Assessing Time-Varying Causal Effect Moderation in the Presence of Cluster-Level Treatment Effect Heterogeneity

Jieru Shi, Zhenke Wu, Walter Dempsey

January 29, 2021

Abstract

Micro-randomized trial (MRT) is a sequential randomized experimental design to empirically evaluate the effectiveness of mobile health (mHealth) intervention components that may be delivered at hundreds or thousands of decision points. The MRT context has motivated a new class of causal estimands, termed "causal excursion effects", for which inference can be made by a weighted, centered least squares approach (Boruvka et al., 2017). Existing methods assume between-subject independence and non-interference. Deviations from these assumptions often occur which, if unaccounted for, may result in bias and overconfident variance estimates. In this paper, causal excursion effects are considered under potential cluster-level correlation and interference and when the treatment effect of interest depends on cluster-level moderators. The utility of our proposed methods is shown by analyzing data from a multi-institution cohort of first year medical residents in the United States. The approach paves the way for construction of mHealth interventions that account for observed social network information.

1 Introduction

In many areas of behavioral science, push notifications in mobile apps that are adapted to continuously collected information on an individual's current context have received a considerable amount of attention. These time-varying adaptive interventions are hypothesized to lead to meaningful short- and long-term behavior change. The assessment of the time-varying effect of such push notifications motivated sequential randomized designs such as the micro-randomized trial (MRT) [1, 2]. In an MRT, interventions are randomized to be potentially delivered at a over a period of time, at hundreds or thousands of decision points. The MRT design enables the estimation of the time-delayed marginal treatment effects of a push notification upon the outcome of interest at multiple time-lags, referred to as "causal excursion effects" [3, 4]. A weighted, centered least squares (WCLS) approach has been developed to consistently estimate causal excursion effects with robust methods of uncertainty assessment. Extensions to balance the number of randomizations by observed or derived individual-level contextual information have also been developed, e.g., a stratified MRT [5].

Existing inferential methods for data collected in an MRT rely on two key assumptions. First, an intervention delivered to an individual is assumed to only impact the outcome of that individual, i.e., non-interference. Second, the method assumes no stochastic dependence among outcomes of different subjects. Deviations from these assumptions, however, may occur when individuals naturally form clusters. Consider, for example, an MRT in which medical interns are recruited from several hospitals and every day potentially receive push notifications designed to improve physical activity, mood, or sleep. First, *interference* may occur when an intern receiving a push notification to increase their steps may then encourage another intern in the same specialty and institution to do the same hence altering the others'

outcomes. Second, even if there is no direct interference, interns in the same specialty and institution are likely to have correlated outcomes, induced by variation in unobserved factors shared within each cluster. For example, interns within a specialty and institution may share similar experiences during their residency which will lead to a higher correlation in their weekly mood scores than interns in different specializations or institutions.

The primary contribution of this paper is the significant extension of data analytic methods for MRTs to account for potential cluster-level correlation and the extension of the causal excursion effect to account for potential interference, i.e., we define both a direct and indirect causal excursion effect. In Section 2, we review the MRT design and the existing method's implicit assumptions. In Section 3, we introduce the moderated causal excursion effect using a group-based conceptualization of potential outcomes [6, 7]. We then generalize WCLS criterion to a cluster-based WCLS (C-WCLS) criterion to estimate both direct and indirect causal excursion effects in Section 4. Finally, we demonstrate that when proximal outcomes exhibit within-cluster correlation but there is an absence of between-cluster treatment effect heterogeneity, our proposal agrees with the traditional approach both in terms of the point estimate and the asymptotic variance. In Section 6, we illustrate the proposed method by estimating the causal excursion effects of push notifications on mood, sleep, and step count for a cohort of first year medical residents in the United States. The paper concludes with a brief discussion.

2 Preliminaries

We first review the design of MRTs and the existing inferential methods based on MRT data. We focus on the assumptions made in the inferential methods to set the stage for our proposed extension.

2.1 Micro-Randomized Trials (MRT)

An MRT consists of a sequence of within-subject decision times $t = 1, \ldots, T$ at which treatment options may be randomly assigned. The resulting longitudinal data for any subject can be summarized as

$$\{O_0, O_1, A_1, O_2, A_2, \ldots, O_T, A_T, O_{T+1}\}$$

where t indexes a sequence of decision points, O_0 is the baseline information, O_t is the information collected between time $t-1$ and t which may include mobile sensor and wearable data, and A_t is the treatment option provided at time t. In an MRT, A_t results from randomization at each decision point; here, we assume this randomization probability depends on the complete observed history $H_t := \{O_0, O_1, A_1, \ldots, A_{t-1}, O_t\}$ and is denoted $p_t(A_t \mid H_t)$. Treatment options are designed to impact a proximal response Y_{t+1} which is a known function of $\{O_t, A_t, O_{t+1}\}$. In this section, we will use HeartSteps [8], an MRT designed to increase physical activity among sedentary adults, as our illustrative example. In HeartSteps [8], $A_t = 1$ denotes sending a tailored activity message designed to increase proximal physical activity and therefore Y_{t+1} is chosen to be total step count in the 30-minutes after the decision time.

Remark 2.1. At some decision points, it is inappropriate for scientific, ethical, or burden reasons to provide treatment. To address this, the observation vector O_t also contains an indicator of availability: $I_t = 1$ if available for treatment and $I_t = 0$ otherwise. Availability at time t is determined before treatment randomization. When an individual is unavailable for treatment ($I_t = 0$), then no treatment is provided ($A_t = 0$). In this paper, for simplicity, individuals are assumed always available for treatment. All methods can be readily extended to account for availability as in **(author?)** [3].

2.2 Estimand and Inferential Method

Following [3], [4], and [5], we focus on the class of estimands that is referred to as "causal excursion effects" defined by

$$\beta(t; s) = \mathbb{E}\left[\mathbb{E}\left[Y_{t+1} \mid H_t, A_t = 1\right] - \mathbb{E}\left[Y_{t+1} \mid H_t, A_t = 0\right] \mid S_t = s\right]$$

where S_t is a time-varying potential effect moderator (or "state") and a deterministic function of the observed history H_t. For example, in the HeartSteps study, S_t may represent subject-specific contextual information such as step count in the previous 30 minutes. See Section 3.2 for how this estimand is derived using potential outcomes. Under the assumption that $\beta(t; s) = f_t(s)^\top \beta^\star$ where $f_t(s) \in \mathbb{R}^q$ is a feature vector depending only on state s and decision point t, a consistent estimator $\hat{\beta}$ can be obtained by minimizing the following weighted and centered least squares (WCLS) criterion:

$$\arg\max_{\alpha,\beta} \mathbb{P}_n \left[\sum_{t=1}^{T} W_t \left(Y_{t+1} - g_t(H_t)'\alpha - (A_t - \tilde{p}_t(1 \mid S_t)) f_t(S_t)^\top \beta \right) \right] \tag{1}$$

where $\mathbb{P}_n$ is defined as the average of a function over the sample, $W_t = \tilde{p}_t(A_t \mid S_t)/p_t(A_t \mid H_t)$ is a weight where the numerator is an arbitrary function with range $(0, 1)$ that only depends on S_t and the denominator is the MRT specified randomization probability, and $g_t(H_t) \in \mathbb{R}^p$ are a set of control variables chosen to help reduce the variance of the estimators and to construct more powerful test statistics. In HeartSteps, for example, a natural control variable is the number of steps in the prior 30 minutes as this variable is likely highly correlated with the proximal response and thus can be used to reduce variance in the estimation of β. See [3] for additional details on how to define and estimate moderated time-varying treatment effects using data from an MRT.

3 Cluster-Level Proximal Treatment Effects

MRTs are implicitly defined with the individual as the unit of analysis. The associated analytics rely on statistical independence among individuals in deriving asymptotic normality; moreover, the treatment effects assume that treatments provided to one individual do not impact another individual. In Section 5, we show that the nominal coverage probability of the 95% confidence interval obtained from these methods decays quickly when the MRT is conducted among individuals with cluster-level treatment effect heterogeneity. Moreover, in such a setting, an interesting scientific question concerns moderation analysis where the moderator is defined at the cluster-level, i.e., average values of a moderator such as mood or step-count across a cluster. Here we consider a cluster-based definition of causal excursion effect and corresponding analyses to address these issues. We start with our motivating example that demonstrates the necessity for considering clustering in the context of MRTs.

3.1 Motivating Example

Intern Health Study (IHS) is a 6-month micro-randomized trial on 1,565 medical interns [9]. Due to high depression rates and levels of stress during the first year of physician residency training, a critical question is whether targeted notifications can be used at the right time to improve mood, increase sleep time, and/or increase physical activity. Enrolled medical interns were randomized weekly to receive either mood, activity, or sleep notifications or receive no notifications for that week (probability 1/4 each). On a week in which notifications are provided, interns were randomized daily to receive a notification with probability 50%. Hence for a notification week, the user received, on average, 3.5 notifications. The exploratory and MRT analyses conducted in this paper focus on weekly randomization; See [9] for further study details.

Medical internships occur in a specific location and within a specific specialty. The 1,565 interns in the MRT represented 321 different residency locations and 42 specialties. Thus,

6

interns are naturally clustered. Table 3.1 presents a standard equally-weighted analysis of variance (ANOVA) applied to the average weekly mood scores with three factors: week-in-study, residency and specialty both interacted with weekly treatment (Txt). Compared with the residual mean square of 1.27, there is considerable excess variation associated with week-in-study (28.44), residency by treatment (5.89) and specialty by treatment (8.41). In other words, the ANOVA suggests proximal responses exhibit potential within-cluster correlation within residency locations and specialties that varies with treatment. In this paper, we focus on defining clusters using specialty and institution information and do not discuss multi-level clusters.

Source	d.f.	Sum Sq.	Mean Sq.
week-in-study	25	711	28.44
Residency:Txt	1091	6424	5.89
Specialty:Txt	48	404	8.41
Residuals	38946	49479	1.27

Table 1: ANOVA decomposition for average weekly mood scores in the IHS.

Three immediate questions may arise:

(1) Does the standard MRT analysis still have good statistical properties, e.g., consistency and valid nominal coverage probability of 95% confidence intervals when participants exhibit within-cluster correlation?

(2) Can one assess potential cluster-level effect moderators using the standard MRT analysis? and,

(3) Can one extend the MRT analysis to assess the effect of treatment provided to one individual on the proximal response of another individual in the same cluster?

For (1), through a simulation study in Section 5, we demonstrate potential deficiency of the standard WCLS method that results in severe under-coverage of nominal 95% confidence intervals. In particular, we use simulation settings that mimic realistic cluster-level treatment heterogeneity. For (2), in IHS, for example, a scientific question may be how group

7

average mood impacts the proximal effect of notifications on average daily mood. Since each individual's mood may be impacted by their own treatments, if the proximal effect is moderated by cluster-level summary variables then this implies the proximal response is a function of past treatment assignment to all individuals in the group. Finally, for (3), in Section 3.2 we extend the causal excursion effect to distinguish between direct and indirect causal excursion effects.

In this paper, we extend WCLS method [3] to account for clustered continuous outcomes and potential within-cluster interference of other subjects' treatments upon one subject's outcome, offering a general framework for answering the three questions above. Here we study micro-randomized trials in which there is a pre-defined cluster-level structure. The cluster structure may result in (A) within-cluster correlation among proximal responses (question 1), (B) treatment impacting individual-level moderators which then impact proximal responses of other individuals (question 2), or (C) treatment interference, i.e., treatment provided to one individual may impact the proximal response of another individual in the same cluster (question 3). We will demonstrate how our proposed method addresses these three issues.

3.2 Proximal Moderated Treatment Effects

The primary question of interest is whether the treatment has a proximal effect. Here, potential outcomes are used to define the proximal effects. Due to potential treatment interference in the cluster MRT setting, we consider two types of moderated effects: a direct and an indirect effect, both on the linear scale.

Consider a cluster of size G. The overbar is used to denote a sequence through a specified treatment occasion; $\bar{a}_{t,j} = (a_{1,j}, \ldots, a_{t,j})$, for instance, denotes the sequence of realized treatments up to and including decision time t for individual $j \in [G] := \{1, \ldots, G\}$ in the cluster. Let $\bar{a}_t = (\bar{a}_{t,1}, \ldots, \bar{a}_{t,G})$ denote the set of realized treatments up to and including decision time t for all individuals in the cluster and $\bar{a}_{t,-j} = \bar{a}_t \backslash \bar{a}_{t,j}$ to denote this set with the jth individual's realized treatments removed. Let $Y_{t+1,j}(\bar{a}_t)$ denote the potential outcome

for individual $j \in [G]$ in the cluster. Note that the potential outcome for individual j may depend on the realized treatments for all members in the cluster.

Direct causal effects. In standard MRTs, the individual is the unit of interest. In our current setting, the cluster is the unit of interest. At the cluster level, our interest lies in the effect of providing treatment versus not providing treatment at time t on a random individual in the group. This can be expressed as a difference in potential outcomes for the proximal response and is given by

$$\frac{1}{G} \sum_{j=1}^{G} \left[Y_{t+1,j}(\bar{a}_{t,-j}, (\bar{a}_{t-1,j}, a_{tj} = 1)) - Y_{t+1,j}(\bar{a}_{t,-j}, (\bar{a}_{t-1,j}, a_{tj} = 0)) \right]. \tag{2}$$

Equation (2) is a *direct causal effect* of treatment compared to no treatment on the proximal response for the same individual, while all previous treatments on the individual and all treatments on every other member of the cluster are fixed to specific values.

The "fundamental problem of causal inference" [10, 11] is that these individual differences cannot be observed. Thus, similar to prior work [5, 3], in this paper averages of potential outcomes are considered in defining treatment effects. Let $S_t(\bar{a}_{t-1})$ denote a vector of potential moderator variables chosen from $H_t(\bar{a}_{t-1})$, the cluster-level history up to decision point t. We write $S_t(\bar{a}_{t-1}) = (S_{t,j}(\bar{a}_{t-1}), S_{t,-j}(\bar{a}_{t-1}))$ to clarify that the potential moderator variables can contain both information on the selected individual as well as other individuals in the cluster. Then moderated direct treatment effect, denoted $\beta(t; s)$, can be defined as

$$\mathbb{E}\left[Y_{t+1,J}(\bar{A}_{t,-J}, (\bar{A}_{t-1,J}, 1)) \right.$$
$$\left. - Y_{t+1,J}(\bar{A}_{t,-J}, (\bar{A}_{t-1,J}, 0)) \mid \left(S_{t,J}(\bar{A}_{t-1}), S_{t,-J}(\bar{A}_{t-1}) \right) = (s_J, s_{-J}) \right]. \tag{3}$$

where J is a uniformly distributed random index defined on $[G]$. The expectation is over both the potential outcomes ($Y_{t+1,J}(\cdot)$), the set of randomized treatments ($\bar{A}_{t,-J}$ and $\bar{A}_{t-1,J}$), and the random index (J). Different choices of variables in $S_t(\bar{A}_{t-1})$ address a variety of scientific questions. A primary analysis, for instance, may focus on the marginal proximal

effects and therefore set $S_t(\bar{A}_{t-1}) = \emptyset$. A second analysis may focus on assessing the effect conditional on variables only related to the individual indexed by J and set $S_t(\bar{A}_{t-1}) = X_{t,J}(\bar{A}_{t-1,J})$, where $X_{t,J}(\bar{A}_{t-1,J}) \subset O_{t,H}(\bar{A}_{t-1,J})$ is a potential moderator of interest. A third analysis may consider moderation by the average value across all individuals $S_t(\bar{A}_{t-1}) = G^{-1} \sum_j X_{t,j}(\bar{A}_{t-1,j})$ or the joint moderation of the individual's mood and the average of the other individuals' moods $S_t(\bar{A}_{t-1}) = \left(X_{t,j}(\bar{A}_{t-1,j}), \frac{1}{G-1} \sum_{j' \neq j} X_{t,j'}(\bar{A}_{t-1,j'}) \right)$.

Indirect causal effects. Of secondary interest is the effect of providing treatment versus not providing treatment to the jth individual at time t on a different individual's proximal response when they do not receive treatment, i.e., within-cluster treatment interference. This *indirect causal effect* is given by

$$\frac{1}{G \cdot (G-1)} \sum_{j' \neq j} \left[Y_{t+1,j}(\bar{a}_{t,-\{j,j'\}}, (\bar{a}_{t-1,j}, a_{t,j} = 0), (\bar{a}_{t-1,j'}, a_{t,j'} = 1)) \right.$$
$$\left. - Y_{t+1,j}(\bar{a}_{t,-\{j,j'\}}, (\bar{a}_{t-1,j}, a_{t,j} = 0), (\bar{a}_{t-1,j'}, a_{t,j'} = 0)) \right].$$

Again, since the individual differences cannot be observed, we consider averages of potential outcomes in defining indirect treatment effects. The moderated indirect treatment effect, denoted $\beta^{(IE)}(t; s)$, can be defined by

$$\mathbb{E}\left[Y_{t+1,J}(\bar{A}_{t,-\{J,J'\}}, (\bar{A}_{t-1,J}, 0), (\bar{A}_{t-1,J'}, 1)) \right.$$
$$\left. - Y_{t+1,J}(\bar{A}_{t,-\{J,J'\}}, (\bar{A}_{t-1,J}, 0), (\bar{A}_{t-1,J'}, 0)) \mid S_t(\bar{A}_{t-1}) \right]. \tag{4}$$

where J' is uniformly distributed random index on the set $[G]\setminus\{J\}$. The expectation is again over both the potential outcomes ($Y_{t+1,J}(\cdot)$), the set of randomized treatments ($\bar{A}_{t,-\{J,J'\}}$, $\bar{A}_{t-1,J}$, and $\bar{A}_{t-1,J'}$), and the random indices (J and J'). Here, the potential moderator can be written as $S_t(\bar{A}_{t-1}) = \left(S_{t,J}(\bar{A}_{t-1}), S_{t,J'}(\bar{A}_{t-1}), S_{t,-\{J,J'\}}(\bar{A}_{t-1}) \right)$ to clarify that the variables can contain both information on the two selected individuals as well as others in the cluster. Note that a similar effect can be defined when the individual does receive treatment. For

now, we focus on the effect defined by (4).

3.3 Estimand Dependence on Treatment Distribution

It is useful to distinguish the above estimands, the traditional MRT estimand [3], and estimands commonly studies in longitudinal treatment effect estimation [12]. In the causal inference literature, the typical estimand is an expected outcome for a particular sequence of treatments, i.e., $\mathbb{E}\left[Y(a_1, \ldots, a_T)\right]$. Such estimands do not depend on the treatment distribution, but are often not of primary interest in the current MRT setting since many sequences of treatment may never be observed in finite samples given the large number of decision points (often hundreds or thousands).

The estimands considered in this paper are most similar to average outcomes under a particular dynamic treatment regime $\mathbb{E}_\pi\left[Y(A_1, \ldots, A_T)\right]$ where π denotes the dynamic treatment regime from which the treatments are drawn [13]. Indeed, for any $A_{u,j}$ not contained in $S_t(\bar{A}_{t-1})$, the direct and indirect effects depend on the distribution of $\{A_{u,j}\}_{u<t, j \in [G]}$. Estimands (2) and (4) marginalize over treatments not contained in $S_t(\bar{A}_{t-1})$ under the micro-randomized distribution. Marginalization over different probabilistic assignment of treatments may yield different results. Therefore, the marginal direct and indirect effects depend on the study protocol. The reason for this is that micro-randomization is meant to gather information on how to optimize the design of intervention components [14]. The marginal formulation of main and moderation effects characterizes contrasts excursions from the current treatment proctocol, mimics analyses used in a factorial design that marginalize over factors including time. Also see **(author?)** [4, Section 8] for additional considerations. Regardless, the effects considered in this paper are causal and depend on the cluster-level treatment distribution. Due to this dependence, we recommend presenting the micro-randomization distribution together with the estimated treatment effects.

3.4 Proximal Effect of Treatment & Observable data: Identification

Causal effects (3) and (4) can be expressed in terms of the observable data under the following standard set of causal inference assumptions [12]:

Assumption 3.1. We assume consistency, positivity, and sequential ignorability:

- Consistency: For each $t \leq T$ and $j \in [G]$, $\{Y_{t+1,j}(\bar{A}_t), O_{t,j}(\bar{A}_{t-1}), A_{t,j}(\bar{A}_{t-1})\} = \{Y_{t+1,j}, O_{t,j}, A_{t,j}\}$. That is, the observed values equal the corresponding potential outcomes;

- Positivity: if the joint density $\{H_t = h, A_t = a\}$ is greater than zero, then $\mathbb{P}(A_t = a_t \mid H_t = h_t) > 0$;

- Sequential ignorability: for each $t \leq T$, the potential outcomes, $\{Y_{2,j}(a_1), O_{2,j}(a_1), A_{2,j}(a_1), \ldots, Y_{T,j}(\bar{a}_{T-1})\}_{j \in [G], \bar{a}_{T-1} \in \{0,1\}^{(T-1) \times G}}$, are independent of $A_{t,j}$ conditional on the observed history H_t.

Sequential ignorability and, assuming all of the randomization probabilities are bounded away from 0 and 1, positivity, are guaranteed for a cluster MRT by design. In a standard MRT, the randomization probabilities may depend on the individual's observed history so $\Pr(A_t = a_t \mid H_t = h_t) = \prod_{j=1}^{G} \Pr(A_{t,j} = a_{t,j} \mid H_{t,j} = h_{t,j})$ and the positivity constraint can be placed on the individual-level randomization probabilities. Here, we allow for the possibility that the randomization probabilities depend on the joint cluster-level history. Consistency is a necessary assumption for linking the potential outcomes as defined here to the data. Since an individual's outcomes may be influenced by the treatments provided to other individuals in the same cluster, consistency holds due to our use of a group-based conceptualization of potential outcomes as seen in [6] and [7].

Lemma 3.2. *Under assumption 3.1, the moderated direct treatment effect $\beta(t; s)$ is equal to*

$$\mathbb{E}\left[\frac{1}{G}\sum_{j=1}^{G}\mathbb{E}\left[Y_{t+1,j} \mid H_t, A_{t,j}=1\right] - \mathbb{E}\left[Y_{t+1,j} \mid H_t, A_{t,j}=0\right] \mid S_t\right]$$

where each expectation is with respect to the distribution of the data collected using the treatment assignment probabilities. Under assumption 3.1, the moderated indirect treatment effect $\beta(t; s)^{(IE)}$ is equal to

$$\mathbb{E}\left[\frac{1}{G(G-1)}\sum_{j\neq j'}\mathbb{E}\left[Y_{t+1,j} \mid H_t, A_{t,j}=0, A_{t,j'}=1\right] - \mathbb{E}\left[Y_{t+1,j} \mid H_t, A_{t,j}=0, A_{t,j'}=0\right] \mid S_t\right].$$

Proof of Lemma 3.2 can be found in the Appendix A.

4 Estimation

In this section, we consider estimation of the two estimands defined in Lemma 3.2 using cluster MRT data.

4.1 Direct Effect Estimation

We make the following assumptions regarding the direct treatment effect specification:

Assumption 4.1. Assume the treatment effect, $\beta(t; S_t)$, of interest satisfies

$$\mathbb{E}\left[\mathbb{E}\left[Y_{t+1,J} \mid H_t, A_{t,J}=1\right] - \mathbb{E}\left[Y_{t+1,J} \mid H_t, A_{t,J}=1\right] \mid S_t\right] = f_t(S_t)^{\top}\beta^{\star}$$

where $f_t(s) \in \mathbb{R}^p$ is a p-dimensional vector function of s and time t.

We now consider inference on the unknown p-dimensional parameter $\beta^{\star}$. To do so, we define the weight $W_{t,j}$ at decision time t for the jth individual as equal to $\frac{\tilde{p}(A_{t,j}|S_t)}{p_t(A_{t,j}|H_t)}$ where $\tilde{p}_t(a \mid S_t) \in (0,1)$ is arbitrary as long as it does not depend on terms in H_t other than S_t,

and $p(A_{t,j} \mid H_t)$ is the marginal probability that individual j receives treatment $A_{t,j}$ given H_t. For now, we consider pre-specified $\tilde{p}_t(a \mid S_t)$ (i.e., does not depend on the observed MRT data). Here we consider an estimator $\hat{\beta}$ of $\beta^\star$ which is the minimizer of the following cluster-based, weighted-centered least-squares (C-WCLS) criterion, minimized over (α, β):

$$\mathbb{P}_M \left[\frac{1}{G_m} \sum_{j=1}^{G_m} \sum_{t=1}^{T} W_{t,j} \left(Y_{t+1,j} - g_t(H_t)'\alpha - (A_{t,j} - \tilde{p}_t(A_{t,j} \mid S_t)) f_t(S_t)^\top \beta \right)^2 \right] \tag{5}$$

where $\mathbb{P}_M$ is defined as the average of a function over the sample, which in this context is the sample of *clusters* rather than the sample of *individuals* as in traditional MRT settings. In Appendix A, we prove the following result.

Lemma 4.2. *Under assumption 4.1, then, given invertibility and moment conditions, the estimator $\hat{\beta}$ that minimizes* (5) *satisfies $\sqrt{M} \left(\hat{\beta} - \beta^\star \right) \to N(0, Q^{-1}WQ^{-1})$ where*

$$Q = \mathbb{E} \left[\sum_{t=1}^{T} \tilde{p}_t(1 \mid S_t)(1 - \tilde{p}_t(1 \mid S_t)) f_t(S_t) f_t(S_t)^\top \right]$$

and

$$W = \mathbb{E} \left[\sum_{t=1}^{T} W_{J,t}\, \epsilon_{J,t}(A_{J,t} - \tilde{p}_t(1 \mid S_t)) f_t(S_t) \times \sum_{t=1}^{T} W_{\tilde{J},t}\, \epsilon_{\tilde{J},t}(A_{\tilde{J},t} - \tilde{p}_t(1 \mid S_t)) f_t(S_t)^\top \right]$$

where $\epsilon_{t,j} = Y_{t+1} - g_t(H_t)'\alpha^\star - (A_t - \tilde{p}_t(1 \mid S_t)) f_t(S_t)^\top \beta^\star$, $\alpha^\star$ minimizes the least-squares criterion $\mathbb{E} \left[G^{-1} \sum_{j=1}^{G} \sum_{t=1}^{T} W_{t,j} \left(Y_{t+1,j} - g_t(H_t)'\alpha \right)^2 \right]$, and both J and $\tilde{J}$ are independent randomly sampled indices from the cluster.

In practice, plug-in estimates $\hat{Q}$ and $\hat{W}$ are used to estimate the covariance structure. Appendix C presents small sample size adjustments that we use in simulation and analysis in this paper. It is clear that from Lemma 4.2 that use of standard MRT data analytic methods, i.e., equation (1), will produce biased estimates for the marginal effect of interest.

4.2 Indirect Effect Estimation

We make the following assumptions regarding the indirect treatment effect specification:

Assumption 4.3. We assume the treatment effect of interest satisfies

$$\mathbb{E}\left[\mathbb{E}\left[Y_{t+1,J} \mid H_t, A_{t,J} = 0, A_{t,J'} = 1\right] - \mathbb{E}\left[Y_{t+1,J} \mid H_t, A_{t,J} = 0, A_{t,J'} = 0\right] \mid S_t\right] = f_t(S_t)^\top \beta^{\star\star}$$

where $f_t(s) \in \mathbb{R}^p$ is a p-dimensional vector function of s and time t.

We now consider inference on the unknown p-dimensional parameter $\beta^{\star\star}$. To do so, we define a new weight $W_{t,j,j'}$ at decision time t for the jth individual as equal to $\frac{\tilde{p}(A_{t,j}, A_{t,j'} \mid S_t)}{p_t(A_{t,j}, A_{t,j'} \mid H_t)}$ where $\tilde{p}_t(a, a' \mid S_t) \in (0, 1)$ is arbitrary as long as it does not depend on terms in H_t other than S_t, and $p(A_{t,j}, A_{t,j'} \mid H_t)$ is the marginal probability that individuals j and j' receive treatments $A_{t,j}$ and $A_{t,j'}$ respectively given H_t. Here we consider an estimator $\hat{\beta}^{(IE)}$ of $\beta^{\star\star}$ which is the minimizer of the following cluster-based weighted-centered least-squares (C-WCLS) criterion, minimized over (α, β):

$$\mathbb{P}_M\left[\frac{1}{G_m(G_m - 1)} \sum_{j \neq j'} \sum_{t=1}^T W_{t,j,j'} \left(Y_{t+1,j} - g_t(H_t)'\alpha - (1 - A_{t,j})(A_{t,j'} - \tilde{p}_t^\star(1 \mid S_t))f_t(S_t)^\top \beta\right)^2\right] \tag{6}$$

where $\tilde{p}_t^\star(1 \mid S_t) = \frac{\tilde{p}_t(0,1 \mid S_t)}{\tilde{p}_t(0,0 \mid S_t) + \tilde{p}_t(0,1 \mid S_t)}$. If an individual's randomization probabilities only depends on their own observed history then $\tilde{p}_t^\star(1 \mid S_{t,j'}) = \tilde{p}_t(1 \mid S_{t,j'})$. In Appendix A, we prove the following result.

Lemma 4.4. *Under assumption 4.3, then, under invertibility and moment conditions, the estimator $\hat{\beta}^{(IE)}$ that minimizes (6) satisfies $\sqrt{M}\left(\hat{\beta}^{(IE)} - \beta^{\star\star}\right) \to N(0, Q^{-1}WQ^{-1})$ where*

$$Q = \mathbb{E}\left[\sum_{t=1}^T (\tilde{p}_t(0,0 \mid S_t) + \tilde{p}_t(0,1 \mid S_t))\tilde{p}_t^\star(1 \mid S_t)(1 - \tilde{p}_t^\star(1 \mid S_t))f_t(S_t)f_t(S_t)^\top\right]$$

and

$$W = \mathbb{E}\Bigg[\sum_{t=1}^{T} W_{t,J,J'} \epsilon_{t,J,J'} (1 - A_{J,t})(A_{t,J'} - \tilde{p}_t^\star(1 \mid S_t)) f_t(S_t)$$

$$\times \sum_{t=1}^{T} W_{t,J,J'} \epsilon_{t,\tilde{J},\tilde{J}'} (1 - A_{\tilde{J},t})(A_{t,\tilde{J}'} - \tilde{p}_t^\star(1 \mid S_t)) f_t(S_t)^\top \Bigg]$$

where $\epsilon_{t,j,j'} = Y_{t+1,j} - g_t(H_t)'\alpha^{\star\star} - (1 - A_{t,j})(A_{t,j'} - \tilde{p}_t^\star(1 \mid S_t)) f_t(S_t)^\top \beta^{\star\star}$, $\alpha^{\star\star}$ *minimizes the least-squares criterion* $\mathbb{E}\left[\frac{1}{G_m(G_m-1)} \sum_{j \neq j'} \sum_{t=1}^{T} W_{t,j,j'} \left(Y_{t+1,j} - g_t(H_t)'\alpha \right)^2 \right]$, *and both* (J, J') *and* $(\tilde{J}, \tilde{J}')$ *are independently, randomly sampled pairs from the cluster (without replacement).*

In Appendix E, the above variances are re-derived in the general setting where the numerators $\tilde{p}_t(a \mid S_t)$ and $\tilde{p}_t(a, a' \mid S_t)$ are estimated using the observed MRT data.

4.3   Connection to the Standard MRT Analysis

A natural question is whether there are any conditions under which the standard MRT analysis presented in Section 2.2 is equivalent to the proposed cluster MRT analysis for the direct effect presented in Section 4. It is clear from Proposition 4.2 that if the cluster-size is one then we recover the estimators and asymptotic theory for standard MRT analyses. Lemma 4.5 below proves that, under certain conditions, even when cluster sizes are greater than one, an equivalence of point estimates and asymptotic variances is guaranteed.

Lemma 4.5. *Consider the direct effect when the moderator is defined on the individual (i.e., $S_{j,t}$), and the randomization probabilities only depend on the individual's observed history, i.e., $p(A_{j,t} \mid H_t) = p(A_{j,t} \mid H_{j,t})$. If cluster size is constant across clusters (i.e., $G_m \equiv G$), then the point estimates from (1) and (3) are equal for any sample size. Moreover, if*

$$\mathbb{E}\left[\mathbb{E}\left[\epsilon_{j,t} \epsilon_{j',t'} \mid H_{j,t}, A_{j,t} = a, H_{j',t'}, A_{j',t'} = a' \right] \mid S_{t,j}, S_{t,j'} \right] = \psi(S_{t,j}, S_{t,j'}), \tag{7}$$

for some function ψ, i.e., the cross-term is constant in a and a', where $\epsilon_{j,t}$ is the error defined

in Lemma 4.2 then the estimators share the same asymptotic variance.

Proof of Lemma 4.5 can be found in Appendix B. Here, we consider the following class of random effect models to help with interpretation of the sufficient condition (7). Specifically, for participant j at decision time t suppose the generative model for the proximal outcome is

$$Y_{t+1,j} = g_t(H_{t,j})'\alpha + \underbrace{Z_{t,j}'b_g}_{(I)} + (A_{t,j} - p_t(1 \mid H_{t,j}))(f_t(H_{t,j})'\beta + \underbrace{Z_{t,j}'\tilde{b}_g}_{(II)}) + e_{t,j}$$

where (I) and (II) are random effects with design matrix $Z_{t,j}$, $\mathbb{E}[f_t(H_{t,j})'\beta \mid S_{t,j}] = f_t(S_{t,j})'\beta$, and $e_{t,j}$ is a participant-specific error term. The treatment effect conditional on the complete observed history and the random effects is $f_t(H_{t,j})'\beta + Z_{t,j}'\tilde{b}_g$, which implies the marginal causal effect is $f_t(S_{t,j})'\beta$ so Assumption 4.1 holds. Random effects in (I) allow for cluster-level variation in baseline values of the proximal outcome, while random effects in (II) allow for cluster-level variation in the fully-conditional treatment effect. Given the above generative model, sufficient condition (7) holds if $\tilde{b}_g \equiv 0$, i.e., when the treatment effect does not exhibit cluster-level variation. To see this, note that the errors $\epsilon_{t,j}$ and $\epsilon_{t,j'}$ as defined in Lemma 4.2 do not depend on treatment if $\tilde{b} = 0$. The cross-term is non-zero ($\psi \not\equiv 0$) because there is within-cluster variation due to random effects in (I). For this reason, we refer to (7) as a treatment-effect heterogeneity condition, i.e., when clusters exhibit treatment-effect heterogeneity this induces marginal residual correlation to depend on a and a' which means the proposed approach is necessary for assessing direct effects rather than standard MRT analyses. In Section 5, we conduct simulations to support this calculation.

5 Simulations

To evaluate the performance of the proposed cluster-level proximal treatment effect estimator, we extend the simulation setup in [3]. Consider a standard MRT where the randomization probability $p_t(1|H_t)$ is known. The observation vector O_t is a single state variable

17

$S_t \in \{-1, 1\}$ at each decision time t. Assume the response, Y_{t+1}, follows a linear model given by:

$$
\begin{aligned}
Y_{t+1} =& \theta_1 \{ S_t - \mathbb{E}\left[S_t | A_{t-1}, H_{t-1} \right] \} + \theta_2 \{ A_{t-1} - p_{t-1}(1|H_{t-1}) \} \\
& + \{ A_t - p_t(1|H_t) \}(\beta_{10} + \beta_{11} S_t) + e_{t+1}
\end{aligned}
\tag{8}
$$

The randomization probability is given by $p_t(1|H_t) = \mathrm{expit}(\eta_1 A_{t-1} + \eta_2 S_t)$ where $\mathrm{expit}(x) = (1 + \exp(-x))^{-1}$; the state dynamics are given by $\mathbb{P}(S_t = 1|A_{t-1}, H_{t-1}) = \mathrm{expit}(\xi A_{t-1})$ (note $A_0 = 0$), and the independent error term satisfies $e_t \sim \mathrm{N}(0, 1)$ with $\mathrm{Corr}(e_u, e_t) = 0.5^{|u-t|/2}$. To keep similar to [3], in scenarios below, we set $\theta_1 = 0.8, \theta_2 = 0, \xi = 0, \eta_1 = -0.8, \eta_2 = 0.8, \beta_{10} = -0.2$, and $\beta_{11} = 0.2$. It is readily seen that because $\xi = 0$, $\mathbb{E}[S_t] = 0.5 \times (-1) + 0.5 \times 1 = 0$. The marginal proximal effect is then

$$
\mathbb{E}\left[\mathbb{E}\left[Y_{t+1}|A_t = 1, H_t \right] - \mathbb{E}\left[Y_{t+1}|A_t = 0, H_t \right] \right] = \beta_{10} + \beta_{11}\mathbb{E}\left[S_t \right] = \beta_{10}.
$$

The marginal treatment effect is thus constant in time and is given by $\beta_1^\star = \beta_{10} = -0.2$.

In the following, we consider four simulation scenarios that extend the generative model described above. The first three scenarios concern estimation of the marginal proximal treatment effect $\beta_1^\star$. In each case, when the weighted and centered method is used, we set $f_t(S_t) = 1$ in (8) (i.e., $S_{1t} = \emptyset$). We report average $\hat{\beta}_1$ point estimates, standard deviation (SE) and root mean squared error (RMSE) of $\hat{\beta}_1$, and 95% confidence interval coverage probabilities (CP) across 1000 replicates. We vary the number of clusters and cluster size. Confidence intervals are constructed based on standard errors that are corrected for the estimation of weights and/or small samples. We compare the proposed cluster-based method (C-WCLS) to the approach by [3] (WCLS).

Simulation Scenario I. The first scenario concerns the estimation of β_1^* when an important individual-level moderator exists and proximal outcomes share a random cluster-level intercept term that does not interact with treatment. In the data generative model, we

incorporate a cluster-level random-intercept $e_g \sim \mathcal{N}(0, 0.5)$, so that

$$Y_{t+1,j} = (-0.2 + 0.2 \cdot S_{t,j}) \times (A_{t,j} - p_t(1|H_{t,j})) + 0.8 S_{t,j} + e_g + e_{t+1,j}.$$

Table 2 presents the results, which shows both WCLS and the proposed C-WCLS approach achieve unbiasedness and proper coverage. This is in line with our theoretical results in Section 4.3 stating asymptotic equivalence of the two procedures under no cluster-level treatment heterogeneity.

Simulation Scenario II. In the second scenario, we extend the above generative model to include a random cluster-level intercept term that interacts with treatment by considering the linear model

$$Y_{t+1,j} = (-0.2 + b_g + 0.2 \cdot S_{t,j}) \times (A_{t,j} - p_t(1|H_{t,j})) + 0.8 \cdot S_{t,j} + e_g + e_{t+1,j} \tag{9}$$

where $b_g \sim \mathcal{N}(0, 0.1)$ is a random-intercept term within the treatment effect per cluster. Table 2 presents the results. This demonstrates that if cluster-level random effects interact with treatment (i.e. $b_g \not\equiv 0$), then both methods produce unbiased estimates of β_1 but only the proposed method achieves the nominal 95% coverage probability. To further demonstrate this, Figure 1 presents nominal coverage as a function of the ratio of the variance of the random effect b_g over the variance of the random effect e_g and group size respectively. This shows that the coverage probability of the WCLS method decays rapidly while the proposed method achieves the nominal 95% coverage probability for all choices of the variance of b_g and when the group size increases. Note that even when group size is 5 (i.e., small groups), the nominal coverage drops to 80%.

Simulation Scenario III. In the third scenario, we assume the treatment effect for an individual depends on the average state of all individuals in the cluster. Therefore, we define
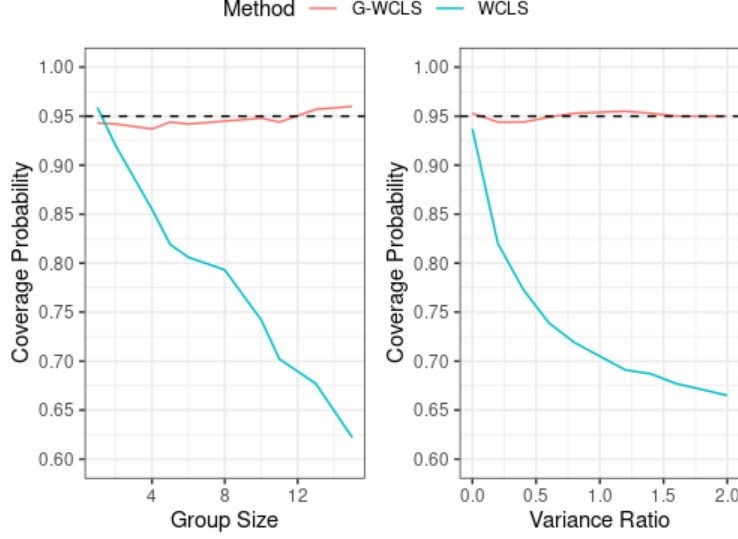
Figure 1: C-WCLS offers more valid 95% confidence intervals than WCLS in Scenario II. Observed coverage probability varies by group size (G) (total sample size fixed) (**left**), and the variance of random intercept b_g as a fraction of the variance of the random error term e_g (**right**).

the cluster-level moderator $\bar{S}_{t,g} = \frac{1}{G_g} \sum_{j=1}^{G_g} S_{t,j}$ and consider the generative model:

$$Y_{t+1,j} = (-0.2 + b_g + 0.2 \cdot \bar{S}_{t,g}) \times (A_{t,j} - p_t(1|H_{t,j})) + 0.8S_{t,j} + e_g + e_{t+1,j}$$

The proposed estimator again achieves the nominal 95% coverage probability while the WCLS method does not (see Scenario III, Table 2).

Simulation Scenario IV. The fourth scenario considers the indirect effect. For individual j at decision point t, we construct the total effect as $TE_{j,t} = \sum_{j' \neq j}\{A_{t,j'} - \tilde{p}_{t,j'}(1 \mid H_t)\}(\beta_{20} + \beta_{21}S_{t,j'})$, where $\beta_{20} = -0.1$ and $\beta_{21} = 0.2$. The generative model is given by:

$$Y_{t+1,j} = (-0.2 + b_g + 0.2 \cdot \bar{S}_{t,g}) \times \{A_{t,j} - p_t(1|H_{t,j})\} + 0.8S_{t,j} + TE_{j,t} + e_g + e_{t+1,j}$$

This model implies a marginal indirect effect equal to $\beta_1^{(IE)} = \beta_{20} = -0.1$. Table 3 presents the results. We see that the proposed indirect estimator exhibited nearly no bias and achieved the nominal coverage probability.

20

Scenario	Estimator	# of Clusters	Cluster Size	Estimate	SE	RMSE	CP
I	C-WCLS	25	10	-0.198	0.035	0.036	0.945
	WCLS			-0.200	0.036	0.035	0.956
	C-WCLS	25	25	-0.199	0.022	0.023	0.948
	WCLS			-0.199	0.023	0.022	0.958
	C-WCLS	50	10	-0.198	0.025	0.027	0.935
	WCLS			-0.198	0.026	0.026	0.944
	C-WCLS	50	25	-0.198	0.016	0.016	0.950
	WCLS			-0.198	0.016	0.017	0.937
	C-WCLS	100	10	-0.199	0.018	0.018	0.949
	WCLS			-0.198	0.018	0.019	0.949
	C-WCLS	100	25	-0.198	0.011	0.012	0.941
	WCLS			-0.199	0.011	0.012	0.944
II	C-WCLS	25	10	-0.199	0.070	0.076	0.935
	WCLS			-0.201	0.041	0.076	0.710
	C-WCLS	25	25	-0.196	0.065	0.071	0.933
	WCLS			-0.200	0.026	0.065	0.557
	C-WCLS	50	10	-0.200	0.051	0.049	0.957
	WCLS			-0.200	0.029	0.052	0.723
	C-WCLS	50	25	-0.200	0.047	0.049	0.947
	WCLS			-0.199	0.019	0.048	0.555
	C-WCLS	100	10	-0.198	0.036	0.035	0.955
	WCLS			-0.198	0.021	0.037	0.718
	C-WCLS	100	25	-0.198	0.033	0.035	0.942
	WCLS			-0.199	0.013	0.032	0.583
III	C-WCLS	25	10	-0.199	0.070	0.073	0.948
	WCLS			-0.202	0.041	0.074	0.734
	C-WCLS	25	25	-0.196	0.066	0.069	0.931
	WCLS			-0.200	0.026	0.068	0.563
	C-WCLS	50	10	-0.198	0.051	0.052	0.941
	WCLS			-0.199	0.029	0.052	0.742
	C-WCLS	50	25	-0.199	0.047	0.048	0.946
	WCLS			-0.200	0.018	0.048	0.561
	C-WCLS	100	10	-0.200	0.036	0.037	0.946
	WCLS			-0.200	0.021	0.037	0.740
	C-WCLS	100	25	-0.201	0.034	0.033	0.956
	WCLS			-0.198	0.013	0.034	0.555

Table 2: Group weighted-centered least squares (C-WCLS) and weighted-least squares estimator (WCLS) comparison over Scenario I, II, and III.

Scenario	# of Clusters	Cluster Size	Estimate	SD	RMSE	CP
	25	10	-0.097	0.028	0.029	0.958
	25	25	-0.101	0.017	0.019	0.942
IV	50	10	-0.097	0.020	0.021	0.953
	50	25	-0.100	0.013	0.013	0.942
	100	10	-0.097	0.015	0.015	0.943
	100	25	-0.100	0.009	0.009	0.944

Table 3: Group weighted-centered least squares (C-WCLS) for estimation of indirect effects.

6 Case Study: Intern Health Study

The Intern Health Study (IHS) was a 6-month MRT on 1,565 medical interns where four types of weekly notification - mood, activity, sleep, or none – were randomly assigned with equal probability to each subject [9] (see Section 3.1). In IHS, 273 institutions and 14 specialties were observed. Here, we assess the effect of the three types of notifications (mood, activity, and sleep) compared to no notifications on the weekly average of self-reported mood scores and log step-count for the population of interns. See Appendix D for an additional analysis of log sleep minutes. Due to high levels of missing data, weekly average mood scores and were computed from multiply imputed daily self-reported scores. Similar imputation was performed for daily log step-counts and daily log sleep minutes. See [9] for further details.

Let $t = 1, \ldots, T$ denote the weekly decision points at which the individual is randomized to the various types of notifications. Because of the form of the intervention, all participants were available for this intervention throughout the study; i.e., $I_t \equiv 1$. The proximal outcome $Y_{j,t+1}$ is weekly mood score, which is reported on a Likert scale taking values from 1 to 10 (higher scores mean better mood). We collapse notifications to a binary variable, i.e., $A_{t,j} = 1$ if the individual was assigned to receive either mood, activity, or sleep notifications on week t; otherwise, $A_{t,j} = 0$ and no notifications were sent that week. At any occasion t, an individual's notification randomization probabilities were only dependent of their past observed history $H_{t,j}$. For simplicity, we start with clusters beingare constructed based on medical specialty. The average cluster size was 120; the first and third quartile were 56 and 140 respectively, with maximum and minimum sizes of 353 and 24. For every individual in

each cluster at each decision point we compute the average prior weekly mood score for all others in the cluster, denoted $\bar{Y}_{-j,t}$ for the jth individual in the cluster. We consider two moderation analyses that can both be expressed as

$$\beta(t; S_t) = \beta_0 + \beta_1 \cdot Y_{j,t-1} + \beta_2 \bar{Y}_{-j,t-1}.$$

The first set of moderation analyses considers the standard moderation analysis where only individual-level moderators are included (i.e., $\beta_2 = 0$), with or without accounting for cluster-level moderation effect heterogeneity. Table 4 presents the results and compares our proposed approach against the WCLS approach from [3]. In this case, the effects do not change too much for the average weekly mood analysis; however, the significant effect of messages on weekly log step count under the traditional MRT analysis becomes insignificant when accounting for cluster effects.

The second moderation analysis lets β_2 be a free parameter, enabling novel moderation analyses that accounts for the average weekly previous mood score of other individuals. Table 4 presents the results. Here, we see that the constant term β_0 becomes negative but insignificant while the new term β_2 is positive and significant. The results suggest the average mood of others in the cluster may moderate the effect of notifications. Therefore, the impact of a notification is larger when the average mood score of others in the specialty is large while the individual's score is low. Similar results hold for the log step-count analysis.

Finally, we consider indirect moderation effect analyses. In this analysis, clusters are constructed based on medical specialty and institution. This was done as interference was only likely when interns are in close geographic proximity. Here, we consider the marginal indirect effect (e.g., no moderators) both when the individual did not receive the intervention and when the individual did receive an intervention at decision time t. Table 5 presents the results. In this case, the estimated indirect effects are weaker than the direct effects by a factor of approximately 10. Even a weak effect may be unexpected as none of the content

Outcome	Setting	Variables	Estimate	Std. Error	t-value	p-value
Mood	WCLS	β_0	0.369	0.086	4.268	0.000
		β_1	-0.055	0.011	-4.822	0.000
	C-WCLS	β_0	0.350	0.103	3.401	0.001
		β_1	-0.053	0.014	-3.868	0.000
	C-WCLS	β_0	-0.238	0.282	-0.842	0.401
		β_1	-0.054	0.014	-3.973	0.000
		β_2	0.083	0.037	2.241	0.026
Steps	WCLS	β_0	0.729	0.295	2.472	0.015
		β_1	-0.037	0.015	-2.484	0.015
	C-WCLS	β_0	0.622	0.384	1.618	0.108
		β_1	-0.031	0.019	-1.580	0.117
	C-WCLS	β_0	-2.095	1.248	-1.678	0.094
		β_1	-0.034	0.019	-1.745	0.083
		β_2	0.143	0.062	2.330	0.020

Table 4: Moderation analysis for the effect of notifications on average weekly mood scores and log step counts respectively in IHS. WCLS estimates use the estimator in [3], while the C-WCLS estimates use our proposed estimator.

in the push notifications was aimed at impacting other individuals' behavior. The fact that there is any signal for the impact on proximal mood implies the scientific team may wish to consider interventions that target these indirect effects more explicitly, either in the framing of the messages or in messaged aimed at impacting the overall group. For the other proximal outcomes, we see limited evidence of an indirect effect. This implies that the scientific team, when building an optimal JITAI for these behaviors, may decide to ignore the indirect effects and instead solely focus on its impact on the individual who receives this push notification.

Outcome	Variables	Estimate	Std. Error	t-value	p-value
Mood	$\tilde{\beta}_0$	0.015	0.017	0.901	0.184
	$\tilde{\beta}_1$	0.027	0.019	1.411	0.079
Steps	$\tilde{\beta}_0$	0.023	0.051	0.454	0.325
	$\tilde{\beta}_1$	0.052	0.060	0.871	0.192
Sleep	$\tilde{\beta}_0$	0.034	0.035	0.980	0.164
	$\tilde{\beta}_1$	0.034	0.035	0.980	0.164

Table 5: Moderation analysis for the indirect effect of notifications on average weekly mood scores, log step counts, and log sleep minutes respectively in IHS. Coefficient $\tilde{\beta}_0$ represents the indirect effect under $A_{j,t} = 0$, while the coefficient $\tilde{\beta}_1$ represents the indirect effect under $A_{j,t} = 1$.

7 Discussion

Here we consider the causal excursion effect in the presence of a priori known clusters. We have extended the causal excursion effect to naturally account for cluster information. In particular, both direct and indirect excursion effects have been formalized in the context of MRT to account for potential interference. The effects described in this paper are most important when using MRT data to build optimized JITAIs for deployment in an mHealth package. Specifically, the estimation procedure for the direct excursion effect accounts for within-cluster correlation in the proximal outcomes which helps the scientific team avoid making erroneous conclusions about intervention effectiveness using standard MRT methods. Moreover, estimation of indirect effects allows the scientific team to answer questions about impact of interventions on other members of the same cluster. Use of these methods provides empirical evidence for the scientific team to include or exclude intervention components that may have had unanticipated second order effects, or potentially lead to novel ways to improve the intervention component by revising the intervention to more explicitly account for cluster-level interference. While this work represents a major step forward in the analysis of micro-randomized trial data, further work is required. Specifically, extensions to be considered future work include accounting for overlapping communities and/or network (rather than cluster-only) structure [15, 16], accounting for general non-continuous proximal outcomes such as binary or count outcomes [4], penalization of the working model to allow for high-dimensional moderators, and a method to use the proposed approach to form warm-start policies at the individual level while accounting for group level information [17].

References

[1] I. Nahum-Shani, S.N. Smith, B.J. Spring, L.M. Collins, K. Witkiewicz, A. Tewari, and S.A. Murphy. Just-in-time adaptive interventions (JITAIs) in mobile health: Key components and design principles for ongoing health behavior support. *Annals of Behavioral*

Medicine, 2017.

[2] P. Klasnja, E.B. Hekler, S. Shiffman, A. Boruvka, D. Almirall, A. Tewari, and S.A. Murphy. Microrandomized trials: An experimental design for developing just-in-time adaptive interventions. *Health Psychol*, 34:1220–1228, 2015.

[3] A. Boruvka, D. Almirall, K. Witkiewitz, and S.A. Murphy. Assessing time-varying causal effect moderation in mobile health. To appear in the *Journal of the American Statistical Association*, 2017.

[4] Tianchen Qian, Hyesun Yoo, Predrag Klasnja, Daniel Almirall, and Susan Murphy. Estimating time-varying causal excursion effect in mobile health with binary outcomes. *Biometrika*, 09 2020. asaa070.

[5] W. Dempsey, P. Liao, and S.A. Murphy. The stratified micro-randomized trial design: Sample size considerations for testing nested causal effects of time-varying treatments. *Annals of Applied Statistics*, 14(2):661–684, 2020.

[6] G. Hong and S. W. Raudenbush. Evaluating kindergarten retention policy. *Journal of the American Statistical Association*, 101(475):901–910, 2006.

[7] T. J. Vanderweele, G. Hong, S.M. Jones, and J.L. Brown. Mediation and spillover effects in group-randomized trials: A case study of the 4Rs educational intervention. *Journal of the American Statistical Association*, 108(502):469–482, 2013.

[8] Predrag Klasnja, Shawna Smith, Nicholas J Seewald, Andy Lee, Kelly Hall, Brook Luers, Eric B Hekler, and Susan A Murphy. Efficacy of contextually tailored suggestions for physical activity: a micro-randomized optimization trial of HeartSteps. *Annals of Behavioral Medicine*, 53(6):573–582, 09 2018.

[9] Timothy NeCamp, Srijan Sen, Elena Frank, Maureen A Walton, Edward L Ionides, Yu Fang, Ambuj Tewari, and Zhenke Wu. Assessing real-time moderation for develop-

ing adaptive mobile health interventions for medical interns: Micro-randomized trial. *Journal of medical Internet research*, 22(3):e15033, March 2020.

[10] DB. Rubin. Bayesian inference for causal effects: The role of randomization. *The Annals of Statistics*, 6(1):34–58, 1978.

[11] J. Pearl. Causal inference in statistics: An overview. *Statistics Surveys*, 3:96–146, 2009.

[12] J. Robins. A new approach to causal inference in mortality studies with a sustained exposure period-application to control of the healthy worker survivor effect. *Mathematical Modelling*, 7(9):1393–1512, 1986.

[13] S A Murphy, M J van der Laan, J M Robins, and Conduct Problems Prevention Research Group. Marginal mean models for dynamic regimes. *Journal of the American Statistical Association*, 96(456):1410–1423, 2001. PMID: 20019887.

[14] L.M. Collins. *Optimization of Behavioral, Biobehavioral, and Biomedical Interventions*. Springer International Publishing, 2018.

[15] E. Ogburn and T. VanderWeele. Causal diagrams for interference. *Statistical Science*, 29(4):559–578, 2014.

[16] Georgia Papadogeorgou, Fabrizia Mealli, and Corwin M. Zigler. Causal inference with interfering units for cluster and population level treatment allocation programs. *Biometrics*, 75(3):778–787, 2019.

[17] Daniel J. Luckett, Eric B. Laber, Anna R. Kahkoska, David M. Maahs, Elizabeth Mayer-Davis, and Michael R. Kosorok. Estimating dynamic treatment regimes in mobile health using v-learning. *Journal of the American Statistical Association*, 115(530):692–706, 2020. PMID: 32952236.

[18] LA. Mancl and T.A. DeRouen. A covariance estimator for GEE with improved small-sample properties. *Biometrics*, 57(1):126–134, 2001.

A   Technical Details

Proof. Proof of Lemma 3.2 We establish Lemma 3.2 for the direct effect (3). For $a_s \in \{0,1\}^G$, we consider

$$\mathbb{E}\left[\left(\prod_{s=1}^{t-1} p_s(a_s|H_s(\bar{a}_{s-1}))\right)\left(\prod_{j'\neq j} p_t(a_{t,j'} \mid H_t(\bar{a}_{t-1}))\right) Y_{t+1,j}(\bar{a}_{t,-j}, (\bar{a}_{t-1,J}, a))\mathbf{1}_{S_t(\bar{a}_t)=s}\right]$$

$$\mathbb{E}\left[\left(\prod_{s=1}^{t-1} p_s(a_s|H_s(\bar{a}_{s-1}))\right)\left(\prod_{j'\neq j} p_t(a_{t,j'} \mid H_t(\bar{a}_{t-1}))\right) \mathbf{1}_{S_t(\bar{a}_t)=s}\mathbb{E}\left[Y_{t+1,j}(\bar{a}_{t,-j}, (\bar{a}_{t-1,j}, a)) \mid H_t(\bar{a}_{t-1})\right]\right]$$

since the history H_t includes the moderator variable S_t at time t. By consistency, $H_t(\bar{A}_{t-1}) = H_t$ so the above is equal to

$$\mathbb{E}\left[\left(\prod_{s=1}^{t-1} p_s(a_s|H_s)\right)\left(\prod_{j'\neq J} p_t(a_{t,j'} \mid H_t)\right) \mathbf{1}_{S_t(\bar{a}_t)=s}\mathbb{E}\left[Y_{t+1,j}(\bar{a}_{t,-j}, (\bar{a}_{t-1,j}, a)) \mid H_t\right]\right] \qquad (10)$$

Sequential ignorability implies that

$$\mathbb{E}\left[Y_{t+1,j}(\bar{a}_{t,-j}, (\bar{a}_{t-1,j}, a)) \mid H_t\right]$$

$$= \mathbb{E}\left[Y_{t+1,j}(\bar{a}_{t,-j}, (\bar{a}_{t-1,j}, a)) \mid H_t, A_{t,j} = a\right]$$

Summing over all potential outcomes and normalizing yields

$$\mathbb{E}\left[\sum_{\bar{a}_{t-1}} \frac{\left(\prod_{s=1}^{t-1} p_s(a_s|H_s)\right)\left(\prod_{j'\neq j} p_t(a_{t,j'} \mid H_t)\right)\mathbf{1}_{S_t(\bar{a}_t)=s}}{\mathbb{E}[\left(\prod_{s=1}^{t-1} p_s(a_s|H_s)\right)\left(\prod_{j'\neq j} p_t(a_{t,j'} \mid H_t)\right)\mathbf{1}_{S_t(\bar{a}_t)=s}]}\mathbb{E}\left[Y_{t+1,j}(\bar{a}_{t,-j}, (\bar{a}_{t-1}, a)) \mid H_t, A_{t,j} = a\right]\right]$$

$$= \mathbb{E}\left[\sum_{\bar{a}_{t-1}} \frac{\left(\prod_{s=1}^{t-1} p_s(a_s|H_s)\right)\left(\prod_{j'\neq j} p_t(a_{t,j'} \mid H_t)\right)\mathbf{1}_{S_t(\bar{a}_t)=s}}{\mathbb{E}[\left(\prod_{s=1}^{t-1} p_s(a_s|H_s)\right)\left(\prod_{j'\neq j} p_t(a_{t,j'} \mid H_t)\right)\mathbf{1}_{S_t(\bar{a}_t)=s}]}\mathbb{E}\left[Y_{t+1,j}(\bar{a}_{t,-j}, (\bar{a}_{t-1}, a)) \mid H_t, A_t = a\right] \mid S_t = s\right]$$

$$= \mathbb{E}\left[\mathbb{E}\left[Y_{t+1,j} \mid H_t, A_t = a\right] \mid S_t = s\right].$$

In the final equation, the outer expectation is with respect to the history H_t conditional on $S_t = s$. That is, over *both* past treatments A_s and past observations O_s for $s < t$ as well

as over current treatments for $A_{t,j'}$ for $j' \neq j$. The above shows

$$\mathbb{E}\left[Y_{t+1}(\bar{A}_{t,-j}, (\bar{A}_{t-1,j}, a)) \mid S_t(\bar{a}_t) = s\right] = \mathbb{E}\left[\mathbb{E}\left[Y_{t+1,j} \mid H_t, A_{t,j} = a\right] \mid S_t = s\right].$$

Averaging over individuals in the group $j \in [G]$ group size completes the proof. The proof for the indirect effect follows the exact same structure. □

A.1 Lemma 4.2

We next provide a detailed proof of asymptotic normality and consistency for the weighted-centered least squares estimator.

Proof of consistency for direct and indirect effects. The solutions $(\hat{\alpha}, \hat{\beta})$ that minimize equation (5) are consistent estimators for the solutions that minimize the following

$$\mathbb{E}\left[\frac{1}{G}\sum_{j=1}^{G}\sum_{t=1}^{T}W_{t,j}\left(Y_{t+1,j} - g_t(H_t)'\alpha - (A_{t,j} - \tilde{p}_t(1\mid S_t))f_t(S_t)'\beta\right)^2\right]$$

Differentiating the above equation with respect to α yields a set of p estimating equations.

$$0_{q'} = \mathbb{E}\left[\frac{1}{G}\sum_{j=1}^{G}\sum_{t=1}^{T}W_{t,j}\left(Y_{t+1,j} - g_t(H_t)'\alpha - (A_{t,j} - \tilde{p}_t(1\mid S_t))f_t(S_t)'\beta\right)g_t(H_t)\right]$$

We note that

$$\mathbb{E}\left[W_{t,J}(A_{t,J} - \tilde{p}_t(1\mid S_t))f_t(S_t)'\beta \mid H_t\right] = 0.$$

Therefore, we have,

$$0_p = \mathbb{E}\left[\frac{1}{G}\sum_{j=1}^{G}\sum_{t=1}^{T}(g_t(H_t)\mathbb{E}\left[W_{t,j}Y_{t+1,j}\right] - g_t(H_t)g_t(H_t)'\alpha)\right]$$

$$\Rightarrow \alpha = \mathbb{E}\left[\frac{1}{G}\sum_{j=1}^{G}\sum_{t=1}^{T}g_t(H_t)g_t(H_t)'\right]^{-1}E\left[\frac{1}{G}\sum_{j=1}^{G}\sum_{t=1}^{T}g_t(H_t)\mathbb{E}\left[W_{t,j}Y_{t+1,j}\right]\right]$$

We note that

$$\mathbb{E}\left[W_{t,J}(A_{t,J} - \tilde{p}_t(1\,|\,S_t))g_t(H_t)'\alpha\right] = 0, \quad \text{and}$$

$$\mathbb{E}\left[W_{t,J}(A_{t,J} - \tilde{p}_t(1\,|\,S_t))Y_{t+1,J}\right] = \tilde{p}_t(1\,|\,S_t)(1 - \tilde{p}_t(1\,|\,S_t))\beta(t; S_t)$$

Now differentiating with respect to β yields

$$0_q = \mathbb{E}\left[\sum_{t=1}^{T} W_{t,J}\left(Y_{t+1,J} - g_t(H_t)'\alpha - (A_{t,J} - \tilde{p}_t(1\,|\,S_t))f_t(S_t)'\beta\right)(A_{t,J} - \tilde{p}_t(1\,|\,S_t))f_t(S_t)\right]$$

$$0_q = \mathbb{E}\left[\sum_{t=1}^{T} \tilde{p}_t(1\,|\,S_t)(1 - \tilde{p}_t(1\,|\,S_t))\left(\beta(t; S_t) - f_t(S_t)'\beta^\star\right)f_t(S_t)\right]$$

Then we have

$$\beta^\star = \mathbb{E}\left[\sum_{t=1}^{T} \tilde{p}_t(1\,|\,S_t)(1 - \tilde{p}_t(1\,|\,S_t))f_t(S_t)f_t(S_t)'\right]^{-1}\mathbb{E}\left[\sum_{t=1}^{T} \tilde{p}_t(1\,|\,S_t)(1 - \tilde{p}_t(1\,|\,S_t))f_t(S_t)\beta(t; S_t)\right]$$

Under assumption 4.1, we have that $\beta = \beta^\star$ which guarantees consistency.

We next consider the indirect effect estimator. Recall that

$$\tilde{p}_t^\star(1\,|\,S_t) = \frac{\tilde{p}_t(0, 1\,|\,S_t)}{\tilde{p}_t(0, 0\,|\,S_t) + \tilde{p}_t(0, 1\,|\,S_t)}$$

is the replacement for $\tilde{p}(1\,|\,S_t)$ in the direct effect for centering. If we make the assumption that $\tilde{p}_t(0, 1\,|\,S_t) = \tilde{p}_t(0\,|\,S_t)\tilde{p}_t(1\,|\,S_t)$ then $\tilde{p}_t^\star(1\,|\,S_t) = \tilde{p}_t(1\,|\,S_t)$; however, we provide the proof in complete generality. The estimates that minimize equation (6) are consistent estimators for the solutions that minimize the following

$$\mathbb{E}\left[\sum_{t=1}^{T} W_{t,J,J'}\left(Y_{t+1,J} - g_t(H_t)'\alpha - (1 - A_{t,J})(A_{t,J'} - \tilde{p}_t^\star(1\,|\,S_t))f_t(S_t)'\beta\right)^2\right]$$

Differentiating the above equation with respect to α yields a set of p estimating equations.

$$0_p = \mathbb{E}\left[\sum_{t=1}^{T} W_{t,J,J'}\left(Y_{t+1,J} - g_t(H_t)'\alpha - (1 - A_{t,J})(A_{t,J'} - \tilde{p}_t^\star(1 \mid S_t))f_t(S_t)'\beta\right)g_t(H_t)\right]$$

We note that

$$\mathbb{E}\left[W_{t,J,J'}(1 - A_{t,J})(A_{t,J'} - \tilde{p}_t^\star(1 \mid S_t))f_t(S_t)'\beta \mid H_t\right] = 0.$$

Therefore, we have,

$$0_p = \mathbb{E}\left[\sum_{t=1}^{T}(g_t(H_t)\mathbb{E}\left[W_{t,J,J'}Y_{t+1,J}\right] - g_t(H_t)g_t(H_t)'\alpha)\right]$$

$$\Rightarrow \alpha = \mathbb{E}\left[\frac{1}{G}\sum_{j=1}^{G}\sum_{t=1}^{T}g_t(H_t)g_t(H_t)'\right]^{-1} E\left[\sum_{t=1}^{T}g_t(H_t)\mathbb{E}\left[W_{t,J,J'}Y_{t+1,J}\right]\right]$$

First, we show that

$$\mathbb{E}\left[W_{t,J,J'}(1 - A_{t,J})(A_{t,J'} - \tilde{p}_t^\star(1 \mid S_t)) \mid H_t\right]$$

$$= \sum_{a' \in \{0,1\}} \mathbb{E}\left[\tilde{p}_t(0, a' \mid S_t)(a' - \tilde{p}_t^\star(1 \mid S_t)) \mid H_t, A_t = 0, A_{t,J'} = a'\right]$$

$$= \tilde{p}_t(0, 1 \mid S_t)(1 - \frac{\tilde{p}_t(0, 1 \mid S_t)}{\tilde{p}_t(0, 0 \mid S_t) + \tilde{p}_t(0, 1 \mid S_t)}) - \tilde{p}_t(0, 0 \mid S_t)\frac{\tilde{p}_t(0, 1 \mid S_t)}{\tilde{p}_t(0, 0 \mid S_t) + \tilde{p}_t(0, 1 \mid S_t)} = 0$$

and

$$\mathbb{E}\left[W_{t,J,J'}(1 - A_{t,J})(A_{t,J'} - \tilde{p}_t^\star(1 \mid S_t))^2 \mid H_t\right]$$

$$= \tilde{p}_t(0, 1 \mid S_t)\left(\frac{\tilde{p}_t(0, 0 \mid S_t)}{\tilde{p}_t(0, 0 \mid S_t) + \tilde{p}_t(0, 1 \mid S_t)}\right)^2 + \tilde{p}_t(0, 0 \mid S_t)\left(\frac{\tilde{p}_t(0, 1 \mid S_t)}{\tilde{p}_t(0, 0 \mid S_t) + \tilde{p}_t(0, 1 \mid S_t)}\right)^2$$

$$= (\tilde{p}_t(0, 0 \mid S_t) + \tilde{p}_t(0, 1 \mid S_t))\tilde{p}_t^\star(1 \mid S_t)(1 - \tilde{p}_t^\star(1 \mid S_t)).$$

This implies that

$$\mathbb{E}\left[W_{t,J,J'}(1 - A_{t,J})(A_{t,J'} - \tilde{p}_t^\star(1 \mid S_t))g_t(H_t)'\alpha\right] = 0, \quad \text{and}$$

$$\frac{\mathbb{E}\left[W_{t,J,J'}(1 - A_{t,J})(A_{t,J'} - \tilde{p}_t^\star(1 \mid S_t))Y_{t+1,J}\right]}{\tilde{p}_t(0,0 \mid S_t) + \tilde{p}_t(0,1 \mid S_t)} = \tilde{p}_t^\star(1 \mid S_t)(1 - \tilde{p}_t^\star(1 \mid S_t))\beta^{(IE)}(t; S_t).$$

Now differentiating with respect to β yields

$$0_q = \mathbb{E}\left[\sum_{t=1}^{T}(\tilde{p}_t(0,0 \mid S_t) + \tilde{p}_t(0,1 \mid S_t))\tilde{p}_t^\star(1 \mid S_t)(1 - \tilde{p}_t^\star(1 \mid S_t))\left(\beta^{(IE)}(t; S_t) - f_t(S_t)'\beta^{\star\star}\right)f_t(S_t)\right]$$

Under assumption 4.3, we have that $\beta = \beta^{\star\star}$ which guarantees consistency. $\qquad\square$

Proof of Asymptotic Normality. We now consider the issue of asymptotic normality. First, let

$$\epsilon_{t,j} = Y_{t+1,j} - g_t(H_t)'\alpha^\star - (A_{t,j} - \tilde{p}_t(1 \mid S_t))f_t(S_t)'\beta^\star,$$

$\hat{\theta} = (\hat{\alpha}, \hat{\beta})$, and $\theta^\star = (\alpha^\star, \beta^\star)$. Since $S_t \subset H_t$ define $h_{t,j}(H_t)' = (g_t(H_t)', (A_{t,j} - \tilde{p}_t(1 \mid S_t))f_t(S_t)')$. Then

$$\sqrt{M}(\hat{\theta} - \theta^\star) = \sqrt{M}\left\{\mathbb{P}_M\left(\frac{1}{G_m}\sum_{j=1}^{G_m}\sum_{t=1}^{T}W_{t,j}h_{t,j}(H_t)h_{t,j}(H_t)'\right)^{-1}\left[\mathbb{P}_M\left(\frac{1}{G_m}\sum_{j=1}^{G_m}\sum_{t=1}^{T}W_{t,j}Y_{t+1,j}h_{t,j}(H_t)\right)\right.\right.$$

$$\left.\left. - \mathbb{P}_M\left(\frac{1}{G_m}\sum_{j=1}^{G_m}\sum_{t=1}^{T}W_{t,j}h_{t,j}(H_t)h_{t,j}(H_t)'\right)\theta^\star\right]\right\}$$

$$= \sqrt{M}\left\{E\left[\frac{1}{G}\sum_{j=1}^{G}\sum_{t=1}^{T}W_{t,j}h_{t,j}(H_t)h_{t,j}(H_t)'\right]^{-1}\right.$$

$$\left.\left[\mathbb{P}_M\left(\frac{1}{G_m}\sum_{j=1}^{G_m}\sum_{t=1}^{T}W_{t,j}\epsilon_{t,j}h_{t,j}(H_t)\right)\right]\right\} + o_p(\mathbf{1})$$

By definitions of $\alpha^\star$ and $\beta^\star$ and the previous consistency argument

$$E\left[\frac{1}{G}\sum_{j=1}^{G}\sum_{t=1}^{T}W_{t,j}h_{t,j}(H_t)h_{t,j}(H_t)'\right] = 0$$

Then under moments conditions, we have asymptotic normality with variance Σ_θ given by

$$\Sigma_\theta = E\left[\frac{1}{G}\sum_{j=1}^{G}\sum_{t=1}^{T}W_{t,j}h_{t,j}(H_t)h_{t,j}(H_t)'\right]^{-1}$$

$$E\left[\frac{1}{G}\sum_{j=1}^{G}\sum_{t=1}^{T}W_{t,j}\epsilon_{t,j}h_{t,j}(H_t) \times \frac{1}{G}\sum_{j=1}^{G}\sum_{t=1}^{T}W_{t,j}\epsilon_{t,j}h_{t,j}(H_t)'\right]$$

$$E\left[\frac{1}{G}\sum_{j=1}^{G}\sum_{t=1}^{T}W_{t,j}h_{t,j}(H_t)h_{t,j}(H_t)'\right]^{-1}$$

Due to centering, the expectation of the matrix $W_{t,J}h_{t,J}(H_t)h_{t,J}(H_t)'$ is block diagonal and the sub-covariance matrix Σ_β can be extracted and is equal to

$$\Sigma_\beta = \left[\sum_{t=1}^{T}E[(A_{t,J}-\tilde{p}_t(1\mid S_t))^2 W_{t,J}f_t(S_t)f_t(S_t)']\right]^{-1}$$

$$\cdot E\left[\sum_{t=1}^{T}W_{t,J}\epsilon_{t,J}(A_{t,J}-\tilde{p}_t(1\mid S_t))f_t(S_t) \times \sum_{t=1}^{T}W_{t,J}\epsilon_{t,J}(A_{t,J}-\tilde{p}_t(1\mid S_t))f_t(S_t)'\right]$$

$$\cdot \left[\sum_{t=1}^{T}E[(A_{t,J}-\tilde{p}_t(1\mid S_t))^2 W_{t,J}f_t(S_t)f_t(S_t)']\right]^{-1}$$

as desired.

We next consider asymptotic normality in the indirect setting. First, let

$$\epsilon_{t,j,j'} = Y_{t+1,j} - g_t(H_t)'\alpha^{\star\star} - (1-A_{t,j})(A_{t,j'}-\tilde{p}_t^\star(1\mid S_t))f_t(S_t)'\beta^{\star\star},$$

$\hat{\theta} = (\hat{\alpha}, \hat{\beta})$, and $\theta^\star = (\alpha^{\star\star}, \beta^{\star\star})$. Since $S_t \subset H_t$ define $h_{t,j,j'}(H_t)' = (g_t(H_t)', (1-A_{t,j})(A_{t,j'}-$

$\tilde{p}_t^\star(1\,|\,S_t))f_t(S_t)')$. Then $\sqrt{M}(\hat{\theta}-\theta^{\star\star})$ equals

$$
\sqrt{M}\Bigg\{\mathbb{P}_M\Bigg(\frac{1}{G_m\cdot(G_m-1)}\sum_{j=1}^{G_m}\sum_{j'\neq j}\sum_{t=1}^{T}W_{t,j,j'}h_{t,j,j'}(H_t)h_{t,j,j'}(H_t)'\Bigg)^{-1}
$$
$$
\Bigg[\mathbb{P}_M\Bigg(\frac{1}{G_m\cdot(G_m-1)}\sum_{j=1}^{G_m}\sum_{j'\neq j}\sum_{t=1}^{T}W_{t,j}Y_{t+1,j}h_{t,j,j'}(H_t)\Bigg)
$$
$$
-\mathbb{P}_M\Bigg(\frac{1}{G_m\cdot(G_m-1)}\sum_{j=1}^{G_m}\sum_{j'\neq j}\sum_{t=1}^{T}W_{t,j,j'}h_{t,j,j'}(H_t)h_{t,j,j'}(H_t)'\Bigg)\theta^\star\Bigg]\Bigg\}
$$
$$
=\sqrt{M}\Bigg\{E\Bigg[\frac{1}{G(G-1)}\sum_{j=1}^{G}\sum_{j'\neq j}\sum_{t=1}^{T}W_{t,j,j'}h_{t,j,j'}(H_t)h_{t,j,j'}(H_t)'\Bigg]^{-1}
$$
$$
\Bigg[\mathbb{P}_M\Bigg(\frac{1}{G_m\cdot(G_m-1)}\sum_{j=1}^{G_m}\sum_{j'\neq j}\sum_{t=1}^{T}W_{t,j,j'}\epsilon_{t,j,j'}h_{t,j,j'}(H_t)\Bigg)\Bigg]\Bigg\}+o_p(\mathbf{1})
$$

By definitions of $\alpha^{\star\star}$ and $\beta^{\star\star}$ and the previous consistency argument

$$
E\Bigg[\frac{1}{G\cdot(G-1)}\sum_{j=1}^{G}\sum_{j'\neq j}\sum_{t=1}^{T}W_{t,j,j'}h_{t,j,j'}(H_t)h_{t,j,j'}(H_t)'\Bigg]=0
$$

Then under moments conditions, we have asymptotic normality with variance Σ_θ given by

$$
\Sigma_\theta=E\Bigg[\frac{1}{G\cdot(G-1)}\sum_{j=1}^{G}\sum_{j'\neq j}\sum_{t=1}^{T}W_{t,j,j'}h_{t,j,j'}(H_t)h_{t,j,j'}(H_t)'\Bigg]^{-1}
$$
$$
E\Bigg[\frac{1}{G(G-1)}\sum_{j=1}^{G}\sum_{j'\neq j}\sum_{t=1}^{T}W_{t,j,j'}\epsilon_{t,j,j'}h_{t,j,j'}(H_t)\times\frac{1}{G(G-1)}\sum_{j=1}^{G}\sum_{j'\neq j}\sum_{t=1}^{T}W_{t,j}\epsilon_{t,j,j'}h_{t,j,j'}(H_t)'\Bigg]
$$
$$
E\Bigg[\frac{1}{G\cdot(G-1)}\sum_{j=1}^{G}\sum_{j'\neq j}\sum_{t=1}^{T}W_{t,j,j'}h_{t,j,j'}(H_t)h_{t,j,j'}(H_t)'\Bigg]^{-1}
$$

Due to centering, the expectation of the matrix $W_{t,J,J'}h_{t,J,J'}(H_t)h_{t,J,J'}(H_t)'$ is block diagonal

and the sub-covariance matrix Σ_β can be extracted and is equal to

$$\Sigma_\beta = \left[\sum_{t=1}^{T} E[(1 - A_{t,J})(A_{t,J'} - \tilde{p}_t^\star(1 \mid S_t))^2 W_{t,J,J'} f_t(S_t) f_t(S_t)'] \right]^{-1}$$

$$\cdot E\left[\sum_{t=1}^{T} W_{t,J,J'} \epsilon_{t,J,J'} (1 - A_{t,J})(A_{t,J'} - \tilde{p}_t(1 \mid S_t)) f_t(S_t) \right.$$

$$\times \left. \sum_{t=1}^{T} W_{t,\tilde{J},\tilde{J}'} \epsilon_{t,\tilde{J},\tilde{J}'} (1 - A_{t,\tilde{J}})(A_{t,\tilde{J}'} - \tilde{p}_t^\star(1 \mid S_t)) f_t(S_t)' \right]$$

$$\cdot \left[\sum_{t=1}^{T} E[(1 - A_{t,J})(A_{t,J'} - \tilde{p}_t(1 \mid S_t))^2 W_{t,J,J'} f_t(S_t) f_t(S_t)'] \right]^{-1}$$

as desired. □

B   Proof of Lemma 4.5

Proof. Consider the W-matrix for the direct effect asymptotic variance,

$$\frac{1}{G^2} \sum_{t,t'} \sum_{j,j'} \mathbb{E}\left[W_{t,j} \epsilon_{t,j} (A_{t,j} - \tilde{p}_t(1 \mid S_t)) W_{t',j'} \epsilon_{t',j'} (A_{t',j'} - \tilde{p}_t(1 \mid S_{t'})) f_t(S_t) f_{t'}(S_{t'})' \right]$$

$$= \frac{1}{G^2} \sum_{t,t'} \sum_{j,j'} \mathbb{E}\left[W_{t,j} \epsilon_{t,j} (A_{t,J} - \tilde{p}_t(1 \mid S_t)) W_{t',j'} \epsilon_{t',j'} (A_{t',j'} - \tilde{p}_t(1 \mid S_{t'})) f_t(S_t) f_{t'}(S_{t'})' \right]$$

Consider the cross-terms with $j \neq j'$ and without loss of generality assume $t \geq t'$, then

$$\mathbb{E}\left[\sum_{a,a'} \tilde{p}_t(a \mid S_t)(a - \tilde{p}_t(1 \mid S_t)) \tilde{p}_{t'}(a' \mid S_{t'})(a' - \tilde{p}_{t'}(1 \mid S_{t'})) \right.$$

$$\mathbb{E}\left[\mathbb{E}\left[\epsilon_{t,j} \epsilon_{t',j'} \mid H_{t,j}, A_{t,j} = a, H_{t',j'}, A_{t',j'} = a' \right] \mid S_t, S_{t'} \right] f_t(S_t) f_{t'}(S_{t'})' \right].$$

Under the assumption of the error cross-term being constant in a and a' we can re-write the above as:

$$= \mathbb{E}\left[\sum_{a,a'} \tilde{p}_t(a \mid S_t)(a - \tilde{p}_t(1 \mid S_t))\tilde{p}_{t'}(a' \mid S_{t'})(a' - \tilde{p}_{t'}(1 \mid S_{t'}))\psi(S_t, S_{t'})f_t(S_t)f_{t'}(S_{t'})'\right]$$

$$= \mathbb{E}\left[\psi(S_t, S_{t'})f_t(S_t)f_{t'}(S_{t'})' \underbrace{\left(\sum_{a,a'} \tilde{p}_t(a \mid S_t)(a - \tilde{p}_t(1 \mid S_t))\tilde{p}_{t'}(a' \mid S_{t'})(a' - \tilde{p}_{t'}(1 \mid S_{t'}))\right)}_{=0}\right]$$

$$= \mathbb{E}\left[\psi(S_t, S_{t'})f_t(S_t)f_{t'}(S_{t'})' \cdot 0\right] = 0.$$

Therefore, we have that the W-matrix simplifies to

$$\mathbb{E}\left[\sum_{t=1}^{T} W_{t,J}\epsilon_{t,J}(A_{t,J} - \tilde{p}_t(1 \mid S_t))f_t(S_t) \times \sum_{t=1}^{T} W_{t,J}\epsilon_{t,J}(A_{t,J} - \tilde{p}_t(1 \mid S_t))f_t(S_t)'\right]$$

$$= \mathbb{E}\left[\frac{1}{G}\sum_{j=1}^{G}\left[\sum_{t=1}^{T} W_{t,J}\epsilon_{t,J}(A_{t,J} - \tilde{p}_t(1 \mid S_t))f_t(S_t) \times \sum_{t=1}^{T} W_{t,J}\epsilon_{t,J}(A_{t,J} - \tilde{p}_t(1 \mid S_t))f_t(S_t)'\right]\right]$$

$$= \mathbb{E}\left[\sum_{t=1}^{T} W_t\epsilon_t(A_t - \tilde{p}_t(1 \mid S_t))f_t(S_t) \times \sum_{t=1}^{T} W_t\epsilon_t(A_t - \tilde{p}_t(1 \mid S_t))f_t(S_t)'\right]$$

which is the W matrix as in the standard MRT analysis. $\square$

C Small sample size adjustment for covariance estimation

The robust sandwich covariance estimator [18] for the entire variance matrix is given by $Q^{-1}\Lambda Q^{-1}$. The first term, Q, is given by

$$\left(\sum_{m=1}^{M}\sum_{j=1}^{G_m} D_{j,m}^T W_{j,m} D_{j,m}\right)$$

where $D_{j,m}$ is the model matrix for individual j in group g associated with equation (5), and $W_{j,m}$ is a diagonal matrix of individual weights. The middle term Λ is given by

$$\sum_{m=1}^{M}\sum_{i,j=1}^{G_m} D'_{i,m}W_{i,m}(I_{i,m}-H_{i,m})^{-1}e_{i,m}e'_{j,m}(I_{j,m}-H_{j,m})^{-1}W_{j,m}D_{j,m}$$

where I_i is an identity matrix of correct dimension, e_i is the individual-specific residual vector and

$$H_{j,m}=D_{j,m}\left(\sum_{m=1}^{M}\sum_{j=1}^{G_m}D'_{j,m}W_{j,m}D_{j,m}\right)^{-1}D'_{j,m}W_{j,m}$$

From $Q^{-1}\Lambda Q^{-1}$ we extract $\hat{\Sigma}_\beta$.

D Additional analysis of IHS

Setting	Variables	Estimate	Std. Error	t-value	p-value
WCLS	β_0	1.325	0.350	3.782	0.000
	β_1	-0.068	0.017	-3.916	0.000
C-WCLS	β_0	1.193	0.454	2.627	0.011
	β_1	-0.061	0.023	-2.696	0.009
C-WCLS	β_0	-1.912	1.379	-1.386	0.171
	β_1	-0.067	0.023	-2.948	0.004
	β_2	0.162	0.065	2.487	0.015

Table 6: Moderation analysis for the effect of notifications on average weekly log sleep minutes in IHS. WCLS estimates use the estimator in [3], while the C-WCLS estimates use our proposed estimator.

E Additional on the indirect effect

Weights used in the estimation of the indirect effect is a natural extension of [3]. As in Section 4.2, the weight $W_{t,j,j'}$ at decision time t for the jth individual is equal to $\frac{\tilde{p}(A_{t,j},A_{t,j'}|S_t)}{p_t(A_{t,j},A_{t,j'}|H_t)}$ where $\tilde{p}_t(a,a'\mid S_t)\in(0,1)$ is arbitrary as long as it does not depend on terms in H_t other than S_t, and $p(A_{t,j},A_{t,j'}\mid H_t)$ is the marginal probability that individuals j and j' receive

treatments $A_{t,j}$ and $A_{t,j'}$ respectively given H_t.

In the simulation, the treatment individuals j and j' receive $A_{t,j}$ and $A_{t,j'}$ are mutually independent conditioning on the previous history. thus, the denominator of $W_{t,j,j'}$ can be factorized into:

$$p(A_{t,j}, A_{t,j'} \mid H_t) = p(A_{t,j} \mid H_t)p(A_{t,j'} \mid H_t)$$

Besides, the numerator of $W_{t,j,j'}$ is defined as the empirical frequency of the treatment pair (a, a'), which takes the value from $\{(0,0),(0,1),(1,0),(1,1)\}$. Here we denote it as

$$\tilde{p}_t(A_{t,j}, A_{t,j'} \mid S_t) = \hat{p}_t(A_{t,j}, A_{t,j'} \mid S_t)$$

Therefore, the weight we used in the simulation is constructed as:

$$W_{t,j,j'} = \frac{\hat{p}_t(A_{t,j}, A_{t,j'} \mid S_t)}{p(A_{t,j} \mid H_t)p(A_{t,j'} \mid H_t)}$$

When the numerators are estimated using the observed data, the variance-covariance must account for this. Throughout we allow for the setting in which individuals are not always available. For completeness we provide results for a more general estimating function which can be used with observational (non-randomized A_t) treatments, under the assumption of sequential ignorability and assuming the data analyst is able to correctly model and estimate the treatment probability $p(A_{t,j}, A_{t,j'}|H_t)$. We indicate how the results are simplified by use of data from an MRT.

Denote the parameterized treatment probability by $p_t(a, a'|H_t; \eta)$ (with parameter η); note η is known in an MRT. Denote the parameterized numerator of the weights by $\tilde{p}_t(a, a'|S_t; \rho)$ (with parameter ρ). The proof below allows the data analyst to use a $\tilde{p}_t$ with an estimated parameter $\tilde{\rho}$ or to pre-specify ρ as desired. We use a superscript of $\star$ to denote limiting values of estimated parameters (e.g. $\eta^\star, \rho^\star$). Then the more general version of the estimating

equation $U_W(\alpha, \beta; \hat{\eta}, \hat{\rho})$ is:

$$\sum_{t=1}^{T} \left(Y_{t+1,J} - g_t(H_t)^\top \alpha - (1 - A_{t,J})(A_{t,J'} - \tilde{p}_t^\star(1 \mid S_t; \hat{\rho})) f_t(S_t)' \beta \right) I_{t,J} I_{t,J'} W_{t,J,J'}(A_{t,J}, A_{t,J'}, H_t; \hat{\eta}, \hat{\rho}) \times$$

$$\begin{pmatrix} g_t(H_t) \\ (1 - A_{t,J})(A_{t,J'} - \tilde{p}_t^\star(1 \mid S_t; \hat{\rho})) f_t(S_t) \end{pmatrix}$$

Note $W_{t,J,J'}$ in the body of the paper is replaced here by $W_{t,J,J'}(A_t, H_t; \hat{\eta}, \hat{\rho})$, and $\hat{\eta}$, $\hat{\rho}$ are estimators.

Treatment Probability Model: If the data is observational then we assume: $p_t(a, a' \mid H_t; \eta)$ is a correctly specified model for $\mathbb{P}(a, a' \mid H_t, I_{t,J} = 1, I_{t,J'} = 1)$. Let $\eta^\star$ be the true value of η; that is, $Pr(A_{t,J}, A_{t,J'} \mid H_t, I_{t,J} = 1, I_{t,J'} = 1) = p_t(a, a' \mid H_t; \eta^\star)$ Assume that the estimator of η, say $\hat{\eta}$, satisfies $\mathbb{P}_n U_D(\hat{\eta}) = 0$ and $\sqrt{n}(\hat{\eta} - \eta^\star) = \mathbb{E}\left[\dot{U}_D(\eta^\star)\right]^{-1} \mathbb{P}_n U_D(\eta^\star) + o_P(1)$. Thus $\sqrt{n}(\hat{\eta} - \eta^\star)$ converges in distribution to a mean zero, normal random vector with variance-covaraince matrix given by $\mathbb{E}\left[\dot{U}_D(\eta^\star)\right]^{-1} \mathbb{E}[U_D(\eta^\star)^{\otimes 2}] \left(\mathbb{E}\left[\dot{U}_D(\eta^\star)\right]^{-1}\right)'$, which has finite entries. Assume that $\mathbb{P}_n(\dot{U}_D(\hat{\eta}))$ is a consistent estimator of $\mathbb{E}(\dot{U}_D(\eta^\star))$. Assume there exists finite constants, $b_D > 0$ and $B_D < 1$ such that each $b_D < p_t(a, a' \mid H_t; \eta^\star) < B_D$ a.s.

If the data analyst elects to use a parameterized and estimated $\tilde{p}_t(1 \mid S_t, \hat{\rho})$, then we assume:

Numerator of Weights Probability Model: Suppose the estimator $\hat{\rho}$ solves an estimating equation: $\mathbb{P}_n U_N(\rho) = 0$. Assume that, for a finite value of ρ, say $\rho^\star$ and $\sqrt{n}(\hat{\rho} - \rho^\star) = \mathbb{E}\left[\dot{U}_N(\rho^\star)\right]^{-1} \sqrt{n}(\mathbb{P}_n - P) U_N(\rho^\star) + o_P(1)$ where the matrix, $\mathbb{E}\left[\dot{U}_N(\rho^\star)\right]$ is positive definite. Assume $\sqrt{n}(\mathbb{P}_n - P) U_N(\rho^\star)$ converges in distribution to a mean zero, normal random vector with variance-covariance matrix given by $\mathbb{E}[U_N(\rho^\star)^{\otimes 2}]$ which has finite entries. Assume that $\mathbb{P}_n \dot{U}_N(\hat{\rho})$ is a consistent estimator of $\mathbb{E}[\dot{U}_N(\rho^\star)]$. Assume $0 < \rho^\star < 1$.

Proof. The solution to $\mathbb{P}_n U_W(\alpha, \beta; \hat{\eta}, \hat{\rho}) = 0$ gives the estimator:

$$\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = \left\{ \mathbb{P}_n \dot{U}_W(\hat{\eta}, \hat{\rho}) \right\}^{-1} \mathbb{P}_n \sum_{t=1}^{T} I_{t,J} I_{t,J'} W_{t,J,J'}(A_{t,J}, A_{t,J'}, H_t; \hat{\eta}, \hat{\rho}) Y_{t+1,J} \times$$

$$\begin{pmatrix} g_t(H_t) \\ (1 - A_{t,J})(A_{t,J'} - \tilde{p}_t^{\star}(1 \mid S_t)) f_t(S_t) \end{pmatrix}$$

where

$$\dot{U}_W(\hat{\eta}, \hat{\rho}) = \sum_{t=1}^{T} I_{t,J} I_{t,J'} W_{t,J,J'}(A_{t,J}, A_{t,J'}, H_t; \hat{\eta}, \hat{\rho}) \begin{pmatrix} g_t(H_t) \\ (1 - A_{t,J})(A_{t,J'} - \tilde{p}_t^{\star}(1 \mid S_t)) f_t(S_t) \end{pmatrix}^{\otimes 2}$$

Define

$$\begin{pmatrix} \alpha' \\ \beta' \end{pmatrix} = \left\{ \mathbb{E}\left[\dot{U}_W(\eta^{\star}, \rho^{\star}) \right] \right\}^{-1} \mathbb{E}\left[\sum_{t=1}^{T} I_{t,J} I_{t,J'} W_{t,J,J'}(A_{t,J}, A_{t,J'}, H_t; \eta^{\star}, \rho^{\star}) Y_{t+1,J} \right.$$

$$\left. \begin{pmatrix} g_t(H_t) \\ (1 - A_{t,J})(A_{t,J'} - \tilde{p}_t^{\star}(1 \mid S_t)) f_t(S_t) \end{pmatrix} \right]$$

Then standard statistical arguments can be used to show that $\sqrt{n}(\hat{\alpha} - \alpha', \hat{\beta} - \beta')$ converges in distribution to a normal, mean zero, random vector with variance-covariance matrix given by:

$$\left\{ \mathbb{E}\left[\dot{U}_W(\eta^{\star}, \rho^{\star}) \right] \right\}^{-1} \Sigma_W(\alpha', \beta'; \eta^{\star}, \rho^{\star}) \left\{ \mathbb{E}\left[\dot{U}_W(\eta^{\star}, \rho^{\star}) \right] \right\}^{-1}$$

where

$$\Sigma_W(\alpha, \beta; \eta, \rho) = \mathbb{E}\Big[\Big(U_W(\alpha, \beta; \eta, \rho) + \Sigma_{W,D}(\alpha, \beta; \eta, \rho)\big\{\mathbb{E}[\dot{U}_D(\eta)]\big\}^{-1} U_D(\eta) +$$
$$\Sigma_{W,N}(\alpha, \beta; \eta, \rho)\big\{\mathbb{E}[\dot{U}_N(\rho)]\big\}^{-1} U_N(\rho)\Big)^{\otimes 2}\Big]$$

with

$$\Sigma_{W,D}(\alpha, \beta; \eta, \rho) = \mathbb{E}\Big[\sum_{t=1}^{T}(Y_{t+1,J} - g_t(H_t)'\alpha - (1 - A_{t,J})(A_{t,J'} - \tilde{p}_t^\star(1 \mid S_t; \rho))f_t(S_t)'\beta)\, I_{t,J} I_{t,J'}$$

$$W_{t,J,J'}(A_{t,J}, A_{t,J'}, H_t; \eta, \rho) \begin{pmatrix} g_t(H_t) \\ (1 - A_{t,J})(A_{t,J'} - \tilde{p}_t^\star(1 \mid S_t; \rho))f_t(S_t) \end{pmatrix} \left(\frac{d\log p_t^\star(A_{t,J'} \mid H_t; \eta)}{d\eta}\right)'\Big],$$

and

$$\Sigma_{W,N}(\alpha, \beta; \eta, \rho) = \mathbb{E}\Big[\sum_{t=1}^{T}(Y_{t+1,J} - g_t(H_t)'\alpha - (1 - A_{t,J})(A_{t,J'} - \tilde{p}_t^\star(1 \mid S_t; \rho))f_t(S_t)'\beta)\, I_{t,J} I_{t,J'}$$

$$W_{t,J,J'}(A_{t,J}, A_{t,J'}, H_t; \eta, \rho) \begin{pmatrix} g_t(H_t) \\ (1 - A_{t,J})(A_{t,J'} - \tilde{p}_t^\star(1 \mid S_t; \rho))f_t(S_t) \end{pmatrix} \left(\frac{d\log \tilde{p}_t^\star(A_{t,J'} \mid S_t; \rho)}{d\rho}\right)'\Big]$$

$$+ \mathbb{E}\Big[\sum_{t=1}^{T}(Y_{t+1,J} - g_t(H_t)'\alpha - (1 - A_{t,J})(A_{t,J'} - \tilde{p}_t^\star(1 \mid S_t; \rho))f_t(S_t)'\beta)\, I_{t,J} I_{t,J'}$$

$$W_{t,J,J'}(A_{t,J}, A_{t,J'}, H_t; \eta, \rho) \begin{pmatrix} \mathbf{0} \\ -(1 - A_{t,J})\tilde{p}_t^\star(1 \mid S_t; \rho)f_t(S_t) \end{pmatrix} \left(\frac{d\log \tilde{p}_t^\star(1 \mid S_t; \rho)}{d\rho}\right)'\Big]$$

$$+ \mathbb{E}\Big[\sum_{t=1}^{T}(1 - A_{t,J})\tilde{p}_t^\star(1 \mid S_t; \rho)f_t(S_t)'\beta I_{t,J} I_{t,J'} W_{t,J,J'}(A_{t,J}, A_{t,J'}, H_t; \eta, \rho)$$

$$\begin{pmatrix} g_t(H_t) \\ (1 - A_{t,J})(A_{t,J'} - \tilde{p}_t^\star(1 \mid S_t; \rho))f_t(S_t) \end{pmatrix} \left(\frac{d\log \tilde{p}_t^\star(1 \mid S_t; \rho)}{d\rho}\right)'\Big]$$

In our simulation, an individual's randomization probabilities only depends on their observed history, then $\tilde{p}_t^\star(1 \mid S_{t,j'}) = \tilde{p}_t(1 \mid S_{t,j'})$. Since the data is from an MRT (we know

41

p_t) and we pre-specify (not estimate) $\tilde{p}_t$, then $\Sigma_W = \mathbb{E}\left[(U_W(\alpha,\beta))^{\otimes 2}\right]$ greatly simplifying the variance-covaraince matrix.

A consistent estimator of the variance-covariance matrix is given by:

$$\left\{\mathbb{P}_n\left[\dot{U}_W(\hat{\eta},\hat{\rho})\right]\right\}^{-1}\hat{\Sigma}_W(\hat{\alpha},\hat{\beta};\hat{\eta},\hat{\rho})\left\{\mathbb{P}_n\left[\dot{U}_W(\hat{\eta},\hat{\rho})\right]\right\}^{-1} \tag{11}$$

where

$$\hat{\Sigma}_W(\alpha,\beta;\eta,\rho) = \mathbb{P}_n\Bigg[\Bigg(U_W(\alpha,\beta;\eta,\rho) + \hat{\Sigma}_{W,D}(\alpha,\beta;\eta,\rho)\left\{\mathbb{P}_n[\dot{U}_D(\eta)]\right\}^{-1}U_D(\eta) +$$
$$\hat{\Sigma}_{W,N}(\alpha,\beta;\eta,\rho)\left\{\mathbb{P}_n[\dot{U}_N(\rho)]\right\}^{-1}U_N(\rho)\Bigg)^{\otimes 2}\Bigg]$$

with

$$\hat{\Sigma}_{W,D}(\alpha,\beta;\eta,\rho) = \mathbb{P}_n\Bigg[\sum_{t=1}^{T}(Y_{t+1,J} - g_t(H_t)'\alpha - (1-A_{t,J})(A_{t,J'} - \tilde{p}_t^{\star}(1\,|\,S_t;\rho))f_t(S_t)'\beta)\,I_{t,J}I_{t,J'}$$

$$W_{t,J,J'}(A_t,H_t;\eta,\rho)\begin{pmatrix} g_t(H_t) \\ (1-A_{t,J})(A_{t,J'} - \tilde{p}_t^{\star}(1\,|\,S_t;\rho))f_t(S_t) \end{pmatrix}\left(\frac{d\log p_t^{\star}(A_{t,J'}|H_t;\eta)}{d\eta}\right)'\Bigg],$$

and

$$\hat{\Sigma}_{W,N}(\alpha,\beta;\eta,\rho) = \mathbb{P}_n\Big[\sum_{t=1}^T (Y_{t+1,J} - g_t(H_t)'\alpha - (1-A_{t,J})(A_{t,J'} - \tilde{p}_t^\star(1\,|\,S_t;\rho))f_t(S_t)'\beta)\,I_{t,J}I_{t,J'}$$

$$W_{t,J,J'}(A_{t,J}, A_{t,J'}, H_t; \eta, \rho) \begin{pmatrix} g_t(H_t) \\ (1-A_{t,J})(A_{t,J'} - \tilde{p}_t^\star(1\,|\,S_t;\rho))f_t(S_t) \end{pmatrix} \left(\frac{d\log\tilde{p}_t^\star(A_{t,J'}|S_t;\rho)}{d\rho}\right)'\Big]$$

$$+ \mathbb{P}_n\Big[\sum_{t=1}^T (Y_{t+1,J} - g_t(H_t)'\alpha - (1-A_{t,J})(A_{t,J'} - \tilde{p}_t^\star(1\,|\,S_t;\rho))f_t(S_t)'\beta)\,I_{t,J}I_{t,J'}$$

$$W_{t,J,J'}(A_{t,J}, A_{t,J'}, H_t; \eta, \rho) \begin{pmatrix} \mathbf{0} \\ -(1-A_{t,J})\tilde{p}_t^\star(1\,|\,S_t;\rho)f_t(S_t) \end{pmatrix} \left(\frac{d\log\tilde{p}_t^\star(1|S_t;\rho)}{d\rho}\right)'\Big]$$

$$+ \mathbb{P}_n\Big[\sum_{t=1}^T (1-A_{t,J})\tilde{p}_t^\star(1\,|\,S_t;\rho)f_t(S_t)'\beta I_{t,J}I_{t,J'}W_{t,J,J'}(A_{t,J}, A_{t,J'}, H_t; \eta, \rho)$$

$$\begin{pmatrix} g_t(H_t) \\ (1-A_{t,J})(A_{t,J'} - \tilde{p}_t^\star(1\,|\,S_t;\rho))f_t(S_t) \end{pmatrix} \left(\frac{d\log\tilde{p}_t^\star(1|S_t;\rho)}{d\rho}\right)'\Big]$$

It remains to show that $\beta' = \beta^{\star\star}$. Since $\mathbb{E}[U_W(\alpha', \beta'; \eta^\star, \rho^\star)] = 0$,

$$0 = \mathbb{E}\sum_{t=1}^T \left(Y_{t+1,J} - g_t(H_t)^\top\alpha' - (1-A_{t,J})(A_{t,J'} - \tilde{p}_t^\star(1\,|\,S_t;\rho^\star))f_t(S_t)^\top\beta'\right) I_{t,J}I_{t,J'}$$

$$W_{t,J,J'}(A_{t,J}, A_{t,J'}, H_t; \eta^\star, \rho^\star)(1-A_{t,J})(A_{t,J'} - \tilde{p}_t^\star(1\,|\,S_t;\rho^\star))f_t(S_t)$$

$$= \mathbb{E}\sum_{t=1}^T \left(\mathbb{E}\left[Y_{t+1}\,|\,A_{t,J}, A_{t,J'}, H_t, I_{t,J}I_{t,J'} = 1\right] - g_t(H_t)^\top\alpha' - (1-A_{t,J})(A_{t,J'} - \tilde{p}_t^\star(1\,|\,S_t;\rho^\star))f_t(S_t)^\top\beta'\right)$$

$$I_{t,J}I_{t,J'}W_{t,J,J'}(A_{t,J}, A_{t,J'}, H_t; \eta^\star, \rho^\star)(1-A_{t,J})(A_{t,J'} - \tilde{p}_t^\star(1\,|\,S_t;\rho^\star))f_t(S_t)$$

$$= \mathbb{E}\sum_{t=1}^T \sum_{a'\in\{0,1\}} \left(\mathbb{E}\left[Y_{t+1}\,|\,A_{t,J} = 0, A_{t,J'} = a', H_t, I_{t,J}I_{t,J'} = 1\right] - g_t(H_t)^\top\alpha' -\right.$$

$$\left.(a' - \tilde{p}_t^\star(1\,|\,S_t;\rho^\star))f_t(S_t)^\top\beta'\right) I_{t,J}I_{t,J'}\tilde{p}_t(0, a'\,|\,S_t;\rho^\star)(a' - \tilde{p}_t^\star(1\,|\,S_t;\rho^\star))f_t(S_t)$$

where the last equality averages out over $A_{t,J'}$. The above simplifies to:

$$0 = \mathbb{E} \sum_{t=1}^{T} \sum_{a' \in \{0,1\}} \Big(\mathbb{E}\left[Y_{t+1} \mid A_{t,J} = 0, A_{t,J'} = a', H_t, I_{t,J} I_{t,J'} = 1\right] - g_t(H_t)^\top \alpha' -$$

$$(a' - \tilde{p}_t^\star(1 \mid S_t; \rho^\star)) f_t(S_t)^\top \beta' \Big) I_{t,J} I_{t,J'} \tilde{p}_t(0, a' \mid S_t; \rho^\star)(a' - \tilde{p}_t^\star(1 \mid S_t; \rho^\star)) f_t(S_t)$$

$$= \mathbb{E} \sum_{t=1}^{T} \Big(\mathbb{E}\left[Y_{t+1} \mid A_{t,J} = 0, A_{t,J'} = 1, H_t, I_{t,J} I_{t,J'} = 1\right] - g_t(H_t)^\top \alpha' -$$

$$(1 - \tilde{p}_t^\star(1 \mid S_t; \rho^\star)) f_t(S_t)^\top \beta' \Big) I_{t,J} I_{t,J'} \tilde{p}_t(0, 1 \mid S_t; \rho^\star)(1 - \tilde{p}_t^\star(1 \mid S_t; \rho^\star)) f_t(S_t)$$

$$+ \mathbb{E} \sum_{t=1}^{T} \Big(\mathbb{E}\left[Y_{t+1} \mid A_{t,J} = 0, A_{t,J'} = 0, H_t, I_{t,J} I_{t,J'} = 1\right] - g_t(H_t)^\top \alpha' -$$

$$(-\tilde{p}_t^\star(1 \mid S_t; \rho^\star)) f_t(S_t)^\top \beta' \Big) I_{t,J} I_{t,J'} \tilde{p}_t(0, 0 \mid S_t; \rho^\star)(-\tilde{p}_t^\star(1 \mid S_t; \rho^\star)) f_t(S_t)$$

$$= \mathbb{E} \sum_{t=1}^{T} \Big(\mathbb{E}\left[Y_{t+1} \mid A_{t,J} = 0, A_{t,J'} = 1, H_t, I_{t,J} I_{t,J'} = 1\right] - \mathbb{E}\left[Y_{t+1} \mid A_{t,J} = 0, A_{t,J'} = 0, H_t, I_{t,J} I_{t,J'} = 1\right]$$

$$- f_t(S_t)^\top \beta' \Big) f_t(S_t) \gamma(\eta^\star, \rho^\star) I_{t,J} I_{t,J'}$$

where $\gamma(\eta^\star, \rho^\star) = \tilde{p}_t(0, 1 \mid S_t)(1 - \tilde{p}_t^\star(1 \mid S_t; \rho^\star)) = \tilde{p}_t(0, 0 \mid S_t) \tilde{p}_t^\star(1 \mid S_t; \rho^\star)$. From this we obtain:

$$0 = \mathbb{E} \sum_{t=1}^{T} f_t(S_t) \gamma(\eta^\star, \rho^\star) I_{t,J} I_{t,J'} \Big(\mathbb{E}\Big[\mathbb{E}\left[Y_{t+1} \mid A_{t,J} = 0, A_{t,J'} = 1, H_t, I_{t,J} I_{t,J'} = 1\right] -$$

$$\mathbb{E}\left[Y_{t+1} \mid A_{t,J} = 0, A_{t,J'} = 0, H_t, I_{t,J} I_{t,J'} = 1\right] \mid S_t, I_{t,J} I_{t,J'} = 1\Big] - f_t(S_t)^\top \beta' \Big)$$

Thus

$$\beta' = \left[\mathbb{E}\dot{U}_W(\eta^\star, \rho^\star) \right]_{(2,2)}^{-1} \mathbb{E}\Big[\sum_{t=1}^{T} f_t(S_t) \gamma(\eta^\star, \rho^\star) I_{t,J} I_{t,J'} \mathbb{E}\Big[\mathbb{E}\left[Y_{t+1} \mid A_{t,J} = 0, A_{t,J'} = 1, H_t, I_{t,J} I_{t,J'} = 1\right] -$$

$$\mathbb{E}\left[Y_{t+1} \mid A_{t,J} = 0, A_{t,J'} = 0, H_t, I_{t,J} I_{t,J'} = 1\right] \mid S_t, I_{t,J} I_{t,J'} = 1\Big] \Big]$$

$$(12)$$

where

$$\left[\mathbb{E}\dot{U}_W(\eta^\star, \rho^\star)\right]_{(2,2)} = \mathbb{E}\sum_{t=1}^{T} f_t(S_t)f_t(S_t)^\top \gamma(\eta^\star, \rho^\star)I_{t,J}I_{t,J'}$$

$\square$

F   Code to Generate Simulation Results

The R code used to generate the simulation experiments and case study results in this paper can be obtained here (https://github.com/Herashi/MRT).