamma}, F_{A|L}, b, g\right) - E\left[R^2|L\right]\right]\Big\} \end{aligned} \tag{9.18}$$

is a uniform $1-\alpha$ asymptotic confidence for the true $\gamma(L)$, where z_{α} is the upper α quantile of a standard normal so $z_{.05} = 1.64$. Thus

$$C_{d_{op}}(1-\alpha) = \left\{d^*(L) = I\left\{\gamma^*(L) > 0\right\}; \gamma^*(L) \in C_{\gamma}(1-\alpha)\right\}$$

is a uniform $1-\alpha$ asymptotic confidence for the true $d_{op}(L) = I\{\gamma(L) > 0\}$.

To understand what the confidence set $C_\gamma(1-\alpha)$ looks like note that it follows from the above normal distributional result that
$P^{conf}_{n_{conf}}\left[Loss\left(\widehat{\gamma}, F_{A|L}, b, g\right) - E\left[R^2|L\right]\right]$
$= P^{conf}_{n_{conf}}\left[g^2(L)\left[\{\widehat{\gamma}(L) - \gamma(L)\}^2 var[A|L]\right]\right] + \tau n^{-1/2} Z + o_p\left(n^{-1/2}\right)$
where Z is independent standard normal.

Thus our interval (9.18) has the asymptotic expansion

$$C_\gamma(1-\alpha) \tag{9.19a}$$
$$= \left\{ \begin{array}{c} \gamma^*(L); P^{conf}_{n_{conf}}\left\{g^2(L)\left[\{\widehat{\gamma}(L) - \gamma^*(L)\}^2 var[A|L]\right]\right\} < n^{-1/2}\tau(z_\alpha + Z) \\ + P^{conf}_{n_{conf}}\left\{g^2(L)\left[\{\widehat{\gamma}(L) - \gamma(L)\}^2 var[A|L]\right]\right\} \end{array} \right\}$$
$$= \left\{ \begin{array}{c} \gamma^*(L); P^{conf}_{n_{conf}}\left[var[A|L]\, g^2(L)\{\gamma^*(L) - \gamma(L)\}^2 + 2\{\gamma(L) - \widehat{\gamma}(L)\}\right] \\ < n^{-1/2}\tau(z_\alpha + Z) + o_p\left(n^{-1/2}\right) \end{array} \right\} \tag{9.19b}$$

¿From (9.19a) we see that if we choose $g^2(L) = \{var[A|L]\}^{-1}$, then

$$C_\gamma(1-\alpha) = \left\{ \begin{array}{c} \gamma^*(L); P^{conf}_{n_{conf}}\left[\{\widehat{\gamma}(L) - \gamma^*(L)\}^2\right] < n^{-1/2}\tau(z_\alpha + Z) + \\ P^{conf}_{n_{conf}}\left[\{\widehat{\gamma}(L) - \gamma(L)\}^2\right] \end{array} \right\}.$$

If $P^{conf}_{n_{conf}}\left[\{\widehat{\gamma}(L) - \gamma(L)\}^2\right] = O_p\left(n^{-1/2}\right)$ then our confidence region $C_\gamma(1-\alpha)$ is contained in a ball centered on $\widehat{\gamma}(L)$ of radius $n^{-1/4}$ and we cannot obtain a smaller radius even if $P^{conf}_{n_{conf}}\left[\{\widehat{\gamma}(L) - \gamma(L)\}^2\right]$ is smaller than $O_p\left(n^{-1/2}\right)$.[The proof that our confidence region is contained in a ball of radius $n^{-1/4}$ uses the fact that the dependence of τ on γ^* is negligible.] Li (1988) shows that no uniform $1-\alpha$ asymptotic confidence ball for the $\gamma(L)$ can have radius less than $n^{-1/4}$. An $n^{-1/4}$ rate rather than the usual $n^{-1/2}$ rate is the price we pay for admitting that we cannot be certain that we have specified a correct finite dimensional parametric model for $\gamma(A, L) = A\gamma(L)$. The radius of our interval depends on how good an estimator $\widehat{\gamma}(L)$ is of the true $\gamma(L)$. If $O_p\left(n^{-1/2}\right) < P^{conf}_{n_{conf}}\left[\{\widehat{\gamma}(L) - \gamma(L)\}^2\right] = O_p\left(n^{-\beta}\right)$ for $0 < \beta < 1/2$, our intervals have radius $n^{-\beta/2}$, while if $o_p(1) < P^{conf}_{n_{conf}}\left[\{\widehat{\gamma}(L) - \gamma(L)\}^2\right] = O_p(1)$ so that $\widehat{\gamma}(L)$ was so poor an estimate that it was inconsistent, our intervals have radius $O(1)$.

A better choice of the user suppied function $g(L)$ can be made. Recall that we can chose $g(L)$ after knowing our candidate $\widehat{\gamma}(L)$ since $\widehat{\gamma}(L)$ is regarded as fixed in the analysis. Now to obtain a precise interval for $d_{op}(L) = I(\gamma(L) > 0)$, we only need to provide a very narrow interval for $\gamma(L)$ at values of L where it is near zero. Thus we can choose $g(L)$ large at those L

where our candidate $\widehat{\gamma}(L)$ is near zero, although we cannot go too far in this direction without inflating τ, which, by inflating the overall interval length will counteract the benefits of sharpening $g(L)$.

The above approach however is not altogether satisfactory. Not only does the serious problem of extrapolating our confidence interval for $d_{op}(L)$ to values of l not realized in the confidence group remain, but, in addition, we have swept one important difficulty under the rug. To compute the interval (9.18) requires that we know $\sigma^2(L) = var[Y - \gamma(L,A)|L]$ to be able to calculate the term $E\left(R^2 \mid L\right)$ Otherwise, we require an $n^{1/2} - consistent$ of estimator of $\sigma^2(L)$, say computed from the validation sample, for our interval to have the above stated properties (Baraud, 2000). This requirement is a prohibitive limitation since $\sigma^2(L)$ cannot be estimated without an estimate of the residual $Y - \gamma(L,A)$; but $\gamma(L,A)$ is unknown and is in fact the quantity we are constructing an interval for. In the next subsection we show how one can attack both these problems at once.

The Random L Case: We once again take L to be random with full support. Again we follow Robins and van der Vaart (2003). We assume $p(a|l)$ and $f(l)$ are known. The case where these are unknown is considered in Robins and van der Vaart (2004). We first assume we can specify a linear model $\theta^{*,T}W^* = \sum_{s=1}^{S^*}\theta_s^* W_s^*$ for $\xi(L) = \xi(L,\gamma) = g(L)\gamma(L)var[A|L]^{1/2}$ with $g(L)$ a user chosen function and with user chosen regressors $W_s^* = w_s^*(L)$ where S^* is sufficiently large that the approximation bias $inf_{\theta^*=(\theta_s^*;s=1,\ldots S^*)}E\left[\left\{\xi(L) - \sum_{s=1}^{S^*}\theta_s^* w_s^*(L)\right\}^2\right]$ will be small compared to the width of any confidence region for $\xi(L)$ and thus for $\gamma(L)$. By the population version of Gram-Schmidt orthogonalization we can replace $\gamma^*(L;\theta^*) \equiv \theta^{*,T}W^*$ by $\gamma(L;\theta) \equiv \theta^T W = \sum_{s=1}^{S}\theta_s W_s$ for $S \leq S^*$ such that $E[W_s W_p] = 0$ for $p \neq s$, $E[W_s W_s] = 1$ for $p = s$ and for each θ^* there exists a θ such that $\gamma(L;\theta) = \gamma^*(L;\theta^*)$ for all L. With $\widehat{\gamma}(L)$ again our cross-validated selected estimator of $\gamma(L)$ and $\widehat{\xi}(L) = g(L)\widehat{\gamma}(L)\{var(A|L)\}^{1/2}$, let

$\xi\left(L;\widehat{\theta}\right) = \widehat{\theta}^T W = \sum_{s=1}^{S}\widehat{\theta}_s W_s = E\left[\widehat{\xi}(L)W\right]E\left[WW^T\right]^{-1}W$ and $\xi\left(L;\theta^\dagger\right) = \theta^{\dagger T}W = \sum_{s=1}^{S}\theta_s^\dagger W_s = E[\xi(L)W]E\left[WW^T\right]^{-1}W$ be the population least squares projections onto our model. Note $\theta_s^\dagger = E[\xi(L)W_s]$, $E\left[WW^T\right]^{-1}$ is the identity, and
$\xi\left(L;\theta^\dagger\right) = \theta^{*,\dagger T}W^*$, where
$\theta^{*,\dagger} = \arg\min_{\theta^*=(\theta_s^*;s=1,\ldots S^*)}E\left[\left\{\xi(L) - \sum_{s=1}^{S^*}\theta_s^* w_s^*(L)\right\}^2\right]$. We only consider the functions
$\gamma(L;\theta) = \xi(L;\theta)/\left[g(L)\{var(A|L)\}^{1/2}\right]$ as candidates for $\gamma(L)$ based on our assumption, that S^* was chosen large enough to control approximation bias. Our approach to constructing a $(1-\alpha)$ confidence interval for $\theta^\dagger$ and thus for $\gamma\left(L;\theta^\dagger\right)$ is as follows.

We will find a conservative uniform asymptotic $(1-\alpha)$ confidence interval of the form $C_{\theta^\dagger}(1-\alpha) = \left\{\theta; \sum_{s=1}^{S}\left(\widehat{\theta}_s - \theta_s\right)^2 < Q^2\right\}$ for $\theta^\dagger$ where the random variable Q^2 is chosen as described below. By orthonormality of the W_s, $\sum_{s=1}^{S}\left(\widehat{\theta}_s - \theta_s\right)^2 = E\left[\left\{\xi(L;\theta) - \xi\left(L;\widehat{\theta}\right)\right\}^2\right] =$

$E[g^2(L)\left[\gamma(L;\theta) - \gamma\left(L;\widehat{\theta}\right)\right]^2 \{var(A|L)\}]$ so our confidence interval can also be written as the set
$C_{\gamma\left(\cdot;\theta^\dagger\right)}(1-\alpha) = \left\{\gamma(\cdot;\theta); E[g^2(L)\left[\gamma(L;\theta) - \gamma\left(L;\widehat{\theta}\right)\right]^2 \{var(A|L)\}] < Q^2\right\}$.
Finally we obtain the interval

$$C_{d_{op}}(1-\alpha) = \left\{d^*(L) = I\{\gamma(L;\theta) > 0\}; \gamma(\cdot;\theta) \in C_{\gamma\left(\cdot;\theta^\dagger\right)}(1-\alpha)\right\}$$

We next describe how we obtain Q.

Let $M = M\left(g, F_{A|L}\right) = g(L)\{Y\{A - E[A|L]\}\}\{var(A|L)\}^{-1/2}$. Note the estimator $\widetilde{\theta}_s = P^{conf}_{n_{conf}}[MW_s]$ has mean $E\left[g(L)\gamma(L)\{var(A|L)\}^{1/2}W_s\right] = \theta^\dagger_s$.

Further the estimator
$\mathcal{R}_s\left(\widehat{\theta}\right) = \sum\sum_{\{i\neq j, i,j=1,\ldots,n_{conf}\}}\sum_{s=1}^{S}\left(M_i W_{s,i} - \widehat{\theta}_s\right)\left(M_j W_{s,j} - \widehat{\theta}_s\right)$ has mean $\sum_{s=1}^{S}\left(\widehat{\theta}_s - \theta^\dagger_s\right)^2$ conditional on the estimation and validation sample data (Laurent, 1996). Robins and van der Vaart show that we can take $Q^2 = \mathcal{R}_s\left(\widehat{\theta}\right) - c\left(S/(1-\alpha)n^2_{conf}\right)^{1/2}$ for $\alpha < 1/2$ where the constant c is explicit and given in Robins and van der Vaart.

Remark: Conditional on $\widehat{\theta}$, $R_s\left(\widehat{\theta}\right) - \sum_{s=1}^{S}\left(\widehat{\theta}_s - \theta^\dagger_s\right)^2$ is the second order estimation influence function IF^{est}_2 for the functional $\sum_{s=1}^{S}\left(\widehat{\theta}_s - \theta^\dagger_s\right)^2$ in the model satisfying the sole restrictions that $f(l)$ and $f(a \mid l)$ are known (Robins and van der Vaart, 2004).

Sequential Decisions:

We now generalize some of the above results on using cross validation to select the optimal treatment regime to the sequential decision problem. We only cover areas where the natural generalization from single to sequential decisions is either non apparent or somewhat suprising.

Voting Rule Additions: A natural additional candidate, the "vote regime", exists in the sequential problem as in the single decision problem if we have available the candidate functions $\widehat{\gamma}^j\left(\overline{L}_m, \overline{A}_m\right)$ for the optimal function $\gamma\left(\overline{L}_m, \overline{A}_m\right) = \gamma^{\overline{d}_{op}, \overline{d}^*}\left(\overline{L}_m, \overline{A}_m\right)$ rather than simply the associated candidate

regimes $d_{op}^{\widehat{\gamma}^j}$. Specifically given J candidates $\widehat{\gamma}^j\left(\overline{L}_m, \overline{A}_m\right)$ the "vote regime" d_{op}^{vote} is defined recursively as follows :
$d_{op,0}^{vote}\left(l_0\right) = \arg\max_{a_0^*}\left(\sum_j I\left(\arg\max_{a_0}\widehat{\gamma}^j\left(\overline{l}_0, \overline{a}_0\right) = a_0^*\right)\right)$,
$d_{op,m}^{vote}\left(\overline{l}_m, \overline{a}_{m-1} = d_{op,m-1}^{vote}\left(\overline{l}_{m-1}\right)\right)$
$= \arg\max_{a_m^*}\left(\sum_j I\left(\arg\max_{a_m}\widehat{\gamma}^j\left(\overline{l}_m, \overline{a}_{m-1} = d_{op,m-1}^{vote}\left(\overline{l}_{m-1}\right), a_m\right) = a_m^*\right)\right)$
that selects the most recommended treatment given your value of $\overline{l}_m$ and that you have followed the vote regime through $\overline{l}_{m-1}$. Here, for simplicity, we have discounted the possibility of ties.

Unbiased Estimation of Expected Utility in Sequential Decision Problems:

Suppose that each A_m is binary but the number of time periods K is large, say 400, as would not be untypical in many epidemiologic settings. Now, under sequential randomization, for any candidate optimal regime $\overline{d}_{op}^c$ based on a candidate optimal blip function $c\left(\overline{l}_m, \overline{a}_m\right)$, $m = 0, ..., K$, an unbiased estimator of $E\left[Y_{\overline{d}_{op}^c}\right]$ with the $p\left(a_m|\overline{l}_m, \overline{a}_{m-1}\right)$ known is
$P_{n^{val}}^{val}\left[YI\left\{\overline{A}_K = \overline{d}_{op}^c\left(L_{K-1}\right)\right\}/\prod_{m=0}^K p\left(A_m|\overline{L}_m, \overline{A}_{m-1}\right)\right]$ with
$I\left\{\overline{A}_K = \overline{d}_{op}^c\left(\bar{L}_{K-1}\right)\right\}$ the indicator that a subject followed the regime $\overline{d}_{op}^c$ through the end of the study. But for almost all regimes $\overline{d}_{op}^c$, there will be at most a few subjects and often none who followed the regime through K and thus the variance of the estimator is so large as to be useless. Again we might consider replacing the deterministic regime $\overline{d}_{op}^c$ with a random regime $\overline{p}^c$ in which given $\overline{L}_m = \overline{l}_m$ $\overline{A}_{m-1} = \overline{a}_{m-1}$, we treat with a_m drawn from a conditional density p_m^c with $p_m^c\left(a_m|\overline{l}_m, \overline{a}_{m-1}\right)$ equal to $1-\sigma$ if $a_m = \overline{d}_{op}^c\left(\overline{l}_m, \overline{a}_{m-1}\right)$ and to σ otherwise for σ small and then estimate $E\left[Y_{\overline{p}^c}\right]$ with the unbiased estimator $\widehat{E}\left[Y_{\overline{p}^c}\right] = P_{n^{val}}^{val}\left[Y\prod_{m=0}^K p_m^c\left(A_m|\overline{L}_m, \overline{A}_{m-1}\right)/\prod_{m=0}^K p\left(A_m|\overline{L}_m, \overline{A}_{m-1}\right)\right]$.
We must not choose σ too small in order to prevent $\widehat{E}\left[Y_{\overline{p}^c}\right]$ from being a highly variable estimate of $E\left[Y_{\overline{p}^c}\right]$. Now consider the case where $c\left(\overline{l}_m, \overline{a}_m\right)$ is $\widehat{\gamma}^j\left(\overline{l}_m, \overline{a}_m\right) = \gamma^j\left(\overline{l}_m, \overline{a}_m, \widehat{\psi}^j\right)$ based on a fit of the model $\gamma^j\left(\overline{l}_m, \overline{a}_m, \psi^j\right)$ to the estimation sample data so $\overline{d}_{op}^{\widehat{\gamma}^j}\left(\overline{l}_m, \overline{a}_{m-1}\right) = \arg\max_{a_m}\gamma^j\left(\overline{l}_m, \overline{a}_m, \widehat{\psi}^j\right)$. It seems wise to choose $\sigma = \sigma^j\left(\overline{l}_m, \overline{a}_{m-1}\right)$ as a function of $\left(\overline{l}_m, \overline{a}_{m-1}\right)$ and j. Specifically if the lower confidence limit for the absolute value $|\widehat{\gamma}^j\left(\overline{l}_m, \overline{a}_{m-1}, a_m = 1\right)|$ is far from 0 we could choose $\sigma^j\left(\overline{l}_m, \overline{a}_{m-1}\right)$ to be 0 (or very nearly 0) and (ii) if it includes 0 we choose $\sigma^j\left(\overline{l}_m, \overline{a}_{m-1}\right)$ to be large (but less than 1/2). We then treat a new subject beginning at time of HIV infection t_0 with the regime $\overline{p}^{\widehat{\gamma}^{\hat{j}_0}}$ that maximizes $\widehat{E}\left[Y_{\overline{p}^{\widehat{\gamma}^j}}\right]$ over the J candidates.

However for a subject who receives care in the community until the time t_k from infection with HIV and who presents to our optimal treatment clinic at that time with history $\left(\overline{L}_k, \overline{A}_{k-1}\right)$, we would treat with the regime $\underline{p}_k^{\widehat{\gamma}^{\hat{j}_k}}$

that maximizes the unbiased estimator
$P_{n^{val}}^{val}\left[Y\prod_{m=k}^{K}p_m^c\left(A_m|\overline{L}_m,\overline{A}_{m-1}\right)/\prod_{m=k}^{K}p\left(A_m|\overline{L}_m,\overline{A}_{m-1}\right)\right]$ of
$E\left[Y_{\overline{A}_{k-1},\underline{p}_k^{\widehat{\gamma}^j}}\right]=E\left\{E\left[Y_{\overline{A}_{k-1},\underline{p}_k^{\widehat{\gamma}^j}}|\overline{L}_k,\overline{A}_{k-1}\right]\right\}$. The reason that the regime $\underline{p}_k^{\widehat{\gamma}^{\widehat{j}_k}}$ can differ from the regime $\underline{p}_k^{\widehat{\gamma}^{\widehat{j}_0}}$ is that (i) the observed (i.e. community) data distribution of $\left(\overline{L}_k,\overline{A}_{k-1}\right)$ differs from the distribution of the counterfactual variables $\left(\overline{L}_{\overline{p}^{\widehat{\gamma}^{\widehat{j}_0}},k},\overline{A}_{\overline{p}^{\widehat{\gamma}^{\widehat{j}_0}},k-1}\right)$ that would obtain upon following the regime $\overline{p}^{\widehat{\gamma}^{\widehat{j}_0}}$ from time t_0 and (ii) because none of the candidate regimes $\overline{p}^{\widehat{\gamma}^j}$ are considered optimal for each value of $\left(\overline{L}_k,\overline{A}_{k-1}\right)$, the best choice among them from time k onwards will appropriately depend on the distribution of $\left(\overline{L}_k,\overline{A}_{k-1}\right)$ that will exist at time k. However there is a philosophical conundrum associated with this argument. Consider a subject i who receives care in the community until the time t_k and then presents to our optimal treatment clinic with history $\left(\overline{l}_{k,i},\overline{a}_{k-1,i}\right)$ that is strongly compatible with following regime $\overline{p}^{\widehat{\gamma}^{\widehat{j}_0}}$ in the sense $a_{m,i}=\overline{d}_{op,m}^{\widehat{\gamma}^{\widehat{j}_0}}\left(\overline{l}_{m,i},\overline{a}_{m-1,i}\right)$ for $m=0,...,k-1$. That is the patient would have the same data $\left(\overline{l}_{k,i},\overline{a}_{k-1,i}\right)$ had he been treated deterministically with $\overline{d}_{op}^{\widehat{\gamma}^{\widehat{j}_0}}$ beginning at time t_0. In that case why should the subject be viewed as a random member of a population that has the observed data distribution of $\left(\overline{L}_k,\overline{A}_{k-1}\right)$ rather than as a random member of a population that has $\left(\overline{L}_k,\overline{A}_{k-1}\right)$ distributed as $\left(\overline{L}_{\overline{p}^{\widehat{\gamma}^{\widehat{j}_0}},k},\overline{A}_{\overline{p}^{\widehat{\gamma}^{\widehat{j}_0}},k-1}\right)$. This conundrum is very analogous to the conundrum of the individual patient treatment decison discussed above and is a well recognized difficulty that arises in many guises in any frequentist theory of inference.

A Problem in the Generalization of Corollary 9.1: We first provide a partial generalization of Corollary (9.1) to the setting of optimal regime estimation in a sequential decision problem under the assumption of no unmeasured confounders. We then discuss problems with the partial generalization.

We first define the quantity $H_{\text{mod},m}^{\underline{d}_{op,m},\underline{d}_m^*}\left(c_m\right)$ in analogy to Eq. (3.19) except we substitute $c_m\left(\overline{L}_m,\overline{A}_m\right)$ for $\gamma^{\overline{d}_{op},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m,\psi\right)$ and do not specify a model for $\gamma^{\overline{d}_{op},\overline{d}^*}\left(\overline{L}_j,\overline{A}_j\right)$:

$$H_{\text{mod},m}^{\underline{d}_{op,m},\underline{d}_m^*}\left(c_m\right)=Y-c_m\left(\overline{L}_m,\overline{A}_m\right)- \tag{9.20}$$
$$\sum_{j=m+1}^{K}\left\{\gamma^{\overline{d}_{op},\overline{d}^*}\left(\overline{L}_j,\overline{A}_{j-1},d_{op,j}\left(\overline{L}_j,\overline{A}_{j-1}\right)\right)-\gamma^{\overline{d}_{op},\overline{d}^*}\left(\overline{L}_j,\overline{A}_j\right)\right\},$$
$$\text{where } d_{op,m}\left(\overline{L}_m,\overline{A}_{m-1}\right)=\arg\max_{a_m\in\mathcal{A}_m}\gamma^{\overline{d}_{op},\overline{d}^*}\left(\overline{L}_m,\overline{A}_{m-1},a_m\right).$$

Let $H_{\text{mod},m}^{\underline{d}_{op,m},\underline{d}_m^*}\left(\gamma_m^{\overline{d}_{op},\overline{d}^*}\right)$ be $H_{\text{mod},m}^{\underline{d}_{op,m},\underline{d}_m^*}\left(c_m\right)$ with $\gamma^{\overline{d}_{op},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$ replacing $c_m\left(\overline{L}_m,\overline{A}_m\right)$

Corollary 9.2: Assume (2.5) holds. Then recursively for $m = K, K-1, ..., 0$, we have the following: if $E^* \left[H^{\underline{d}_{op,m}, \underline{d}^*_m}_{\text{mod},m} \left(\gamma^{\overline{d}_{op}, \overline{d}^*}_m \right) | \overline{L}_m, \overline{A}_{m-1} \right]$
$= E \left[H^{\underline{d}_{op,m}, \underline{d}^*_m}_{\text{mod},m} \left(\gamma^{\overline{d}_{op}, \overline{d}^*}_m \right) | \overline{L}_m, \overline{A}_{m-1} \right]$ or
$E^* \left[c\left(\overline{L}_m, \overline{A}_m \right) | \overline{L}_m, \overline{A}_{m-1} \right] = E \left[c\left(\overline{L}_m, \overline{A}_m \right) | \overline{L}_m, \overline{A}_{m-1} \right]$ for all $c\left(\overline{L}_m, \overline{A}_m \right)$, then given any function $g\left(\overline{L}_m, \overline{A}_{m-1} \right)$ that is non-zero w.p1, $\gamma^{\overline{d}_{op}, \overline{d}^*}\left(\overline{L}_m, \overline{A}_m \right)$ is the unique function $c\left(\overline{L}_m, \overline{A}_m \right)$ minimizing

$$E \left[g^2 \left(\overline{L}_m, \overline{A}_{m-1} \right) \left\{ H^{\underline{d}_{op,m}, \underline{d}^*_m}_{\text{mod},m} \left(c_m \right) - E^* \left[H^{\underline{d}_{op,m}, \underline{d}^*_m}_{\text{mod},m} \left(c_m \right) | \overline{L}_m, \overline{A}_{m-1} \right] \right\}^2 \right] \tag{9.21}$$

subject to $c\left(\overline{L}_m, \overline{A}_m \right) = 0$ if $A_m = 0$.

The proof is completely analogous to that of Corollary 9.1 and thus is omitted. Further it holds if we replace d_{op} by any other regime d. Now in practice one can see that to use this result to find $\gamma^{\overline{d}_{op}, \overline{d}^*}\left(\overline{L}_m, \overline{A}_m \right)$ one must precede recursively and have already found $\gamma^{\overline{d}_{op}, \overline{d}^*}\left(\overline{L}_{m+1}, \overline{A}_{m+1} \right)$. One might wonder whether a simultaneous minimization might work. For example, define $H^{\overline{d}, \overline{d}^*}\left(\overline{c} \right)$ to be the vector with components

$$H^{\overline{d}, \overline{d}^*}_m \left(\overline{c}_m \right) = Y - c_m \left(\overline{L}_m, \overline{A}_m \right) - \sum_{j=m+1}^{K} \left\{ c_j \left(\overline{L}_j, \overline{A}_{j-1}, d^{c_j}_{op,j} \left(\overline{L}_j, \overline{A}_{j-1} \right) \right) - c_j \left(\overline{L}_j, \overline{A}_j \right) \right\}$$

and $V\left(\overline{c} \right)$ to be the vector with components

$$V_m \left(\overline{c}_m \right) = g \left(\overline{L}_m, \overline{A}_{m-1} \right) \left\{ H^{\overline{d}, \overline{d}^*}_m \left(\overline{c}_m \right) - E \left[H^{\overline{d}, \overline{d}^*}_m \left(\overline{c}_m \right) | \overline{L}_m, \overline{A}_{m-1} \right] \right\}$$

Note one might have hoped that it would be the case that $V\left(\overline{c} \right)^T B V\left(\overline{c} \right)$ is minimzed at $c_m \left(\overline{L}_m, \overline{A}_m \right) = \gamma^{\overline{d}_{op}, \overline{d}^*}\left(\overline{L}_m, \overline{A}_m \right)$ for $m = 0, 1, .., K$ for some positive definite matrix B. However a straightforward calculation shows this is not the case.

The inability to do simultaneous minimization has some very unpleasant implications when we acknowledge that none of our models $\gamma^j \left(\overline{l}_m, \overline{a}_m; \psi^j_m \right)$, $j = 1, ..., J$, $m = K, ..., 0$ for the optimal regimen are correct, where we assume the ψ^j_m are variation independent as m varies to facilitate sequential fitting.). At each time m we will consider the J candidates $\widehat{\gamma}^j_m \left(\overline{l}_m, \overline{a}_m \right) = \gamma^j \left(\overline{l}_m, \overline{a}_m; \widehat{\psi}^j_m \right)$ for $\gamma^j \left(\overline{l}_m, \overline{a}_m \right)$. Suppose, for subjects who receive care in the community until the time t_K from infection and who presents to our optimal treatment clinic with history $\left(\overline{L}_K, \overline{A}_{K-1} \right)$, we chose $\widehat{\gamma}^{\widehat{j}_K} \left(\overline{l}_K, \overline{a}_K \right)$ minimizing, at $m = K$,

$$P^{val}_{n^{val}} \left[g^2 \left(\overline{L}_m, \overline{A}_{m-1} \right) \left\{ H^{\underline{d}_{op,m}, \underline{d}^*_m}_{\text{mod},m} \left(\widehat{\gamma}^j_m \right) - E^* \left[H^{\underline{d}_{op,m}, \underline{d}^*_m}_{\text{mod},m} \left(\widehat{\gamma}^j_m \right) | \overline{L}_m, \overline{A}_{m-1} \right] \right\}^2 \right]$$

for an agreed upon function $g\left(\overline{L}_K, \overline{A}_{K-1}\right)$. Then for subjects who begin following our final estimate $\widehat{d}_{op}$ of the optimal regime beginning at t_0 we would often wish to chose $\widehat{\widehat{\gamma}^{j_0}}\left(\overline{l}_K, \overline{a}_K\right)$ minimizing, at $m = K, P_{n^{val}}^{val}[g^2\left(\overline{L}_m, \overline{A}_{m-1}\right) \times$

$$\left\{H_{\text{mod},m}^{\underline{d}_{op,m},\underline{d}_m^*}\left(\widehat{\gamma}_m^j\right) - E^*\left[H_{\text{mod},m}^{\underline{d}_{op,m},\underline{d}_m^*}\left(\widehat{\gamma}_m^j\right)|\overline{L}_m, \overline{A}_{m-1}\right]\right\}^2 w\left(\overline{L}_m, \overline{A}_{m-1}\right)]$$

with $w\left(\overline{l}_K, \overline{a}_{K-1}\right) = f_{\overline{L}_{\widehat{d}_{op,K}}, \overline{A}_{\widehat{d}_{op,K-1}}}\left(\overline{l}_K, \overline{a}_{K-1}\right) / f_{\overline{L}_K, \overline{A}_{K-1}}\left(\overline{l}_K, \overline{a}_{K-1}\right)$. But our final estimate $\widehat{d}_{op}$, much less an estimate of $f_{\overline{L}_{\widehat{d}_{op,K}}, \overline{A}_{\widehat{d}_{op,K-1}}}\left(\overline{l}_K, \overline{a}_{K-1}\right)$, is unknown to us at the time we are estimating $\widehat{\widehat{\gamma}^{j_0}}\left(\overline{l}_K, \overline{a}_K\right)$ since we estimate $\widehat{d}_{op}$ beginning with occassion K. Thus it is not at all straightforward to estimate an optimal regime using Corollary 9.2, suggesting that, in practice, we should use the cross-validation methods described in the previous subsection.

10 Appendix 1:

A1.1: Exceptional Laws:

In this Appendix we prove that under certain exceptional laws F_O (i) no regular estimators of the parameter $\psi^\dagger$ of an drSNMM $\gamma^{\overline{0}}\left(\overline{l}_m, \overline{a}_m, \psi^\dagger\right)$ for $\gamma^{\overline{d}_{op}}\left(\overline{l}_m, \overline{a}_m\right)$ exists and (ii) although $n^{1/2}$-consistent non-regular estimators exist, they will generally be asymptotically biased. As a consequence our estimators of $E\left[Y_{\overline{d}_{op}}\right]$ may also be non-regular and asymptotically biased.

In the interest of concreteness and with essentially no loss of generality, we will consider example 1 in the main text with $K = 1$, and $L_0 = 0$ wp1 (so we can disregard L_0), A_0, A_1, L_1 all Bernoulli, and $Y = L_2$ continuous. Thus we observe $O = (A_0, A_1, L_1, Y)$. Hence, dropping the $\overline{0}$ superscript from $\gamma^{\overline{0}}$, we have the model $\gamma\left(\overline{l}_1, \overline{a}_1, \psi\right) = a_1\left(1, l_1, a_0\right)\psi_1 = a_1\left(\psi_{11} + \psi_{21}l_1 + \psi_{31}a_0\right), \gamma\left(\overline{l}_0, \overline{a}_0, \psi\right) = \psi_0 a_0$, where we note $\overline{l}_1 = l_1$. We assume sequential randomization and that $f\left(a_m|\overline{l}_m, \overline{a}_{m-1}\right)$ is known. Let $\mathcal{F}$ denote all laws F_O with compact support and consistent with these assumptions. Then, by Eq. 3.8, we have that

$$\psi_0^\dagger = \tag{A1.1}$$
$$[var\left\{A_0\right\}]^{-1} E\left[\left\{A_0 - E[A_0]\right\}\left\{Y + \left(I\left[\left(1, L_1, A_0\right)\psi_1^\dagger > 0\right] - A_1\right)\left(1, L_1, A_0\right)\psi_1^\dagger\right\}\right] .$$

We will prove that their is no regular estimator of $\psi_0^\dagger$ when $pr\left[\left(1, L_1, A_0\right)\psi_1^\dagger = 0\right] \neq 0$, where before doing so it will be pedagogically useful to study the simplest example of the phenomenon.

A simple normal theory example: Consider estimation of the parameter $\psi^\dagger = \psi\left(\mu^\dagger\right) = \mu^\dagger I\left(\mu^\dagger > 0\right)$ from n i.i.d. observations X_i from a $N\left(\mu^\dagger, 1\right)$

distribution. By definition, a necessary condition for the existence of a regular estimator of $\psi\left(\mu^{\dagger}\right)$ at $\mu^{\dagger}$ is that $\psi(\mu)$ be differentiable at $\mu^{\dagger}$. As $\psi(\mu)$ is not differentiable at $\mu=0$, no regular estimator of $\psi^{\dagger}$ exists at $\mu^{\dagger}=0$ and we say that $\psi^{\dagger}$ is a non-regular parameter.

The MLE $I\left(X_{av}>0\right)X_{av}$ with $X_{\ av}=P_n(X)$ is an efficient RAL estimator of $\psi^{\dagger}$ at any non-zero $\mu^{\dagger}$. To see explicitly that $I\left(X_{av}>0\right)X_{av}$ is non-regular at $\mu^{\dagger}=0$, we compute its limiting distribution under the local data generating process (LDGP) $\mu=kn^{-1/2}$. Now regularity would mean that the limiting distribution of $V(k)$ does not depend on k where
$V(k)=n^{1/2}\left\{I\left(X_{av}>0\right)X_{av}-\psi\left(kn^{-1/2}\right)\right\}$
$=n^{1/2}\left\{I\left(X_{av}>0\right)X_{av}-I\left[kn^{-1/2}>0\right]kn^{-1/2}\right\}$
$=I\left(n^{1/2}\left(X_{av}-kn^{-1/2}\right)>-k\right)n^{1/2}\left(X_{av}-kn^{-1/2}\right)+$
$I\left(n^{1/2}\left(X_{av}-kn^{-1/2}\right)>-k\right)k-I\left[k>0\right]k$
$=I\left(Z>-k\right)Z+I\left(Z>-k\right)k-I\left[k>0\right]k$ where Z is a standard normal deviate. Thus $V(k)$ converges to the $N(0,1)$ distribution as $k\to\infty$, to a degenerate random variable with mass 1 at 0 as $k\to-\infty$, and to the law of $I\left(Z>0\right)Z$ for $k=0$. The asymptotic bias $asybias(k)$ of $I\left(X_{av}>0\right)X_{av}$ as an estimator of $\psi\left(kn^{-1/2}\right)$ is

$$asybias(k)=\left\{E\left[Z|Z>-k\right]+k\right\}pr\left[Z>-k\right]-I\left[k>0\right]k$$

so

$$asybias(0)=E\left[Z|Z>0\right]pr\left[Z>0\right]=1/\sqrt{2\pi}=.707\left(\sqrt{\pi}\right)^{-1}\ .$$

Note the asymptotic bias is bounded for all k. Standard attempts at bias correction such as bootstrapping do not result in an asymptotically unbiased estimator. The exact bias of the MLE $I\left(X_{av}>0\right)X_{av}$ is also $\int_{-n^{1/2}\mu}^{\infty}\left(z+n^{1/2}\mu\right)\phi(z)\,dz-\mu I\left(\mu>0\right)$. The parametric bootstrap estimate (i.e. MLE) of bias is thus
$\int_{-n^{1/2}X_{av}}^{\infty}\left(z+n^{1/2}X_{av}\right)\phi(z)\,dz-X_{av}I\left(X_{av}>0\right)$ so the bootstrap biased corrected estimator of $\mu I\left(\mu>0\right)$ is
$2X_{av}I\left(X_{av}>0\right)-\int_{-n^{1/2}X_{av}}^{\infty}\left(z+n^{1/2}X_{av}\right)\phi(z)\,dz$ which itself has bias
$2\int_{-n^{1/2}\mu}^{\infty}\left(z+n^{1/2}\mu\right)\phi(z)\,dz-\mu I\left(\mu>0\right)-E\left[\int_{-n^{1/2}X_{av}}^{\infty}\left(z+n^{1/2}X_{av}\right)\phi(z)\,dz\right]$.
Thus the asymptotic and exact bias of the bias corrected estimator at $\mu=0$ is

$$\begin{aligned}&2\int_{0}^{\infty}z\phi(z)\,dz-\int_{-\infty}^{\infty}\int_{-x}^{\infty}\left\{z+x\right\}\phi(z)\,\phi(x)\,dzdx\\&=\left(\sqrt{\pi}\right)^{-1}\left(\sqrt{2}-1\right)=.41\left(\sqrt{\pi}\right)^{-1}\end{aligned}$$

which is positive although less than the MLE's exact and asymptotic bias of $.707\left(\sqrt{\pi}\right)^{-1}$.

Consider next the submodel of our previous normal model in which we know that $\mu^{\dagger}>0$. Then for every value of $\mu^{\dagger}$ in the parameter space $(0,\infty)$, both the MLE $I\left(X_{av}>0\right)X_{av}$ and the estimator X_{av} of the now regular

parameter $\psi^\dagger = \psi\left(\mu^\dagger\right) = \mu^\dagger I\left(\mu^\dagger > 0\right) = \mu^\dagger$ are RAL estimators with asymptotic variance 1 i.e. both $n^{1/2}\left\{I\left(X_{av} > 0\right) X_{av} - \psi\left(\mu^\dagger + kn^{-1/2}\right)\right\}$ and $n^{1/2}\left\{X_{av} - \psi\left(\mu^\dagger + kn^{-1/2}\right)\right\}$ converge in distribution to $N\left(0,1\right)$ random variable under the LDGP $\mu = \mu^\dagger + kn^{-1/2}$. However the MLE Wald interval $I\left(X_{av} > 0\right) X_{av} \pm z_{\alpha/2} n^{-1/2}$, in contrast to the interval $X_{av} \pm z_{\alpha/2} n^{-1/2}$, is neither an exact nor a uniform asymptotic $(1-\alpha)$ conservative confidence interval for $\mu^\dagger$, although it is a non-uniform asymptotic $(1-\alpha)$ confidence interval where we have used the following definitions.

Some Definitions: Suppose we observe n i.i.d copies of a random variable O whose distribution F_O lies in a set $\mathcal{F} = \left\{F\left(\psi,\rho\right); \left(\psi,\rho\right) \in \Psi \times \mathcal{R}\right\}$ of distributions indexed by a finite dimesional parameter of interest ψ and, a possibly infinite-dimesional, variation-independent nuisance parameter ρ. We shall need the following from Robins and Ritov (1997). Below, we abbreviate $\sup_{(\psi,\rho)\in\Psi\times\mathcal{R}}$ by $\sup_{(\psi,\rho)}$.

Definition: An estimator $\widehat{\psi}_n$ (with n indexing sample size) is uniformly regular Gaussian (URG) with uniform asymptotic variance $\sigma^2\left(\psi,\rho\right)$ if

$$\sup_{(\psi,\rho)} \mid \Pr_{(\psi,\rho)}\left[n^{\frac{1}{2}}\left(\widehat{\psi}_n - \psi\right) < t\right] - \Phi\left(t; \sigma^2\left(\psi,\rho\right)\right) \mid \to 0 \text{ as } n \to \infty \quad \text{(A1.2)}$$

where $\Phi\left(t;\sigma^2\right)$ is the cumulative distribution function of a normal random variable with mean zero and variance σ^2. If $\widehat{\psi}_n$ is a uniformly asymptotic linear estimator of ψ (i.e. the $o_p\left(1\right)$ term in the definition of an asymptotically linear estimator is uniformly $o_p\left(1\right)$ over all laws in $\Psi \times \mathcal{R}$), then $\widehat{\psi}_n$ is URG. However, $\widehat{\psi}_n$, a regular asymptotic linear (RAL) estimator, does not imply $\widehat{\psi}_n$ is URG.

Definition: The estimator $\widehat{\psi}_n$ is uniformly asymptotically normal and unbiased (UANU) for ψ if there exists a sequence $\sigma_n^2\left(\psi,\rho\right)$ such that the z-statistic $n^{\frac{1}{2}}\left(\widehat{\psi}_n - \psi\right)/\sigma_n\left(\psi,\rho\right)$ converges uniformly to a $N\left(0,1\right)$ random variable, i.e.

$$\sup_{(\psi,\rho)} \mid \Pr_{(\psi,\rho)}\left[n^{\frac{1}{2}}\left(\widehat{\psi}_n - \psi\right)/\sigma_n\left(\psi,\rho\right) < t\right] - \Phi\left(t;1\right) \mid \to 0 \text{ as } n \to \infty\,. \quad \text{(A1.3)}$$

$\widehat{\psi}_n$ URG implies $\widehat{\psi}_n$ UANU but the converse is false. However, if $\widehat{\psi}_n$ is UANU and $\sigma_n\left(\psi,\rho\right)$ converges uniformly to $\sigma\left(\psi,\rho\right)$ i.e.

$$\sup_{(\psi,\rho)} \mid \sigma\left(\psi,\rho\right)/\sigma_n\left(\psi,\rho\right) - 1 \mid \to 0 \text{ as } n \to \infty \quad \text{(A1.4)}$$

the $\widehat{\psi}_n$ is URG. Furthermore, if $\widehat{\psi}_n$ is UANU and there exists an estimator $\widehat{\sigma}_n$ of $\sigma_n\left(\psi,\rho\right)$ such that $\sigma_n\left(\psi,\rho\right)/\widehat{\sigma}_n$ converges to one uniformly in probability, i.e. for all $\varepsilon > 0$

$$\sup_{(\psi,\rho)} \Pr_{(\psi,\rho)}\left[\mid 1 - \sigma_n\left(\psi,\rho\right)/\widehat{\sigma}_n \mid > \varepsilon\right] \to 0 \text{ as } n \to \infty \quad \text{(A1.5)}$$

then, by the uniform version of Slutzky's Theorem, the t-statistic $n^{\frac{1}{2}}\left(\widehat{\psi}_n - \psi\right)/\widehat{\sigma}_n$ converges uniformly to a $N(0,1)$ random variable, and thus the "Wald" interval $C_n \equiv \widehat{\psi}_n \pm z_{\alpha/2}\widehat{\sigma}_n/\sqrt{n}$ is a uniform asymptotic $1-\alpha$ confidence interval for ψ where $z_{\alpha/2}$ is the $\alpha/2$ quantile of a standard normal distribution, and we have the following definition.

Definition: C_n is a uniform asymptotic $1-\alpha$ confidence interval for ψ if $\sup_{(\psi,\rho)} |\Pr_{(\psi,\rho)}[\psi \in C_n] - (1-\alpha)| \to 0$ as $n \to \infty$.

Definiton: C_n is a conservative uniform asymptotic $1-\alpha$ confidence interval for ψ if

$$\liminf_n \inf_{(\psi,\rho)} \left\{\Pr_{(\psi,\rho)}[\psi \in C_n] - (1-\alpha)\right\} \geq 0 \text{ as } n \to \infty$$

.

Note a uniform asymptotic $1-\alpha$ confidence interval is a conservative uniform asymptotic $1-\alpha$ confidence interval. We required uniformity in our definition of an asymptotic confidence interval to be consistent with the usual definition of a non-asymptotic confidence interval. Specifically, by definition, for each sample size n, a conservative $1-\alpha$ (non-asymptotic) confidence interval C_n satisfies that for all $(\psi,\rho) \in \Psi \times \mathcal{R}$,

$$\Pr_{(\psi,\rho)}[\psi \in C_n] \geq 1-\alpha.$$

Our definition of a uniform asymptotic confidence interval satisfies the following consistency condition: if there is no conservative (non-asymptotic) $1-\alpha$ confidence interval for ψ whose length converges to 0 in probability as $n \to \infty$, then no conservative uniform asymptotic $1-\alpha$ confidence interval for ψ exists whose length converges to 0 in probability as $n \to \infty$; in contrast, there may still exist a conservative asymptotic $1-\alpha$ confidence interval for ψ whose length converges to 0 in probability as $n \to \infty$, where we have the following.

Definition: C_n is a conservative asymptotic $1-\alpha$ confidence interval for ψ if for all $(\psi,\rho) \in \Psi \times \mathcal{R}$, $\liminf_n \left\{\Pr_{(\psi,\rho)}[\psi \in C_n] - (1-\alpha)\right\} \geq 0$ as $n \to \infty$

If $\widehat{\psi}$ is UANU and $\sigma_n(\psi,\rho)/\widehat{\sigma}_n$ converges uniformly to one, then the $\widehat{\psi}_n \pm z_{\alpha/2}\widehat{\sigma}_n/\sqrt{n}$ will be a uniform asymptotic $(1-\alpha)$ confidence interval for ψ , even if $\widehat{\psi}_n$ is not URG. If $\widehat{\psi}_n$ is UANU but not URG, then even if $\sigma_n(\psi,\rho) \to \sigma(\psi,\rho)$ as $n \to \infty$ for all (ψ,ρ), this convergence cannot be uniform. Further, under mild regularity conditions, when $\widehat{\psi}_n$ is UANU, the non-parametric bootstrap estimator $\widehat{\sigma}_n$ of the standard error of $n^{\frac{1}{2}}\left(\widehat{\psi}_n - \psi\right)$ will satisfy (A1.5). Hence, if $\widehat{\psi}_n$ is UANU, then a Wald interval centered on $\widehat{\psi}_n$ and using a bootstrap estimate of the standard error will be an asymptotic uniform $(1-\alpha)$ Wald confidence interval for ψ

The simple normal example continued: Returning to the simple normal example with $\mu^\dagger \in (0,\infty)$, the Wald interval $I(X_{av} > 0) X_{av} \pm z_{\alpha/2} n^{-1/2}$ is not a uniform asymptotic $(1-\alpha)$ conservative confidence interval for $\mu^\dagger$ because $I(X_{av} > 0) X_{av}$, in contrast to X_{av}, is not URG. Indeed

$I(X_{av} > 0) X_{av}$ is not even UANU and thus cannot center a uniform asymptotic $(1-\alpha)$ confidence interval even were its standard error estimated with the bootstrap. The reason that $I(X_{av} > 0) X_{av}$ is not UANU is that, at each sample size n, there exists a $\mu^\dagger \in (0, \infty)$ depending on n that is sufficiently close to 0 that $I(X_{av} > 0) X_{av}$ is significantly biased upwards as an estimator of $\psi^\dagger = \mu^\dagger$. However, if we took the parameter space for $\mu^\dagger$ to be (σ, ∞) for a fixed $\sigma > 0$, then $I(X_{av} > 0) X_{av}$ is UANU and can center a uniform asymptotic $(1-\alpha)$ confidence interval. Returning to the case where the parameter space for $\mu^\dagger$ is the entire real line so no UANU estimator of $\psi^\dagger$ exists, we can nonetheless construct a conservative non-asymptotic $(1-\alpha)$ confidence interval for $\psi^\dagger$ that is also a conservative uniform asymptotic $(1-\alpha)$ confidence interval by intersecting the usual interval $X_{av} \pm z_{\alpha/2} n^{-1/2}$for $\mu^\dagger$ with the non-negative real line to obtain $\left\{X_{av} \pm z_{\alpha/2} n^{-1/2}\right\} \cap [0, \infty)$.

Return to our optimal-regime SNMM example. In this example, the data distribution can be parametrized as follows: $f\left(Y, \overline{L}_1, \overline{A}_1; \psi_1, \theta\right) =$
$f\left(\delta(\psi_1, \theta_1) | \overline{L}_1, \overline{A}_1; \theta_2\right) f\left(A_1 | \overline{L}_1, A_0\right) f(L_1 | A_0; \theta_3) f(A_0)$, where $\theta =$
$(\theta_1, \theta_2, \theta_3)$, $\delta(\psi_1, \theta_1) = Y - A_1(1, L_1, A_0)\psi_1 -$
$\left\{q(L_1, A_0; \theta_1) - \int q(L_1, A_0; \theta_1) dF(L_1 | A_0; \theta_3)\right\}$ and θ_1 indexes all functions $q(L_1, A_0; \theta_1)$ of (L_1, A_0) that satisfy $q(0, A_0; \theta_1) = 0, \theta_2$ indexes all conditional densities $f\left(u | \overline{L}_1, \overline{A}_1; \theta_2\right)$ satisfying the conditional mean zero restriction $\int u f\left(u | \overline{L}_1, \overline{A}_1; \theta_2\right) = 0$, and θ_3 indexes all conditional densities $f(L_1 | A_0; \theta_3)$. In this parametrization we view $\psi_0^\dagger$ as a function of $\left(\psi_1^\dagger, \theta^\dagger\right)$ determined by equation (A1.1). In the following discussion $\psi_0^\dagger$ and $\psi_1^\dagger$ are analogous to $\psi^\dagger$ and $\mu^\dagger$ respectively in our normal example. Consider a regular parametric submodel with θ fixed at its true value $\theta^\dagger$ and ψ_1 a free parameter with true value $\psi_1^\dagger$. Then $\psi_0(\psi_1) = [var\{A_0\}]^{-1} \times$
$E_{\psi_1, \theta^\dagger}\left[\{A_0 - E[A_0]\}\{Y_{K+1} + (I[(1, L_1, A_0)\psi_1 > 0] - A_1)(1, L_1, A_0)\psi_1\}\right]$ and $\psi_0(\psi_1)$ is a differentiable function of ψ_1 at the truth $\left(\psi_1^\dagger, \theta^\dagger\right)$ w.p.1. if and only if the event $(1, L_1, A_0)\psi_1^\dagger = 0$ has probability 0 under $f\left(Y, \overline{L}_1, \overline{A}_1; \psi_1^\dagger, \theta^\dagger\right)$. It follows that if $pr_{\psi_1^\dagger, \theta^\dagger}\left[(1, L_1, A_0)\psi_1^\dagger = 0\right] \neq 0$, then ψ_0 is not a pathwise differentiable parameter and thus no regular estimator of ψ_0 exists under the data generating process $\left(\psi_1^\dagger, \theta^\dagger\right)$. If $pr_{\psi_1^\dagger, \theta^\dagger}\left[(1, L_1, A_0)\psi_1^\dagger = 0\right] = 0$ the closed form estimator estimator $\widetilde{\psi}_0$ of Sec. 4.2 is regular at $\left(\psi_1^\dagger, \theta^\dagger\right)$.

This raises the question as to the asymptotic distribution and asymptotic mean of $\widetilde{\psi}_0$ when
$pr_{\psi_1^\dagger, \theta^\dagger}\left[(1, L_1, A_0)\psi_1^\dagger = 0\right] \neq 0$, i.e. under an exceptional law. [Analagous results apply to the non-closed form estimators of $\psi_0^\dagger$ discussed earlier; the advantage of considering the simpler $\widetilde{\psi}_0$ is that its behavior is relatively transparent.] Abbreviate $U_{\text{mod},m}^{\overline{d}_{op}, \overline{0}}$ of Section 4.2 to U_m.

Recall

$$\widetilde{\psi}_0 = \widetilde{I}_0^{-1}$$
$$P_n\left[\left\{s_0\left(A_0\right) - E[s_0\left(A_0\right)]\right\}\left\{Y + \left(I\left[\left(1, L_1, A_0\right)\widetilde{\psi}_1 > 0\right] - A_1\right)\left(1, L_1, A_0\right)\widetilde{\psi}_1\right\}\right]$$

with $\widetilde{I}_0 = P_n\left[\left\{s_0\left(A_0\right) - E[s_0\left(A_0\right)]\right\}A_0\right]$. Recall $\widetilde{\psi}_0$ solves

$$0 = n^{1/2}P_n\left[U_0\left(\psi_0, s_0, \widetilde{\psi}_1\right)\right]$$

where

$$\begin{aligned} & U_0\left(\psi_0, s_0, \widetilde{\psi}_1\right) \qquad \text{(A1.6)} \\ & = \left\{s_0\left(A_0\right) - E[s_0\left(A_0\right)]\right\} \times \\ & \left\{Y - \psi_0 A_0 + \left(I\left[\left(1, L_1, A_0\right)\widetilde{\psi}_1 > 0\right] - A_1\right)\left(1, L_1, A_0\right)\widetilde{\psi}_1\right\}. \end{aligned}$$

Now

$$\begin{aligned} & n^{1/2}P_n\left[U_0\left(\psi_0, s_0, \widetilde{\psi}_1\right)\right] \\ & = n^{1/2}P_n\left[U_0\left(\psi_0, s_0, \psi_1^\dagger\right)\right] + n^{1/2}P_n\left[\Delta_0\left(s_0, \widetilde{\psi}_1, \psi_1^\dagger\right)\right], \end{aligned}$$

where
$\Delta_0\left(s_0, \widetilde{\psi}_1, \psi_1^\dagger\right) = I\left[\left(1, L_1, A_0\right)\psi_1^\dagger = 0\right]\left\{s_0\left(A_0\right) - E[s_0\left(A_0\right)]\right\} \times$
$\left(I\left[\left(1, L_1, A_0\right)\left(\widetilde{\psi}_1 - \psi_1^\dagger\right) > 0\right] - A_1\right)\left(1, L_1, A_0\right)\left(\widetilde{\psi}_1 - \psi_1^\dagger\right) +$
$I\left[\left(1, L_1, A_0\right)\psi_1^\dagger \neq 0\right]\left\{s_0\left(A_0\right) - E[s_0\left(A_0\right)]\right\} \times$

$$\left[\begin{array}{c} \left(I\left[\left(1, L_1, A_0\right)\widetilde{\psi}_1 > 0\right] - A_1\right)\left(1, L_1, A_0\right)\widetilde{\psi}_1 - \\ \left(I\left[\left(1, L_1, A_0\right)\psi_1^\dagger > 0\right] - A_1\right)\left(1, L_1, A_0\right)\psi_1^\dagger \end{array}\right].$$

To be concrete suppose that $\sum_{j=1}^3 \psi_{j1}^\dagger = 0$, but none of the $\psi_{j1}^\dagger$ are zero so $\left(1, L_1, A_0\right)\psi_1^\dagger = 0 \Longleftrightarrow L_1 = A_0 = 1$ which we assume happens with positive probability so the law is exceptional. Then, we have $n^{1/2}P_n\left[\Delta_0\left(s_0, \widetilde{\psi}_1, \psi_1^\dagger\right)\right] =$

$$n^{1/2}P_n\left[\begin{array}{c} I\left[\left(1, L_1, A_0\right)\psi_1^\dagger = 0\right]\left\{s_0\left(A_0\right) - E[s_0\left(A_0\right)]\right\} \times \\ \left(I\left[\sum_{j=1}^3\left(\widetilde{\psi}_{j1} - \psi_{j1}^\dagger\right) > 0\right] - A_1\right)\sum_{j=1}^3\left(\widetilde{\psi}_{j1} - \psi_{j1}^\dagger\right) \end{array}\right] +$$
$$n^{1/2}P_n\left[\begin{array}{c} I\left[\left(1, L_1, A_0\right)\psi_1^\dagger \neq 0\right]\left\{s_0\left(A_0\right) - E[s_0\left(A_0\right)]\right\} \times \\ \left(I\left[\left(1, L_1, A_0\right)\psi_1^\dagger > 0\right] - A_1\right)\left(1, L_1, A_0\right)\left(\widetilde{\psi}_1 - \psi_1^\dagger\right) \end{array}\right] + o_p\left(1\right).$$

To obtain this result we used the fact that, conditional on $\{(A_{0i}, L_{1i}); i = 1, ..., n\}$, when $(1, l_1, a_0)\psi_1^\dagger \neq 0$, $I\left[(1, l_1, a_0)\widetilde{\psi}_1 > 0\right] = I\left[(1, l_1, a_0)\psi_1^\dagger > 0\right] + o_p(1)$ under any $F_O \in \mathcal{F}$. To prove this fact we reasoned as follows. Because, as argued below, with probability going to one, $\widetilde{\psi}_1$ is $n^{1/2} - consistent$ for $\psi_1^\dagger$ conditional on $\{(A_{0i}, L_{1i}); i = 1, ..., n\}$, it follows that whenever $(1, l_1, a_0)\psi_1^\dagger \neq 0$, $(1, l_1, a_0)\widetilde{\psi}_1$ and $(1, l_1, a_0)\psi_1^\dagger$ will have the same sign except on a set $C(l_1, a_0)$ with conditional probability given $\{(A_{0i}, L_{1i}); i = 1, ..., n\}$ converging to 0. That is

$$\text{for all } \left\{F_O; F_O \in \mathcal{F},\ (1, l_1, a_0)\psi_1^\dagger(F_O) \neq 0\right\}, \tag{A1.7}$$

$$pr_{F_O}\left[\{O_i, i = 1, ..., n\} \in C(l_1, a_0) \mid \{(A_{0i}, L_{1i}); i = 1, ..., n\}\right] \to 0 \text{ as } n \to \infty.$$

However it is important to note for later reference that this convergence is not uniform, i.e.,

$$\sup_{F_O \in \widetilde{Q}} pr\left[\{O_i, i = 1, ..., n\} \in C(l_1, a_0) \mid \{(A_{0i}, L_{1i}); i = 1, ..., n\}\right] \tag{A1.8}$$

$$\nrightarrow 0 \text{ as } n \to \infty,$$

$\widetilde{Q} = \left\{F_O; F_O \in \mathcal{F},\ (1, l_1, a_0)\psi_1^\dagger(F_O) \neq 0\right\}$, because for each sample size n and each (l_1, a_0), there exists a $F_O \in \mathcal{F}$ with $(1, l_1, a_0)\psi_1^\dagger \neq 0$ but $(1, l_1, a_0)\psi_1^\dagger$ within $O\left(n^{-1/2}\right)$ of 0. It follows that due to the $O\left(n^{-1/2}\right)$ fluctuations in $\widetilde{\psi}_1$, $(1, l_1, a_0)\widetilde{\psi}_1$ and $(1, l_1, a_0)\psi_1^\dagger$ will have different signs with probability substantially greater than 0.

To proceed in our analysis of $\widetilde{\psi}_0$, we need to analyze $\widetilde{\psi}_1$. By a Taylor expansion $\widetilde{\psi}_1$ is a RAL estimator so that

$$n^{1/2}\left(\widetilde{\psi}_1 - \psi_1^\dagger\right) = n^{1/2}P_n[IF_1] + o_p(1),$$

where

$$\begin{aligned} IF_1 &= (IF_{11}, IF_{21}, IF_{31})^T \\ &= I_1^{-1}(s_1)\, U_1\left(\psi_1^\dagger, s_1\right), \\ I_1(s_1) &= -\partial E\left[U_1\left(\psi_1^\dagger, s_1\right)\right]/\partial\psi_1 \end{aligned}$$

Thus, by another Taylor expansion of around $\widetilde{\psi}_1$ around $\psi_1^\dagger$,
$n^{1/2}P_n\left[\Delta_0\left(s_0, \widetilde{\psi}_1, \psi_1^\dagger\right)\right] = E\left[A_0 L_1\left\{s_0(A_0) - E[s_0(A_0)]\right\}\right] \times$
$I\left[n^{1/2}P_n\left[\sum_{j=1}^3 IF_{j1}\right] > 0\right] n^{1/2}P_n\left[\sum_{j=1}^3 IF_{j1}\right] -$
$E\left[A_1 A_0 L_1\left\{s_0(A_0) - E[s_0(A_0)]\right\}\right] n^{1/2}P_n\left[\sum_{j=1}^3 IF_{j1}\right] +$
$E\left[I\left[(1, L_1, A_0)\psi_1^\dagger \neq 0\right]\left\{s_0(A_0) - E[s_0(A_0)]\right\} \times\right.$

$\left(I\left[(1,L_1,A_0)\,\psi_1^\dagger>0\right]-A_1\right)(1,L_1,A_0)\Big]\,n^{1/2}P_n\left[IF_1\right]+o_p\left(1\right)$. Note in terms of the estimation of $\psi_1^\dagger,\{(A_{0i},L_{1i})\,;i=1,...,n\}$ is ancillary, so $\widetilde{\psi}_1$ is $n^{1/2}-$ *consistent* for $\psi_1^\dagger$ conditional on $\{(A_{0i},L_{1i})\,;i=1,...,n\}$.

Let $Z=\left(Z_0,Z_1^T\right)^T=\left(Z_0,Z_{11},Z_{21},Z_{31}\right)^T$ be MVN with mean 0 and variance equal to that of $\left(U_0,IF_1^T\right)^T=\left(U_0,IF_{11},IF_{21},IF_{31}\right)^T$ with $U_0=U_0\left(\psi_0^\dagger,s_0,\psi_1^\dagger\right)$. Let $Z_{1+}=Z_{11}+Z_{21}+Z_{31}$. Then it follows from the expansion above that, under our assumption that for the F_O generating the data $A_0L_1=1\Leftrightarrow I\left[(1,L_1,A_0)\,\psi_1^\dagger=0\right]$, $n^{1/2}\left(\widetilde{\psi}_0-\psi_0^\dagger\right)$ converges in law to the distribution of $I_0^{-1}\left(s_0\right)\times$

$$\begin{array}{c}Z_0+E\left[A_0L_1\left\{s_0\left(A_0\right)-E[s_0\left(A_0\right)]\right\}\right]I\left[Z_{1+}>0\right]Z_{1+}-\\ \left\{E\left[A_1A_0L_1\left\{s_0\left(A_0\right)-E[s_0\left(A_0\right)]\right\}\right]Z_{1+}\right\}+\\ E\left[\left\{1-A_0L_1\right\}\left\{s_0\left(A_0\right)-E[s_0\left(A_0\right)]\right\}\left(I\left[(1,L_1,A_0)\,\psi_1^\dagger>0\right]-A_1\right)(1,L_1,A_0)\right]Z_1\end{array}$$

where

$$I_0\left(s_0\right)=-\partial E\left[U_0\left(\psi_0,s_0,\psi_1^\dagger\right)\right]/\partial\psi_0=E\left[\left\{s_0\left(A_0\right)-E[s_0\left(A_0\right)]\right\}A_0\right]\ .$$

Because of the term $E\left[A_0L_1\left\{s_0\left(A_0\right)-E[s_0\left(A_0\right)]\right\}\right]I\left[Z_{1+}>0\right]Z_{1+}$, $\widetilde{\psi}_0$, although $n^{1/2}$-consistent, is neither asymptotically normal nor asymptotically unbiased. The asymptotic bias $asybias\left(0\right)$ is
$I_0^{-1}\left(s_0\right)\left\{E\left[A_0L_1\left\{s_0\left(A_0\right)-E[s_0\left(A_0\right)]\right\}\right]\right\}E\left[Z_{1+}|Z_{1+}>0\right]pr\left[Z_{1+}>0\right]$. Thus

$$asybias\left(0\right)=I_0^{-1}\left(s_0\right)E\left[A_0L_1\left\{s_0\left(A_0\right)-E[s_0\left(A_0\right)]\right\}\right]\left\{var\left[Z_{1+}\right]\right\}^{1/2}/\sqrt{2\pi}\ .$$

To see that the estimator $\widetilde{\psi}_0$ is non-regular when $\psi_{11}^\dagger+\psi_{21}^\dagger+\psi_{31}^\dagger=0$, (implied by our assumption $I\left[(1,L_1,A_0)\,\psi_1^\dagger=0\right]\Longleftrightarrow L_1=A_0=1$), consider the local data generating process $\left(\psi_1^T,\theta^\dagger\right)=\left(\psi_{11},\psi_{21}^\dagger,\psi_{31}^\dagger,\theta^\dagger\right)$ with $\psi_{11}=\psi_{11}^\dagger+kn^{-1/2}$. Then
$n^{1/2}P_n\left[U_0\left(\psi_0,s_0,\widetilde{\psi}_1\right)\right]=n^{1/2}P_n\left[U_0\left(\psi_0,s_0,\psi_1\right)\right]+n^{1/2}P_n\left[\Delta_0\left(s_0,\widetilde{\psi}_1,\psi_1\right)\right]$,
$\Delta_0\left(s_0,\widetilde{\psi}_1,\psi_1\right)=I\left[(1,L_1,A_0)\,\psi_1^\dagger=0\right]\left\{s_0\left(A_0\right)-E[s_0\left(A_0\right)]\right\}\times$
$\left(I\left[(1,L_1,A_0)\left(\widetilde{\psi}_1-\psi_1^\dagger\right)>0\right]-A_1\right)(1,L_1,A_0)\left(\widetilde{\psi}_1-\psi_1^\dagger\right)\widetilde{\psi}_1$
$-I\left[(1,L_1,A_0)\,\psi_1^\dagger=0\right]\left\{s_0\left(A_0\right)-E[s_0\left(A_0\right)]\right\}\left(I\left[kn^{-1/2}>0\right]-A_1\right)kn^{-1/2}+$
$I\left[(1,L_1,A_0)\,\psi_1^\dagger\neq0\right]\left\{s_0\left(A_0\right)-E[s_0\left(A_0\right)]\right\}\Big[\Big(I\left[(1,L_1,A_0)\,\widetilde{\psi}_1>0\right]-A_1\Big)\times$
$(1,L_1,A_0)\,\widetilde{\psi}_1-\left(I\left[(1,L_1,A_0)\,\psi_1>0\right]-A_1\right)(1,L_1,A_0)\,\psi_1\Big]$.

Thus $n^{1/2}P_n\left[\Delta_0\left(s_0,\widetilde{\psi}_1,\psi_1\right)\right]=$

$$n^{1/2}P_n\left[\begin{array}{c}A_0L_1\left\{s_0\left(A_0\right)-E[s_0\left(A_0\right)]\right\}\times\\ \left(I\left[\sum_{j=1}^3\left(\widetilde{\psi}_{1j}-\psi_{1j}^\dagger\right)>0\right]-A_1\right)\sum_{j=1}^3\left(\widetilde{\psi}_{1j}-\psi_{1j}^\dagger\right)\end{array}\right]$$

$$-n^{1/2}P_n\left[I\left[(1,L_1,A_0)\,\psi_1^\dagger=0\right]\{s_0(A_0)-E[s_0(A_0)]\}\left(I[k>0]-A_1\right)kn^{-1/2}\right]+$$
$$n^{1/2}P_n\left[\begin{array}{c} I\left[(1,L_1,A_0)\,\psi_1^\dagger\neq 0\right]\{s_0(A_0)-E[s_0(A_0)]\}\times \\ \left(I\left[(1,L_1,A_0)\,\psi_1^\dagger>0\right]-A_1\right)(1,L_1,A_0)\left(\widetilde{\psi}_1-\psi_1\right)\end{array}\right]$$
$$+o_p(1).$$

Now $\widetilde{\psi}_1$ is a RAL estimator so that $n^{1/2}\left(\widetilde{\psi}_1-\psi_1\right)$ has the same limiting distribution under $\left(\psi_1,\theta_1^\dagger\right)$ as $n^{1/2}\left(\widetilde{\psi}_1-\psi_1^\dagger\right)$ under $\left(\psi_1^\dagger,\theta_1^\dagger\right)$. Therefore noting
$\left(I\left[\sum_{j=1}^3\left(\widetilde{\psi}_{1j}-\psi_{1j}^\dagger\right)>0\right]-A_1\right)\sum_{j=1}^3\left(\widetilde{\psi}_{1j}-\psi_{1j}^\dagger\right)=$
$\left(I\left[\sum_{j=1}^3\left(\widetilde{\psi}_{1j}-\psi_{1j}\right)>-kn^{-1/2}\right]-A_1\right)\left[\sum_{j=1}^3\left(\widetilde{\psi}_{1j}-\psi_{1j}\right)+kn^{-1/2}\right]$, let
$Z=\left(Z_0,Z_1^T\right)^T=(Z_0,Z_{11},Z_{12},Z_{13})^T$ have the same distribution as above.

Then $n^{1/2}\left(\widetilde{\psi}_0-\psi_0\left(\psi_1^T,\theta^\dagger\right)\right)$ converges in law to the distribution of
$I_0^{-1}(s_0)\times$
$$\left[\begin{array}{c} Z_0+E\left[A_0L_1\{s_0(A_0)-E[s_0(A_0)]\}\right]I[Z_{1+}>-k]\,[Z_{1+}+k] \\ -E\left[A_1A_0L_1\{s_0(A_0)-E[s_0(A_0)]\}\right][Z_{1+}+k] \\ +E\left[I\left[(1,L_1,A_0)\,\psi_1^\dagger=0\right]\{s_0(A_0)-E[s_0(A_0)]\}\left(I[k>0]-A_1\right)\right]k+Z_1\times \\ E\left[I\left[(1,L_1,A_0)\,\psi_1^\dagger\neq 0\right]\{s_0(A_0)-E[s_0(A_0)]\}\times\right. \\ \left.\left(I\left[(1,L_1,A_0)\,\psi_1^\dagger>0\right]-A_1\right)(1,L_1,A_0)\right]\end{array}\right]$$

which can be written $I_0^{-1}(s_0)\times$
$$\left[\begin{array}{c} Z_0+E\left[A_0L_1\{s_0(A_0)-E[s_0(A_0)]\}\right]I[Z_{1+}>-k]\,[Z_{1+}+k] \\ -E\left[A_1A_0L_1\{s_0(A_0)-E[s_0(A_0)]\}\right]Z_{1+} \\ +E\left[A_0L_1\{s_0(A_0)-E[s_0(A_0)]\}\right]k\left(I[k>0]\right)+Z_1\times \\ E\left[\{1-A_0L_1\}\{s_0(A_0)-E[s_0(A_0)]\}\left(I\left[(1,L_1,A_0)\,\psi_1^\dagger>0\right]-A_1\right)(1,L_1,A_0)\right].\end{array}\right]$$
Thus, the mean of the limiting distribution [i.e the asymptotic bias, $asybias(k)$, of $\widetilde{\psi}_0$] is
$$I_0^{-1}(s_0)\left\{\begin{array}{c} E\left[A_0L_1\{s_0(A_0)-E[s_0(A_0)]\}\right]E\left[Z_{1+}+k|Z_{1+}>-k\right]pr\left[Z_{1+}>-k\right] \\ +E\left[A_0L_1\{s_0(A_0)-E[s_0(A_0)]\}\right]k\left(I[k>0]\right)\end{array}\right\}$$

Since the limiting distribution depends on k, $\widetilde{\psi}_0$ is not regular at exceptional laws. It follows that the nominal $(1-\alpha)$ Wald interval centered on $\widetilde{\psi}_0$ is not a conservative $(1-\alpha)$ uniform asymptotic confidence interval. Appropriate alternative methods for construction of uniform asymptotic confidence intervals for the entire vector $\psi^\dagger$ are discussed in Section 4.3 and for subvectors such as $\psi_0^\dagger$ in section 5.1.

Now suppose one objected to the above example by arguing that it is apriori unlikely that $\sum_{j=1}^3\psi_{j1}^\dagger=0$, when none of the $\psi_{j1}^\dagger$ are zero as such a fortuitous cancellation of parameter values would be apriori unlikely. If we apriori excluded such unlikely laws from our model then the only remaining exceptional laws would be those corresponding to the null hypothesis $\psi_{j1}^\dagger=0$ for $j=1,2,3$ that says treatment at time 1 has no effect. Suppose, however,

that it was known from other considerations that this null hypothesis was false. Then we are led to consider the submodel of our original model in which we impose the additional apriori assumption that $(1, l_1, a_0)\,\psi_1^\dagger \neq 0$ for any (l_1, a_0). This model has no exceptional laws. In this setting, $\psi_0^\dagger$ is a regular parameter and $\widetilde{\psi}_0$ is a RAL estimator at all laws in the model. However we now argue that $\widetilde{\psi}_0$ is not UANU. Thus a nominal $(1-\alpha)$ Wald interval centered on $\widetilde{\psi}_0$ is not a conservative $(1-\alpha)$ uniform asymptotic confidence interval, although it is a $(1-\alpha)$ non-uniform asymptotic interval.

To see why we revisit our derivation of the large sample distribution of $\widetilde{\psi}_0$ except now $I\left[(1, L_1, A_0)\,\psi_1^\dagger = 0\right]$ takes the value zero with probability one. Recall that in our derivation we proved that, asymptotically,

$$n^{1/2} P_n \left[\Delta_0\left(s_0, \widetilde{\psi}_1, \psi_1^\dagger\right)\right] = n^{1/2} P_n \left[\{s_0(A_0) - E[s_0(A_0)]\} \times \right.$$
$$\left\{\left(I\left[(1, L_1, A_0)\,\widetilde{\psi}_1 > 0\right] - A_1\right)(1, L_1, A_0)\,\widetilde{\psi}_1 - \right.$$
$$\left.\left.\left(I\left[(1, L_1, A_0)\,\psi_1^\dagger > 0\right] - A_1\right)(1, L_1, A_0)\,\psi_1^\dagger\right\}\right]$$

did not have the random $\widetilde{\psi}_1$ within an indicator function by using the fact that $I\left[(1, l_1, a_0)\,\widetilde{\psi}_1 > 0\right] = I\left[(1, l_1, a_0)\,\psi_1^\dagger > 0\right] + o_p(1)$ under any $F_O \in \mathcal{F}$ leading to $n^{1/2} P_n \left[\Delta_0\left(s_0, \widetilde{\psi}_1, \psi_1^\dagger\right)\right] =$

$$n^{1/2} P_n \left[\{s_0(A_0) - E[s_0(A_0)]\}\left(I\left[(1, L_1, A_0)\,\psi_1^\dagger > 0\right] - A_1\right)(1, L_1, A_0)\left(\widetilde{\psi}_1 - \psi_1^\dagger\right)\right]$$

$+o_p(1)$. But in (A1.8), we showed that these $o_p(1)$ terms were not uniform because for each n there exist $F_O \in \mathcal{F}$ with $(1, l_1, a_0)\,\psi_1^\dagger \neq 0$ but $(1, l_1, a_0)\,\psi_1^\dagger$ within $O\left(n^{-1/2}\right)$ of 0. As a consequence $n^{1/2} P_n \left[\Delta_0\left(s_0, \widetilde{\psi}_1, \psi_1^\dagger\right)\right]$ and thus $n^{1/2}\left(\widetilde{\psi}_0 - \psi_0^\dagger\right)$, although asymptotically normal, are not UANU because we cannot uniformly remove $\widetilde{\psi}_1$ from within an indicator function. However, if we further reduced our model by assuming $\left|(1, l_1, a_0)\,\psi_1^\dagger\right| > \sigma > 0$ for some fixed σ and all (l_1, a_0), then $n^{1/2}\left(\widetilde{\psi}_0 - \psi_0^\dagger\right)$ would be UANU and a nominal $(1-\alpha)$ Wald interval centered on $\widetilde{\psi}_0$ would be $(1-\alpha)$ uniform asymptotic confidence interval.

What does all this asymptotics imply for practical finite sample inference in non toy examples? I believe the take home message is roughly as follows. (The task of backing up the statements in this paragraph through further theoretical work and simulation experiments remains to be done.) Consider the uniform asymptotic confidence interval $C_{op}(1-\alpha)$ for the entire vector $\psi^\dagger$ discussed in Section 4.3. Let $H_{m,i} = card\left\{d_{op,m}\left(\overline{L}_{m,i}, \overline{A}_{m-1,i}, \psi\right); \psi \in C_{op}(1-\alpha)\right\}$ be

the number of potential optimal treatment strategies at time m for subject i based on his observed data $\left(\overline{L}_{m,i}, \overline{A}_{m-1,i}\right)$ that are consistent with his observed data. If the fraction p_{op} of the $H_{m,i}$ in the set $\{H_{m,i}; i = 1, ..., n,\ m = 1, ..., K\}$ that exceed 1 is moderate to large (say, $p_{op} > .05$) then inferences based on ordinary Wald point and interval estimates are unreliable for frequentist inference and the methods of sections 4.3 and 5.1 should be used instead. In terms of our toy model, the fraction of the $H_{m,i}$ exceeding 1 is an informal upper bound on how often $I\left[(1, l_1, a_0)\,\widetilde{\psi}_1 > 0\right]$ might differ from $I\left[(1, l_1, a_0)\,\psi_1^\dagger > 0\right]$ over the set of $\psi_1^\dagger$ consistent with the observed data (as determined via the confidence interval $C_{op}(1-\alpha)$). If p_{op} is small our finite sample inferences should agree with those based on an asymptotics that assumes $\left|(1, l_1, a_0)\,\psi_1^\dagger\right| > \sigma > 0$ and thus inference based on Wald intervals and asymptotic normality should be trustworthy. The governing idea is that we not worry about values of $\psi^\dagger$ that are incompatible with the data.

Indeed I believe a successful strategy with a potential for enormous savings in computing time is as follows. Compute the closed form estimate $\widetilde{\psi}$ and then the locally efficient one step update $\widetilde{\psi}^{(1)}$ of Section 4.2 and use $\widetilde{\psi}^{(1)}$ to center a Wald intervsal $C_{Wald-onestep}(1-\alpha)$. Compute the fraction $p_{Wald-onestep}$ of the $H_{m,i}$ that exceed 1 but now using the easy to compute $C_{Wald-onestep}(1-\alpha)$ in place of the hard to compute $C_{op}(1-\alpha)$. Then base inferences on usual Wald statistics under the assumption of joint multivariate normality if $p_{Wald-onestep}$ is small. If $p_{Wald-onestep}$ is not small use the methods in Secs 4.3 and 5.1. Since $C_{Wald-onestep}(1-\alpha)$ differs from $C_{op}(1-\alpha)$ by at most $O\left(n^{-1/2}\right)$ the hope is that the qualitative sizes of $p_{Wald-onestep}$ and p_{op} will be the same even when p_{op} is large and thus quantitative inferences based on $C_{Wald-onestep}(1-\alpha)$ are inappropriate. Here is an example where I believe this strategy would save lots of effort.

Suppose we modify our toy model such that L_1 is continuous and real-valued with distribution that is absolutely continuous wrt Lesbegue measure and approximately uniform on $(-5, 5)$. Suppose the true but unknown values of $\psi_{21}^\dagger$ and $\psi_{11}^\dagger$ are 1 and $\psi_{31}^\dagger = 0$ is known to be 0. Then the unknown parameters of our drSNMM model are $\left(\psi_{11}^\dagger, \psi_{21}^\dagger, \psi_0^\dagger\right)$. Thus $(1, L_1, A_0)\,\psi_1^\dagger = 0$ if and only if $\psi_{11}^\dagger + \psi_{21}^\dagger L_1 = 0$. Now the event $L_1 = -\psi_{11}^\dagger/\psi_{21}^\dagger = -1$ has probability zero. Further the event that $\left|\psi_{11}^\dagger + \psi_{21}^\dagger L_1\right| < O\left(n^{-1/2}\right)$ has probability $O\left(n^{-1/2}\right)$. Since $\left(\widetilde{\psi}_1 - \psi_1^\dagger\right) = O\left(n^{-1/2}\right)$, $I\left[(1, l_1, a_0)\,\widetilde{\psi}_1 > 0\right]$ will differ from $I\left[(1, l_1, a_0)\,\psi_1^\dagger > 0\right]$ with probability $O\left(n^{-1/2}\right)$ and as $O\left(n^{-1/2}\right)$ is also the radius of $C_{Wald-onestep}(1-\alpha)$, we will find that $p_{Wald-onestep}$ is also $O\left(n^{-1/2}\right)$. Thus, with sample size n sufficiently large that the preceding calculations (which depended only on rates and not on constants) are valid approximations, Wald inferences centered on the one step estimator are valid

and our diagnostic $p_{Wald-onestep}$ will have revealed this, preventing the need for a more difficult analysis.

A1.2: Locally Degenerate Distribution of a Stochastic Process in Our Toy Example :

Our goal is to show that for $\left(\psi-\psi^{\dagger}\right)$ of $O\left(n^{-1/2}\right)$, $n^{1/2}P_n\left[U_0\left(\psi_0,s_0,\psi_1\right)\right]=$ $n^{1/2}P_n\left[U_0\left(\psi_0^{\dagger},s_0,\psi_1^{\dagger}\right)\right]+\kappa\left(\psi,\psi^{\dagger},F_O\right)+o_p\left(1\right)$ where the $o_p\left(1\right)$ is uniform in $F_{O.}$ and $\kappa\left(\psi,\psi^{\dagger},F_O\right)$ is non-random. First note that if $\left(1,L_1,A_0\right)\psi_1$ and $\left(1,L_1,A_0\right)\psi_1^{\dagger}$ are not both positive or both negative, then, by continuity, the function $\left(1,L_1,A_0\right)\psi_1$ of ψ_1 has a zero at some point, say $\psi_1^*\left(L_1,A_0,\psi_1,\psi_1^{\dagger}\right)$, on the line connecting ψ_1 and $\psi_1^{\dagger}$. Further by $\left(\psi-\psi^{\dagger}\right)$=$O\left(n^{-1/2}\right)$, $\psi-\psi_1^*\left(L_1,A_0,\psi_1,\psi_1^{\dagger}\right)$ and $\psi_1^*\left(L_1,A_0,\psi_1,\psi_1^{\dagger}\right)-\psi^{\dagger}$ are $O_p\left(n^{-1/2}\right)$ Now from A1.6 we have

$$
\begin{aligned}
&n^{1/2}P_n\left[U_0\left(\psi_0,s_0,\psi_1\right)\right]-n^{1/2}P_n\left[U_0\left(\psi_0^{\dagger},s_0,\psi_1^{\dagger}\right)\right]\\
&=n^{1/2}\left(\psi_0-\psi_0^{\dagger}\right)P_n\left[A_0\left\{s_0\left(A_0\right)-E[s_0\left(A_0\right)]\right\}\right]+\\
&n^{1/2}P_n\left[\Delta_0\left(s_0,\psi_1,\psi_1^{\dagger}\right)\right]
\end{aligned}
$$

But

$$
\begin{aligned}
&n^{1/2}P_n\left[\Delta_0\left(s_0,\psi_1,\psi_1^{\dagger}\right)\right]\\
&=-n^{1/2}P_n\left[\left\{s_0\left(A_0\right)-E[s_0\left(A_0\right)]\right\}A_1\left(1,L_1,A_0\right)\right]\left(\psi_1-\psi_1^{\dagger}\right)\\
&+n^{1/2}P_n\left[\Delta_0^*\left(s_0,\psi_1,\psi_1^{\dagger}\right)\right]
\end{aligned}
$$

where $n^{1/2}\Delta_0^*\left(s_0,\psi_1,\psi_1^{\dagger}\right)=$
$n^{1/2}\left[\left(I\left[\left(1,L_1,A_0\right)\psi_1>0\right]\right)\left(1,L_1,A_0\right)\psi_1-\left(I\left[\left(1,L_1,A_0\right)\psi_1^{\dagger}>0\right]\right)\left(1,L_1,A_0\right)\psi_1^{\dagger}\right]\}=$
$I\left[\left(1,L_1,A_0\right)\psi_1>0\right]I\left[\left(1,L_1,A_0\right)\psi_1^{\dagger}>0\right]\left(1,L_1,A_0\right)n^{1/2}\left\{\psi_1-\psi_1^{\dagger}\right\}+$
$I\left[\left(1,L_1,A_0\right)\psi_1>0\right]I\left[\left(1,L_1,A_0\right)\psi_1^{\dagger}\leq 0\right]\left(1,L_1,A_0\right)n^{1/2}\left\{\psi_1-\psi_1^*\left(L_1,A_0,\psi_1,\psi_1^{\dagger}\right)\right\}+$
$I\left[\left(1,L_1,A_0\right)\psi_1\leq 0\right]I\left[\left(1,L_1,A_0\right)\psi_1^{\dagger}>0\right]\left(1,L_1,A_0\right)n^{1/2}\left\{\psi_1^*\left(L_1,A_0,\psi_1,\psi_1^{\dagger}\right)-\psi_1^{\dagger}\right\}$
is $O_p\left(1\right)$.

Thus $P_n\left[n^{1/2}\Delta_0^*\left(s_0,\psi_1,\psi_1^{\dagger}\right)\right]=E\left[n^{1/2}\Delta_0^*\left(s_0,\psi_1,\psi_1^{\dagger}\right)\right]+o_p\left(1\right)$ by the uniform law of large numbers.

Thus we have proved the claimed result with $\kappa\left(\psi,\psi^{\dagger},F_O\right)=E\left[n^{1/2}\Delta_0^*\left(s_0,\psi_1,\psi_1^{\dagger}\right)\right]$
$+E\left[A_0\left\{s_0\left(A_0\right)-E[s_0\left(A_0\right)]\right\}\right]n^{1/2}\left(\psi_0-\psi_0^{\dagger}\right)-$
$E\left[\left\{s_0\left(A_0\right)-E[s_0\left(A_0\right)]\right\}A_1\left(1,L_1,A_0\right)\right]n^{1/2}\left(\psi_1-\psi_1^{\dagger}\right)$

Consider the special case where F_O is an exceptional law so $(1, L_1, A_0)\,\psi_1^\dagger = 0$. Assume as above, $(1, L_1, A_0)\,\psi_1^\dagger = 0 \Leftrightarrow L_1 = A_0 = 1$. Then $E\left[n^{1/2}\Delta_0^*\left(s_0, \psi_1, \psi_1^\dagger\right)\right] =$

$$E\left\{\begin{array}{c} I\left[\sum_j \psi_{j1} > 0\right] E\left[L_1 A_0\,(1, L_1, A_0)\right] + \\ E\left[\{\,1 - L_1 A_0\}\, I\left[\,(\,1, L_1, A_0)\;\;\psi_1^\dagger > 0\right](\,1, L_1, A_0)\right] \end{array}\right\} n^{1/2}\left\{\psi_1 - \psi_1^\dagger\right\}.$$

Thus $\partial E\left[U_0\left(\psi_0, s_0, \psi_1\right)\right]/\partial \psi_1^T = DER_{01}\left(\psi\right) =$

$$E\left\{\begin{array}{c} I\left[\sum_j \psi_{j1} > 0\right] E\left[L_1 A_0\,(1, L_1, A_0)\right] + \\ E\left[\{\,1 - L_1 A_0\}\, I\left[\,(\,1, L_1, A_0)\;\;\psi_1^\dagger > 0\right](\,1, L_1, A_0)\right] \end{array}\right\}$$

$$-E\left[\{s_0\left(A_0\right) - E[s_0\left(A_0\right)]\}\, A_1\left(1, L_1, A_0\right)\right]$$

which converges to different limits depending on whether $\sum_j \psi_{j1}$ decreases from above or increases from below to $\sum_j \psi_{j1}^\dagger = 0$ so $DER_{01}\left(\psi^\dagger\right)$ is undefined. Note $\partial E\left[U_0\left(\psi_0, s_0, \widetilde{\psi}_1\right)\right]/\partial \psi_1^T$ will have variance $O\left(1\right)$ since $\sum_j \widetilde{\psi}_{j1}$ takes on both positive and negative values with positive probability as $n \to \infty$.

Even when the data are not generated under an exceptional law so $DER_{01}\left(\psi^\dagger\right)$ exists, the convergence of $DER_{01}\left(\psi\right)$ to $DER_{01}\left(\psi^\dagger\right)$ is non-uniform since, given any sequence $kn^{-1/2}$, there exists at each sample size n, ψ and $\psi^\dagger$ with $\sum_j \psi_{j1}^\dagger$ sufficiently close to 0 and $||\psi - \psi^\dagger|| = kn^{-1/2}$, such that $DER_{01}\left(\psi\right) - DER_{01}\left(\psi^\dagger\right) = O\left(1\right)$ because $\sum_j \psi_{j1}$ and $\sum_j \psi_{j1}^\dagger$ have different signs. However the Lebesgue measure of the set of $\psi^\dagger$ that has this property will decrease as n increases so that posterior associated with a smooth prior (that does not change with sample size) will be quadratic for n large.

11 Appendix 2:

Proof of Lemma 5.1: For some $\psi_2^* \in \left[\psi_2, \widehat{\psi}_2\right]$

$$\begin{aligned} & n^{1/2} P_n\left[U_a\left(\psi_1, \widehat{\psi}_2\right)\right] \\ & = n^{1/2} P_n\left[U_a\left(\psi_1, \psi_2\right)\right] + n^{1/2} P_n\left[\partial U_a\left(\psi_1, \psi_2\right)/\partial \psi_2\right]\left(\widehat{\psi}_2 - \psi_2\right) + \\ & n^{1/2} P_n\left[\partial^2 U_a\left(\psi_1, \psi_2^*\right)/\partial^2 \psi_2\right]\left(\widehat{\psi}_2 - \psi_2\right)^2 \\ & = n^{1/2} P_n\left[U_a\left(\psi_1, \psi_2\right)\right] + n^{1/2} P_n\left[\partial U_a\left(\psi_1, \psi_2\right)/\partial \psi_2\right]\left(\widehat{\psi}_2 - \psi_2\right) + o_p\left(1\right) \end{aligned}$$

Similiarly

$$\begin{aligned} n^{1/2} P_n\left[U_b\left(\psi_1, \widehat{\psi}_2\right)\right] = {} & n^{1/2} P_n\left[U_b\left(\psi_1, \psi_2\right)\right] + \\ & n^{1/2} P_n\left[\partial U_b\left(\psi_1, \psi_2\right)/\partial \psi_2\right]\left(\widehat{\psi}_2 - \psi_2\right) + o_p\left(1\right) \end{aligned}$$

Also

$$P_n\left[\partial U_a\left(\psi_1,\widehat{\psi}_2\right)/\partial\psi_2\right]\left\{P_n\left[\partial U_b\left(\psi_1,\widehat{\psi}_2\right)/\partial\psi_2\right]\right\}^{-1}n^{1/2}P_n\left[U_b\left(\psi_1,\psi_2\right)\right]$$
$$=P_n\left[\partial U_a\left(\psi_1,\psi_2\right)/\partial\psi_2\right]\left\{P_n\left[\partial U_b\left(\psi_1,\psi_2\right)/\partial\psi_2\right]\right\}^{-1}n^{1/2}P_n\left[U_b\left(\psi_1,\psi_2\right)\right]+o_p\left(1\right)$$

Hence,

$$\begin{aligned}
&n^{1/2}P_n\left[U_1\left(\psi_1,\widehat{\psi}_2\right)\right]-n^{1/2}P_n\left[U_1\left(\psi_1,\psi_2\right)\right]\\
&=n^{1/2}P_n\left[\partial U_a\left(\psi_1,\psi_2\right)/\partial\psi_2\right]\left(\widehat{\psi}_2-\psi_2\right)-\\
&P_n\left[\partial U_a\left(\psi_1,\widehat{\psi}_2\right)/\partial\psi_2\right]\left\{P_n\left[\partial U_b\left(\psi_1,\widehat{\psi}_2\right)/\partial\psi_2\right]\right\}^{-1}n^{1/2}\times\\
&P_n\left[\partial U_b\left(\psi_1,\psi_2\right)/\partial\psi_2\right]\left(\widehat{\psi}_2-\psi_2\right)+o_p\left(1\right)\\
&=n^{1/2}P_n\left[\partial U_a\left(\psi_1,\psi_2\right)/\partial\psi_2\right]\left(\widehat{\psi}_2-\psi_2\right)\\
&-P_n\left[\partial U_a\left(\psi_1,\psi_2\right)/\partial\psi_2\right]\left\{P_n\left[\partial U_b\left(\psi_1,\psi_2\right)/\partial\psi_2\right]\right\}^{-1}n^{1/2}\times\\
&P_n\left[\partial U_b\left(\psi_1,\psi_2\right)/\partial\psi_2\right]\left(\widehat{\psi}_2-\psi_2\right)+o_p\left(1\right)\\
&=o_p\left(1\right)
\end{aligned}$$

12 Appendix 3:

Proof of Theorem 7.1: $\Leftarrow$ By induction in reverse time order.

Case 1: $m=K$. Let $a_K=d_K\left(\overline{L}_K,\overline{A}_{K-1}\right)$ .Then $E\left[Y_{\overline{A}_{K-1},a_K}|\overline{L}_K,\overline{A}_{K-1}\right]=$
$E\left[Y_{\overline{A}_{K-1},a_K}|\overline{L}_K,\overline{A}_{K-1},A_K=a_K\right]f\left(a_K|\overline{L}_K,\overline{A}_{K-1}\right)+$
$\left\{1-f\left(a_K|\overline{L}_K,\overline{A}_{K-1}\right)\right\}E\left[Y_{\overline{A}_{K-1},a_K}|\overline{L}_K,\overline{A}_{K-1},A_k\neq a_K\right]$
$=\left\{E\left[Y_{\overline{A}_{K-1},d_K^*}|\overline{L}_K,\overline{A}_{K-1},A_K=a_K\right]+\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_K,\overline{A}_{K-1},a_K\right)\right\}\times$
$f\left(a_K|\overline{L}_K,\overline{A}_{K-1}\right)+\left\{1-f\left(a_K|\overline{L}_K,\overline{A}_{K-1}\right)\right\}\times$
$\left\{E\left[Y_{\overline{A}_{K-1},d_K^*}|\overline{L}_K,\overline{A}_{K-1},A_K\neq a_K\right]+\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_K,\overline{A}_{K-1},a_K\right)-\ r^{\underline{d}_K,d_K^*}\left(\overline{L}_K,\overline{A}_{K-1}\right)\right\}$
$=E\left[Y_{\overline{A}_{K-1},d_K^*}|\overline{L}_K,\overline{A}_{K-1}\right]+\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_K,\overline{A}_{K-1},a_K\right)-$
$\left\{1-f\left(a_K|\overline{L}_K,\overline{A}_{K-1}\right)\right\}\ r^{\underline{d}_K,d_K^*}\left(\overline{L}_K,\overline{A}_{K-1}\right)$
$=E\left\{E\left[Y_{\overline{A}_{K-1},A_K}-\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_K,\overline{A}_K\right)|\overline{L}_K,\overline{A}_{K-1},A_K\right]|\overline{L}_K,\overline{A}_{K-1}\right\}-$
$\left\{1-f\left(a_K|\overline{L}_K,\overline{A}_{K-1}\right)\right\}\ r^{\underline{d}_K,d_K^*}\left(\overline{L}_K,\overline{A}_{K-1}\right)=E\left[H_K\left(\gamma^{\overline{d},\overline{d}^*}\right)|\overline{L}_K,\overline{A}_{K-1}\right]$

where the first equality is by the law of total probability, the second by the definition of $\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_K,\overline{A}_{K-1},a_K\right)$ and $r^{\underline{d}_m,d_m^*}\left(\overline{L}_K,\overline{A}_{K-1}\right)$, the third by the

law of total probability, the fourth by the definition of $\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_K,\overline{A}_K\right)$, and the 5th by the definition of $H_K\left(\gamma^{\overline{d},\overline{d}^*}\right)$.

Case 2: $m<K$. Let $a_m=d_m\left(\overline{L}_m,\overline{A}_{m-1}\right)$.

Then $E\left[Y_{\overline{A}_{m-1},a_m,\underline{d}_{m+1}}|\overline{L}_m,\overline{A}_{m-1}\right]$

$=E\left[Y_{\overline{A}_{m-1},a_m,\underline{d}_{m+1}}|\overline{L}_m,\overline{A}_{m-1},A_m=a_m\right]f\left(a_m|\overline{L}_m,\overline{A}_{m-1}\right)$

$+\left\{1-f\left(a_m|\overline{L}_m,\overline{A}_{m-1}\right)\right\}E\left[Y_{\overline{A}_{m-1},a_m,\underline{d}_{m+1}}|\overline{L}_m,\overline{A}_{m-1},A_m\neq a_m\right]$

$=\left\{E\left[Y_{\overline{A}_{m-1},d_m^*,\underline{d}_{m+1}}|\overline{L}_m,\overline{A}_{m-1},A_m=a_m\right]+\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_{m-1},a_m\right)\right\}\times$

$f\left(a_m|\overline{L}_m,\overline{A}_{m-1}\right)+\left\{1-f\left(a_m|\overline{L}_m,\overline{A}_{m-1}\right)\right\}\times$

$\left\{E\left[Y_{\overline{A}_{m-1},d_m^*,\underline{d}_{m+1}}|\overline{L}_m,\overline{A}_{m-1},A_m\neq a_m\right]+\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_{m-1},a_m\right)\right.$

$\left.-\ r^{\underline{d}_m,d_m^*}\left(\overline{L}_m,\overline{A}_{m-1}\right)\right\}$

$=E\left[Y_{\overline{A}_{m-1},d_m^*,\underline{d}_{m+1}}|\overline{L}_m,\overline{A}_{m-1}\right]+\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_{m-1},a_m\right)-$

$\left\{1-f\left(a_m|\overline{L}_m,\overline{A}_{m-1}\right)\right\}\ r^{\underline{d}_m,d_m^*}\left(\overline{L}_m,\overline{A}_{m-1}\right)=$

$$E\left\{E\left[Y_{\overline{A}_{m-1},A_m,\underline{d}_{m+1}}-\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)|\overline{L}_m,\overline{A}_{m-1},A_m\right]|\overline{L}_m,\overline{A}_{m-1}\right\} \tag{A3.1}$$

$$-\left\{1-f\left(a_m|\overline{L}_m,\overline{A}_{m-1}\right)\right\}\ r^{\underline{d}_m,d_m^*}\left(\overline{L}_m,\overline{A}_{m-1}\right)$$

where the first equality is by the law of total probability, the second by the definition of $\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_{m-1},a_m\right)$ and $r^{\underline{d}_m,d_m^*}\left(\overline{L}_m,\overline{A}_{m-1}\right)$, and the third by the law of total probability, the fourth by the definition of $\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$.

But $E\left[Y_{\overline{A}_{m-1},A_m,\underline{d}_{m+1}}-\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)|\overline{L}_m,\overline{A}_{m-1},A_m\right]$

$=\int E\left[Y_{\overline{A}_{m-1},A_m,\underline{d}_{m+1}}|\overline{L}_{m+1},\overline{A}_{m-1},A_m\right]dF\left(L_{m+1}|\overline{L}_m,\overline{A}_m\right)-\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$

$=\int E\left[H_{m+1}^{\underline{d}_{m+1}}\left(\gamma^{\overline{d},\overline{d}^*}\right)|\overline{L}_{m+1},\overline{A}_m\right]dF\left(L_{m+1}|\overline{L}_m,\overline{A}_{m,}\right)-\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$

$=E\left[H_{m+1}^{\underline{d}_{m+1}}\left(\gamma^{\overline{d},\overline{d}^*}\right)|\overline{L}_m,\overline{A}_m\right]-\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$ where the first equality is by the law of total probability, the second by the induction hypothesis, and the last by the definition of $H_{m+1}^{\underline{d}_{m+1}}\left(\gamma^{\overline{d},\overline{d}^*}\right)$. We complete the proof by plugging this last expression back into (A3.1)

$\Rightarrow$ By contradiction. Let m^* be the largest value of m such that $\Delta\left(\overline{L}_m,\overline{A}_m\right)\equiv\gamma^{\overline{d}^**}\left(\overline{L}_m,\overline{A}_m\right)-\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_m,\overline{A}_m\right)$ is a function of A_m with positive probability. By the assumption that $\gamma^{\overline{d}^**}\left(\overline{L}_m,\overline{A}_m\right)=0$ if $A_m=0$, we are guaranteed that $m^*\geq 0$. It follows that $0=E\left[H_{m^*}^{\underline{d}_{m^*}}\left(\gamma^{\overline{d},\overline{d}^*}\right)|\overline{L}_{m^*},\overline{A}_{m^*}\right]$ but $E\left[H_{m^*}^{\underline{d}_{m^*}}\left(\gamma^{\overline{d}^**}\right)|\overline{L}_{m^*},\overline{A}_{m^*}\right]=\gamma^{\overline{d}^**}\left(\overline{L}_{m^*},\overline{A}_{m^*}\right)-\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_{m^*},\overline{A}_{m^*}\right)\neq 0$ w.p1.

Corollary A3.1: $E\left[H_m^{\underline{d}_m}\left(\gamma^{\overline{d},\overline{d}^*}\right)-\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_{m-1},\overline{A}_{m-1}\right)|\overline{L}_{m-1},\overline{A}_{m-1}\right]=$
$E\left[Y_{\overline{A}_{m-2},d_m^*,\underline{d}_m}|\overline{L}_{m-1},\overline{A}_{m-1}\right]$.

Proof: $E\left[H_m^{\underline{d}_m}\left(\gamma^{\overline{d},\overline{d}^*}\right) - \gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_{m-1},\overline{A}_{m-1}\right)|\overline{L}_{m-1},\overline{A}_{m-1}\right]$
$= E\left[E\left[Y_{\overline{A}_{m-1},\underline{d}_m}|\overline{L}_m,\overline{A}_{m-1}\right] - \gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_{m-1},\overline{A}_{m-1}\right)|\overline{L}_{m-1},\overline{A}_{m-1}\right]$
$= E\left[Y_{\overline{A}_{m-1},\underline{d}_m} - \gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_{m-1},\overline{A}_{m-1}\right)|\overline{L}_{m-1},\overline{A}_{m-1}\right]$
$= E\left[Y_{\overline{A}_{m-2},d_m^*,\underline{d}_m}|\overline{L}_{m-1},\overline{A}_{m-1}\right]$

where the 1st equality is by theorem 7.1, the second by the law of total probability, and the third by the definition of $\gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_{m-1},\overline{A}_{m-1}\right)$.

Proof of Theorem 7.2:
$\Leftarrow E\left[H_m^{\underline{d}_m}\left(\gamma^{\overline{d},\overline{d}^*}\right) - \gamma^{\overline{d},\overline{d}^*}\left(\overline{L}_{m-1},\overline{A}_{m-1}\right) - q^{\overline{d},\overline{d}^*}\left(\overline{L}_{m-1},\overline{A}_{m-1}\right)|\overline{L}_{m-1},\overline{A}_{m-1}\right]$
$= E\left[Y_{\overline{A}_{m-2},d_m^*,\underline{d}_m} - q^{\overline{d},\overline{d}^*}\left(\overline{L}_{m-1},\overline{A}_{m-1}\right)|\overline{L}_{m-1},\overline{A}_{m-1}\right] =$
$E\left[Y_{\overline{A}_{m-2},d_m^*,\underline{d}_m}|\overline{L}_{m-1},\overline{A}_{m-2},A_{m-1}=0\right]$ where the 1st equality is by Corollary A.1 and the second by the definition of $q^{\overline{d},\overline{d}^*}\left(\overline{L}_{m-1},\overline{A}_{m-1}\right)$.

$\Rightarrow$trivial

Proof of Theorem 7.4:$\Leftarrow$By induction

Case 1: $m = K$: Trivial since $Y_{\overline{A}_{m-1},a_m,\underline{d}_{m+1}}$ does not depend on $\underline{d}_{m+1}$.

Case 2: Assume it is true for $m+1$ and we shall prove it for m. By the induction hypothesis for $\underline{d}_{m+1} \in \underline{\mathcal{D}}_{m+1}$

$$E\left[Y_{\overline{A}_{m-1},a_m,\underline{d}_{op,m+1}} - Y_{\overline{A}_{m-1},a_m,\underline{d}_{m+1}}|\overline{L}_{m+1},\overline{A}_{m-1},A_m=a_m\right] \geq 0$$

Thus $E\left[Y_{\overline{A}_{m-1},a_m,\underline{d}_{op,m+1}} - Y_{\overline{A}_{m-1},a_m,\underline{d}_{m+1}}|\overline{L}_m,\overline{A}_{m-1},A_m=a_m\right] \geq 0$ after integrating over L_{m+1}. By the backward induction feasibility assumption this implies $E\left[Y_{\overline{A}_{m-1},a_m,\underline{d}_{op,m+1}} - Y_{\overline{A}_{m-1},a_m,\underline{d}_{m+1}}|\overline{L}_m,\overline{A}_{m-1}\right] \geq 0$. Hence since $d_{op,m}\left(\overline{L}_m,\overline{A}_{m-1}\right) = \arg\max_{a_m\in\mathcal{A}_m} E\left[Y_{\overline{A}_{m-1},a_m,\underline{d}_{op,m+1}}|\overline{L}_m,\overline{A}_{m-1}\right]$, we conclude that $\underline{d}_{op,m}$ maximizes $E\left[Y_{\overline{A}_{m-1},\underline{d}_m}|\overline{L}_m,\overline{A}_{m-1}\right]$ over all regimes $\underline{d}_m \in \underline{\mathcal{D}}_m$. The proof in the other direction is trivial.

Bibliography:

Baraud, Y. (2002). Confidence balls in Gaussian regression, (to appear).

Bertsekas, D.P. and Tsitsiklis, J.N. (1996). *Neuro-dynamic programming.* Belmont MA: Athena Scientific.

Bickel, P.J. and Ritov, Y. (1988). Estimating integrated squared density derivatives: sharp best order of convergence estimates. *Sankya Ser.*A 50: 381-393.

Bickel, P.J., Klaassen, C., Ritov, Y., and Wellner, J. (1993). Efficient and adapted estimation for semiparametric models. Johns Hopkins, Baltimore.

Cowell, R.G., Dawid, A.P., Lauritzen, S.L, and Spiegelhalter, D.J. (1999). Probabilistic networks in expert systems, New York: Springer-Verlag.

Donald, S.G. and Newey, W.K. (1994). Series estimation of the semilinear models. Journal of Multivariate Analysis, 5): 30-40.

Gill, R.D. and Robins, J.M. (2001). Causal inference for complex longitudinal data: the continuous case. *Annals of Statistics*, 29(6): 1785-1811

Hoffman, M. and Lepski, O. (2002). Random rates and anisotropic regression (with discussion and rejoinder). Annals of Statistics, 30: 325-396.

Li, KC. (1989). Honest Confidence Regions for Nonparametric Regression. The Annals of Statistics, 17(3):1001-1008.

Laurent, B. (1996). Efficient estimation of integral functionals of a density. Annals of Statistics, 24(2): 659 -681.

Laurent, B. and Massart, P. (2000). Adaptive estimation of a quadratic functional by model selection. *Annals of Statistics*, 28(5): 1302-1338.

Murphy, Susan. (2003). Optimal dynamic treatment regimes. *Journal of the Royal Statistical Society B,* 65(2):331-355.

Ritov, Y. and Bickel, P. (1990) Achieving information bounds in non- and semi- parametric models. Annals of Statistics, 18: 925-938.

Robins, J.M. (1986). A new approach to causal inference in mortality studies with sustained exposure periods - Application to control of the healthy worker survivor effect. *Mathematical Modelling,* 7:1393-1512

Robins, J.M. (1994). Correcting for non-compliance in randomized trials using structural nested mean models. C*ommunications in Statistics,* 23:2379-2412.

Robins, J.M. (1997). Causal Inference from Complex Longitudinal Data. Latent Variable Modeling and Applications to Causality. *Lecture Notes in Statistics* (120), M. Berkane, Editor. NY: Springer Verlag, pp. 69-117.

Robins, J.M. (1998a). Correction for non-compliance in equivalence trials. *Statistics in Medicine,* 17:269-302.

Robins, J.M., (1998b) Structural nested failure time models. **Survival Analysis,** P.K. Anderson and N. Keiding, Section Editors. *The Encyclopedia of Biostatistics***.** P. Armitage and T. Colton, Editors. Chichester, UK: John Wiley & Sons. pp 4372-4389.

Robins, J.M. (1999). Marginal Structural Models versus Structural Nested Models as Tools for Causal Inference. *Statistical Models in Epidemiology: The Environment and Clinical Trials.* M.E. Halloran and D. Berry, Editors, IMA Volume 116, NY: Springer-Verlag, pp. 95-134.

Robins, J.M. (2000). Robust estimation in sequentially ignorable missing data and causal inference models. *Proceedings of the American Statistical Association Section on Bayesian Statistical Science* 1999, pp. 6-10.

Robins, J.M., Greenland, S. and Hu F-C. (1999). Rejoinder to Comments on "Estimation of the causal effect of a time-varying exposure on the marginal mean of a repeated binary outcome." *Journal of the American Statistical Association, Applications and Case Studies,* 94:708-712.

Robins, J.M. and Ritov, Y. (1997). Toward a curse of dimensionality appropriate (CODA) asymptotic theory for semi-parametric models. *Statistics in Medicine,* 16:285-319.

Robins J.M. and Rotnitzky A. (2001). Comment on the Bickel and Kwon article, "Inference for semiparametric models: Some questions and an answer" *Statistica Sinica,* 11(4):920-936. ["On Double Robustness."]

Robins, J.M. and Rotnitzky, A. (2003). Direct effects structural nested mean models. *Annals of Statistics,* (under review).

Robins, J.M., Rotnitzky, A. and Scharfstein, D. (1999a). Sensitivity Analysis for Selection Bias and Unmeasured Confounding in Missing Data and Causal Inference Models. In: *Statistical Models in Epidemiology: The Environment and Clinical Trials*. Halloran, E. and Berry, D., eds. IMA Volume 116, NY: Springer-Verlag, pp. 1-92.

Robins J.M., Rotnitzky A., van der Laan M. (2000). Comment on "On Profile Likelihood" by Murphy SA and van der Vaart AW. *Journal of the American Statistical Association – Theory and Methods,* 95(450):431-435.

Robins, J.M., Scheines, R., Spirtes, P., and Wasserman, L.(2003).Uniform consistency in causal inference. Biometrika, 90(3):491-515.

Robins, J.M. and Wasserman L. (1997). Estimation of Effects of Sequential Treatments by Reparameterizing Directed Acyclic Graphs. Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence, Providence Rhode Island, August 1-3, 1997. Dan Geiger and Prakash Shenoy (Eds.), Morgan Kaufmann, San Francisco, pp. 409-420.

Robins, J.M. and van der Vaart, A.W. (2003). Non parametric confidence sets by cross-validation. (Technical Report).

Robins, J.M. and van der Vaart, A.W. (2004). A unified approach to estimation in non-semiparametric models using higher order influence functions. (Technical Report)

Small, C.G. and McLeish, D. (1994). Hilbert space methods in probability and statistical inference. New York: Wiley.

Sutton, R.S. and Barto, A.G. (1998). Reinforcement learning: An introduction. Cambridge, MA: MIT Press.

van der Laan, M. and Dudoit (2003) Asymptotics of cross-validated risk estimation in model selection and performance assessment, revised for publication in Annals of Statistics.

van der Laan, M.J., Murphy, S., and Robins, J.M. (2003). Marginal structural nested models. (to be submitted).

van der Laan M.J., Robins JM (1998). Locally efficient estimation with current status data and time-dependent covariates. *Journal of the American Statistical Association*, 93:693-701.

van der Laan, M. and Robins, J.M. (2002). *Unified methods for censored longitudinal data and causality.* Springer-Verlag.

van der Vaart, A.W. (1991). On differentiable functionals. Annals of Statistics, 19:178-204.

Waterman, R.P. and Lindsay, B.G. (1996). Projected score methods for approximating conditional scores. *Biometrika*, 83(1): 1-13.

Wegkamp, M. (2003) Model selection in nonparametric regression. *Annals of Statistics*, 31(1):252-273.

Acknowledgements: I would like to thank Thomas Richardson for his very careful reading of the manuscript and many helpful suggestions, Andrea Rotnitzky for useful discussions, and Aad van der Vaart for his codevelopment of an important part of the statistical methodology used in this paper. This methodology is described in detail in our joint papers. This work was supported in part by NIH grant A1-32475.