Dynamic Treatment Regimes and Interference in Dyadic Networks: A Joint Optimization Approach

by

Marzieh Mussavi Rizi

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Statistics - Biostatistics

Waterloo, Ontario, Canada, 2023

© Marzieh Mussavi Rizi 2023

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	Wenqing He Professor, Dept. Statistics and Actuarial Science, Western University
Supervisor(s):	Joel A. Dubin Professor, Dept. Statistics and Actuarial Science, University of Waterloo
	Michael P. Wallace Associate Professor, Dept. Statistics and Actuarial Science, University of Waterloo
Internal Member:	Cecilia Cotton Associate Professor, Dept. Statistics and Actuarial Science, University of Waterloo
	Liqun Diao Assistant Professor, Dept. Statistics and Actuarial Science, University of Waterloo
Internal-External Member:	Mark Ferro Associate Professor, School of Public Health Science, University of Waterloo

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Identifying interventions that are optimally tailored to each individual is of significant interest in various fields, in particular precision medicine. Dynamic treatment regimes (DTRs) employ sequences of decision rules that utilize individual patient information to recommend treatments. However, the assumption that an individual's treatment does not impact the outcomes of others, known as the *no interference* assumption, is often challenged in practical settings. For example, in infectious disease studies, the vaccine status of individuals in close proximity can influence the likelihood of infection. Imposing this assumption when it, in fact, does not hold, may lead to biased results and impact the validity of our resulting DTR optimization. In this thesis, we extend the estimation method of dynamic weighted ordinary least squares (dWOLS), a doubly robust and easily implemented approach for estimating optimal DTRs, to incorporate the presence of interference. Specifically, we develop new methodologies to optimize DTRs in the presence of interference for both binary and continuous treatments. Through comprehensive simulations and analysis of the Population Assessment of Tobacco and Health (PATH) data, we demonstrate the performance of the proposed joint optimization strategy compared to the current state-of-the-art conditional optimization methods. Furthermore, we extend the dWOLS method to accommodate multiple outcomes and patient-specific costs, enhancing its flexibility and applicability in complex health contexts.

Acknowledgements

I would like to express my sincere gratitude to my supervisors, Professors Joel Dubin and Michael Wallace, for their invaluable guidance and mentorship throughout the entire thesis journey. Their support and expertise have been instrumental in shaping and refining my research. Additionally, I would like to extend my thanks to Professors Cecilia Cotton and Liqun Diao for their dedicated service on my examining committee. I am also grateful to Professor Wenqing He and Mark Ferro for serving as the external examiners. Their insights and feedback have greatly enriched the quality of my work.

Table of Contents

Examining Committee Membership ii

Author's Declaration iii

Abstract iv

Acknowledgements v

List of Figures x

List of Tables xii

1 Introduction 1

1.1 Precision Medicine and Dynamic Treatment Regimens 1

1.2 Causal Inference and Interference 3

1.3 DTR and Multiple Outcomes 5

1.4 Motivating Example 7

1.5 Objectives and Thesis Outline 8

2 Methodological Background 10

2.1 Causal Inference in Observational Studies 10

2.2 Optimal DTRs Estimation . . . 12
2.3 dWOLS Estimation Method for a K-stage DTR without Interference . . . 13
2.4 Dyadic Data Structure . . . 16

3 DTRs on Dyadic Networks with Binary Treatments 18
3.1 Introduction . . . 18
3.2 Methodology . . . 20
3.2.1 Notation . . . 20
3.2.2 Causal framework . . . 22
3.2.3 Modelling assumptions . . . 23
3.2.4 Modelling a K-stage dynamic treatment regime in a dyadic network 25
3.3 Estimation and Determination of Optimal Treatment Regime . . . 28
3.3.1 Balancing weights . . . 29
3.3.2 Estimation of model parameters . . . 30
3.3.3 Optimal treatment assignment and the dyad-health function . . . 32
3.4 Simulation Study . . . 34
3.4.1 Simulation I: double robustness . . . 34
3.5 Application: PATH Study . . . 37
3.5.1 Results . . . 40
3.6 Discussion . . . 42

4 DTRs on Dyadic Networks with Continuous Treatments 45
4.1 Introduction . . . 45
4.2 Methodology . . . 47
4.2.1 Notation and assumptions . . . 47
4.2.2 dWOLS estimation method for a K-stage DTR in a dyadic network 49

4.3 Simulation Study . . . 54
4.3.1 Results . . . 55
4.4 Application: PATH Study . . . 58
4.4.1 Analysis . . . 59
4.5 Discussion . . . 62

5 DTRs and Multiple Outcomes **64**
5.1 Introduction . . . 64
5.2 Methodology . . . 65
5.2.1 Notation . . . 65
5.2.2 Constrained dWOLS for multiple outcomes . . . 66
5.3 Simulation . . . 68
5.3.1 Simulation 1: the impact of constrained optimization . . . 69
5.3.2 Simulation 2: double robustness . . . 70
5.4 Discussion . . . 71

6 Discussion **75**
6.1 Conclusion . . . 75
6.2 Future Work . . . 76

References **78**

APPENDICES **88**

A Chapter 3 Appendix **89**
A.1 Double Robustness . . . 89
A.1.1 Treatment-free model correctly specified, treatment models misspecified . . . 90

A.1.2 Treatment-free model misspecified, treatment models correctly specified . . . 90
A.1.3 Interference models and weights estimation . . . 92
A.2 General Interference Model for a Two-stage DTR . . . 94
A.3 Simulation . . . 97
A.3.1 Single-stage regime . . . 97
A.3.2 Two-stage regime . . . 98

B Chapter 4 Appendix **101**
B.1 Double Robustness Proof . . . 101
B.2 Extra Simulation Results . . . 102
B.3 Weights . . . 109

List of Figures

3.1 Model 1: General Interference Model 23

3.2 Model 2: Restricted Model 24

3.3 Model 3 (Independent Treatment Model 25

3.4 Dyadic-blip function parameter estimates for 1000 simulated datasets . . . 35

3.5 Estimated expected population outcome under conditional and joint optimization 37

3.6 Boxplot of nicotine exposure 42

4.1 Illustration of the general shape of the outcome model as function of treatments 52

4.2 The estimates for the blip parameter ψ_1 for different sample sizes. 56

4.3 The estimates for ψ_1 under different model specifications 57

4.4 Boxplots of the total tobacco dependence scores 62

5.1 The distribution of outcomes $Y_{(1)}$ and $Y_{(2)}$ under different optimization . . 70

5.2 Estimates of the blip parameter ψ_0 under different model misspecification, Simulation 2 72

5.3 Estimates of the blip parameter ψ_1 under different model misspecification, Simulation 2 73

A.1 General Interference Model for two stages 95

A.2 Dyadic-blip function parameter estimates visa dyadic-dWOLS approach . . 97

A.3 Dyadic-blip function parameter estimates visa dyadic-dWOLS approach: first stage 99

A.4 Dyadic-blip function parameter estimates visa dyadic-dWOLS approach: second stage 100

B.1 Dyadic-blip parameter estimates for 1000 simulated datasets via the dyadic-dWOLS approach treatment-free models are misspecified and the treatments are normal. 103

B.2 Dyadic-blip parameter estimates for 1000 simulated datasets via the dyadic-dWOLS approach treatment-free models are misspecified and the treatments are normal. 104

B.3 Dyadic-blip parameter estimates for 1000 simulated datasets via the dyadic-dWOLS approach treatment models are misspecified and the treatments are normal. 105

B.4 Dyadic-blip parameter estimates for 1000 simulated datasets via the dyadic-dWOLS approach treatment models are misspecified and the treatments are normal. 106

List of Tables

2.1 An example of standard dyadic data structure 17

3.1 Dyadic-Blip Estimates 41

4.1 Weight functions implemented in the dyadic-dWOLS. 56

4.2 RMSE for different model misspecification 58

4.3 Dyadic-Blip Estimates 60

A.1 Mean, standard error, bias, root mean squared error of the blip estimators 96

B.1 RMSE: Normal treatment, treatment free model is misspecified. Sample size is 1000 dyads and the RSME are multiplied by 10^2 107

B.2 Log normal treatment. Treatment model is misspecified. Sample size is 1000 dyads and the RSME are multiplied by 10^2 107

B.3 RMSE: Log normal treatment, treatment free model is misspecified. Sample size is 1000 dyads and the RSME are multiplied by 10^2 108

B.4 Log Normal treatment, both treatment and treatment-free models are correct. Sample size is 1000 dyads and the RSME are multiplied by 10^2 108

B.5 RMSE: Log normal treatment, both models are misspecified. Sample size is 1000 dyads and the RSME are multiplied by 10^2. 109

B.6 Weight functions implemented in the dyadic-dWOLS 110

Chapter 1

Introduction

1.1 Precision Medicine and Dynamic Treatment Regimens

Treatment of a chronic disease or condition often involves a series of decisions for which multiple long-term strategies are possible. The decision of which regime of treatments to pursue is often made by a medical professional, who chooses an appropriate treatment regime based on metrics that, historically, emphasize the illness. *Precision medicine* is an approach that aims to improve upon the traditional method of treatment selection by taking into account additional relevant factors. The basic idea is that the efficacy of a particular treatment may be influenced by certain individual-level characteristics of each person, such that the effectiveness of the treatment will vary across individuals with different values of these characteristics. By incorporating relevant information into the treatment decision, precision medicine enables clinicians to select the most suitable treatment for each individual with greater precision. The focus on individual-level characteristics is why precision medicine is sometimes also known as personalized medicine [1].

Models developed for analyzing the relationship between treatments and individual-level characteristics are known as *dynamic treatment regimes* (DTRs). A DTR is a sequence of decision rules each of which considers individual-level information thought to impact the effectiveness of the treatments under consideration [2]. A central focus of DTR researchers

is devising methodologies for the estimation of *optimal* DTRs [3]. What is meant by optimal is contextual, but a typical approach is to consider it in terms of the individual's expected health outcome under the estimated regime.

Within the general framework of causal inference, DTR estimation can be approached through two main methodologies: classification- based and regression-based methods. Classification-based methods, also known as direct methods or value search methods, directly estimate the marginal mean outcome of a treatment regime and identify the optimal DTR by maximizing the estimated value across all possible DTRs [4]. In contrast, regression-based methods, often referred to as indirect methods, model and estimate the conditionally expected outcome to derive the optimal DTR, aiming to maximize the expected outcome.

Regression-based methods are favored for their foundation in regression techniques and ease of implementation. One common regression-based method, known as Q-learning [5, 6], originates from the reinforcement learning literature. This approach transforms the problem of estimating the decision rules of a DTR into a sequence of computationally frugal optimization problems. Although this method is relatively straightforward to implement, it suffers from a lack of robustness to model misspecification.

An approach that improves on this shortcoming of Q-learning is G-estimation [7]. G-estimation is similar to Q-learning in many ways (see for example [8]), but poses some key differences which, most notably, give it a property called *double robustness*. In particular, this property provides estimation with robustness against nuisance models misspecification. The robustness of G-estimation, and the flexibility to accommodate nonlinear modelling strategies, makes it an attractive procedure for estimating optimal DTRs. The main downside of the methodology stems from the complexity, making description and implementation more cumbersome than the previously specified techniques.

Of particular interest for our purposes is the *dynamic weighted ordinary least squares* (dWOLS) approach [9]. This methodology combines the strengths of G-estimation and Q-learning to offer the property of double-robustness while remaining comparatively straightforward to implement. The dWOLS algorithm has been extended to accommodate continuous treatments, as well as time-to-event outcomes [10, 11].

DTRs have shown great promise in application. For example, some recent studies have

implemented DTRs in the context of treatments for cancer [12], epilepsy [13], depression [14], schizophrenia [15], HIV infection [16], smoking cessation [8], and childhood amblyopia [17].

1.2 Causal Inference and Interference

The statistical framework underlying the study of DTRs relies on fundamental concepts of causal inference. In this thesis, we adopt the potential outcome framework. Most causal analyses depend on the assumption that each participant possesses a single potential outcome under each possible treatment assignment. A potential outcome, sometimes called counterfactual outcome, is then the outcome that an individual would have if, possibly counter to fact, they were to have a specific treatment. The outcome is considered *potential* because it is only realized (i.e. observed) if the individual is administered the associated treatment. This is one of the foundational principles of the counterfactual approach to inference, where the comparison of observed outcomes to unobservable counterfactual outcomes plays a central role in assessing causal relationships. When assessing causal effects, one of the main assumptions is that the outcome of an individual does not depend on the treatment assignments of other individuals [18]. This is more commonly known as the *no interference* assumption, introduced by Cox [18]. Rubin [19] combined the assumption of no interference with the assumption of "no hidden versions of treatment" and called it the *stable unit treatment value assumption* (SUTVA). The no interference assumption, while quite useful in classical scenarios such as in clinical trials, appears implausible and often detrimental in settings where more complex interactions exist between the individuals under study.

For example, the interference arises in the study of infectious diseases. In this scenario, an individual's treatment (i.e. vaccination status) directly impacts their likelihood of infection; however, so does the vaccination status of those people around them, as this directly correlates with the ability of the virus to spread. Generalizing this scenario, public health interventions may impact those in close physical or social proximity to those recipients of the intervention, even if they have not received the intervention themselves, or further, are even unaware that someone around them has received the intervention. Hence,

the assumptions of the potential outcomes framework need to be reworked when the no interference assumption does not hold [20].

Halloran and Struchiner [21] are among the earlier examples of researchers accommodating the effect of interference in causal inference. They propose individual-level causal estimands that account for interference by letting the potential outcome for any individual depend on the vector of treatment assignments of other individuals. However, even in the simple scenario of a binary treatment and a binary outcome, assuming there are N people in the population, there are 2^N possible treatment vectors and 2^N possible potential outcomes. Causal inference in the presence of so many potential outcomes appears difficult, if not impossible, without the introduction of additional assumptions and constraints.

One possible additional assumption which simplifies the problem, sometimes to manageable levels, is the assumption that individuals can be partitioned into groups (or clusters), upon which one assumes that interference exists among the members of the same group, but where no interference extends beyond the confines of a given group. Sobel [22] referred to this assumption as *partial interference*. Under the assumption of partial interference, cluster-randomized trials are sometimes designed to rule out the presence of interference between treated and untreated clusters [23–25]. Another design implemented to incorporate partial interference is the sequential two-stage randomization design, which has been proposed to estimate both treatment and spillover effects and consists of two stages. In the first stage, groups are randomized to different treatment allocation strategies, and in the second stage, individuals are randomized to treatment or control conditional on the strategy assigned to their group in the first stage [26].

Although more general than no interference, there are still many settings where partial interference is not a plausible model for interaction within the population. In these cases, the literature typically turns to the most general form of interference, aptly called *general interference*. Recent approaches for assessing treatment effects in the presence of general interference, include Rosenbaum [27], Toulis and Kao [28], Ugander et al. [29], Eckles et al. [30], Sofrygin and van der Laan [31], Aronow et al. [32], and Liu et al. [33] among others. These general interference methods typically assume one finite population of N individuals. For each individual, a set of other individuals is defined who are permitted or assumed to interfere with that individual. Different names have been used to describe these sets, including interference sets [33], neighbourhoods [28, 29], and friends of the individuals [31].

In the context of randomized experiments, Aronow et al. [32] consider the presence of arbitrary interference, defining an exposure mapping that relates the individual experimental treatment assignments to the exposures received by individuals depending on with whom they interact and other underlying features of the population. Ugander et al. [29], and Eckles et al. [30] also considered randomized experiments on networks, where treatment of one individual could have impact on neighbouring individuals, and estimate causal effects under this general interference framework.

Another way to model the inclusion of interference is through observational network data, which is an emerging approach in causal inference. Liu et al. [33] focused on direct and indirect effects based on Bernoulli allocation strategies similar to those defined by Tchetgen and VanderWeele [34]. Although their estimators allow for general forms of interference, asymptotic results are proved under partial interference and clustered data. Sofrygin et al. [35] proposed a maximum-likelihood–based approach called *targeted maximum likelihood estimator* (TMLE), to estimate treatment and interference effects and proved asymptotic results under the assumption of independent and identically distributed data. Finally, Ogburn et al. [36] extended this TMLE estimator to allow for dependence due to contagion, where one's treatment affects the outcome of another individual indirectly by changing the treated individual's own outcome and derived asymptotic results. In a recent study Forastiere et al. [37] develop new covariate adjustment methods that lead to valid estimates of treatment and interference effects in observational studies on networks.

1.3 DTR and Multiple Outcomes

The majority of studies in the field of DTR have primarily focused on optimizing a single scalar outcome. However, in many clinical decision settings, there is a need to balance multiple, potentially competing, outcomes to provide comprehensive care. For instance, in the treatment of depression, it becomes important to maximize efficacy while minimizing adverse effects such as sleep disturbance, sexual dysfunction, and weight gain. Similarly, in bipolar depression, managing depressive symptoms needs to be balanced with the risk of inducing mania [38] or when treating schizophrenia, there exists a trade-off between the side effects of medication and its efficacy [39]. A cancer patient may seek to strike a

balance between the effectiveness of chemotherapy, the financial burden, and the impact on their quality of life [40].

Recent work on precision medicine with multiple outcomes encompasses several approaches, including (i) the construction of *composite outcomes*, (ii) the utilization of *set-valued treatment regimes*, and (iii) the application of *constrained estimation methods*. A composite outcome combines multiple outcomes into a single measure, enabling analysis and optimization using established methodologies. However, constructing a high-quality composite outcome without considering individuals' preferences could lead to misspecification of the composite outcome. Such misspecification of the composite outcome can significantly impact the accuracy of estimated dynamic treatment regimes [4]. To address this, Butler et al. [39] proposed a method in the single-decision setting that incorporates patient preferences using item response theory and Q-learning to optimize the patient-preferred composite outcome. Their proposed composite outcome is a convex combination of two outcomes. Building upon the same class of composite outcomes, Luckett et al. [38] studied situations where clinicians have imperfect information about patient preferences and estimated the corresponding preferences as a function of patient-level covariates and the probability of optimal treatment assignment [38].

Set-valued treatment regimes have been proposed as another approach to address multiple outcomes [4, 41, 42]. Such treatment regimes, map the patient's current information to a subset of treatment options that are not uniformly worse across all outcomes. By selecting a single treatment from the recommended set based on patient preferences and clinical judgment, set-valued treatment regimes incorporate patient preferences without directly eliciting them. While set-valued treatment regimes consider competing outcomes, such as treatment efficacy and the risk of adverse events, they do not directly handle constraints.

Another set of methodologies for estimating optimal treatment regimes with multiple outcomes aim to maximize a primary outcome while constraining a secondary outcome within an acceptable region [43, 44]. For example, Linn et al. [43] proposed maximizing the expectation of one outcome while incorporating constraints on features of the other outcomes' distributions. By formulating the estimation as a constrained optimization problem, they provide a framework that explicitly accounts for competing outcomes. This approach enables the exploration of trade offs among treatment strategies by adjusting the constraints and re-estimating the optimal regime.

1.4 Motivating Example

Despite the fact that interference among individuals is a fairly active field of study among causal inference researchers, few studies have been conducted to apply these developed methods to DTR. Most of the existing methods used for estimating optimal DTRs assume the treatment of one individual has no impact on the outcome of any other individual under consideration. However, this assumption may not hold true in all scenarios where DTR modelling is applied. One notable example is smoking behavior. It has been well-documented that smoking behavior tends to cluster within families. Moreover, the effectiveness of smoking cessation interventions can be heavily influenced by social factors. Studies have shown that family members play a significant role in successful smoking cessation. For instance, research suggests that cancer patients who quit smoking tobacco face greater difficulty remaining abstinent if other family members are also tobacco users [45].

Couples who smoke may encounter unique challenges in their smoking cessation journey. In cases where smoking tendencies are not synchronized between partners, achieving desired outcomes becomes more difficult [46]. Having a partner who was consistently a non smoker during the period of study, was associated with a greater likelihood of quitting smoking [47]. Furthermore, a supportive social network has been identified as a crucial element for achieving successful smoking cessation outcomes [48]. Such support can provide encouragement, accountability, and practical assistance during the quitting process. Understanding the dynamics of social support and its contribution to smoking cessation is essential for tailoring interventions to promote positive health behaviors [49]. Thus, incorporating an understanding of the impact of couple dynamics on smoking behavior can enhance the design and implementation of interventions aimed at smoking cessation.

Furthermore, it is commonplace in smoking behavior studies to assess multiple outcomes, including nicotine dependence and attempts to quit. This aspect presents an intriguing avenue for exploration within our research. Understanding the potential influence of optimal decision-making in smoking behaviors on other outcomes becomes vital in designing effective smoking cessation interventions. Moreover, studies have demonstrated that the smoking behaviors of couples can extend their impact beyond smoking itself, affecting outcomes such as relationship satisfaction [50, 51]. This additional dimension adds further significance to the study of the multiple outcome framework within dyads, where

interference exists and may impact various (and possibly competing) outcomes of both individuals.

These social complexities of smoking behavior, motivates us to investigate DTR and possible interference effect in this setting. Through a detailed analysis of the Population Assessment of Tobacco and Health (PATH) dataset, in the households with two tobacco users, we consider (i) whether the adoption of e-cigarettes by an individual impacts their nicotine exposure, and (ii) whether the nicotine exposure of an individual impacts their tobacco dependence.

1.5 Objectives and Thesis Outline

Despite the growing literature on accommodating causal inference techniques to include interference, implementing these ideas into the DTR framework is still a relatively underdeveloped area of research. In this thesis, we extend the existing methodology of dWOLS to incorporate interference in the form of *dyadic network structures*–a network in which each individual is connected with exactly one other individual to form a dyad, or a pair.

The remainder of this thesis is structured as follows: Chapter 2 provides the methodological and theoretical background, including notation, as well as a brief discussion of estimating optimal DTRs under the dWOLS method without interference.

In Chapter 3, We extend the dWOLS method to accommodate the presence of interference within a dyadic-network population setting. Additionally, an appropriate outcome model is formalized, and the estimation of an optimal decision rule in the dyadic-network context is described. Furthermore, we demonstrate via simulation that this new dyadic decision rule performs better than existing methods at estimating the correct optimal treatment assignments when interference exists within dyads. The effectiveness of these decision rules is illustrated by their application to the Population Assessment of Tobacco and Health (PATH) dataset, showcasing its improved performance compared to existing methods in estimating the optimal treatment assignments when within-dyad interference exists.

Chapter 4 then discusses the necessary changes that allow extending the methodology of Chapter 3 to the cases when the treatment is continuous. In this chapter we propose

one of the first efforts at a DTR methodology which allows for comprehensive modelling of continuous treatments alongside participants interference. Moreover, our method enables the modelling of varying degrees of exposure to spillover effects, by considering both symmetric and asymmetric network connections, and those characterized by heterogeneous intensities. The proposed methodology aims to produce doubly robust estimates of the treatment and spillover effects of a continuous treatment, and quantify the corresponding optimal decision rule. We demonstrate the performance of the proposed model through a detailed analysis of the PATH dataset. Specifically, in households with two tobacco users, we examine whether the nicotine exposure of an individual impacts their tobacco dependence and whether a spillover effect exists.

Chapter 5 addresses the need to consider multiple outcomes within DTRs and extends the dWOLS methodology to tackle this important aspect. The decision-making process in healthcare often involves considering not only the primary outcome but also potential costs or side effects associated with different treatments. Motivated by the recognition that optimal decision-making should account for multiple dimensions of treatment effectiveness, we explored the incorporation of multiple outcomes in the context of DTRs. Chapter 5 contributes to the field by extending the dWOLS method to enable the optimization of treatment decisions when multiple outcome measures are involved.

Chapter 6 of this thesis offers a discussion of the findings and implications of the research conducted in earlier chapters. It also provides insights into potential future directions for further exploration that extend beyond the scope of the thesis. The appendixes contain additional details, including proofs and supplementary simulation results.

Chapter 2

Methodological Background

2.1 Causal Inference in Observational Studies

The study of causality has been considered from different perspectives within different disciplines. One particular point of view is the potential outcome's framework which several researchers advocated, including Rubin [52, 53] and Robins [54]. Another is the directed acyclic graph-centric approach of Pearl [55]. Each perspective finds frequent use within the literature, where adoption is typically motivated by the ease of implementation. We adopt the potential outcomes framework in the research in this thesis. The remainder of this section is dedicated to a review of the potential outcomes framework and its associated modelling assumptions.

The study of causality has been considered from different perspectives within different disciplines. One particular point of view, as discussed in Section 1.2, is the potential outcome's framework, which has received advocacy from several researchers, including Rubin [52, 53] and Robins [54]. Another perspective is the *directed acyclic graph*-centric approach of Pearl [55]. Directed acyclic graphs (DAGs) are graphical models that depict variables and their causal relationships while ensuring there are no cycles or feedback loops. The Pearl model employs DAGs for the identification of causality.

In the context of precision medicine, sometimes referred to as precision health, the potential outcome is a random variable defined in the following way. Consider the set

of all possible treatment options $\mathcal{A}$. It is assumed that there exists a random variable, $Y(a)$ whose value represents a random patient's outcome if they were, possibly contrary to reality, to be assigned treatment $a \in \mathcal{A}$. For example, in a binary treatment scenario where $\mathcal{A} = \{0, 1\}$, each individual has two potential outcomes, denoted by $Y(0)$ and $Y(1)$, which correspond to the outcome that we would observe if they were to be administered the control or the treatment, respectively. We refer to these generally as potential outcomes, since, for each individual we typically observe one of these values, although we could potentially observe either, depending on which treatment they receive.

Suppose we have a population of individuals indexed by i where $\mathcal{I}$ is the index set of the population. Let A be the random variable indicating a binary treatment assignment and X be the random variable generating the individual covariates. We wish to determine the impact of the treatment A on an outcome Y. Underlying this causal problem is the conceptualization that there exist random variables $\{A, X, Y(1), Y(0), Y\}$ with some joint distribution. We wish to deduce the distribution of the potential outcomes $Y(1)$ and $Y(0)$ from the distribution of (A, X, Y), using the observable data $(A_i, X_i, Y_i)_{i\in\mathcal{I}}$. To make this task achievable, we work under the three central assumptions of causal inference, namely, the stable unit treatment value assumption (SUTVA)[56], no unmeasured confounders (NUC), and positivity. We now list the conventional specification of these three assumptions in turn and denote them as assumptions A1, A2, and A3, respectively, below.

The stable unit treatment value assumption (SUTVA) states that (i) the potential outcome of each unit is unaffected by the treatment assignment of any other unit, and (ii) there are no different versions of each treatment level. The first assumption of SUTVA, which is known as the no interference assumption, was introduced by Cox [18]. The second part assumes that there are not multiple versions of a treatment. This is called the consistency assumption. The consistency assumption states that the observed outcome of an individual i, $i \in \mathcal{I}$, who is known to have received treatment a_i, is exactly equal to the potential outcome of that individual corresponding to treatment a_i, regardless of the conditions under which that treatment was administered. The consistency assumption, is

usually stated as follows,

$$Y_i = \sum_{a \in \mathcal{A}} Y_i(a) \mathrm{I}\{A_i = a\}, \ \forall a \in \mathcal{A} \ , \forall i \in \mathcal{I}. \tag{A1}$$

The no unmeasured confounders (NUC) assumption, also referred to as the strong ignorability assumption, is typically stated as,

$$\{Y_i(a)\}_{a \in \mathcal{A}} \perp\!\!\!\perp A_i \mid X_i, \quad i \in \mathcal{I}. \tag{A2}$$

Intuitively, the NUC assumption states that the set of measured covariates X contains all the information that influences both treatment and the potential outcomes. Hence, conditioning on these makes the treatment assignment independent of the set of potential outcomes. For a binary treatment A and an individual i, $i \in \mathcal{I}$, A2 is equivalent to, $P[A_i = a_i | Y_i(0), Y_i(1), X_i] = P[A_i = a_i | X_i]$. That is, treatment assignment does not depend on the latent value of the potential outcomes themselves.

Lastly, the positivity assumption assumes that, within each level of the covariate X, there is a nonzero probability of receiving each possible treatment. That is, for treatment A and covariate X, the positivity assumption requires that

$$P(A_i = a_i | x_i) > 0, \qquad \forall a_i \in \mathcal{A} \quad , i \in \mathcal{I}, x \in \mathcal{X}. \tag{A3}$$

Note that under the assumptions A1-A3 we may write the probability of a potential outcome $Y(a)$ as, $P(Y(a) \mid X) = P(Y(a) \mid A = a, X) = P(Y \mid A = a, X)$, so that we have now related the distribution of the potential outcome to a distribution involving only the observable data, where the first equivalence follows from A2 and the second equivalence follows from A1. Utilizing this distribution enables us to estimate the causal effects of interest.

2.2 Optimal DTRs Estimation

In the framework of DTR, it is posited that patients undergo multiple stages of treatment. This process commences with the initiation of treatment, followed by subsequent treatment

decisions made at predetermined intervals, taking into account the current information available for each patient.

In accordance with the established DTR literature, we employ $\mathcal{K} = \{1, 2, ..., K\}$ to index the stages of treatment and $\mathcal{I} = \{1, 2, ..., n\}$ to index individuals. The set of treatment options available at stage k is denoted by $\mathcal{A}_k$. Given our consideration of binary treatment, $\mathcal{A}_k$ is isomorphic to $\{0, 1\}$. Prior to the treatment at stage k, a series of covariate measurements are assumed to be taken for each individual. These covariates are represented by X_{ik}. We define H_{ik} as the collection of measurements taken prior to treatment at stage k, which serve to inform the subsequent treatment decision. This history, H_{ik}, for individual i at stage k, encompasses both current and past stage covariates, along with the values of previous treatments. Observed values of the history will be denoted by lowercase notation, h_{ik}. At the end of the K-stage treatment regime a health outcome Y is measured. We assume Y is continuous and that the higher values of Y are preferred.

In a typical DTR with a single outcome, the goal is to find the collection of decision rules which maximizes the expected outcome of Y. We use $d = d_k(\cdot)_{k \in \mathcal{K}}$ to denote the set of K decision rules, where each d_k is a map, $d_k : \mathcal{H}_k \to \mathcal{A}_k$, taking an observed history $h_i \in \mathcal{H}_k$ to a recommended kth stage treatment $A_{ik} \in \mathcal{A}_k$ for that individual. Thus, d defines a treatment regime which can be used to systematically treat patients, and we define $\mathcal{D}$ to be the space of all possible treatment regimes.

2.3 dWOLS Estimation Method for a K-stage DTR without Interference

The dWOLS methodology is a regression-based method for estimating optimal decision rules. In its original specification (see, Wallace and Moodie [9]), it was assumed that the outcome of interest is continuous, the treatments were restricted to be binary and no interference existed within the population. Under these assumptions, it was shown that the dWOLS methodology consistently estimates a sequence of optimal decision rules for the maximization of the population expected outcome via maximization of the expected health outcome of each individual.

More specifically, let the optimal treatment of each individual i, at stage k, be denoted by a_{ik}^{opt}. Estimation of the sequence of optimal decision rules with dWOLS begins at the last stage K. Here, it estimates the optimal stage K decision rule by considering the effect of the stage K treatment on the expected counterfactual outcome, $E\left[Y_i(a_1, \ldots, a_K) \mid h_{iK}\right]$. Having achieved this, the method then moves on to stage $K-1$. The optimal stage $K-1$ decision rule is estimated by considering the effect of the stage $K-1$ treatment, and its interactions with tailoring variables, on the expected outcome, assuming that the stage K treatment is optimal, $E\left[Y_i(a_1, \ldots a_{iK}^{opt}) \mid h_{i(K-1)}\right]$. This requires the construction of a potentially counterfactual outcome labelled *pseudo outcome* (see, Wallace and Moodie [9]) for each individual, under the assumption they received optimal treatment a_K^{opt} at the Kth stage. This can be done by adding a positive quantity to the observed outcome of individuals who did not receive their optimal stage K treatment. This process continues until a decision rule has been estimated for all stages.

To formalize the dWOLS estimation process, we postulate an outcome model consisting of two components: the *treatment-free* component, and the *blip* components. The treatment-free component characterizes the relationship between the tailoring covariates and the outcome in the absence of treatment, while the blip component characterizes the treatment effect, including both main effects and interactions between treatments and tailoring covariates, at each stage. At the last stage, K, dWOLS assumes a model of the form,

$$E[Y_i \mid \mathbf{h}_{iK}, \mathbf{a}_i] = f_K(\mathbf{h}_{iK}; \boldsymbol{\beta}_{iK}) + \gamma_K(\mathbf{h}_{ik}, a_{iK}; \boldsymbol{\Psi}_K), \tag{2.1}$$

where f_K is the stage K treatment-free model with parameter $\boldsymbol{\beta}_K$, and γ_K is the stage K blip model with parameter $\boldsymbol{\Psi}_K$. The blip component identifies the optimal treatment rule by characterizing the treatment effect on the outcome. Hence, the optimal treatment allocation at stage K can be found by $\underset{a_{iK}}{\operatorname{argmax}} E[Y(a_{iK}) \mid \mathbf{h}_{iK}] = \underset{a_{iK}}{\operatorname{argmax}} \left[\gamma_k(\mathbf{h}_{iK}, a_{iK}; \boldsymbol{\Psi}_K)\right]$.

Once the optimal treatment for stage K has been chosen, the pseudo outcome of stage $K-1$ for each individual under the optimal treatment assignment a_{iK}^{opt} is constructed as $\tilde{y}_{i(K-1)} = y_i + [\gamma_K(\mathbf{h}_{iK}, a_{iK}^{\text{opt}}) - \gamma_K(\mathbf{h}_{ik}, a_{iK})]$. We now proceed to regress pseudo outcome of stage $K-1$ on the $f_{(K-1)}$ and $\gamma_{(K-1)}$ and estimate $a_{i(K-1)}^{\text{opt}}$. We repeat this process, until the optimal treatment of the first stage is estimated. The pseudo outcome of $\tilde{y}_{i(k-1)} =$

$\tilde{y}_{ik} + [\gamma_k(\mathbf{h}_{ik}, a_{ik}^{\text{opt}}) - \gamma_K(\mathbf{h}_{ik}, a_{ik})]$. Note that $\tilde{y}_{iK}$ at the last stage K is defined as y_i, the observed outcome of the individual i. The decision rule at each stage can be written in terms of the steps provided in Algorithm 1.

Algorithm 1: Decision Rule $d_k(\mathbf{h}_k)$ for stage k

Result: Optimal Treatment Decision
if $\gamma_k(\mathbf{h}_k, 1) > 0$ **then**
  Choose $a_k^{\text{opt}} = 1$
else
  Choose $a_k^{\text{opt}} = 0$.
end

As discussed, the blip component of the model is crucially linked to the identification of the optimal treatment regime. Therefore, it is imperative that the estimation procedure used to fit the proposed model offers desirable properties in terms of the resulting estimates for $\mathbf{\Psi}_k$ at each stage k. Indeed, assuming we can specify balancing weights which eliminate any dependence between the history and the treatment at a given stage, the dWOLS framework will provide consistent estimation of the blip parameters. Along with specifying the balancing weights, dWOLS additionally requires modelling the treatment allocation.

In the original dWOLS paper by Wallace and Moodie [9], for a continuous outcome in a binary treatment setting where $\pi(x) = P(A = 1|x)$, it is shown that weights $w(a, x)$ satisfying the *balancing condition* of

$$\pi(\boldsymbol{x})w(1, \boldsymbol{x}) = (1 - \pi(\boldsymbol{x}))w(0, \boldsymbol{x}), \tag{2.2}$$

will lead to consistent estimation of the regression parameter $\mathbf{\Psi}_k$, even if at most one of the nuisance models (treatment-free model and treatment model) is incorrectly specified. Moreover, Wallace and Moodie [9] introduced a balancing weight of the form:

$$w(a, \boldsymbol{x}) = \big|a - E[A \mid X = \boldsymbol{x}]\big|. \tag{2.3}$$

We refer to these weights as *absolute value weights*. They allocate higher weights to individuals whose observed treatment significantly differs from what is expected. In an independent work, Li et al. [57] introduced a class of balancing weights known as *overlapping*

weights and discussed their properties. It is worth noting that the absolute value weights introduced by Wallace and Moodie [9] are overlapping weights for binary treatment.

2.4 Dyadic Data Structure

Our contributions to the DTRs literature in this thesis operate in the context of populations exhibiting a dyadic network structure. In this section, we provide the details of what it means for a population to exhibit this structure and its impact on the assumptions of causal inference, and therefore the estimation of a DTR.

A population exhibits a *dyadic-network structure* if each individual in the population forms a pair with one other member of the population. We call these pairs, dyads. What makes these dyads important in the context of DTRs, and causal inference in general, is that we allow interference to exist between the members of a dyad. We still maintain the assumption that no interference exists outside, i.e., between dyads. This kind of interference is then, intuitively, called partial interference, since interference is only presumed evident in restricted portions of the population.

Dyads arise in two types: exchangeable and distinguishable. When the members are *distinguishable*, each member serves a role that cannot be filled by the other. For example, in familial dyads consisting of sister pairs, the older sister might play a vastly different role from that played by the younger one, and this is an inherent structure of the dyad—these roles cannot be sensibly exchanged. We say that the data are *exchangeable* if the members of dyads are not distinguishable; for example, in twin studies, identical twins could be considered as exchangeable. The decision of whether or not dyad members are exchangeable is both empirical and theoretical and depends on different factors, such as available covariates, the purpose of analysis, and others. In other words, if the members of a dyad do not differ in any meaningful way with respect to modelling assumptions, then they are exchangeable. We label the members of a dyad as member 1 and member 2. A different way to word this is to say that if we could randomly label the dyad members to member 1 and member 2 without losing important information that describes the dyad, they are exchangeable. In describing the dyadic data, we distinguish between those baseline covariates, which are characteristics of the pair–and thus necessarily common between the

	Wave 1					Wave 2					
	Member 1		Shared	Member 2		Member 1			Shared	Member 2	
Household ID	S	E	H	S	E	S	E	Y	H	S	E	Y
H1	1	0	55,000	1	1	0	1	1	60,000	1	0	3
H2	0	0	32,000	0	1	1	1	2	30,000	1	1	3

Table 2.1: An example of a standard dyadic data structure for two individuals from the same household in a smoking cessation study. "S" indicates smoking status, "E" represents whether they use e-cigarettes, and "H" is the household's income. The outcome variable Y represents the tobacco dependence score.

members of a given pair, such as the household income of individuals from the same household–and baseline covariates, which are characteristics of an individual. When we are in the framework of exchangeable members, it is common to label members with generic labels. Herein we label the members of a dyad as member 1 and member 2. We index the members of a given dyad using $\mathcal{J} = \{1, 2\}$. In a study with n independent dyads, we index each dyad by i, $i \in \mathcal{I} = \{1, ..., n\}$.

In a study with n independent dyads, the sample consists of $2n$ individuals in total. The observed data for these n dyads are assumed to be drawn independently (in other words, dyads are presumed independent of one another) and are identically distributed. One possible way to structure dyadic data is to think of it as a set of n pairs and, as a result, allocate a row to each observed dyad as demonstrated in Table 2.1.

Chapter 3

DTRs on Dyadic Networks with Binary Treatments

3.1 Introduction

Most of the foregoing methods for estimating optimal DTRs assume that the treatment of one individual has no impact on the outcome of any other individual under consideration. This is the so-called no interference assumption of classical causal inference, a crucial component of the SUTVA [56]. However, it is important to recognize that this assumption may not hold in all situations where DTR modelling is desired. For instance, when targeting smoking cessation, the analysis of DTR faces challenges that arise from the potential interference between individuals [45–49].

Recently, attempts have been made to extend DTRs to situations in which the no interference assumption is relaxed. One of the first works on estimating an optimal treatment regime in the presence of interference with a regression-based population model was proposed by Su et al. [58]. The authors characterize network interference by including neighbours' covariates and treatment status in each individual's outcome model. However, the model does not assume any interaction between individuals' treatments. That leads to an optimal decision rule that ultimately does not depend on the treatment status of one's neighbours. A more recent study considers a DTR framework in which the treat-

ment status and covariate values of the neighbours of each individual are gathered into a scalar-valued exposure interference term, which is then incorporated at the modelling stage [59]. However, the neighbours of each individual are not being considered for optimal treatment under this model, hence no interference exists between those for whom optimal treatment decisions are being made. The process of estimating optimal treatment assignment is therefore not impacted by this modification, so that estimation may proceed with the usual dWOLS algorithm. Additionally, this approach can never find the global optimal solution for a population because it dismisses at least half of the population at the estimation stage.

In this thesis, we propose a methodology for the estimation of optimal dynamic treatment regimes within a population in which interference among those being considered for treatment exists. In particular, we consider the context of a population exhibiting a dyadic-network structure, where each individual in the sample forms a dyad (i.e. pair) with another individual in the sample. Interference is then permitted to exist among those individuals belonging to the same dyad, but not between dyads. In a population with such an interference structure, each individual is influenced by two treatments: their own prescribed treatment and the treatment of their dyadic partner. It follows that each treatment is then associated with two individuals, and therefore optimal value of the treatment can only be obtained when the resulting ramifications on both parties are considered. Estimation of the optimal treatment regime then proceeds through, without loss of generality, maximization of a dyad outcome through the *dyad-health function*, a contextual measure of the health status of those comprising a dyad. The model, therefore, accounts for the interference of treatment statuses within the population in estimation, and, to our knowledge, is the first model to do so in the DTR literature. In addition, the introduction of the dyad-health function provides practitioners with the flexibility to accommodate ethical and fairness concerns, such as prioritizing marginalized groups or individuals within dyads who suffer from more serious health problems.

We demonstrate the importance of the innovations provided by the proposed model through a detailed analysis of the Population Assessment of Tobacco and Health (PATH) dataset. In particular, in the households with two tobacco users, we consider whether the adoption of e-cigarettes by an individual impacts their nicotine exposure, and whether there exists an interference effect, i.e., when one member of a tobacco using household

adopts e-cigarette use, does it impact the nicotine exposure in the other members?

The remainder of this chapter is structured as follows: Section 3.2 provides the methodological and theoretical background, including notation, and sets up the modelling framework for a K-stage DTR on a dyadic-network population. In Section 3.3, we discuss the details of estimating the proposed model, including the treatment and outcome models, as well as the optimal treatment decisions for the dyadic-network DTR. Section 3.4 contains two simulation studies. The first demonstrates that the proposed methodology retains the double-robustness property of dWOLS, while the second compares the performance of the proposed model with existing methodologies on data sampled from a dyadic-network population structure. Section 3.5 contains our analysis of PATH dataset, and Section 3.6 concludes the paper with some discussion.

3.2 Methodology

3.2.1 Notation

Suppose we have a sample of $2n$ individuals who have completed a K-stage treatment regime. Further suppose that individuals of the sample interact with each other, and that an individual's outcome is dependent with whom they interact. Such behaviour encapsulates the well-known notion of interference in the causal inference literature. We assume here that interference takes a special form, specifically of a dyadic network, which is to say, each individual in the sample forms a dyad (i.e. pair) with another individual in the sample, with whom they interact. Under this assumption, interference is restricted to occur only within these dyads. Hence, the sample may be considered analogously as a sample of n independent dyads. We index the members of a given dyad using $\mathcal{J} = \{1, 2\}$. In general, when the member of interest is member j we use the notation j^* to represent the measurements corresponding to its dyadic partner. For simplicity, we will define quantities in terms of member 1 when the quantity is member-specific. The analogous quantity for member 2 can be found by simply replacing all the 1s with 2s in these definitions, and vice versa.

As previously covered in Chapter 2, we let $\mathcal{K} = \{1, 2, ..., K\}$ index the treatment stages, and we will let $\mathcal{I} = \{1, 2, ..., n\}$ index the sampled dyads (rather than individuals). Consider

an arbitrary dyad i at an arbitrary stage k. We use $\mathcal{A}_k$ to denote the set of treatment options available at stage k. Typically, binary treatments are considered, hence $\mathcal{A}_k$ is isomorphic to $\{0, 1\}$ and we will, without loss of generality, use this representation going forward. The random variable $\mathbf{A}_{ik} = (A_{i1k}, A_{i2k})$ denotes the treatment of dyad i at stage k. This consists of A_{ijk}, the treatment of the members of a dyad i at stage k, i.e., A_{i1k} and A_{i2k}.

We assume that a set of covariate measurements are taken on each individual prior to treatment at stage k. We denote these covariates by X_{ijk}. Additionally, we let Z_{ik} denote the vector of shared covariates among the dyad members as of stage k. This shared vector may contain, e.g., household income, living conditions, geographic location, and so on. Although covariates are assumed to be measured at each stage k, any measurements preceding treatment at that stage can also be used to inform the decision. Accordingly, we define $\mathbf{h}_{ijk}$ to be the history of measurements relevant for informing the treatment decision of dyad member j at stage k, through the current stage k itself. The history may consist of current and past stage covariates of either member of the dyad, previous treatment values of either, the shared covariates, or functions thereof. We use $\mathbf{h}_{ik} = (\mathbf{h}_{i1k}, \mathbf{h}_{i2k})$ to analogously denote history at the dyad level at stage k.

Following the Kth stage, i.e. after the treatment regime has finished, we assume a measurement is made on each individual for the purpose of assessing the efficacy of the treatment regime. We refer to this measurement simply as the outcome, and we use Y_{ij} to represent the outcome associated with member j. We further assume that the outcome is continuous and higher values are better. Under interference, the individual-level potential outcome is a function of the full sequence of dyad treatments $\mathbf{A} := \mathbf{A}_{ik \in \mathcal{K}}$. Occasionally it will be necessary to make this dependence explicit, and in such a case we will use the more explicit notation $Y_{ij}(\mathbf{A})$. Thus, for a given dyad of i, $Y_{ij}(\mathbf{a})$ represents the outcome of member j that would be observed if the dyadic treatments observed over the K stages corresponds to the sequence $\mathbf{a}$. Note that this function is defined for all possible dyad treatment sequences in the space $\mathcal{A}_1^2 \otimes \mathcal{A}_2^2 \otimes \cdots \otimes \mathcal{A}_K^2 =: \mathcal{A}^2$, and hence takes a value for all possible treatment assignments among the dyad pair, not only what was observed in reality.

Let g be a bivariate function which takes as its argument a dyad outcome $Y = (Y_1, Y_2)$, and, based on this input, returns a value summarizing the overall health of the dyad in a

contextually relevant manner. We call this g the *dyad-health function*. It follows that we may obtain our desired dyad-level optimal treatment regime by (without loss of generality) maximizing g over $\mathcal{A}^2$ rather than maximizing Y.

Specifically, let $d = d_k(\mathbf{h})_{k\in\mathcal{K}}$ be the set of K decision rules comprising the dyadic DTR, where $d_k(\mathbf{h}_{ik})$ determines the prescribed dyad treatment A_{ik} at stage k given the dyad history $\mathbf{h}_{ik}$. Let $\mathcal{D}$ be the space of all possible sets of K decision rules. Our goal is to find the collection of decision rules which maximizes the expected dyad health,

$$\arg\max_{d\in\mathcal{D}} \mathbb{E}\left(g\Big\{Y\left[d_k(\mathbf{H}_k)_{k\in\mathcal{K}}\right]\Big\}\right),$$

where the expected value is taken over the population of dyads.

Consider choosing g to be the average of the elements of Y, that is, $g(Y) = 0.5(Y_1+Y_2)$. Under the proposed approach, this choice of g recovers the oft-used maximization of the population expected value, with the caveat that it now accounts for dyad-level interaction (because it technically maximizes the average of pairwise averages). Although this is a natural and easily understood choice, we will discuss in Section 3.3.3 that this choice may not always be the most ethical or contextually appropriate.

3.2.2 Causal framework

To allow for the estimation of causal effects in the presence of interference at the dyad level, updates to the typical causal inference assumptions of consistency, no unmeasured confounders, and positivity, are required (see, A1-A3). In this section, we specify the required changes in turn. Note that each assumption is specified at the dyad level and is assumed to hold for all dyads in the population. We hence drop the subscript i denoting an arbitrary dyad in the following assumption definitions.

For any stage $k \in \mathcal{K}$, and any observable dyad history $\mathbf{h}_k$, we make the analogy to the consistency assumption as follows. Let $\mathcal{A}^2$ be the space of all sequences of dyad-level treatments over a K-stage regime and $\mathcal{Y}_j = \{Y_j(\mathbf{a}) \mid \mathbf{a} \in \mathcal{A}^2_{\mathcal{K}}\}$ be the associated set of potential outcomes for member $j \in \mathcal{J}$. We assume that the observed outcome Y_j can be specified as, $Y_j = \sum_{\mathbf{a}\in\mathcal{A}^2} Y_j(\mathbf{a})\,\mathrm{I}\{\mathbf{a}\},\ j \in \mathcal{J}$.

For the analog to the no unmeasured confounders assumption, define $\mathcal{Y} = \{(Y_1(\mathbf{a}), Y_2(\mathbf{a})) \mid \mathbf{a} \in \mathcal{A}^2\}$ as the set of dyad-level potential outcomes, and $\mathcal{H}$ as the space of observable histories over a K-stage treatment regime. We must assume that, conditional on an observed sequence of histories $\mathbf{h} \in \mathcal{H}$, the observed sequence of dyad treatments $\mathbf{a}$ is assigned independently of the set of potential outcomes of that dyad. That is, $\mathcal{Y} \perp\!\!\!\perp \mathbf{A} \mid \mathbf{h}$.

Furthermore, we assume that within each level of the shared or unshared covariates, any possible dyad-level treatments can occur. That is, for dyad treatment $\mathbf{a}_k = (a_{1k}, a_{2k})$ we assume, $\mathbb{P}\left(\mathbf{a}_k \,\middle|\, \mathbf{h}_k\right) > 0$, $\forall \mathbf{a}_k \in \mathcal{A}_k^2$ and $\mathbf{h}_k$ in the support of $\mathbf{h}_k$. This is equivalent to the standard positivity assumption.

In terms of modelling, we can sometimes impose a simplifying assumption called *exchangeability*. If exchangeability holds, we have that $(Y_1, X_1, X_2, Z) \overset{d}{=} (Y_2, X_2, X_1, Z)$. When exchangeability holds, we say that the dyad members are exchangeable.

3.2.3 Modelling assumptions

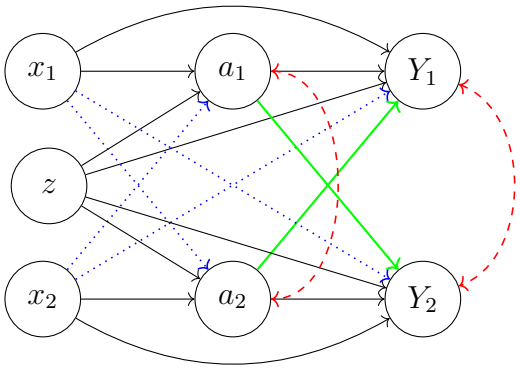


Figure 3.1: Model 1 (General Interference Model): The treatments and individual characteristics of each member can influence both the treatment and the outcome of the other member. A directed arrow signifies a potential causal effect of the first variable on the second, while a bidirectional arc between variables suggests the presence of unobserved factors that could be shared causes of the two variables.

The causal framework outlined in Section 3.2.2 facilitates the development of causal models which enable the estimation and optimization of treatment effects within a dyadic-

network population. The most general formulation of our modelling assumptions for a single stage are represented succinctly in the causal diagram of Figure 3.1. All arrows in Figure 3.1 represent dependence, however colors have been used to distinguish between dependence structures of different contextual or modelling importance. For example, blue arrows have been used to show causal impact of the individual covariates of one member of the dyad on the treatment or outcome of the other, while green arrows specifically denote treatment interference. Additionally, red arrows have been used to signify association, which can occur between the treatments or the outcomes of the dyad members because of a common cause. The generalization of Model 1 for a two-stage DTR is shown as an example in Figure A.1.

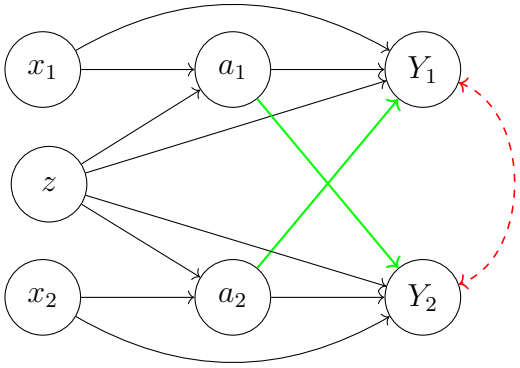


Figure 3.2: Model 2 (Restricted Model): The individual covariates of a member do not exert a causal impact on the other member, and the treatments of members are conditionally independent given the observed covariates.

In practice, there are scenarios that can be represented by simpler interference structures. For example, we might assume that the individual covariates of one member do not have an impact on the treatment or outcome of the other member, and that the treatment assignment of dyad members is independent from each other, given their shared and individual covariates. Figure 3.2 illustrates the causal graph representing such scenarios. Such context can present substantial simplifications to the estimation of model components, which is discussed in more detail in Section 3.3.2.

For comparison purposes, we also present the causal model assumed by Jiang et al. [59], which is shown in Figure 3.3. This represents a significant simplification compared to the

proposed modelling approach. In this simplified approach, the treatment of member 1, the only member included in the model, is independent of the treatment of the other member. This simplification contrasts with the proposed interference modelling approaches, as only member 1 from each dyad is explicitly modeled, and the individual covariates and treatment of the member 2 are only assumed to potentially have a causal impact on the outcome of member 1.

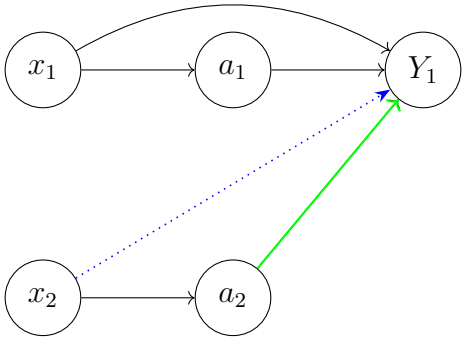


Figure 3.3: Model 3 (Independent Treatment Model): This model focuses solely on member 1. The treatment and covariates of member 2 have a causal impact on the outcome of member 1.

3.2.4 Modelling a K-stage dynamic treatment regime in a dyadic network

To formalize the dyadic-network DTR model, we start by considering the final treatment decision, which is made at stage K. Note that for simplicity, we will omit the subscript i that represents a specific dyad throughout this section. The goal is to choose the treatment which maximizes the dyad outcome (or some function thereof, see: Section 3.3.3). In general, the model for the expected value of the outcome at stage K for an individual j in a given dyad, can be written as,

$$E[Y_j \mid \mathbf{h}_{jK}, \mathbf{a}_K] = f_{jK}(\mathbf{h}_{jK}; \boldsymbol{\beta}_{jK}) + \gamma_{jK}(\mathbf{h}_{jK}, \mathbf{a}_K; \boldsymbol{\Theta}_{jK}), \quad j \in \mathcal{J}, \tag{3.1}$$

where f_{jK} is the stage K treatment-free model of individual j with parameter $\boldsymbol{\beta}_{jK}$, and γ_{jK} is the stage K blip model with parameter $\boldsymbol{\Theta}_{jK}$, where each component of equation (3.1) is

specified below. Equation (3.1) is a general model specification used in much of the DTR literature and is not unique to the dyadic-network interference structure context. Both of the model components serve a particular purpose which can be understood intuitively in the context of personalized medicine. For example, one may interpret the treatment-free component as modelling the outcome that would occur if an individual were to go untreated at the Kth stage of the regime. This role is unchanged when the model is specified for the dyadic-network context.

Complementing the role of the treatment-free component, the blip component models the way in which the dyad treatment level influences an individual dyad member's outcome. Although this role also does not change in the dyadic-network interference setting, the assumption of interference among the dyad members does imply that the blip should be modelled as a function of the dyad treatment, rather than just the individual's treatment. This subtle but important change is what distinguishes the dyadic-network DTR from earlier DTR methodologies. To emphasize this point, we refer to the blip component in this context hereafter as the *dyadic-blip* component.

The functional form of each model component can be chosen relatively arbitrarily, although one often chooses a linear form as a middle ground between flexibility and ease of implementation. For instance, a linear form for the dyadic-blip leads to,

$$\gamma_{jK}(\mathbf{h}_{jK}, \mathbf{a}_K; \boldsymbol{\Theta}_{jK}) = \mathbf{h}_{jK}^\top (a_{jK}\boldsymbol{\psi}_{jK} + a_{j^*k}\boldsymbol{\eta}_{jK} + a_{jK}a_{j^*K}\boldsymbol{\xi}_{jK}),$$

where we have assumed for simplicity in presentation that each component of the blip depends on the entire history $\mathbf{h}_{jK}$, which is not truly a requirement of the model. We specify $\boldsymbol{\Theta}_{jK} = (\boldsymbol{\psi}_{jK}, \boldsymbol{\eta}_{jK}, \boldsymbol{\xi}_{jK})$ to be the concatenation of all the component parameters. Under this formulation, the dyadic-blip is decomposed into three sub-components: one corresponding to the treatment of member j, which we call the *direct blip*, one corresponding to the treatment of their dyadic partner, which we call the *indirect blip*, and one corresponding to the interaction between both treatments, which we call the *interaction blip*. With this form, the components of $\boldsymbol{\Theta}_{jK}$ are directly interpretable in terms of understanding the relationship between the history and dyad level treatment as it pertains to expected potential outcome.

Utilizing this outcome model, we intend to fit an optimal dynamic treatment regime,

which can then be used to optimal treatment assignments to members of an arbitrary population dyad. Typically, an optimal treatment refers to that treatment which maximizes $E(Y(\mathbf{a}))$ over all possible treatment assignments. However, because the dyad-level potential outcome is vector-valued, maximization of its value is not well-defined unless constraints are imposed. There is not one unique way to specify these constraints, and indeed, many plausible constraints are possible. Differing constraints are imposed through specification of the dyad-health function g. We will discuss these ideas further in the Section 3.3.3—for now, we specify the dyad-health function to be $g(Y) = 0.5(Y_1 + Y_2)$. At stage K, then, the optimal dyad-level treatment can be found by solving,

$$\mathbf{a}_K^{\text{opt}} = \arg\max_{\mathbf{a}\in\mathcal{A}_K} \left\{ \frac{1}{2} \left[E(Y_1 \mid \mathbf{h}_{1k}, \mathbf{a}) + E(Y_2 \mid \mathbf{h}_{2k}, \mathbf{a}) \right] \right\}.$$

This idea of optimizing the objective function jointly over the space of possible treatment vectors $\mathcal{A}_K$ is a key feature that sets our method apart from existing literature regarding DTRs and interference. For comparison, the conditional optimization approach proposed by Jiang et al. [59] would solve

$$\mathbf{a}_K^{\text{opt}} = \arg\max_{a_1\in\{0,1\}} E(Y_1 \mid \mathbf{h}_{1k}, a_1, a_2).$$

Although our objective function of the maximization includes two terms, it is still possible to express the decision rule in a simple manner utilizing only the dyadic-blip components. The treatment-free component does not change with the dyad level treatment, hence it acts as an additive constant in the maximization step and may therefore be ignored. As a result, the maximization step can be equivalently written,

$$\mathbf{a}_K^{\text{opt}} = \arg\max_{\mathbf{a}\in\mathcal{A}_K} \left\{ \frac{1}{2} \left[\gamma_{1K}(\mathbf{h}_{1K}, \mathbf{a}) + \gamma_{2K}(\mathbf{h}_{2K}, \mathbf{a}) \right] \right\}, \tag{3.2}$$

so that we are only concerned with the sum of the dyad members' dyadic-blip components.

Once the optimal treatment has been chosen, the next step is to estimate how much the expected dyad outcome would be improved under this potentially counterfactual treatment assignment. From the assumed model, the expected outcome for each dyad member under

the optimal treatment assignment at the final stage can be expressed as,

$$\tilde{y}_j = y_j + [\gamma_{jK}(\mathbf{h}_{jK}, \mathbf{a}_K^{\text{opt}}) - \gamma_{jK}(\mathbf{h}_{jK}, \mathbf{a}_K)], \quad j \in \mathcal{J}. \tag{3.3}$$

Intuitively, we are correcting for the effect of the observed treatment, which was potentially suboptimal, by subtracting off its effect while simultaneously adding on the effect of the optimal treatment. If the optimal treatment and the observed treatment coincide, the added quantity will be zero. The bracketed term in Equation (3.3) is therefore sometimes referred to as the *regret*, since it quantifies, in terms of the individual's health, how much we regret treating them suboptimally.

The quantity $\tilde{y}_j$ is also known as the pseudo-outcome [9], and is a crucial component of the model. Indeed, as we will discuss in the next section, having finished modelling of the Kth stage and ascertaining the Kth stage decision rule, we then move backward to stage $K-1$ and repeat the process: define a model for the outcome, fit the parameters, and maximize the dyad-health function. The difference between fitting at stage K and stage $(K-1)$ is that the latter needs to account for the fact that more treatments will be assigned in the future. However, due to the backward nature of our estimation algorithm, we already know how to treat optimally at the stage K. We can account for this at stage $K-1$ by fitting the outcome model on the pseudo-outcomes $\tilde{y}_j$ rather than the observed outcomes y_j. This process continues for each k in $\mathcal{K}$ until all stages have been addressed, after which we will have obtained K decision rules with which we may inform treatment decisions of future patients at each stage of the regime.

3.3 Estimation and Determination of Optimal Treatment Regime

It is clear that the dyadic-blip component of the model is crucially linked to identification of the optimal treatment regime. It is therefore imperative that the estimation procedure used to fit the proposed model offers desirable properties in terms of the resulting estimates for $\boldsymbol{\Theta}_k$ at each stage k. One such approach is the (generalized) dWOLS procedure [9, 10].

Indeed, assuming we can specify balancing weights which eliminate any dependence

between the history and the treatment at a given stage, the dWOLS framework will provide consistent estimation of the dyadic-blip parameters. Along with specifying these weights, implementation of this estimation framework additionally requires modelling of the propensity score. An important property of this procedure is that it is doubly-robust with respect to model misspecification—consistent estimation of the dyadic-blip parameters is still guaranteed if one of either the propensity score model or the treatment-free model component is misspecified. This is a desirable property, and motivates our use of this estimation procedure for fitting the proposed model.

3.3.1 Balancing weights

For each treatment stage, the algorithm fits the specified model using a weighted dataset. The weights need to be specified such that they remove any dependence between the history and the observed treatment assignment, i.e., they must offset any differences in the treatment assignment mechanism which are due to the value of the measured dyad history.

We denote the kth stage weights by $w(\mathbf{a}_k, \mathbf{h}_k)$ to emphasize that the weights are a function which depends on the history and observed treatment assignment. To find weights which achieve the desired effect, it is common to specify that they satisfy a particular balancing condition. In the seminal dWOLS work [9], a balancing condition for the case of a binary treatment was presented (please see Equation 2.2 for more details). Later, in the development of the generalized dWOLS methodology [10], the balancing condition was extended to the case of continuous and categorical treatments. When dyadic-interference exists, each individual is exposed to a bivariate treatment. We now present a theorem detailing a balancing condition for treatments arising in this form.

Theorem 3.3.1. *A weighted ordinary least-squares regression based on the model,*

$$E[Y_j \mid \mathbf{h}_j, \mathbf{a}] = f_j(\mathbf{h}_j; \boldsymbol{\beta}_j) + \gamma_j(\mathbf{h}_j, \mathbf{a}; \boldsymbol{\Theta}_j),$$

will give consistent estimates for the dyadic-blip parameters $\boldsymbol{\Theta}$ provided that the weights satisfy,

$$\pi(\mathbf{a} \mid \mathbf{h}) w(\mathbf{a}, \mathbf{h}) = \pi(\mathbf{a}' \mid \mathbf{h}) w(\mathbf{a}', \mathbf{h}), \tag{3.4}$$

for any $\mathbf{a} \neq \mathbf{a}' \in \mathcal{A}^2$*, where* $\pi(\mathbf{a} \mid \mathbf{h}) = \mathbb{P}(\mathbf{A} = \mathbf{a} \mid \mathbf{h})$ *is the joint propensity score.*

A proof of this theorem is provided in the Appendix A.1.2. Details on the different types of weights satisfying this balancing condition, and their estimation methods, are presented in Appendix A.1.3. In particular, under the restricted interference model (Model 2 presented in Figure 3.2), we extend the absolute value weights defined by $w(a, x) = |a - E(A|X = x)|$ (see, Equation 2.3) to dyadic-network settings. We call these these new weights as *dyadic-absolute value weights*, and they have the form,

$$w_{ij}(\mathbf{a}, \mathbf{h}_{ij}) = \Big|a_{ij} - E(A_{ij} \mid \mathbf{h}_{ij})\Big| \cdot \Big|a_{ij*} - E(A_{ij*} \mid \mathbf{h}_{ij})\Big|,$$

where $E(A_{ij} \mid \mathbf{h}_{ij})$ is the expected probability that member j of dyad i receives treatment, given their observed history $\mathbf{h}_{ij}$, for each $j \in \mathcal{J}$. Intuitively, these weights assign greater importance to individuals whose own observed treatments or their dyadic partner observed treatments significantly differ from what is expected according to the treatment model.

3.3.2 Estimation of model parameters

Having decided on an appropriate weight function, we proceed with estimation of the model parameters. Estimation takes place in K stages, one for each stage of treatment. The procedure at each stage is exactly the same, so we will outline the details of how it may be carried out at an arbitrary stage k, which necessarily includes how we take this information backward to allow for the completion of the same steps at stage $k - 1$.

Recall that we specify a linear model for the outcome which takes the form,

$$E(Y_j \mid \mathbf{h}_{jk}, \mathbf{a}_k) = \mathbf{h}_{jk}^\top \boldsymbol{\beta}_{jk} + \mathbf{h}_{jk}^\top (a_{jk}\boldsymbol{\psi}_{jk} + a_{j*k}\boldsymbol{\eta}_{jk} + a_{jk}a_{j*k}\boldsymbol{\xi}_{jk}), \tag{3.5}$$

where we have, without loss of generality, used $\mathbf{h}_{jk}$ to represent the tailoring covariates included in each term. As previously stated, this is not a requirement of the model and in general each component may depend on only a subset of the covariates included in the full history $\mathbf{h}_{jk}$. To estimate the parameters, we solve the following weighted generalized

estimating equation,

$$U_k(\boldsymbol{\beta}_k, \boldsymbol{\Theta}_k) = \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} w_{ijk}\mathbf{S}(\mathbf{h}_{ijk}, \mathbf{a}_{ik}) \cdot \left(Y_{ij} - \mathbf{h}_{ijk}^\top\boldsymbol{\beta}_k - \mathbf{h}_{ijk}^\top(a_{ijk}\boldsymbol{\psi}_{jk} + a_{ij*k}\boldsymbol{\eta}_{jk} + a_{ijk}a_{ij*k}\boldsymbol{\xi}_{jk})\right) = \mathbf{0}. \tag{3.6}$$

Here we have used the notation $\mathbf{S}(\cdot,\cdot)$ to represent the score vector of the specified outcome model with respect to the model parameters, which in the case of the model presented in Equation (3.5), takes the form,

$$\mathbf{S}(\mathbf{h}_{ijk}, \mathbf{a}_{ik}) = (\mathbf{h}_{ijk}, a_{ijk}\mathbf{h}_{ijk}, a_{ij*k}\mathbf{h}_{ijk}, a_{ijk}a_{ij*k}\mathbf{h}_{ijk}).$$

Since our model is linear, solving this GEE is equivalent to estimation via weighted ordinary least-squares regression. We choose to present estimation in the more general GEE framework, however, to demonstrate that alternative outcome models can be accommodated, according the context of the data under consideration.

Upon solving Equation (3.6), we obtain model parameter estimates, which we denote by $(\hat{\boldsymbol{\beta}}_k, \hat{\boldsymbol{\Theta}}_k)$. These are then utilized to formulate the optimal treatment decision rule for stage k, which is derived from the maximization presented in Equation (3.2).

From this expression we find that the decision rule can be written generally in terms of the steps provided in Algorithm 2.

Algorithm 2: Decision Rule $d_k(\mathbf{h}_{1k}, \mathbf{h}_{2k})$ for stage k

Result: Optimal Treatment Decision

if $\gamma_k(\mathbf{h}_{1k}, \mathbf{a}) + \gamma_k(\mathbf{h}_{2k}, \mathbf{a}) > 0$ *for at least one* $\mathbf{a} \in \mathcal{A}_k$ **then**

Choose $\mathbf{a}_k^{\text{opt}}$ such that

$$\gamma_k(\mathbf{h}_{1k}, \mathbf{a}_k^{\text{opt}}) + \gamma_k(\mathbf{h}_{2k}, \mathbf{a}_k^{\text{opt}}) \geq \gamma_k(\mathbf{h}_{1k}, \mathbf{a}') + \gamma_k(\mathbf{h}_{2k}, \mathbf{a}')$$

for all $\mathbf{a}' \in \mathcal{A}_k$

else

Choose $\mathbf{a}_k^{\text{opt}} = (0, 0)$.

end

The decision rule takes into account the effect of the chosen dyad treatment on both members of the dyad, in particular, the chosen treatment may no longer be the one which leads to the individual maximum for each member. However, when this occurs, whatever is lost by the member not attaining their individual maximum is less than what is gained by the other member. Hence, the dyad members may end up compromising their own outcome for the "greater good" of the dyad.

Having obtained our estimated decision rule, estimation at stage k is complete, and the algorithm moves backward to stage $k-1$. To proceed with estimation at stage $k-1$ the appropriate pseudo-outcome needs to be computed. This can be done according to,

$$\tilde{y}_{ijk-1} = \tilde{y}_{ijk} + [\gamma_k(\mathbf{h}_{ijk}, \mathbf{a}_{ijk}^{\text{opt}}; \hat{\boldsymbol{\Theta}}_k) - \gamma_k(\mathbf{h}_{ijk}, \mathbf{a}_{ijk}; \hat{\boldsymbol{\Theta}}_k)], \quad i \in \mathcal{I},\, j \in \mathcal{J},\, k = 2, ..., K,$$

where we let $\tilde{y}_{ijK} = y_{ij}$ the outcome observed following the stage K. Thus, the pseudo-outcome is defined recursively, where, at each stage, we "add back" the individual level regret to account for the potentially suboptimal treatment decision that was initially observed at that stage.

3.3.3 Optimal treatment assignment and the dyad-health function

In the classic optimal DTR scenario, we seek the regime d^{opt} such that,

$$E[Y(d^{\text{opt}})] \geq E[Y(d)],$$

for all possible regimes d. For example, in a two-stage regime, the optimal regime can be found by solving

$$\max_{a_1 \in \mathcal{A}_1} E\left[\max_{a_2 \in \mathcal{A}_2} E\Big[Y \mid H_1 = h_2, \mathcal{A}_2 = a_2\Big] \mid H_1 = h_1, \mathcal{A}_1 = a_1\right]$$

iteratively, starting with the inner most maximization problem and working outward. This is the approach of Q-learning, which can be extended to an arbitrary number of stages by further breaking down the inner most conditional expected value appropriately; see, for example, the book by Tsiatis et al. [60].

As discussed in Section 3.2.4, we cannot directly apply this approach when the outcome of interest is vector-valued, since the maximization of a vector is not well-defined. This motivates the introduction of the dyad-health function $g : \mathbb{R}^2 \mapsto \mathbb{R}$, which we assume takes the outcome values Y of a particular dyad and returns a value which is to be interpreted as a measure of the dyad's overall health. Without loss of generality, it is assumed that larger values of $g(Y)$ are associated with greater dyad health. It follows that in the case of a dyadic-network population structure, the associated optimal treatment regime d^{opt} can be found according to,

$$\max_{\mathbf{a}_1 \in \mathcal{A}_1} E\left[\max_{\mathbf{a}_2 \in \mathcal{A}_2} E\Big[g(Y) \mid \mathbf{H}_2 = \mathbf{h}_2, \mathbf{A}_2 = \mathbf{a}_2\Big] \mid \mathbf{H}_1 = \mathbf{h}_1, \mathbf{A}_1 = \mathbf{a}_1\right]. \tag{3.7}$$

At first glance, this seems to imply that modelling of the random variable $g(Y)$ is required, but this would be undesirable for a number of reasons. One of the more critical arguments against the modelling of $g(Y)$ directly is that the impact of the optimal treatment on the health of each individual in the dyad often becomes unidentifiable in such scenarios. For example, consider the choice $g(Y) = 0.5(Y_1 + Y_2)$, which we model directly using a parametric linear model, i.e. $E[g(Y)|\mathbf{H} = \mathbf{h}] = \mathbf{h}^\top \boldsymbol{\beta}$. This approach seems very natural, and in such a case we may indeed still proceed to find the optimal dyad-level treatment regime d^{opt} without issue. However, we will not be able to say anything about how the optimally assigned treatment impacts the individual health levels of the dyad members, as these are not identifiable under the assumed model. Hence, such an approach moves away from the ethos of personalized medicine, in which optimal treatment regime estimation is rooted. Further, such an approach may lead to ethical concerns, as outright maximization of g may lead to an "optimal" regime in which the health of one dyad member is sacrificed for a sufficiently large gain in the health of the other to increase $g(Y)$. We therefore do not pursue this approach here, instead proceeding with a model for Y.

By focusing on the modelling of Y, we run the risk that the conditional expected value of $g(Y)$ is difficult to handle in practice. Indeed, one needs to derive the distribution of $g(Y)$ from the distribution of Y in order to compute the conditional expectations of Equation (3.7). Since no distributional assumptions are made on the error term in the specification of Y, this distribution cannot be determined in general. One could proceed by addressing exactly this issue, that is, by making distributional assumptions of the errors. This may

unnecessarily restrict the scope of the model, hence we prefer to make some restrictions on the form of g. We suppose that g is such that it is monotonic (non-decreasing) in both of its arguments. That is, for $c > 0$ we have,

$$g(y_1 + c, y_2 + c) \geq g(y_1 + c, y_2) \geq g(y_1, y_2), \quad \text{and,}$$
$$g(y_1 + c, y_2 + c) \geq g(y_1, y_2 + c) \geq g(y_1, y_2).$$

With this assumption, and the modelling assumption for Y given in Equation 3.1 it is straightforward to show that maximization of the conditional expected value of $g(Y)$ is equivalent to the maximization of g composed with the conditional expected value of Y. That is, $\max_{\mathbf{a}\in\mathcal{A}} E(g(Y) \mid \mathbf{h}, \mathbf{a}) = \max_{\mathbf{a}\in\mathcal{A}} g(E(Y \mid \mathbf{h}, \mathbf{a}))$. For any such choice of g, we find our optimal decision rule at stage k by maximizing $g(E[Y \mid \mathbf{h}_k, \mathbf{a}])$ over all possible dyad treatment decisions $\mathbf{a}$.

3.4 Simulation Study

We illustrate key properties of the dyadic-network DTR approach through two simulation studies. Specifically, we investigate the purported double robustness property of the estimated coefficients, and the ability to obtain an improved optimal treatment regime with respect to the conditional optimization approach. The first simulation shows these results for a single stage model, while the second shows the same results for a two-stage model. Since the results of both simulations are similar, we present only the first simulation in this Section. The results of the second simulation can be found in Appendix A.3.2.

3.4.1 Simulation I: double robustness

First, we illustrate the double robustness property regarding consistent estimation of the dyadic-blip parameters. We do this for varying sample sizes, and under different specifications of the treatment-free, and dyadic-blip model. Second, we compare the performance of the dyadic-network DTR with existing DTR methodologies on data exhibiting a dyadic-network interference structure. This illustrates the necessity of the joint optimization approach in contexts where interference exists.

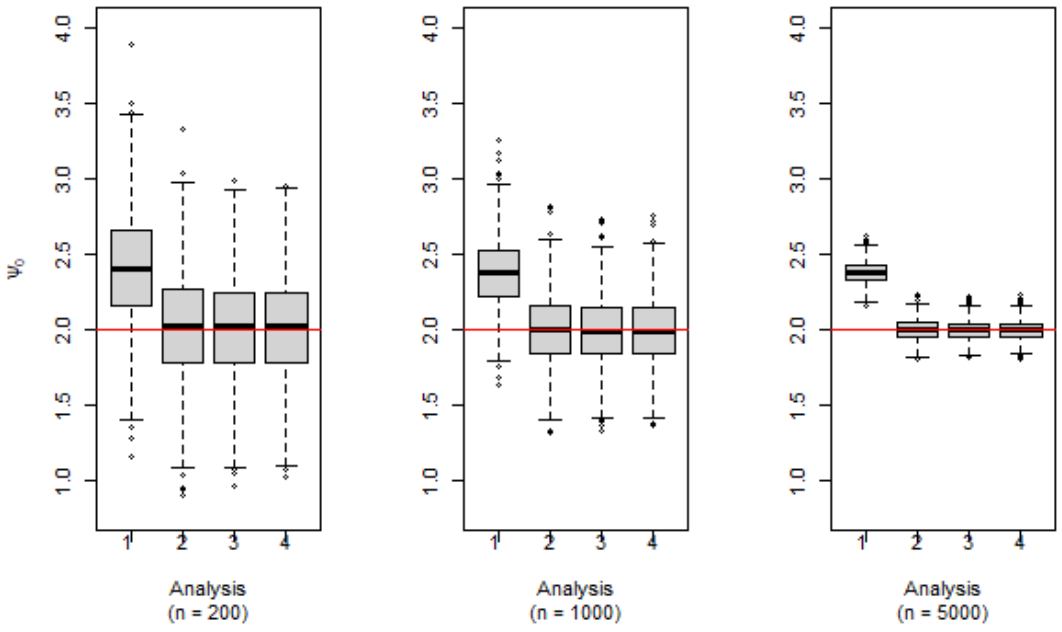


Figure 3.4: Dyadic-blip function parameter estimates for 1000 simulated datasets via our dyadic-dWOLS approach when neither (Analysis 1), one (Analyses 2 and 3) or both (Analysis 4) treatment and treatment-free outcome models are correct.

Since consistency is an asymptotic property, we choose sample size as our simulation parameter. We use the sample sizes $n = 200, 1000$, and 5000 dyads. For each sample size, we generate 1000 datasets, fit the model on the data, and record the observed bias of the estimated dyadic-blip parameters. For each dataset, we also find optimal treatments according to the conditional optimization approach, so that we can compare those results with the proposed methodology.

Our baseline model is generated as follows: for each dyad $i = 1, \ldots, n$ and each member $j = 1, 2$, we generate the patient information, X_{ij}, according to a uniform distribution on the interval $[0, 2]$. The treatment model is a logistic regression on the patient information, $P(A_{ij} = 1 \mid X_{ij} = x_{ij}) = \text{expit}\{1 + x_{ij}\}$. Finally, we define the dyad outcome model to be $Y_{ij} \mid (\mathbf{h}_{ij}, \mathbf{a}_i) = \mu(\mathbf{h}_{ij}, \mathbf{a}_i) + \varepsilon_{ij}$, where $\mu(\mathbf{h}_{ij}, \mathbf{a}_i) = e^{x_{ij}} - x_{ij}^3 + a_{ij}(2 + x_{ij}) - 2a_{ij^*} + a_{ij}a_{ij^*}(1 -$

$2x_{ij}$) and $\varepsilon_{ij} \sim N(0, 1)$.

In order to show double robustness, we show that the model obtains consistent estimators of the dyadic-blip parameters even when one of the model components is misspecifed. Hence, we perform four analyses, one for each possible combination of specifying or misspecifying the treatment-free and treatment models. Analysis 1 denotes the case in which neither component of the model is correctly specified, Analysis 2 denotes the case in which only the treatment-free component is misspecified, Analysis 3 denotes the case where only the treatment model is misspecified, and Analysis 4 denotes the case in which everything is correctly specified. To misspecify the treatment-free component, we omit the terms involving non-linear functions of the patient information, while the treatment model is misspecified through the removal of the intercept term.

Results of the simulation are provided in Figure 3.4. This figure shows three sets of boxplots, which correspond to our three chosen sample size settings. Each boxplot in one subfigure corresponds to one of the four analysis scenarios. From these plots, we can see that the model appears to provide doubly-robust estimation as expected. When at least one of the model components is correctly specified, the dyadic-blip parameter estimates appear to be consistent, which can be surmised from their central tendency of zero and their shrinking variance as the sample size increases.

For the second part of this simulation, we compare the results of the conditional optimization approach and our proposed joint optimization approach [59]. As shown in Figure 3.5 (right), the population expected outcome under the joint optimization decision rule is higher than that of the conditional optimization rule. In fact, maximizing the dyad-health outcome, where dyad health is the average of the outcome of members of a given dyad, will lead to the maximum expected outcome of the population.

There are a couple of important features to discuss in Figure 3.5. First, comparing only the boxplots of member 1, we see that the conditional approach appears superior in terms of improving the health of these members that have dyadic-network DTR. However, the cost of ignoring interference in the population can be seen in the boxplots of the second members, where these members are seen to be suffering relative to their dyadic partners under the conditional approach. In contrast, under the dyadic-network DTR approach, both member 1 and member 2 boxplots seem to be similar. The consequence of "sacrificing" the conditional optimal treatment for the first members for greater improvements in the

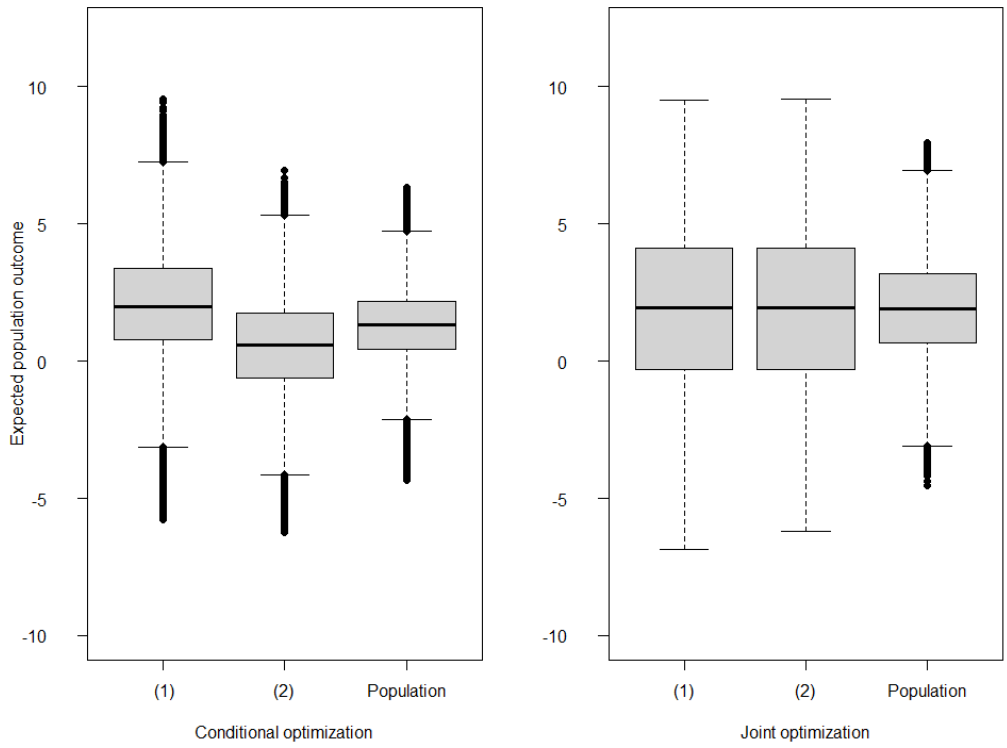


Figure 3.5: Estimated expected population outcome under conditional and joint optimization. The label "(1)" denotes the expected outcome those labelled member 1, while "(2)" denotes the expected outcome of the individuals labelled member 2, while "Population" is the expected population outcome as a result of optimization. Left: Results under conditional optimization Right: Results under the dyadic-network DTR approach.

second member is a better overall population health under the joint optimization approach, as seen by comparing the third boxplots.

3.5 Application: PATH Study

The addictive properties of tobacco products are well-established, with nicotine being the primary compound responsible [61, 62]. In light of this, it is of interest to investigate the

extent to which e-cigarette adoption may reduce total nicotine exposure among pairs of tobacco product users in the United States. To this end, we will analyze data from the Population Assessment of Tobacco and Health (PATH) study [63], which provides a rich source of information on various tobacco-related behaviors and outcomes.

The PATH study is a longitudinal cohort study which was undertaken with the broad aim of accruing information to aid in the assessment of longitudinal patterns of tobacco use behaviours in the United States. A nationally representative sample of households was undertaken using a four-stage, stratified area probability sample design. Sampling began in 2013 with consenting members of each sampled household forming the "Wave 1" cohort. From each of the consenting households, a maximum of two adults and two youth were assessed. Members of the cohort who chose to remain in the study were then contacted annually for further data collection, with the results forming subsequent waves. A replenishment sample was added to the cohort at the outset of Wave 4 and in total, six waves have been measured. More details regarding the logistics of this study can be found in the report of Hyland et al. [63].

Measurements collected consist of answers to a survey regarding tobacco use behaviours, and biospecimens taken from blood and urine samples of study participants who provided the additional requisite consent. The survey can be roughly divided into three main sections. The first section captures various socio-demographic characteristics of each participant, including age, sex, ethnicity, education, employment status, and annual income. The second section focuses on tobacco use, inquiring about the types of products used and frequency of use. Finally, the third section aims to provide context by asking questions about health status, quality of life, and the perceived social role of tobacco among the participants' social circles. Biospecimen samples were used to measure biomarkers of tobacco exposure, such as tobacco-specific nitrosamines, and disease purported to be associated with tobacco use.

To investigate the proposed research question we utilize the PATH dataset to fit a two-stage dyadic network dynamic treatment regime. In accordance with our research question, we restrict our focus to households where two adults were identified as current users of any tobacco product at Wave 1. A current user is defined as any adult who has ever used a tobacco product and now uses it with some degree of regularity [64]. Among these households, we retain for analysis only those households in which both adult members

provided survey responses and biospecimen samples at Wave 2. These conditions yield a sample consisting of 516 households, or 1032 individuals. We randomly assign the label of member 1 and member 2 to each individual in a household.

Our intended response is nicotine exposure, which we propose to quantify using total nicotine equivalents. Total nicotine equivalents involve the sum of molar concentrations of nicotine and its various metabolites. Total nicotine equivalents have been shown to be highly correlated with true nicotine intake, and robust to metabolic differences among individuals [65, 66]. The PATH data provides measurements of total cotinine and total trans-3-hydroxycotinine, which can be summed together to give the measure known as TNE-2 [62]. We choose the log-transformed sum of TNE-2 values at Wave 1 and Wave 2 to be the continuous response for our model. We call this outcome as log total TNE-2.

Our treatment designation is based on e-cigarette use. To be more specific, we consider any adult reported as an e-cigarette user within the year prior to Wave 1 as having received initial treatment (stage 1). Similarly, individuals who are recorded as e-cigarette users between Wave 1 and Wave 2 are considered to have been treated at stage 2. With this setup, PATH study waves correspond to model stages, as defined in Section 3.3. Consequently, our model is a two-stage dynamic treatment regime. For the purpose of comparison, we implement both our joint optimal treatment model and a conditional optimal treatment model as done in previous literature [59].

At each stage, we fit a dyadic-blip model of the form $a_{ijk}(\psi_{0k} + \psi_k^\top X_{ijk}) + a_{ij*k}(\eta_{0k} + \eta_{1k} X_{ij*k}) + \xi a_{ijk} a_{ij*k}$. For the tailoring covariates, at stage 1, X_{ijk} is comprised of biological sex (binary), age (continuous), non-Hispanic (binary), reported heart disease (binary), cigarette smoking status of the individual in the corresponding wave (everyday, some days and not a current smoker), and average cigarettes smoked per day (numeric). If someone is a non-smoker, their average number of cigarettes is zero [67]. The covariate X_{ij^*k} is the same at both stages and is comprised of dyad partners j^*'s cigarette smoking status and average number of cigarettes per day.

Since there is no meaningful way to assign membership labels within the path dataset, we assign labels within each dyad randomly. As a result, member exchangeability is satisfied in this analysis. Accordingly, we fit the outcome and the treatment model on a stacked dataset, combining the member data to fit a single outcome regression.

In each stage the weights were estimated under two interference settings, which we have called M_2 under the restricted model and M_3 under the independent treatment model, respectively (see Appendix A.1.3). The propensity for receiving treatment in each stage exhibits acceptable overlap, and the balance of covariates has been achieved, as indicated by the standardized mean difference of weighted covariates being below 0.1, with the exception of age, which has a standardized mean difference of 0.21.

3.5.1 Results

The dyadic-blip model estimates at each stage for each model are presented in Table 3.1. Recall that our outcome in this case is log-transformed sum of TNE-2 for Wave 1 and Wave 2, and hence lower values correspond to a better health outcome. From this reference point, it is possible to interpret negative blip coefficients as relating to those factors which increase the efficacy of the treatment, while positive coefficients show the opposite relationship. Also, recall that in joint optimization, we optimize the sum of the blips, so cigarette smoking by either member will appear in the optimization equation in two capacities. For example, at stage 1 for Model 2, if both are someday-cigarette smokers, the terms corresponding to cigarette smoking in the blip sum gives $2 * (0.689) + 2 * (0.313) = 2.004$. It follows that, when both members are everyday cigarette smokers, treatment is more likely to be effective. Additionally, if either member is of non-Hispanic origin and is a male, the treatment is more likely to be effective.

Figure 3.6 gives an assortment of boxplots related to the models and their outcomes under optimal treatment. In particular, the first boxplot shows the joint total TNE-2 for all couples as observed in the dataset, while the second and third boxplots show the new outcomes under conditional and dyadic-network DTR modelling, respectively. From the plot, it is clear that the dyadic-network model appears to show an improvement in total TNE-2 values among the sampled households, which can be seen by its noticeably lower central tendency, and skewness toward negative values. The boxplot of the outcomes under the conditional models optimal treatment rule also show a general decrease in total TNE-2 values across the sample, although its central tendency and skewness toward negative values are noticeably less.

Specifically, under the conditional approach, which only optimizes treatments based on

Table 3.1: Dyadic-Blip Estimates

Tailoring covariate		Coefficients			
		Stage 1		Stage 2	
		M_3	M_2	M_3	M_2
ψ_0	E-Cigarette$_1$	-0.084	0.367	-0.746	-0.701
ψ_1	E-Cigarette$_1$: Age$_1$	-0.001	-0.008	0.007	0.012
ψ_2	E-Cigarette$_1$: Sex$_1$ (Female)	0.124	0.078	0.293	0.134
ψ_3	E-Cigarette$_1$: Non-Hispanic$_1$	-0.010	-0.177	-0.135	0.027
ψ_4	E-Cigarette$_1$: Heart Disease$_1$ (No)	-0.839	-0.684	0.189	0.241
ψ_5	E-Cigarette$_1$: Cigarette$_1$ (Not a current smoker)	1.715	1.831	1.257	1.023
ψ_6	E-Cigarette$_1$: Cigarette$_1$ (Some days)	0.689	0.372	-0.187	-0.368
ψ_7	E-Cigarette$_1$: Average Cigarette$_1$ per day	0.015	0.011	0.031	0.019
η_0	E-Cigarette$_2$	-0.498	-0.502	-0.383	-0.316
η_1	E-Cigarette$_2$: Cigarette$_2$ (Not a current smoker)	1.811	1.723	0.550	0.582
η_2	E-Cigarette$_2$: Cigarette$_2$ (Some days)	0.313	0.325	0.150	0.152
η_3	E-Cigarette$_2$: Average Cigarette$_2$ per day	0.013	0.014	0.006	0.005
ξ_0	E-Cigarette$_1$: E-Cigarette$_2$	0.170	0.152	0.569	0.491

Weights were estimated under two interference settings, M_2 under the restricted model and M_3 under the independent treatment model 3.2.3.

the first dyad member, the expected log total TNE-2 values decreased by 0.46 on average across the optimized members. This also indirectly led to a decrease of 0.20 in the non-optimized members. Overall, the population expected log total TNE-2 values decreased on average by 0.33 under this approach, from 3.33 to 3.01. The joint optimization approach of the dyadic-DTR methodology, on the other hand, results in 0.58 reduction in the mean of log sum of TNE-2 of each member and the total of 0.58 reduction for the population of dyads.

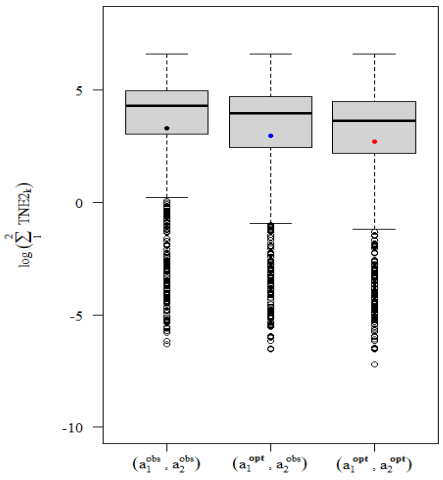


Figure 3.6: Boxplots of log total TNE-2 values in the observed data (Left), under conditional optimization (middle), and under joint optimization (right). In this case both approaches seem to have generally decreased population TNE-2 values, with the dyadic-DTR approach preforming slightly better.

3.6 Discussion

The proposed methodology extends the estimation of dynamic treatment regimes to account for interference when the population arises in the form of a dyadic network.

This is done by modification of the dWOLS algorithm's optimization step through introduction of the dyad-health function. This function takes the outcome values of a particular dyad's members, and uses them to produce a value which, ostensibly, corresponds to dyad health—a value we intend to maximize through treatment assignment. The dyad-health function is specified by the practitioner, and may take multiple forms. In particular, when the dyad-health function is the average of the outcome of the individuals of a dyad, the joint optimization will optimize the population expected outcome. This is in contrast to the conditional optimization methodology introduced previously in the literature which cannot promise this result. Additionally, this feature provides the practitioner flexibility in defining what is meant by optimal as it relates to the outcome of a treatment decision. This

is an important point, because the typical choice of maximizing the population expected outcome can potentially result in undesirable repercussions for certain population members at the individual level, and may therefore be unethical to implement in practice.

The proposed methodology was implemented through a simulation study and an analysis of the PATH data. Through these examples we were able to demonstrate how the models perform. Specifically, when the dyad-health function is chosen to be the average of the dyad outcomes, we demonstrated the claim that the joint optimization approach always outperforms the conditional optimization previously proposed in the literature. Note that this was the sole purpose of the analysis of the PATH dataset—the decision rule estimated therein should not be used to make real policy decisions in this field. In particular, the completed analysis has a few simplifying assumptions that may not hold in practice. For example, it was assumed that dyad members within the PATH study were exchangeable, but the existence of a variable which makes the dyad members distinguishable should be considered and analysis run accordingly and compared. The assumption that the behavior of individuals observed at each wave remained homogeneous, so that they did not undergo any significant changes throughout the one-year interval between waves, is also unlikely to be true generally. While this assumption facilitated our analysis, it is important to recognize that individual behaviors can be influenced by various factors that may evolve or fluctuate over time. Therefore, the generalizability of our findings may be affected, as unaccounted changes in behavior patterns between waves could potentially impact the observed results.

Furthermore, while we were able to showcase the application of the dyadic dWOLS method in a multi-stage setting using two waves of data, it is important to note that the inclusion of additional waves would have provided a more comprehensive understanding of the dynamic nature of the phenomenon under investigation. We were unable to incorporate a larger number of waves into the analysis, due to the limitations imposed by the sample size.

In terms of potential future works in the area of intersection between DTRs and interference, our study motivates several intriguing possibilities. The most natural next step would be to allow for interference among groups consisting of two or more individuals. This, along with allowing for groups of various sizes within the population would extend the model to much more realistic scenarios.

While much of the existing literature on causal inference and interference has primarily focused on binary treatment, exploring interference in the context of continuous treatment could lead to the development of more effective DTR and interference methods. This avenue of investigation would allow for the development of tailored methods specifically designed for use in continuous treatment scenarios.

Another potential extension that merits exploration in the intersection between DTRs and interference involves cases where different sources or types of measurement from one outcome are available. In a more general context, this extension would also encompass situations where a treatment and its potential interference have an impact on multiple health outcomes of an individual. Investigating the effects of interference in such complex scenarios would undoubtedly be challenging, but the insights gained would be valuable in developing more effective intervention strategies that target multiple health outcomes simultaneously. The development of sophisticated analytical approaches that can accommodate the interplay between different health outcomes and the various factors that influence them would enable more personalized and tailored treatment plans, leading to improved patient outcomes and quality of life.

Chapter 4

DTRs on Dyadic Networks with Continuous Treatments

4.1 Introduction

Much work of the early studies in DTR research has primarily focused on binary treatments [11, 68, 69]. However, a significant number of treatments in real-world scenarios are more accurately represented as a continuous variable, such as dose and duration. The concept of utilizing continuous treatment variables in DTR literature has been less explored, primarily due to the added complexity it brings to the optimization process during model fitting. Recent developments have yielded significant progress in this field [10, 70, 71]. Nonetheless, the proposed methods work under the general assumptions of causal inference, such as no interference assumption and developed methods to incorporate interference are currently limited to binary treatment settings [58, 59, 72, 73]. Despite recent advancements in this area, there remains a critical gap in understanding the implications of interference within the context of continuous treatments. In light of this gap, our work aims to bridge this divide by investigating the effects of interference in the utilization of continuous treatment variables within the DTR literature.

In this work we propose one of the first efforts at a DTR methodology which allows for comprehensive modelling of continuous treatments alongside patient interference. The

foundation of the proposed approach rests upon the method that was established in Chapter 3. A key distinction of this chapter is the type of treatment considered; while in Chapter 3 we only examined binary treatments, here we allow for continuous and categorical treatments at both the individual and network level. Moreover, our method enables the modelling of varying degrees of exposure to spillover effects, the effect of one's treatment on the outcome of the other dyad member, by considering both symmetric and asymmetric network connections, and those characterized by heterogeneous intensities. The proposed methodology aims to produce doubly robust estimates of the treatment and spillover effects of a continuous treatment, and quantify the corresponding optimal decision rule. We provide a general formalization of the issue under the potential-outcomes framework, by developing a joint propensity score for the estimation of treatment and spillover effects in observational data [74]. In so doing, we assume that spillover effects are mediated through a dyadic-network structure.

We demonstrate the performance of the proposed model through a detailed analysis of the PATH dataset. Specifically, in households with two tobacco users, we examine whether the nicotine exposure of an individual impacts their tobacco dependence and whether a spillover effect exists. In other words, we investigate whether the nicotine exposure of an individual influences the tobacco dependence of other members.

The structure of this chapter is as follows: Section 4.2 presents the methodological and theoretical background, including notation, and establishes the modeling framework for a K-stage DTR on a dyadic-network population when the treatment is continuous. Section 4.2 focuses on the estimation details of the proposed model, encompassing the treatment and outcome models, as well as the optimal treatment decisions for the dyadic-network DTR. In Section 4.3, we conduct simulation studies to showcase that the proposed methodology maintains the double-robustness property of dWOLS. Section 4.4 features our analysis of the PATH dataset. Lastly, the chapter concludes with a discussion in Section 4.5.

4.2 Methodology

In this section, we introduce notation and provide a brief overview of the dWOLS estimation method. Then we present our proposed methodology for incorporating dyadic interference in the continuous treatment setting.

4.2.1 Notation and assumptions

In Chapter 3, we introduced a set of notations for dyadic-dWOLS when considering a binary treatment. Here, we will build upon those notations, which remain largely the same, with the primary distinction being related to the treatment variable. Our sample of a dyadic-network consists of n pairs of individuals, or $2n$ individuals in total.

We assume that over a certain period of time, the $2n$ individuals underwent K treatment decision points, which we refer to as stages. We let $\mathcal{K} = \{1, 2, ..., K\}$ index the treatment stages, $\mathcal{I} = \{1, 2, ..., n\}$ index the sampled dyads, and $\mathcal{J} = \{1, 2\}$ index the members of a particular dyad. When we need to refer specifically to a quantity related to a particular dyad member, we will use the subscript j to indicate that it belongs to member j, and the subscript j^* to indicate that it belongs to member j's dyadic partner.

For each individual j in dyad i, we use A_{ijk} to denote the treatment of that individual at stage k, for arbitrary k, which has support $\mathcal{A}_{jk}$. That is, we assume member j of each dyad receives the same treatment options. We allow A_{ijk} to be a continuous, categorical, or binary variable, and we allow this property to change with both stage, k, and dyad member, j. We use $\mathbf{A}_{ik} = (A_{i1k}, A_{i2k})$ to denote the treatments received by the two members of dyad i, $i \in \mathcal{I}$. We define this as the dyad treatment. Clearly, $\mathbf{A}_{ik} \in \mathcal{A}_k$ where $\mathcal{A}_k = \mathcal{A}_{1k} \times \mathcal{A}_{2k}$. Finally, $\mathbf{A}_i = (\mathbf{A}_1, \ldots, \mathbf{A}_k)$ denotes the sequence of dyad treatments across all k stages of the treatment regime which has corresponding support $\mathcal{A}$.

At each stage k, it is assumed measurements are taken on each individual prior to treatment. We denote these covariates as X_{ijk}, and use $\mathbf{X}_{ik} = (X_{i1k}, X_{i2k})$ to denote the collection of measured covariates for dyad i. We also define Z_{ik} as the vector of measured covariates shared by the members of dyad i at stage k, which we assume are distinct from those covariates composing $\mathbf{X}_{ik}$. We use $\mathbf{h}_{ijk}$ to represent any measurements relevant for

informing the treatment decision (including individual or shared covariates and previous treatments) of individual j in dyad i at stage k. We call this the individual's history at stage k, and we likewise define the dyad history at stage k as $\mathbf{h}_{ik} = (\mathbf{h}_{i1k}, \mathbf{h}_{i2k})$.

Following completion of the Kth stage, i.e. after the treatment regime has finished, we assess the efficacy of the treatment regime by making a measurement on the health status of interest for each individual, denoted by Y_{ij}. We let $\mathcal{Y}$ denote the space of possible outcomes. It is assumed that the outcome is continuous and that larger values of the outcome are preferable. Under interference, the individual-level potential outcome is a function of the full sequence of dyad treatments $\mathbf{A}$. Occasionally it will be necessary to make this dependence explicit, and in such a case we will use the more explicit notation $Y_{ij}(\mathbf{A})$ for the outcome of individual j in dyad i, while $Y(\mathbf{A}) = \{Y_{i1}(\mathbf{A}), Y_{i2}(\mathbf{A})\}$ denotes the dyad outcome.

Our model operates under the potential outcomes framework introduced in Chapter 1 [53]. This assumes that each person has an associated outcome function which maps the treatment space into the outcome space. This is the $Y_i(\cdot)$ which we defined previously. The notion of potential outcomes follows from the fact that any of these outcomes could potentially be observed, depending on the treatment that is ultimately assigned. The potential outcomes framework works by making a few assumptions which allow for the estimation of causal effects when working with observational data. From the perspective of dyadic-network DTRs, the assumptions of consistency, no unmeasured confounders, and positivity still hold. However, the assumption of no interference will be relaxed. We assume that the potential outcomes of a particular dyad are independent of the treatment assignment of individuals belonging to other dyads. The interference within the dyad may exist under this definition, whereas interference across dyads is ruled out. That is,

Assumption 4.2.1 (Dyadic-Interference)**.** Suppose $Y_i(\mathbf{a}_i)$ is the outcome of dyad i after having received treatment $\mathbf{a}_i$ over a K stage regime. Then, $Y_i(\mathbf{a}_i) \mid \mathbf{a}_{i'} = Y_i(\mathbf{a}_i)$ for all $i' \neq i$, and for all $\mathbf{a}_i \in \mathcal{A}$.

A potential outcome is a value of the response that would result if a participant were to receive a specified treatment, not necessarily the same as that they received in the study. To modify the potential outcome to permit the presence of dyadic-network interference the observable outcome for an individual is a function of the treatment of themselves and

their corresponding dyadic pair, i.e $Y_{ij}(\mathbf{A})$ for $\mathbf{A} \in \mathcal{A}$. In the event that all treatments are continuous, the outcome is referred to as the dose–response surface.

In terms of modelling, we can sometimes impose a simplifying assumption called exchangeability. If exchangeability holds, we have that $(Y_1, X_1, X_2, Z) \stackrel{d}{=} (Y_2, X_2, X_1, Z)$. When exchangeability holds, we say that the dyad members are exchangeable.

4.2.2 dWOLS estimation method for a K-stage DTR in a dyadic network

Under the assumption of no interference, the sequence of treatment decisions that maximizes the outcome of interest for an individual or subgroups (as reflected on the covariates), will result in optimization of population outcome as well. However, as discussed in Chapter 3, when interference exists, optimization methods that aim to achieve an individual's optimal treatment may not even result in increasing the population outcome on average, let alone optimizing it. Thus, when interference exists, it is therefore crucial to clarify the goal of the DTR and identify the target population we aim to benefit from the treatment plan. This is achieved by defining a dyad health outcome function, $g\left[\mathbf{Y}(\mathbf{a})\right]$ which is a surrogate measure of the each dyad's overall health. In so defining, we account for the assumed interference within the population. Optimization of this dyad health function rather than the individual health outcomes then optimizes the population outcome.

In the simplest case, the dyad health function, combines the effect of treatment on each individual and its effect on another outcome (spillover effect) into a single scalar utility score. In a more complex way, the dyadic health function could aim to maximize an individual's outcome subject to constraints on the risk of the other individual. With the objective of maximizing the expected population outcome, we define the dyadic health function as the average of the outcome of each member of a dyad, $g\left[\mathbf{Y}(\mathbf{a})\right] = 0.5(Y_1(\mathbf{a}) + Y_2(\mathbf{a}))$.

When it comes to computational aspects, our approach extends the dWOLS algorithm and therefore relies on backward induction. To formalize the dyadic-network DTR model, we modify the individual's outcome model to incorporate interference. That is, the outcome model of each member can potentially include the covariates and treatments of the

other member. Therefore, the treatment-free component specifies the relationship between the tailoring covariates and the outcome when none of the members of a dyad received treatment (i.e., $a_1 = a_2 = 0$). The blip component, characterizes the joint-treatment effect, including both the main effects of joint-treatment and the interaction(s) between treatments and the tailoring covariates and is called dyadic-blip. The outcome model for the members of dyad i can be written as,

$$E[Y_{i1} \mid \mathbf{h}_{i1K}, \mathbf{a}_{iK}] = \sum_{k=1}^{K} \left[f_{1k}(\mathbf{h}_{i1k}; \boldsymbol{\beta}_{1k}) + \gamma_{1k}(\mathbf{h}_{i1k}, \mathbf{a}_{ik}; \boldsymbol{\Theta}_{1k}) \right], \quad j = 1,$$
$$E[Y_{i2} \mid \mathbf{h}_{i2K}, \mathbf{a}_{iK}] = \sum_{k=1}^{K} \left[f_{2k}(\mathbf{h}_{i2k}; \boldsymbol{\beta}_{2k}) + \gamma_{2k}(\mathbf{h}_{i2k}, \mathbf{a}_{ik}; \boldsymbol{\Theta}_{2k}) \right], \quad j = 2,$$

where f_k is the stage k treatment-free model, parameterized by $\boldsymbol{\beta}_{jk}$, and γ_{jk} is the corresponding blip model, parameterized with parameter $\boldsymbol{\Theta}_{jk}$, for the member j of a given dyad. When extending the framework to continuous treatment, the dyadic-blip function should typically be parameterized in such a way to ensure that the joint-optimal treatment $\mathbf{a}^{opt}$ could potentially be an interior point of $\mathcal{A}_1 \times \mathcal{A}_2$. For example, the analyst may posit a blip model that includes a quadratic (or higher order) term for the treatments.

We start by considering the final treatment decision, which is made at stage K. With the goal of optimization of the expected outcome of the population, we define the dyad health outcome as the sum of the outcome of two members. Assuming higher values of Y are preferred, the optimization will be a maximization problem. To secure that the optimal treatment relies on the interior point of $\mathcal{A}_1 \times \mathcal{A}_2$ we define a dyadic blip with quadratic terms of joint treatment. Hence the outcome model for dyad i at stage K is,

$$E[Y_{i1} \mid \mathbf{h}_{i1}, \mathbf{a}_i] = \boldsymbol{\beta}_1^\top \mathbf{h}_{i1} + \psi_{10} a_1 + \eta_{10} a_2 + a_1^2 \boldsymbol{\psi}_{11}^\top \mathbf{h}_{i1} + a_2^2 \boldsymbol{\eta}_{11}^\top \mathbf{h}_{i1} + a_1 a_2 \boldsymbol{\xi}_{11}^\top \mathbf{h}_{i1}, \quad j = 1,$$
$$E[Y_{i2} \mid \mathbf{h}_{i2}, \mathbf{a}_i] = \boldsymbol{\beta}_2^\top \mathbf{h}_{i2} + \psi_{20} a_2 + \eta_{20} a_1 + a_2^2 \boldsymbol{\psi}_{21}^\top \mathbf{h}_{i2} + a_1^2 \boldsymbol{\eta}_{21}^\top \mathbf{h}_{i2} + a_2 a_1 \boldsymbol{\xi}_{21}^\top \mathbf{h}_{i2}, \quad j = 2,$$

where we have, without loss of generality, specified each term to be a function of the entire history. Note that, for simplicity, the subscripts of stage K have been omitted from the outcome model equation. The optimal treatment allocation can be found by maximizing,

$$\underset{(a_1,a_2)}{\text{argmax}}\ E[Y_{i1}(a_1,a_2)+Y_{i2}(a_1,a_2)\mid \mathbf{h}_{i1},\mathbf{h}_{i2}] =$$
$$\underset{(a_1,a_2)}{\text{argmax}}\ \Big[\psi_{10}a_1+\eta_{10}a_2+a_1^2\psi_{11}^\top\mathbf{h}_{i1}+a_2^2\eta_{11}^\top\mathbf{h}_{i1}+a_1a_2\xi_{11}^\top\mathbf{h}_{i1}$$
$$+\psi_{20}a_2+\eta_{20}a_1+a_2^2\psi_{21}^\top\mathbf{h}_{i2}+a_1^2\eta_{21}^\top\mathbf{h}_{i2}+a_2a_1\xi_{21}^\top\mathbf{h}_{i2}\Big],$$

subject to the condition that $(a_1,a_2)\in\mathcal{A}_1\times\mathcal{A}_2$. This equation can be written succinctly in matrix notation as,

$$\frac{1}{2}\mathbf{a}^\top\mathbf{Q}_i\mathbf{a}+\boldsymbol{\phi}^\top\mathbf{a},$$

where we have defined $\phi=(\psi_{10}+\eta_{20},\psi_{20}+\eta_{10})$ and,

$$\mathbf{Q}_i=\begin{bmatrix}\psi_{11}^\top\mathbf{h}_{i1}+\eta_{21}^\top\mathbf{h}_{i2} & 1/2(\xi_{11}^\top\mathbf{h}_{i1}+\xi_{21}^\top\mathbf{h}_{i2})\\ 1/2(\xi_{11}^\top\mathbf{h}_{i1}+\xi_{21}^\top\mathbf{h}_{i2}) & \psi_{21}^\top\mathbf{h}_{i1}+\eta_{11}^\top\mathbf{h}_{i2}\end{bmatrix}.$$

As a result, the optimal treatment decision rule can be found by solving the following quadratic programming problem:

$$\begin{aligned}\text{maximize }\frac{1}{2}\mathbf{a}^\top\mathbf{Q}_i\mathbf{a}+\boldsymbol{\phi}^\top\mathbf{a},\quad &\text{subject to}\\ \mathcal{A}=\mathcal{A}_1\times\mathcal{A}_2=\{\mathbf{a}\mid\mathbf{l}\preceq\mathbf{a}\preceq\mathbf{u}\}.&\end{aligned}\tag{4.1}$$

Here, $\mathbf{l}_i$ is the left endpoint of $\mathcal{A}_i$, while $\mathbf{u}_i$ is the corresponding right endpoint, $i=1,2$. Problems such as the one given in Equation (4.1) are commonly called quadratic optimization problems with box constraints (assuming that $\mathcal{A}_1$ and $\mathcal{A}_2$ are intervals). The quadratic form of our model will always have dimension two. It follows that, for each patient we will be maximizing a surface that has one of the three forms given in Figure 4.1. In particular, when $\mathbf{Q}_i$ is negative semi-definite, this optimization problem will be convex. If the maximum is an interior point of $\mathcal{A}$, then the solution can be found by setting the gradient equal to 0. If $\mathbf{Q}_i$ is instead positive semi-definite, the problem is still easily solved by checking all four vertices and choosing the one corresponding to the largest value. Finally, if $\mathbf{Q}_i$ is of indeterminate form, we must maximize on each of the

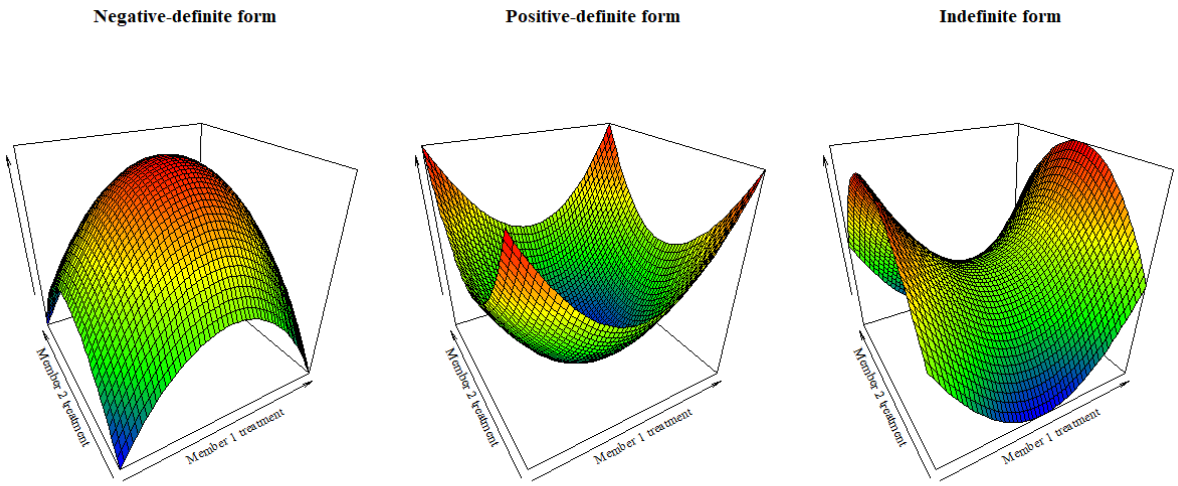


Figure 4.1: Illustration of the general shape of the outcome model as a function of the dyad treatment for definite and indefinite $\mathbf{Q}_i$.

boundary values of the treatment corresponding to the positive eigenvalue. In all cases, the maximization can be completed in closed form so that no numerical solvers are required.

Surface and contour plots then can be used to inform the decision maker (see, for example, Figure 4.1).

Following the same process as dWOLS, once the optimal treatment at last stage has been chosen, the pseudo-outcome for each dyad member under the dyadic-optimal treatment assignment is estimated by,

$$\tilde{y}_{ij(k-1)} = \tilde{y}_{ijk} + [\gamma_{ijk}(\mathbf{h}_{ijk}, \mathbf{a}_{ijk}^{\text{opt}}) - \gamma_{ijk}(\mathbf{h}_{ijk}, \mathbf{a}_{ijk})], \quad j \in \mathcal{J}, \forall i \in \mathcal{I}.$$

Intuitively, we are correcting for the effect of the observed treatment, which was (potentially) suboptimal, by subtracting off its effect while simultaneously adding on the effect of the optimal treatment. If the optimal treatment and the observed treatment coincide, the added quantity will be zero. Note that in the setting without interference this is happening by adding a positive quantity to the observed outcome of the individuals who did

not receive their optimal treatment at the stage k. However, this might not necessarily be the case where interference exists. The optimal decision is a treatment decision that leads to the maximum expected outcome of a dyadic health outcome rather than an individual outcome. We continue this procedure and estimate the optimal joint treatment for each k in $\mathcal{K}$ until all stages have been addressed, after which we will have obtained K decision rules with which we may inform treatment decisions of future patients at each stage of the regime.

We extend the balancing weights of Schulz and Moodie [10] to continuous treatment within dyad interference. Under interference within dyads, balance must be achieved across arms defined not only by the individual treatment but also by their dyadic partner treatment.

Theorem 4.2.2 (Extension of Schulz & Moodie, 2020)**.** *A weighted ordinary least-squares regression based on the model,*

$$E[Y_j \mid \mathbf{h}_j, \mathbf{a}] = f(\mathbf{h}_j; \boldsymbol{\beta}) + \gamma(\mathbf{h}_j, \mathbf{a}; \boldsymbol{\Theta}),$$

will give consistent estimates for the dyadic-blip parameters $\boldsymbol{\Theta}$ *provided that the weights satisfy,*

$$\pi(\mathbf{a} \mid \mathbf{h}) w(\mathbf{a}, \mathbf{h}) = \pi(\mathbf{a}' \mid \mathbf{h}) w(\mathbf{a}', \mathbf{h}) \tag{4.2}$$

for any $\mathbf{a} \neq \mathbf{a}' \in \mathcal{A}$*, where* $\pi(\mathbf{a} \mid \mathbf{h}) = f_{A|\mathbf{h}}(\mathbf{A} = \mathbf{a} \mid \mathbf{h})$ *is the joint propensity score.*

A proof of this theorem follows that by Schulz and Moodie [10] and is provided in Appendix B.1. Theorem 4.2.2 provides a general class of weights which can provide the balancing property. In order to estimate these weights, for the continuous treatment case, we fit a joint treatment model based on the ideas of Del Prete et al. [75]. Their work defines a generalized propensity score for populations exhibiting a weighted network structure. They call this the *joint propensity score*. In our scenario, the network treatment is simply the treatment of the other dyad member.

In its general form we assume that the treatment assignment of those in the same dyad are not independent, then we assume a joint model for the treatments, $(A_1, A_2) \sim f(\mathbf{H}; \theta)$

where $\mathbf{H}$ denotes the dyad history. With the treatment model taken care of, we construct the weights as

$$w_{ijk} = \frac{\pi(a_{ijk}, a_{ij^*k} \mid h_{ijk})^{-1}}{\int \pi(u, v \mid h_{ijk})^{-1} \, dudv}.$$

As it is shown in Appendix B.1, these weights satisfy the balancing condition, leading to double robust estimation property of dyadic-blip parameters.

4.3 Simulation Study

In this section, we investigate the properties of our proposed method via simulation study.

Through simulation, we evaluate the double-robustness of the proposed dyadic-dWOLS estimation strategy by examining different types of model misspecification. Four scenarios are presented: 1) correct specification for both joint treatment and treatment-free models, 2) misspecification of joint treatment model but correct specification of treatment-free model, 3) correct specification of treatment model but misspecification of treatment-free model, and 4) misspecification of both joint treatment and treatment-free models. If the proposed dyadic-dWOLS estimation method is double-robust, we should get consistent estimation of the dyadic-blip parameters in scenarios 1-3, where at least one of the treatment model or treatment free model is correctly specified, but not necessarily in scenario 4, where both are misspecified.

In the first set of simulations, we generate the tailoring covariate x_{ij} from a folded Normal with mean 1 and standard deviation 1 and a shared covariate z_i from a Normal distribution with mean 3 and standard deviation 1 for $j \in 1, 2$. The treatments A_{i1} and A_{i2} each follow a Normal distribution with mean $\alpha_0 + \alpha_1 x_{ij} + \alpha_2 z_i$ and variance 1, and the parameter vector α is set to $(-3, -2, 0.7)$. The treatment-free model is represented by the equation $x_{ij} + \sin(x_{ij}) + z_i$, and the dyadic-blip model is $a_{ij}(\psi_0 + \psi_1 a_{ij} + \psi_2 a_{ij} x_{ij}) + a_{ij^*}(\psi_3 + \psi_4 a_{ij^*} + \psi_5 a_{ij^*} x_{ij}) + \psi_6 a_{ij} a_{ij^*} x_{ij}$ where the parameter vector $\mathbf{\Psi}$ is set to $(-1, -3, -10, -1, -2, -10, 2)$ and the error term e_{ij} is assumed to follow a standard Normal. The treatment-free model is misspecified by the omission of $\sin(x_{ij})$ and the treatment model is misspecified by excluding x_{ij}. The second set of simulations involves heteroscedastic doses, which are generated according to the model $\text{Dose}_{ij} = ce^{A_{ij}}$, where

A_{ij} is determined by a regression on the tailoring covariates $-3 - 0.1x_{ij} + e_{ij}$. Model misspecifications in this study remain the same as the previous simulation.

As discussed by Schulz and Moodie [10], weights such as *inverse probability weights* and *overlapping weights* [57], satisfy the balancing condition for dWOLS in the context of continuous treatment without interference. In this work, we modify these weights to satisfy the balancing condition presented in Theorem 4.2.2 for the dyadic-dWOLS estimation method and refer to them as IPWD and OWD, respectively. It is important to note that Theorem 4.2.2 introduces a sufficient but not necessary condition, indicating that there exist weights, such as *stabilized inverse probability of treatment* weights, which may not satisfy the condition but still achieve balance. Furthermore, we also examine the performance of non-parametric weights, such as *entropy weights*, which are described in Table 4.1. We call these weights STWD and EWD, respectively. Additional details on different types of balancing weights can be found in Appendix B.3.

Note that as Schulz and Moodie [10] discussed, for a continuous treatment, weights derived from the generalized propensity score may be volatile. To address this, we considered binning the continuous treatments according to joint empirical quantiles by

$$\pi\left[(a_{ij}, a_{ij^*}) \mid x_{ij}\right]^{-1} / \sum_{r=1}^{m} \pi\left[(a_{ij}, a_{ij^*})_r \mid x_{ij}\right]^{-1}.$$

In the dyadic-blip model, however, treatments still behave as the continuous variables. The simulation results presented throughout are based on 1000 Monte Carlo replications of different scenarios. The simulations were all carried out using R.

4.3.1 Results

The results of the estimation of dyadic-blip parameter ψ_1, for the simulation setting where the treatments follow the Normal distribution, are presented in Figure 4.2 and Figure 4.3. Figure 4.2 showcases three sets of boxplots representing different sample size settings for estimating the dyadic-blip parameter ψ_1 when the treatment-free model is misspecified. Each boxplot within a subfigure corresponds to a specific type of balancing weights. In Figure 4.3, each sub-figure corresponds to one of the four analysis scenarios under different balancing weights. These plots indicate that the model indeed provides double-robust

Table 4.1: Weight functions implemented in the dyadic-dWOLS.

Weight Function	**Description**
IPWD (inverse probability weighting)	$\frac{1}{\pi(\mathbf{a}\|x)}$
EWD (entropy balancing weights) [76–78]	$\mathrm{H}(\omega_i) = \omega_i \ln(\frac{\omega_i}{q_i})$
STWD (stabilized inverse probability weights) [79]	$\frac{\pi(\mathbf{a})}{\pi(\mathbf{a}\|x)}$

estimation for dyadic-blip parameter ψ_1. As expected, when at least one of the model components is correctly specified, the dyadic-blip parameter estimates appear to be consistent, which can be surmised from their central tendency of zero and their shrinking variance as the sample size increases.

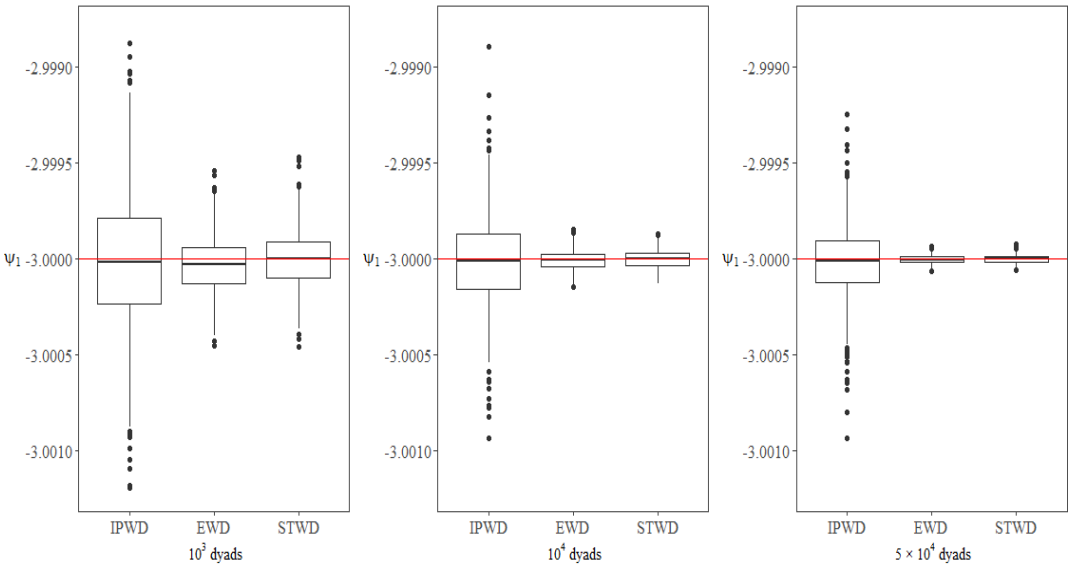


Figure 4.2: The estimates for ψ_1 under different dyad sample sizes and for different types of weights, under the scenario (3) where treatment-free model is misspecified.

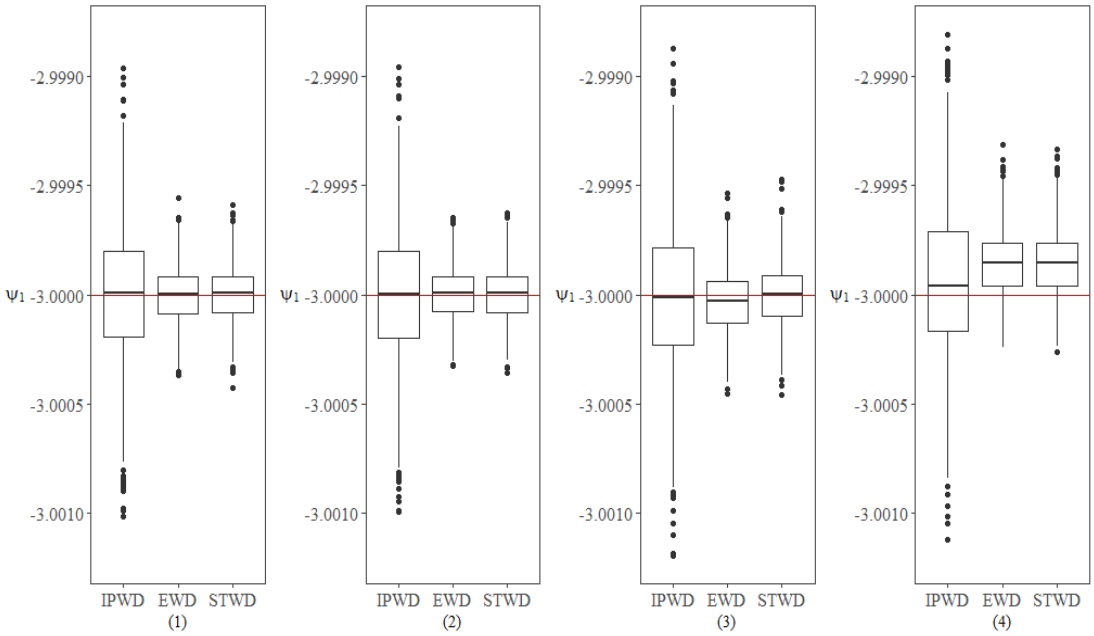


Figure 4.3: The estimates for ψ_1 under different model specifications, (1) correct specification of both models, (2) and (3) when one of the treatment model or treatment-free model is misspecified and (4) when both models are misspecified. The sample size is 1000 dyads and under 1000 Monte Carlo replication.

Furthermore, the table of root mean squared error, RMSE, (Table 4.2) provides insights into the estimation performance of the dyadic-blip parameters across different sample sizes and balancing weights, when treatment-free model is misspecified. First, as the sample size increases the RMSE values generally decrease for all parameters and weights, indicating improved estimation precision with larger sample sizes. Comparing the different balancing weights within each sample size, the IPWD weight tends to yield larger RMSE values compared to EWD and STWD weights. This is expected as dividing by the propensity score can inflate the variance of IPWD estimators. Examining the individual parameters, ψ_0 and ψ_3 generally have higher RMSE values. The findings presented in this analysis align with the results reported by Schulz and Moodie [10] in their study under a no-interference setting. For simplicity, the results for other dyadic-blip parameters under different model specifications, which also exhibit the double-robustness property, have been included in Appendix B.2.

		Dyadic-blip parameters						
Weight	Sample size	ψ_0	ψ_1	ψ_2	ψ_3	ψ_4	ψ_5	ψ_6
IPWD	10^3	0.437	0.035	0.034	0.419	0.033	0.035	0.028
	10^4	0.293	0.022	0.022	0.296	0.020	0.022	0.015
	5×10^4	0.242	0.018	0.017	0.258	0.017	0.017	0.013
EWD	10^3	0.153	0.015	0.016	0.159	0.017	0.018	0.015
	10^4	0.048	0.005	0.005	0.049	0.005	0.006	0.005
	5×10^4	0.021	0.002	0.002	0.022	0.002	0.003	0.002
STWD	10^3	0.162	0.014	0.016	0.159	0.015	0.018	0.015
	10^4	0.051	0.005	0.005	0.050	0.005	0.006	0.005
	5×10^4	0.022	0.002	0.002	0.022	0.002	0.003	0.002

Table 4.2: RMSE when the treatment-free model is misspecified. Sample size is 1000 dyads and the treatments follow Normal distribution. RMSE values are multiplied by 10^2.

4.4 Application: PATH Study

Total Nicotine Equivalents (TNE), which represent the molar sum of nicotine and its metabolites within an individual's system, are increasingly recognized as valuable biomarkers for investigating tobacco dependence. Previous research has shown a correlation between TNE levels and tobacco dependence, suggesting that higher levels of these biomarkers often imply a stronger dependency on tobacco. Moreover, it has been observed that different tobacco products can yield different relationships between TNE and tobacco dependence, underlining the complexity and variability of this dependency[80].

In this paper, we aim to utilize the proposed model to explore the relationship between TNE and tobacco dependence, as well as the existence of a spillover effect. Our goal is to not only contribute to the existing body of knowledge but also to encourage a broader discussion on tobacco dependence beyond the confines of cigarette consumption.

To carry out the proposed study, we turn to the Population Assessment of Tobacco and Health (PATH) dataset. The PATH Study represents a comprehensive longitudinal investigation centered on the monitoring and analysis of tobacco use patterns within the United States.

Initiated in September 2013, the first wave of the PATH Study consisted of a nationally representative household sample of 45,971 individuals, including 32,320 adults and 13,651 youths. The study involved annual follow-ups, spanning up to six waves of data collection, with the most recent wave recorded in 2021. At each wave, participants were asked to fill out a survey and upon providing additional consent, their biospecimens were collected.

4.4.1 Analysis

We make use of the PATH dataset to construct a two-stage dyadic network dynamic treatment regime, which will allow us to study the interactions and effects between paired individuals across the first two waves of the study. For our analysis, we specifically focus on pairs of individuals engaged in ongoing tobacco use. An individual is classified as a "current tobacco user" if they have previously used a tobacco product and continue to do so regularly [64]. For our analysis, we chose a sample comprised of households in which two adults were recognized as current tobacco users during the first wave of data collection, and who additionally had their biospecimens gathered in both the first and second waves. From this specific selection process, we ended up with a total of 516 households, which collectively represented a population of 1032 individuals. Subsequently, within each household, we randomly assigned a unique membership label to each individual, categorizing them as either "member 1' or "member 2'.

We now describe how we intend to formulate a measurable representation of tobacco dependence. To do this, we build a measure based on existing literature, making use of pertinent data from the survey responses from among those included in our chosen sample. Specifically, a tobacco dependence score is crafted based on the responses to twelve targeted questions on the Population Assessment of Tobacco and Health (PATH) survey, originating from the Wisconsin Inventory of Smoking Dependence Motives. We apply a factor analysis model with a single latent variable to these data. The factor scores produced from this model become our measure of tobacco dependence, serving as the dependent variable in our research endeavor. Fitted score values vary from -3 to 3, where higher values of the score pertain to stronger tobacco dependence.

Our intended treatment is total nicotine equivalents, which involves the sum of molar concentrations of nicotine and its various metabolites. Total nicotine equivalents have been

shown to be highly correlated with true nicotine intake, and robust to metabolic differences among individuals [65, 66]. The PATH data provides measurements of total cotinine and total trans-3-hydroxycotinine, which can be summed together to give the measure known as TNE-2 [62]. We choose log-transformed TNE-2 values to be the continuous treatment for our model. total nicotine equivalents (TNE-2, the molar sum of the two major nicotine metabolites)

In terms of covariates, at each model stage, age is considered as the baseline covariate, and smoking status, which indicates whether or not an individual smokes cigarettes, is used as the tailoring covariate X_{ij}. Hence, the stage j dyadic-blip model takes the form, $a_{ij} + (\psi_0 + \psi_1 a_{ij} + \psi_2 x_{ij}) + a_{ij^*}(\psi_3 + \psi_4 + a_{ij^*}\psi_5 x^*_{ij^*}) + \psi_6 a_{ij} a_{ij^*} x_{ij}$. Overlapping weights of the form $w_j = \pi(a_j, a_{j^*}) \mid x_j, x_{j^*})^{-1} / \int \pi(u, v \mid x_j, x_{j^*})^{-1} \, dudv$ are used. Specific details regarding weight estimation are presented in Appendix B.

Table 4.3: Dyadic-Blip Estimates

	Tailoring covariate	Coefficient $\times 10^{-2}$	
		Stage 1	Stage 2
ψ_0	log-TNE$_1$	4.44	-29.41
ψ_1	log-TNE$_1^2$	-1.73	-3.60
ψ_2	log-TNE$_1^2$: Cigarette$_1$	-1.82	2.36
ψ_3	log-TNE$_2$	-4.86	-2.72
ψ_4	log-TNE$_2^2$	0.18	0.51
ψ_5	log-TNE$_2^2$: Cigarette$_1$	3.02	0.47
ψ_6	log-TNE$_1$:log-TNE$_2$: Cigarette$_1$	0.49	0.85

The average dependency score of the population is 0.115 which, as a result of joint optimization, reduced to -0.255. The log-TNE2 treatment level varies among individuals from -0.62 to 0.11.

Dyadic-blip parameter estimates for the fitted two-stage model are given in Table 4.3.

The resulting dyadic-blip at the first stage is then,

$$\begin{aligned}\gamma_1(\mathbf{h}_1, \mathbf{a}_1) =& a_1^2(-1.73 - 1.82\text{Cigarette}_1 + 3.02\text{Cigarette}_2) - 0.42a_1 + \\ & a_2^2(-1.73 - 1.82\text{Cigarette}_2 + 3.02\text{Cigarette}_1) - 0.42a_2 + \\ & 0.49(\text{Cigarette}_1 + \text{Cigarette}_2)a_1a_2.\end{aligned}$$

We therefore find $\hat{\boldsymbol{\phi}} = (-0.42, -0.42)$ at the first stage, and since there are four possible ways to distribute smoking status among the two members of a dyad, there are four possible values of Q. When both dyad members are smokers, i.e. $\text{Cigarette}_1 = 1$ and $\text{Cigarette}_2 = 1$, we have,

$$Q = \begin{bmatrix} -0.53 & 0.49 \\ 0.49 & -0.53 \end{bmatrix},$$

which is a negative definite matrix. It follows that the optimal $\log \text{TNE2}$ values are $(-1.5, -1.5)$. When neither is a smoker, so that $\text{Cigarette}_1 = 0$ and $\text{Cigarette}_2 = 0$, the matrix Q is diagonal with entries -1.73. It follows that the optimal solution in this case is $(-0.24, -0.24)$. Finally, in the case where one dyad member is a smoker and one is not, the matrix Q is of indeterminate form. In such a case, we need to set a boundary condition for the treatment corresponding to the positive eigenvalue of Q. In this case, we choose $\log 0.0002$, which corresponds to the minimum measurable value of the treatment in the study. Under this scenario, the optimal treatment is then -1.29 for the smoker and -8.52 for the non-smoker.

The resulting change in the outcome values under the dyadic-network regime can be seen in Figure 4.4. For completeness, we also show a boxplot of the true outcomes that would be obtained if one were to fit a conditional regime, in the style of [59], although the methodology presented there did not accommodate continuous treatment values. We can see from the boxplots that the dyadic-network DTR approach results in significant improvements in the outcomes, which are, as plotted, negative tobacco dependence scores.

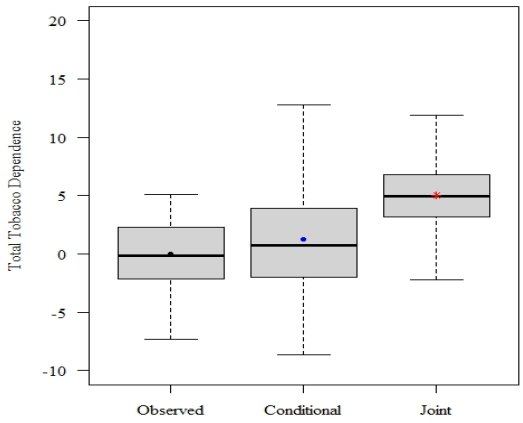


Figure 4.4: Boxplots of the dyad outcomes under the observed data [Left], a conditional optimization approach [Middle] and a dyadic-DTR approach [Right].

4.5 Discussion

We have extended the literature on traditional dynamic treatment regime estimation through the introduction of the dyadic-dWOLS methodology. This approach, in a significant deviation from previous models, takes into account the role of interference within a dyadic network population when treatment is continuous.

The methodology is comprehensive in that it is capable of accommodating different types of interference, specifically those that are symmetric (exchangeable members influencing each other to the same degree) or asymmetric (non-exchangeable members where one member's influence on the other is not reciprocated to the same extent). This adaptability is made possible through the use of a modified optimization step, which utilizes the dWOLS method in combination with a dyad-health function. This feature allows users to clearly define what they consider to be 'optimal' outcomes, thus expanding its applicability across various research contexts. Additionally, the dyadic-dWOLS approach enables the inclusion of ethical considerations into the treatment analysis, making it a robust tool for modern medical research. This flexibility in defining optimal outcomes and ethical parameters empowers researchers to apply the methodology in a variety of treatment research

areas.

Our proposed estimation procedure successfully preserves the double robustness property inherent to the original dWOLS methodology. We achieved this by specifying certain balancing weights—when weights are used which satisfy the conditions stated in Theorem 1, the approach will be doubly-robust. This property was also verified empirically through our simulation study, which showed that estimation was indeed doubly-robust under various weight forms. Through simulations and data analyses, we demonstrate our methodology where treatments of members followed the same distribution type. However, the proposed methodology is designed to be versatile, extending its applicability to situations involving mixed treatments. The evaluation of weight estimations, along with the application of the balancing condition, particularly in scenarios of increased complexity, is a focus of our current research.

We have showcased our methodology through an analysis of the Population Assessment of Health and Tobacco (PATH) data. In this study, we delved into the potential effect of nicotine exposure on tobacco dependence, employing a two-stage dyadic-DTR estimated with the dyadic-dWOLS procedure. In this analysis we purposefully limited the number of tailoring covariates to a select few, based on relevant literature. This decision was made to produce an effective and easy to understand demonstration of the methodology, rather than to offer a specific decision-rule to be used for actual policy-making.

It is crucial to note that our analysis operated under the assumption that nicotine exposure is a potential cause of tobacco dependence. While some studies lend support to this assumption, its validity is still a topic of ongoing debate. There are research findings that suggest a simultaneous influence of tobacco dependence and nicotine exposure, presenting a more complex interplay. Currently, our proposed method is not designed to accommodate simultaneous outcomes. Nevertheless, expanding the methodology to consider simultaneous and multiple outcomes presents an exciting and significant path for future research. These extensions could significantly enhance our understanding of complex health behaviors, such as tobacco dependence, and potentially lead to more refined and effective treatment strategies.

Chapter 5

DTRs and Multiple Outcomes

5.1 Introduction

The estimation of optimal DTRs has primarily focused on single scalar outcomes in the existing literature. However, in clinical decision settings, it is often necessary to consider multiple, potentially competing, outcomes in order to ensure comprehensive care. The current methods proposed to handle multiple outcomes rely on Q-learning, which relies on the fully correct specification of outcome models and suffers from a lack of robustness, especially with respect to model misspecification [5, 6, 8, 12, 81, 82]. In this chapter, we extend the dWOLS estimation method to accommodate multiple outcomes in DTRs. The dWOLS has been shown to be effective in estimating optimal treatment regimes based on a single outcome. By extending dWOLS to handle multiple outcomes, we aim to provide a regression-based approach that balances competing outcomes while considering constraints. The proposed extension not only preserves the double robustness of dWOLS but also allows the clinicians to incorporate patient-specific costs (when available to them) rather than imposing a global population cost.

The remaining of this chapter is as follows: Section 5.2 introduces the notation and estimation of dWOLS for multiple outcomes. In particular, we modify the optimization step to impose the constraint on the intended outcome. Section 5.3 provides simulation studies to demonstrate double robustness property and performance of constrained optimization. We conclude by discussion and future work in section 5.4.

5.2 Methodology

5.2.1 Notation

In the framework of DTR, it is posited that patients undergo multiple stages of treatment. This process commences with the initiation of treatment, followed by subsequent treatment decisions made at predetermined intervals, taking into account the current information available for each patient.

In accordance with the established literature on DTR for a binary treatment setting with a single outcome (as detailed in Section 2.3), we employ $\mathcal{K} = 1, 2, ..., K$ to index the stages of treatment and $\mathcal{I} = 1, 2, ..., n$ to index individual patients. The set of treatment options available at stage k is denoted by $\mathcal{A}_k$. Given our consideration of binary treatment, $\mathcal{A}_k$ is isomorphic to $\{0, 1\}$. Prior to the treatment at stage k, a series of covariate measurements are assumed to be taken for each individual. These covariates are represented by X_{ik}. We define H_{ik} as the collection of measurements taken prior to treatment at stage k, which serve to inform the subsequent treatment decision. This history, H_{ik}, for individual i at stage k, encompasses both current and past stage covariates, along with the values of previous treatments. Observed values of the history will be denoted by lowercase notation, h_{ik}. At the end of K-stage treatment regime a health outcome Y is measured. We assume that the higher values of Y are preferred.

In a typical DTR with a single outcome, the goal is to find the collection of decision rules which maximizes the expected outcome of Y. We use $d = d_k(\cdot)_{k\in\mathcal{K}}$ to denote the set of K decision rules, where each d_k is a map, $d_k : \mathcal{H}_k \to \mathcal{A}_k$, taking an observed history $h_i \in \mathcal{H}$ to a recommended kth stage treatment $A_{ik} \in \mathcal{A}_k$ for that individual. Thus, d defines a treatment regime which can be used to systematically treatment patients, and we define $\mathcal{D}$ to be the space of all possible treatment regimes.

In the multiple outcome setting, we extend our analysis to consider that, upon completion of the Kth stage, multiple outcomes of interest are measured for each patient. This differs from the single outcome scenario where we only observed a single outcome variable Y. Here, we represent the vector of outcomes associated with patient i as $\mathbf{Y}_i = (Y_{(1)i}, \ldots, Y_{(m)i})$. Here, the subscript (m) indicates the different outcomes being observed. The observed value of $\mathbf{Y}_i$ is contingent on the assigned treatment, which we will occasion-

ally denote explicitly using the notation $\mathbf{Y}_i(\mathbf{a})$. In line with standard practices in dWOLS estimation [see, e.g. 9, 11, 60], our proposed method hinges on several key assumptions of causal inference (Assumptions A1 - A3).

5.2.2 Constrained dWOLS for multiple outcomes

In a typical single outcome context, the goal of fitting an optimal DTR is to find the collection of decision rules which maximizes the expected outcome of Y,

$$\arg\max_{d\in\mathcal{D}} \mathbb{E}\left[Y\left(d_k(\mathbf{H}_k)_{k\in\mathcal{K}}\right)\right].$$

However, when multiple outcomes of interest exist, a treatment regime chosen to maximize one particular outcome will not necessary produce desirable changes in the outcomes it fails to consider. Explicitly, if we have two outcomes, denoted by $Y_{(1)}$ and $Y_{(2)}$, we will not necessarily find that the solution to $\arg\max_{d\in\mathcal{D}} \mathbb{E}\left[Y_{(1)}\left(d_k(\mathbf{H}_k)_{k\in\mathcal{K}}\right)\right]$ agrees with the solution to $\arg\max_{d\in\mathcal{D}} \mathbb{E}\left[Y_{(2)}\left(d_k(\mathbf{H}_k)_{k\in\mathcal{K}}\right)\right]$. This could be problematic in many scenarios. For example, we may have that $Y_{(1)}$ is some health outcome of interest, which we prefer to be larger, while $Y_{(2)}$ is patient discomfort, inconvenience, or pain accrued throughout the regime. We use the general term of *cost* for such outcomes. In such a case, unmitigated endeavours to maximize $Y_{(1)}$ may cause $Y_{(2)}$ to reach intolerable levels, rendering the resulting regime ineffective. To maintain regime effectiveness, one should proceed to seek treatments that strike a balance between improving patient health, $Y_{(1)}$, and the impact on patient cost, $Y_{(2)}$. We may formulate such a regime by specifying a constraint at the optimization step which bounds the increase in patient cost. That is, we propose an optimization step of the form,

$$\begin{aligned} &\arg\max_{d\in\mathcal{D}} \mathbb{E}\left[Y_{(1)}\left(d_k(\mathbf{H}_k)_{k\in\mathcal{K}}\right)\right], \\ &\text{subject to } \mu_{Y_{(2)}}\left(d_k(\mathbf{H}_k)_{k\in\mathcal{K}}\right) \leq c_{k\in\mathcal{K}}, \end{aligned} \tag{5.1}$$

where $\mu_{Y_{(2)}}$ is the cost that optimal decision rule $d_k(\mathbf{H}_k)$ of $Y_{(1)}$ imposes on $Y_{(2)}$.

In the same vein of dWOLS for a single outcome, we posit a model for each outcome

Algorithm 3: Constraint optimization decision rule $d_k(\mathbf{h}_k)$ for stage k

Result: Optimal Treatment Decision
if $\gamma_{(1)k}(\mathbf{h}_{(1)k}, 1) > 0$ *and* $\gamma_{(2)k}(\mathbf{h}_{(2)k}, 1) < c$ **then**
  choose $a^{\text{opt}} = 1$
else
  Choose $a^{\text{opt}} = 0$.
end

as:

$$E[Y_{(1)\,i} \mid \mathbf{h}_{(1)ik}, a_{ik}] = \sum_{k=1}^{K} \left[f_{(1)k}(\mathbf{h}_{(1)ik}; \boldsymbol{\beta}_{(1)ik}) + \gamma_{(1)k}(\mathbf{h}_{(1)ik}, a_{ik}; \boldsymbol{\Psi}_{(1)k}) \right], \tag{5.2}$$

$$E[Y_{(2)i} \mid \mathbf{h}_{(2)ik}, a_{iK}] = \sum_{k=1}^{K} \left[f_{(2)k}(\mathbf{h}_{(2)ik}; \boldsymbol{\beta}_{(2)ik}) + \gamma_{(2)k}(\mathbf{h}_{(2)ik}, a_{iK}; \boldsymbol{\Psi}_{(2)k}) \right], \tag{5.3}$$

where $f_{(1)k}$ is the stage k treatment-free model with parameter $\boldsymbol{\beta}_{(1)k}$, and $\gamma_{(1)k}$ is the stage k blip model with parameter $\boldsymbol{\Psi}_{(1)k}$ for the outcome $Y_{(1)}$. The corresponding terms are defined for $Y_{(2)}$ accordingly. The error terms of the outcome model $Y_{(1)}$ and $Y_{(2)}$, could in general, be correlated.

Upon solving Equations (5.2) and (5.3), we obtain model parameter estimates, which we denote by $(\hat{\boldsymbol{\beta}}_{(m)k}, \hat{\boldsymbol{\Psi}}_{(m)k})$ for each outcome $m = \{1, 2\}$. These are then utilized to formulate the optimal treatment decision rule for stage k, which is derived from the maximization presented in Equation (5.1). The decision rule can be written generally in terms of the steps provided in Algorithm 3.

The decision rule takes into account the effect of the chosen treatment on both outcomes, in particular, the chosen treatment may no longer be the one which leads to the maximum for the main outcome $Y_{(1)}$. It will only be the maximum of the $Y_{(1)}$ outcome under the constrain on the corresponding cost, $Y_{(2)}$. However, when this occurs, whatever is lost by the patient not attaining their individual maximum is less than what is gained by decreasing their cost. Hence, the individuals end up compromising their own outcome for the "greater good" of their general health.

Having obtained our estimated decision rule, estimation at stage k is complete, and the

algorithm moves backward to stage $k-1$. To proceed with estimation at stage $k-1$ the appropriate pseudo-outcome needs to be computed. This can be done according to,

$$\tilde{y}_{(1)ik-1} = \tilde{y}_{(1)ik} + [\gamma_{(1)k}(\mathbf{h}_{ik}, a_{ik}^{\text{opt}}; \hat{\mathbf{\Psi}}_{(1)k}) - \gamma_{(1)k}(\mathbf{h}_{(1)ik}, a_{ik}; \hat{\mathbf{\Psi}}_{(1)k})],$$
$$\tilde{y}_{(2)ik-1} = \tilde{y}_{(2)ik} + [\gamma_{(2)k}(\mathbf{h}_{ik}, a_{ik}^{\text{opt}}; \hat{\mathbf{\Psi}}_{(2)k}) - \gamma_{(2)k}(\mathbf{h}_{(2)ik}, a_{ik}; \hat{\mathbf{\Psi}}_{(2)k})], \quad i \in \mathcal{I},\ k = 2, ..., K,$$

where we let $\tilde{y}_{(1)K} = y_{(1)}$ and $\tilde{y}_{(2)K} = y_{(2)}$ be the outcome observed following the stage K. Thus, the pseudo-outcomes are defined recursively, where, at each stage, we "add back" the individual level *regret* to account for the potentially sub-optimal treatment decision that was initially observed at that stage. Unlike the classic dWOLS with a single outcome, here the regret of an optimal decision for the primary outcome $Y_{(1)}$ is not necessarily positive.

As explained in Section 3.2, for a single outcome, we can specify balancing weights which eliminate any dependence between the history and the treatment at a given stage. Along with specifying the balancing weights, dWOLS additionally requires modelling the treatment allocation to estimate balancing weights. Incorporating these weights into an ordinary least squares estimation, the dWOLS framework will yield consistent estimates of the blip parameters. An important property of the dWOLS procedure is its double-robustness with respect to model mispecification.

The balancing approach (Equation 2.2) and the corresponding weights proposed in dWOLS (Equation 2.3), intend to eliminate the possible dependence between the history and treatment of the individuals and do not put restrictions or condition on the outcome. Hence, utilizing the same weights in a weighted ordinarily least square on each outcome will provide consistent estimation of the blip parameters so long as at least one of the treatment model or the treatment-free model component of the corresponding outcomes are correctly specified.

5.3 Simulation

We begin with an exemplification of the constrained optimization of multiple outcomes within the dWOLS framework. This example elucidates how our approach empowers the decision-maker to optimize a primary health outcome, denoted as $Y_{(1)}$, while ensuring that

the cost outcome, denoted by $Y_{(2)}$, associated with the corresponding decision remains below a specified threshold. The first simulation is designed to showcase the impact of introducing a constraint on the expected values at the population level. By comparing the population expected values of $Y_{(1)}$ and $Y_{(2)}$ across the unconstrained and constrained approaches, we aim to illustrate how the constraint alters the impact of the resulting treatment regime. In the second simulation, we aim to demonstrate that the introduction of the added constraint does not compromise the double robustness property of the original dWOLS algorithm. This simulation serves to confirm that our estimation procedure maintains its double robustness, even in the presence of the added constraint.

5.3.1 Simulation 1: the impact of constrained optimization

In this first simulation, we begin with an outcome model for which $Y_{(1)}$ is the primary health outcome of interest and $Y_{(2)}$ is the corresponding cost outcome. We define the outcome model for $Y_{(1)}$ to be, $Y_{(1)} = \beta^T_{y_{(1)}}(1, X_1, X_2) + A_1\psi^T_{y_{(1)1}}(1 + X_1) + A_2\psi^T_{y_{(1)2}}(1 + X_2) + \epsilon_{y_{(1)}}$, where $\beta_{y_{(1)}} = (30, 2, 2)$ and $\psi_{y_{(1)1}} = (4, 0.5)$ and $\psi_{y_{(1)2}} = (4, -1.5)$. Similarly, we specify the outcome model for $Y_{(2)}$ as, $Y_{(2)} = \beta^T_{y_{(2)}}(1, X_1, X_2)+A_1\psi^T_{y_{(2)1}}(1+X_1)+A_2\psi^T_{y_{(2)2}}(1+X_2)+\epsilon_{y_{(2)}}$, where $\beta_{y_{(2)}} = (30, 2, 2)$ and $\psi_{y_{(2)1}} = (3, 3)$ and $\psi_{y_{(2)2}} = (3, 6)$. The error terms are correlated, with $(\epsilon_{y_{(1)}}, \epsilon_{y_{(2)}}) \sim \text{Gaussian}(0, \Sigma)$, where Σ is a diagonal matrix with off-diagonal elements of 0.7.

Using this model, we generate a sample of $n = 1000$ patients, and use the generated data as the basis of our demonstration. In this simulation we intend to demonstrate how the resulting optimal values change when $Y_{(1)}$ is maximized without considering the impact on the cost $Y_{(2)}$, and when it is maximized under a constraint imposed on $Y_{(2)}$. We fit the model using a threshold value, c, which is approximately the median value of the empirical distribution of the blips values for $Y_{(2)}$ at each stage. That is, at stage j we set the threshold value c_j such that, $\frac{1}{n}\sum_{i=1}^{n} \mathbb{1}\{\psi^\top_{y(2)j}H_j < c_j\} = \frac{1}{2}$.

With this setup, the condition on $Y_{(2)}$ is important for around half of the individuals in the sample at any particular stage. Figure 5.1 shows the results of fitting the model on the data under the specified threshold. Figure 5.1(left) shows that, by taking the value of $Y_{(2)}$ into account, we end up with a distribution of optimal $Y_{(1)}$ values which are, on average, lower than those found in the unconstrained approach, as one would expect. Figure 5.1

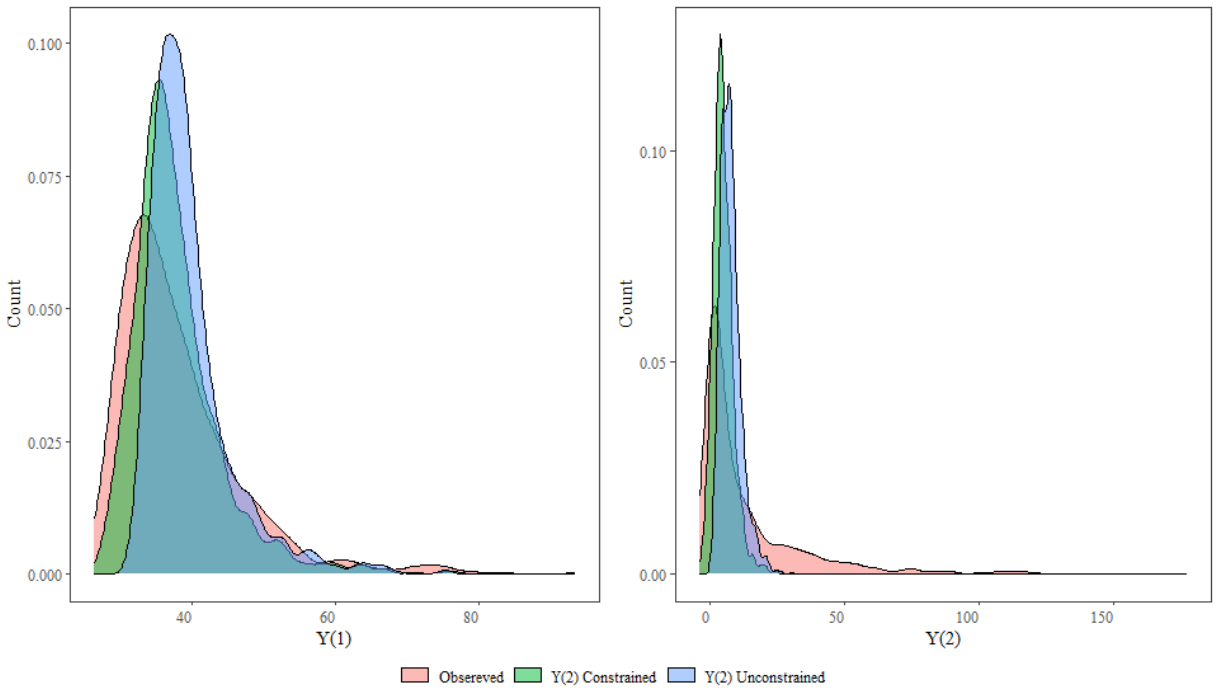


Figure 5.1: The distribution of outcomes $Y_{(1)}$ and $Y_{(2)}$, under unconstrained and constrained optimization under simulation 1.

(right), however, shows the benefit of this approach. The resulting distribution of the $Y_{(2)}$ values associated with the unconstrained optimal $Y_{(2)}$ values has a large right tail, which is undesirable and represents a population of patients under high treatment burden. On the other hand, the distribution of $Y_{(2)}$ values associated with the constrained optimal values shows a lower central tendency, smaller variance, and a shallower right tail. Hence, the burden of treatment has been severely reduced by the constrained approach, with minimal cost to the overall optimal $Y_{(1)}$ values.

5.3.2 Simulation 2: double robustness

In this second simulation, we demonstrate that the proposed algorithm maintains double robustness in estimation of the model blip parameters. To do this, we demonstrate empirically that the model produces consistent estimates of the dyadic-blip parameters when at

least one or fewer of the nuisance models are incorrectly specified.

We consider a single-stage DTR in this study with two patient covariates, X_1 and X_2. X_1 is generated from a Gamma distribution with shape parameter 1 and scale parameter 1/3 while X_2 follows a normal distribution with mean 1 and standard deviation 1. Treatments are assigned based on a binary random variable A, which is assigned according to $P(A = 1 \mid X_1, X_2) = \text{expit}\{-1 + X_1 + X_1^2\}$. The outcome model for $Y_{(1)}$ is defined as $Y_{(1)} = \beta_{y_{(1)}}^T(1, X_2, X_1, \sqrt{X_1}) + A\psi_{y_{(1)}}^T(1 + X_1) + \epsilon_{y_{(1)}}$, where $\beta_{y_{(1)}} = (30, 1, 2)$ and $\psi_{y_{(1)}} = (4, 0.5)$. Similarly, the outcome model for $Y_{(2)}$ is given by $Y_{(2)} = \beta_{y_{(2)}}^T(1, X_2, X_1, \log(X_1 + 1)) + A\psi_{y_{(2)}}^T(1 + X_1) + \epsilon_{y_{(2)}}$, where $\beta_{y_{(2)}} = (1, 1, 0.5, 2)$ and $\psi_{y_{(2)}} = (1, 1)$. The error terms are correlated, with $(\epsilon_{y_{(1)}}, \epsilon_{y_{(2)}}) \sim \text{Gaussian}(0, \Sigma)$, where Σ is a diagonal matrix with off-diagonal elements of 0.7.

To assess the impact of model misspecification, we conducted four analyses, each representing a different combination of correctly specifying or misspecifying the treatment-free and treatment models. Analysis 1 corresponds to the case where neither component of the model is correctly specified. Analysis 2 corresponds to the case where only the treatment-free component is misspecified. In this case, we retained only the X_1 term and excluded all other terms from the treatment-free models for $Y_{(1)}$ and $Y_{(2)}$. Analysis 3 corresponds to the case where only the treatment model for an outcome is misspecified. In this case, the treatment model is misspecified by removing the covariates and retaining only the intercept. Analysis 4 represents the ideal scenario where both the treatment-free and treatment models are correctly specified. Each of these model specification scenarios was replicated for sample sizes of 100, 1000, and 10000 individuals.

Figure 5.2 and 5.3 show the bias of the blip-parameter estimates. As it is shown we get unbiased estimates of the blip parameters when at least one of the treatment or treatment free model are correctly specified. As the sample size increases the variance of the estimates are reduced.

5.4 Discussion

The primary focus of the DTR methodology is to optimize a single outcome measure. However, in various health applications, multiple outcomes are assessed. In situations

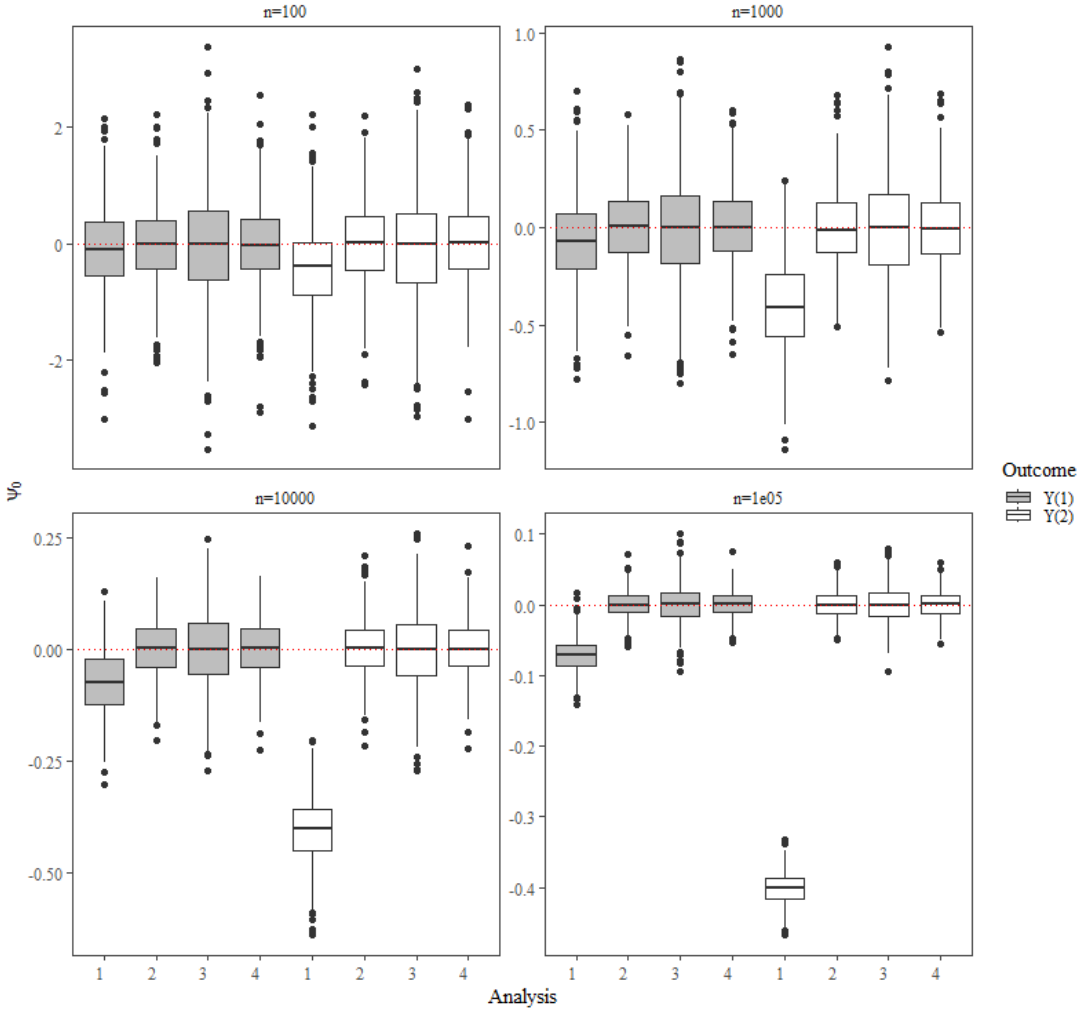


Figure 5.2: Estimates of the blip parameter ψ_0 for both outcome models and under different model specifications, for different sample sizes (Simulation 2) and 1000 simulated datasets. (1) misspecification of both models, (2) and (3) when one of the treatment model or treatment-free model is misspecified and (4) when both models are correctly specified.

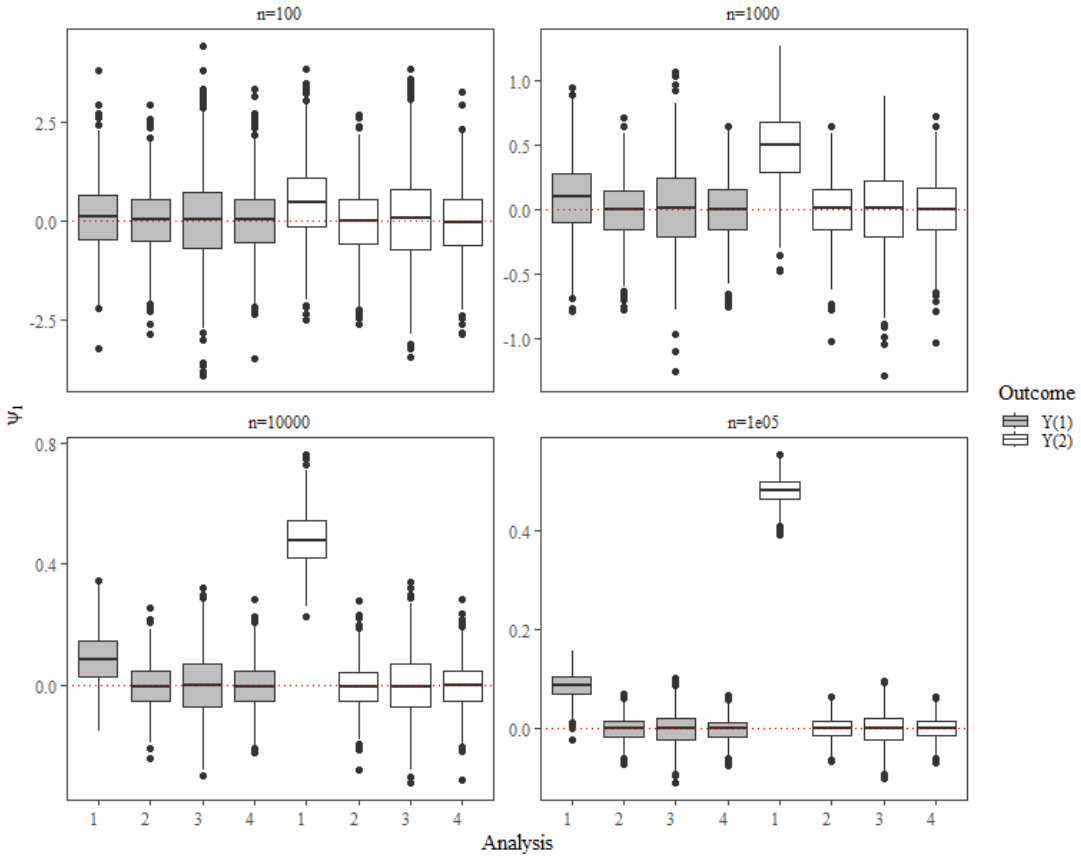


Figure 5.3: Estimates of the blip parameter ψ_1 for both outcome models and under different model specifications, for different sample sizes (Simulation 2) and 1000 simulated datasets. (1) misspecification of both models, (2) and (3) when one of the treatment model or treatment-free model is misspecified and (4) when both models are correctly specified.

where these outcomes are competing measures or involve associated costs, relying solely on a single outcome approach in DTR may lead to suboptimal overall health outcomes for individuals. To address this limitation, the first step in extending the DTR methodology is to revise the optimization goal, as the optimization objective is not straightforward when dealing with multiple outcomes.

In this chapter, we propose an extension of the dWOLS methodology to address the consideration of multiple outcomes, particularly when one of the outcomes represents a measure of cost or side effects. Our proposal involves modifying the optimization strategy to select the optimal treatment strategy that maximizes a primary outcome while incorporating constraints on the cost of the corresponding decision. The extended dWOLS methodology retains its double robustness property of original dWOLS, which is advantageous compared to existing methods that rely on accurate model specifications. Furthermore, it allows the clinician to incorporate the patient-specific costs, when it is available, rather than relying on a global constraint for the cost threshold. While the proposed estimation method in this chapter assumes that both outcomes are continuous, it is worth noting that the dWOLS method has already been developed to handle other types of outcomes. By utilizing these extensions, the proposed constrained optimization approach can accommodate different types of outcomes.

Furthermore, we select the threshold of the cost based on the empirical distribution of the cost outcome. The selection of such threshold and personalizing it could be potential areas of investigation. One approach is to use parametric methods to estimate the individual threshold and then incorporate it into the analysis. Another approach is to estimate patient-specific costs without directly eliciting them by assuming that the observed treatments have already been assigned with the aim of reducing costs. Additionally, an intriguing direction for future research is to expand the methodology to scenarios where each outcome has its own primary treatment. For instance, if a patient has both high blood pressure and elevated cholesterol levels, studying the possibility of interactions between each medication and its effects on the other outcome should be studied. The multiple outcomes dWOLS method would allow for potential interactions between treatments or with other outcomes, creating a more comprehensive and realistic representation of health scenarios.

Chapter 6

Discussion

6.1 Conclusion

The thesis aimed to address the challenges of estimating dynamic treatment regimes (DTRs) in the presence of interference, particularly in dyadic network populations, and propose methodologies for robust and flexible solutions. Chapter 3 focused on developing a methodology for binary treatments in dyadic network populations. The dyadic-dWOLS methodology, incorporating a modified optimization step and a dyad-health function, allowed researchers to define and optimize outcomes tailored to their specific context. Through simulations and analysis of the Population Assessment of Health and Tobacco (PATH) data, we demonstrated the performance and adaptability of the dyadic-dWOLS methodology. Our analysis distinguishes itself by offering a distinctive perspective. We not only investigate the presence of spillover effects, but our approach provides a framework for optimal recommendations tailored to pairs of tobacco users, rather than focusing solely on individual-level interventions.

Building on the foundation of Chapter 3, Chapter 4 extended the dyadic-dWOLS methodology to include continuous treatments. This extension allowed for a broader application of the methodology that applies to various realistic scenarios, such as different dosing levels of treatments. The dyadic-dWOLS methodology successfully accommodated various types of interference, including symmetric and asymmetric patterns. We implemented and evaluated the methodology using the PATH data, specifically examining the

potential impact of nicotine exposure on tobacco dependence. Additionally, we explore and extend various weighting approaches proposed for continuous treatments, evaluating their performance within the context of dyadic-dWOLS.

Inspired by the concept of considering dyads as independent entities, specially in the setting where multiple outcome measures are available, Chapter 5 delved into the under-explored area of incorporating multiple outcomes in the context of dynamic treatment regimes (DTRs). In Chapter 5, we extended the dWOLS methodology to address the consideration of multiple outcomes, specifically when one of the outcomes represents a measure of cost or side effects. Our proposal involved modifying the optimization strategy to select the optimal treatment strategy that maximizes the primary outcome while incorporating constraints on the cost of the corresponding decision. The extended dWOLS methodology retains the double robustness property, which is advantageous compared to existing methods that rely on fully accurate model specifications. Furthermore, it allows the incorporation of the patient-specific costs rather than relying on a global constraint for the cost threshold.

In conclusion, this thesis extended the dWOLS methodology to address the challenges of estimating dynamic treatment regimes in dyadic network populations. Chapter 3 and Chapter 4 introduced the dyadic-dWOLS methodology for binary and continuous treatments, respectively, demonstrating its performance and adaptability using the PATH data. Furthermore, Chapter 5 extended the dWOLS methodology to incorporate multiple outcomes, considering patient-specific costs and retaining the double robustness property. These contributions provide valuable insights and tools for optimizing treatment decisions in complex health contexts, paving the way for more effective personalized interventions.

6.2 Future Work

While this thesis has made significant strides in addressing interference within dyadic-network populations, there are some limitations that open avenues for future research. Firstly, our current framework focuses on dyadic interference, considering interactions between two individuals. Future work could explore the extension of our methodology to include interference among groups of more than two people, allowing for a more com-

prehensive understanding of the complex dynamics within larger social networks. Additionally, our current assumption assumes an unlimited availability of treatment resources. However, it would be valuable to investigate scenarios where there are limitations or resource constraints, as this would provide insights into optimizing treatment strategies under real-world resource constraints. Moreover, our current methodology assumes complete information for both members of the dyad. Future research could address the challenges posed by partial information or missing values for one member or both members, exploring imputation techniques or statistical methods to handle these scenarios effectively. Furthermore, an exciting direction for future work involves extending the dWOLS method for multiple outcomes, particularly in the context of comorbid mental and physical diseases, where each disease may require specific treatment strategies. This would provide a more comprehensive approach to optimizing treatment decisions for individuals with complex health conditions.

Lastly, the incorporation of interference and multiple outcomes within the context of our study is an intriguing topic. We proposed a method that ensures the cost associated with a decision rule remains below a specified threshold, which raises the question of how to effectively manage the cost of decisions for each individual while considering the overall population cost. While we did not impose a restriction on the population cost in our analysis, it is not unrealistic to envision scenarios where both individual and the population cost thresholds are in place. This combination, coupled with the potential presence of interference in the population, presents a complex yet captivating area for further exploration.

References

[1] Michael R Kosorok and Eric B Laber. Precision Medicine. *Annual Review of Statistics and Its Application*, 6(1):263–286, 2019. doi: 10.1146/annurev-statistics-030718-105251. URL https://doi.org/10.1146/annurev-statistics-030718-105251.

[2] Bibhas Chakraborty. *Statistical methods for dynamic treatment regimes*. Springer, 2013.

[3] S A Murphy, E Arjas, C Jennison, A P Dawid, D R Cox, S Senn, R G Cowell, V Didelez, R D Gill, J B Kadane, and J M Robins. Optimal dynamic treatment regimes. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 65 (2):331–355, 2003. doi: 10.1111/1467-9868.00389.

[4] Eric B Laber, Daniel J Lizotte, and Bradley Ferguson. Set-valued dynamic treatment regimes for competing outcomes. *Biometrics*, 70(1):53–61, 2014.

[5] Erica E M Moodie, Bibhas Chakraborty, and Michael S Kramer. Q-learning for estimating optimal dynamic treatment rules from observational data. *Canadian Journal of Statistics*, 40(4):629–645, 2012.

[6] E E M Moodie, N Dean, and Y R Sun. Q-Learning: Flexible Learning About Useful Utilities. *Statistics in Biosciences*, 6(2):223–243, 2014. doi: 10.1007/s12561-013-9103-z.

[7] J M Robins. Optimal structural nested models for optimal sequential decisions. *Proceedings of the Second Seattle Symposium on Biostatistics*, pages 189–326, 2004.

[8] Bibhas Chakraborty, Susan Murphy, and Victor Strecher. Inference for non-regular parameters in optimal dynamic treatment regimes. *Statistical methods in medical research*, 19(3):317–343, 2010. doi: 10.1177/0962280209105013.

[9] Michael P Wallace and Erica E M Moodie. Doubly-robust dynamic treatment regimen estimation via weighted least squares. *Biometrics*, 71(3):636–644, 2015.

[10] Juliana Schulz and Erica E M Moodie. Doubly robust estimation of optimal dosing strategies. *Journal of the American Statistical Association*, pages 1–13, 2020.

[11] Gabrielle Simoneau, Erica E M Moodie, Jagtar S Nijjar, Robert W Platt, Scottish Early Rheumatoid Arthritis Inception Cohort Investigators, and Others. Estimating optimal dynamic treatment regimes with survival outcomes. *Journal of the American Statistical Association*, 115(531):1531–1539, 2020.

[12] Y Zhao, M R Kosorok, and D Zeng. Reinforcement learning design for cancer clinical trials. *Statistics in Medicine*, 28(26):3294–3315, 2009. doi: 10.1002/sim.3720.

[13] Arthur Guez, Robert D Vincent, Massimo Avoli, and Joelle Pineau. Adaptive Treatment of Epilepsy via Batch-mode Reinforcement Learning. In *AAAI*, pages 1671–1678, 2008.

[14] Susan A Murphy, David W Oslin, A John Rush, and Ji Zhu. Methodological challenges in constructing effective treatment sequences for chronic psychiatric disorders. *Neuropsychopharmacology*, 32(2):257–262, 2007.

[15] Susan M Shortreed, Eric Laber, Daniel J Lizotte, T Scott Stroup, Joelle Pineau, and Susan A Murphy. Informing sequential clinical decision-making through reinforcement learning: an empirical study. *Machine learning*, 84(1-2):109–136, 2011.

[16] Damien Ernst, Guy-Bart Stan, Jorge Goncalves, and Louis Wehenkel. Clinical data based optimal STI strategies for HIV: a reinforcement learning approach. In *Proceedings of the 45th IEEE Conference on Decision and Control*, pages 667–672. IEEE, 2006.

[17] Merrick J Moseley, Michael P Wallace, David A Stephens, Alistair R Fielder, Laura C Smith, and Catherine E Stewart. Personalized versus standardized dosing strategies

for the treatment of childhood amblyopia: study protocol for a randomized controlled trial. *Trials*, 16(1):1–7, 2015.

[18] D R Cox. Planning of Experiments' John Wiley and Sons. *Inc., New York*, 1958.

[19] Donald B Rubin. Bayesian inference for causal effects: The role of randomization. *The Annals of statistics*, pages 34–58, 1978.

[20] Michael G Hudgens and M Elizabeth Halloran. Causal vaccine effects on binary postinfection outcomes. *Journal of the American Statistical Association*, 101(473): 51–64, 2006.

[21] M Elizabeth Halloran and Claudio J Struchiner. Causal inference in infectious diseases. *Epidemiology*, pages 142–151, 1995.

[22] Michael E Sobel. What do randomized studies of housing mobility demonstrate? Causal inference in the face of interference. *Journal of the American Statistical Association*, 101(476):1398–1407, 2006.

[23] Alberto Chong, Marco Gonzalez-Navarro, Dean Karlan, and Martin Valdivia. Effectiveness and spillovers of online sex education: Evidence from a randomized evaluation in Colombian public schools. *NBER Working Paper Series*, page 18776, 2013.

[24] Abhijit Vinayak Banerjee, Esther Duflo, Rachel Glennerster, and Dhruva Kothari. Improving immunisation coverage in rural India: clustered randomised controlled evaluation of immunisation campaigns with and without incentives. *Bmj*, 340, 2010.

[25] M J Campbell, A Donner, and N Klar. Developments in cluster randomized trials and statistics in medicine. *Statistics in medicine*, 26(1):2–19, 2007.

[26] Michael G Hudgens and M Elizabeth Halloran. Toward causal inference with interference. *Journal of the American Statistical Association*, 103(482):832–842, 2008.

[27] Paul R Rosenbaum. Interference between units in randomized experiments. *Journal of the American Statistical Association*, 102(477):191–200, 2007.

[28] Panos Toulis and Edward Kao. Estimation of causal peer influence effects. In *International conference on machine learning*, pages 1489–1497. PMLR, 2013.

[29] Johan Ugander, Brian Karrer, Lars Backstrom, and Jon Kleinberg. Graph cluster randomization: Network exposure to multiple universes. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 329–337, 2013.

[30] Dean Eckles, Brian Karrer, and Johan Ugander. Design and analysis of experiments in networks: Reducing bias from interference. *Journal of Causal Inference*, 5(1), 2017.

[31] Oleg Sofrygin and Mark J van der Laan. Semi-parametric estimation and inference for the mean outcome of the single time-point intervention in a causally connected population. *Journal of causal inference*, 5(1), 2017.

[32] Peter M Aronow, Cyrus Samii, and Others. Estimating average causal effects under general interference, with application to a social network experiment. *The Annals of Applied Statistics*, 11(4):1912–1947, 2017.

[33] Lan Liu, Michael G Hudgens, and Sylvia Becker-Dreps. On inverse probability-weighted estimators in the presence of interference. *Biometrika*, 103(4):829–842, 2016.

[34] Eric J Tchetgen Tchetgen and Tyler J VanderWeele. On causal inference in the presence of interference. *Statistical methods in medical research*, 21(1):55–75, 2012.

[35] Oleg Sofrygin, Romain Neugebauer, and Mark J van der Laan. Conducting Simulations in Causal Inference with Networks-Based Structural Equation Models. *arXiv preprint arXiv:1705.10376*, 2017.

[36] Elizabeth L Ogburn, Ilya Shpitser, and Youjin Lee. Causal inference, social networks and chain graphs. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 183(4):1659–1676, 2020.

[37] Laura Forastiere, Edoardo M Airoldi, and Fabrizia Mealli. Identification and estimation of treatment and interference effects in observational studies on networks. *Journal of the American Statistical Association*, 116(534):901–918, 2021.

[38] Daniel J Luckett, Eric B Laber, Siyeon Kim, and Michael R Kosorok. Estimation and optimization of composite outcomes. *Journal of Machine Learning Research*, 22(167): 1–40, 2021.

[39] Emily L Butler, Eric B Laber, Sonia M Davis, and Michael R Kosorok. Incorporating patient preferences into estimation of optimal individualized treatment rules. *Biometrics*, 74(1):18–26, 2018.

[40] Anne Shrestha, Charlene Martin, Maria Burton, Stephen Walters, Karen Collins, and Lynda Wyld. Quality of life versus length of life considerations in cancer patients: a systematic literature review. *Psycho-oncology*, 28(7):1367–1380, 2019.

[41] Daniel J Lizotte, Michael Bowling, and Susan A Murphy. Linear fitted-q iteration with multiple reward functions. *The Journal of Machine Learning Research*, 13(1): 3253–3295, 2012.

[42] Daniel J Lizotte and Eric B Laber. Multi-objective markov decision processes for data-driven decision support. *The journal of machine learning research*, 17(1):7378–7405, 2016.

[43] Kristin A Linn, Eric B Laber, and Leonard A Stefanski. Chapter 15: Estimation of dynamic treatment regimes for complex outcomes: balancing benefits and risks. In *Adaptive treatment strategies in practice: Planning trials and analyzing data for personalized medicine*, pages 249–262. SIAM, 2015.

[44] Yuanjia Wang, Haoda Fu, and Donglin Zeng. Learning optimal personalized treatment rules in consideration of benefit and risk: with an application to treating type 2 diabetes patients with insulin therapies. *Journal of the American Statistical Association*, 113(521):1–13, 2018.

[45] Mary E Cooley, Linda Sarna, Jean K Brown, Roma D Williams, Cynthia Chernecky, Geraldine Padilla, Leda Layo Danao, and David Elashoff. Tobacco use in women with lung cancer. *Annals of Behavioural Medicine*, 33:242–250, 2007.

[46] Augusto Di Castelnuovo, Gianni Quacquaruccio, Maria Benedetta Donati, Giovanni de Gaetano, and Licia Iacoviello. Spousal Concordance for Major Coronary Risk Factors: A Systematic Review and Meta-Analysis. *American Journal of Epidemiology*, 169(1):1–8, 11 2008. ISSN 0002-9262. doi: 10.1093/aje/kwn234. URL https://doi.org/10.1093/aje/kwn234.

[47] Sarah E Jackson, Andrew Steptoe, and Jane Wardle. The influence of partner's behavior on health behavior change: the english longitudinal study of ageing. *JAMA internal medicine*, 175(3):385–392, 2015.

[48] Robert A Schnoll, Michael Malstrom, Calvin James, Randi L Rothman, Suzanne M Miller, John A Ridge, Benjamin Movsas, Michael Unger, Corey Langer, and Melvyn Goldberg. Correlates of tobacco use among smokers and recent quitters diagnosed with cancer. *Patient Education and Counseling*, 46(2):137–145, 2002. ISSN 0738-3991. doi: https://doi.org/10.1016/S0738-3991(01)00157-4. URL https://www.sciencedirect.com/science/article/pii/S0738399101001574.

[49] Katherine Regan Sterba, Vance Rabius, Matthew J. Carpenter, Pamela Villars, Dawn Wiatrek, and Alfred McAlister. Dyadic Efficacy for Smoking Cessation: Preliminary Assessment of a New Instrument. *Nicotine Tobacco Research*, 13(3):194–201, 01 2011. ISSN 1462-2203. doi: 10.1093/ntr/ntq236. URL https://doi.org/10.1093/ntr/ntq236.

[50] Pamela Mudar, Kenneth E Leonard, and Kerry Soltysinski. Discrepant substance use and marital functioning in newlywed couples. *Journal of consulting and clinical psychology*, 69(1):130, 2001.

[51] Gregory G Homish, Kenneth E Leonard, Lynn T Kozlowski, and Jack R Cornelius. The longitudinal association between multiple substance use discrepancies and marital satisfaction. *Addiction*, 104(7):1201–1209, 2009.

[52] Donald B Rubin. Comment: Neyman (1923) and causal inference in experiments and observational studies. *Statistical Science*, 5(4):472–480, 1990.

[53] Donald B Rubin. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology*, 66(5):688, 1974.

[54] James 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(9-12):1393–1512, 1986.

[55] Judea Pearl. *Causality*. Cambridge university press, 2009.

[56] Donald B. Rubin. Randomization analysis of experimental data: The fisher randomization test comment. *Journal of the American Statistical Association*, 75(371): 591–593, 1980. ISSN 01621459. URL http://www.jstor.org/stable/2287653.

[57] Fan Li, Kari Lock Morgan, and Alan M Zaslavsky. Balancing covariates via propensity score weighting. *Journal of the American Statistical Association*, 113(521):390–400, 2018.

[58] Lin Su, Wenbin Lu, and Rui Song. Modelling and estimation for optimal treatment decision with interference. *Stat*, 8(1):e219, 2019.

[59] Cong Jiang, Michael P Wallace, and Mary E Thompson. Dynamic treatment regimes with interference. *Canadian Journal of Statistics*, 2022.

[60] Anastasios A Tsiatis, Marie Davidian, Shannon T Holloway, and Eric B Laber. *Dynamic Treatment Regimes: Statistical Methods for Precision Medicine*. CRC press, 2019.

[61] The health consequences of smoking : nicotine addiction : a report of the surgeon general, 1988. URL https://stacks.cdc.gov/view/cdc/22014. Report.

[62] Neal L Benowitz, Janne Hukkanen, and Peyton Jacob III. Nicotine chemistry, metabolism, kinetics and biomarkers. *Nicotine psychopharmacology*, pages 29–60, 2009.

[63] Andrew Hyland, Bridget K Ambrose, Kevin P Conway, Nicolette Borek, Elizabeth Lambert, Charles Carusi, Kristie Taylor, Scott Crosse, Geoffrey T Fong, K Michael Cummings, David Abrams, John P Pierce, James Sargent, Karen Messer, Maansi Bansal-Travers, Ray Niaura, Donna Vallone, David Hammond, Nahla Hilmi, Jonathan Kwan, Andrea Piesse, Graham Kalton, Sharon Lohr, Nick Pharris-Ciurej, Victoria Castleman, Victoria R Green, Greta Tessman, Annette Kaufman, Charles Lawrence, Dana M van Bemmel, Heather L Kimmel, Ben Blount, Ling Yang, Barbara O'Brien, Cindy Tworek, Derek Alberding, Lynn C Hull, Yu-Ching Cheng, David Maklan, Cathy L Backinger, and Wilson M Compton. Design and methods of the population assessment of tobacco and health (PATH) study. *Tob. Control*, 26(4):371–378, July 2017.

[64] David R Strong, Jennifer Pearson, Sarah Ehlke, Thomas Kirchner, David Abrams, Kristie Taylor, Wilson M Compton, Kevin P Conway, Elizabeth Lambert, Victoria R Green, et al. Indicators of dependence for different types of tobacco product users: Descriptive findings from wave 1 (2013–2014) of the population assessment of tobacco and health (path) study. *Drug and alcohol dependence*, 178:257–266, 2017.

[65] Neal L Benowitz, Katherine M Dains, Delia Dempsey, Lisa Yu, and Peyton Jacob, 3rd. Estimation of nicotine dose after low-level exposure using plasma and urine nicotine metabolites. *Cancer Epidemiol. Biomarkers Prev.*, 19(5):1160–1166, May 2010.

[66] Neal L Benowitz, Gideon St Helen, Natalie Nardone, Lisa Sanderson Cox, and Peyton Jacob. Urine metabolites for estimating daily intake of nicotine from cigarette smoking. *Nicotine Tob. Res.*, 22(2):288–292, February 2020.

[67] Maciej L Goniewicz, Danielle M Smith, Kathryn C Edwards, Benjamin C Blount, Kathleen L Caldwell, Jun Feng, Lanqing Wang, Carol Christensen, Bridget Ambrose, Nicolette Borek, et al. Comparison of nicotine and toxicant exposure in users of electronic cigarettes and combustible cigarettes. *JAMA network open*, 1(8):e185937–e185937, 2018.

[68] DA Stephens. Chapter 7: G-estimation for dynamic treatment regimes in the longitudinal setting. In *Adaptive treatment strategies in practice: Planning trials and analyzing data for personalized medicine*, pages 89–117. SIAM, 2015.

[69] Michael P Wallace, Erica EM Moodie, and David A Stephens. Reward ignorant modeling of dynamic treatment regimes. *Biometrical Journal*, 60(5):991–1002, 2018.

[70] Benjamin Rich, Erica EM Moodie, and David A Stephens. Optimal individualized dosing strategies: A pharmacologic approach to developing dynamic treatment regimens for continuous-valued treatments. *Biometrical Journal*, 58(3):502–517, 2016.

[71] Michael P Wallace, Erica E M Moodie, and David A Stephens. Dynamic Treatment Regimen Estimation via Regression-Based Techniques: Introducing R Package DTRreg. *Journal of Statistical Software, Articles*, 80(2):1–20, 2017. ISSN 1548-7660. doi: 10.18637/jss.v080.i02. URL https://www.jstatsoft.org/v080/i02.

[72] Eli Sherman, David Arbour, and Ilya Shpitser. General Identification of Dynamic Treatment Regimes Under Interference. In *International Conference on Artificial Intelligence and Statistics*, pages 3917–3927. PMLR, 2020.

[73] Davide Viviano. Policy targeting under network interference. *arXiv preprint arXiv:1906.10258*, 2019.

[74] Laura Forastiere, Fabrizia Mealli, Albert Wu, and Edoardo M Airoldi. Estimating causal effects on social networks. In *2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA)*, pages 60–69. IEEE, 2018.

[75] Davide Del Prete, Laura Forastiere, and Valerio Leone Sciabolazza. Causal inference on networks under continuous treatment interference, 2020. URL https://arxiv.org/abs/2004.13459.

[76] Stefan Tübbicke. Entropy balancing for continuous treatments. *Journal of Econometric Methods*, 11(1):71–89, 2021.

[77] Brian G Vegetabile, Beth Ann Griffin, Donna L Coffman, Matthew Cefalu, Michael W Robbins, and Daniel F McCaffrey. Nonparametric estimation of population average dose-response curves using entropy balancing weights for continuous exposures. *Health Services and Outcomes Research Methodology*, 21:69–110, 2021.

[78] Jens Hainmueller. Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political analysis*, 20 (1):25–46, 2012.

[79] James M Robins. Marginal structural models versus structural nested models as tools for causal inference. In *Statistical models in epidemiology, the environment, and clinical trials*, pages 95–133. Springer, 2000.

[80] David R Strong, Eric Leas, Madison Noble, Martha White, Allison Glasser, Kristie Taylor, Kathryn C Edwards, Kevin C Frissell, Wilson M Compton, Kevin P Conway, Elizabeth Lambert, Heather L Kimmel, Marushka L Silveira, Lynn C Hull, Dana van Bemmel, Megan J Schroeder, Kenneth Michael Cummings, Andrew Hyland, June Feng, Benjamin Blount, Lanqing Wang, and Ray Niaura. Validation of the Wave 1

and Wave 2 Population Assessment of Tobacco and Health (PATH) Study Indicators of Tobacco Dependence Using Biomarkers of Nicotine Exposure Across Tobacco Products. *Nicotine Tobacco Research*, 24(1):10–19, 08 2021. ISSN 1469-994X. doi: 10.1093/ntr/ntab162. URL https://doi.org/10.1093/ntr/ntab162.

[81] Y Zhao, D Zeng, M A Socinski, and M R Kosorok. Reinforcement learning strategies for clinical trials in nonsmall cell lung cancer. *Biometrics*, 67(4):1422–1433, 2011. doi: 10.1111/j.1541-0420.2011.01572.x.

[82] B Chakraborty, E B Laber, and Y Zhao. Inference for Optimal Dynamic Treatment Regimes Using an Adaptive m-Out-of-n Bootstrap Scheme. *Biometrics*, 69(3):714–723, 2013. doi: 10.1111/biom.12052.

[83] Jocelyn R Dale. Global cross-ratio models for bivariate, discrete, ordered responses. *Biometrics*, pages 909–917, 1986.

[84] Justin R Williams and Catherine M Crespi. Causal inference for multiple continuous exposures via the multivariate generalized propensity score. *arXiv preprint arXiv:2008.13767*, 2020.

[85] Christian Fong, Chad Hazlett, and Kosuke Imai. Covariate balancing propensity score for a continuous treatment: Application to the efficacy of political advertisements. *The Annals of Applied Statistics*, 12(1):156–177, 2018.

[86] Keisuke Hirano and Guido W Imbens. The propensity score with continuous treatments. *Applied Bayesian modeling and causal inference from incomplete-data perspectives*, 226164:73–84, 2004.

[87] Kosuke Imai and David A Van Dyk. Causal inference with general treatment regimes: Generalizing the propensity score. *Journal of the American Statistical Association*, 99 (467):854–866, 2004.

[88] Edward H Kennedy, Zongming Ma, Matthew D McHugh, and Dylan S Small. Nonparametric methods for doubly robust estimation of continuous treatment effects. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 79(4): 1229–1245, 2017.

APPENDICES

Appendix A

Chapter 3 Appendix

A.1 Double Robustness

We give a proof for the consistency and double-robustness of dyadic dWOLS following Wallace & Moodie [9, 11]. The proof is concerned with the dyadic-blip estimators $\hat{\boldsymbol{\Theta}}_k = (\hat{\psi}_k, \hat{\eta}_k, \hat{\xi}_k)$. To estimate the parameters, we solve the following weighted generalized estimating equation,

$$U_k(\boldsymbol{\beta}_k, \boldsymbol{\Theta}_k) = \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} w_{ijk}\mathbf{S}(\mathbf{h}_{ijk}, \mathbf{a}_{ik}) \cdot (Y_{ij} - \mathbf{h}_{ijk}^\top\boldsymbol{\beta}_k - \mathbf{h}_{ijk}^\top(a_{ijk}\boldsymbol{\psi}_k + a_{ij^*k}\boldsymbol{\eta}_k + a_{ijk}a_{ij^*k}\boldsymbol{\xi}_k)) = \mathbf{0}.$$

Here we have used the notation $\mathbf{S}(\cdot,\cdot)$ to represent the score vector of the specified outcome model with respect to the model parameters, which in the case of the model presented in Equation (3.5), takes the form,

$$\mathbf{S}(\mathbf{h}_{ijk}, \mathbf{a}_{ik}) = (\mathbf{h}_{ijk}, a_{ijk}\mathbf{h}_{ijk}, a_{ij^*k}\mathbf{h}_{ijk}, a_{ijk}a_{ij^*k}\mathbf{h}_{ijk}).$$

For simplicity, the proof is detailed for a single-stage DTR. It is straightforward to extend the reasoning to more than one stage.

A.1.1 Treatment-free model correctly specified, treatment models misspecified

Assuming that the dyadic blip model is correctly specified, when the treatment-free model is correctly specified, $\mathbb{E}\left[Y_{ij} - \mathbf{h}_{ij}^{\top}(a_{ij}\boldsymbol{\psi} + a_{ij^*}\boldsymbol{\eta} + a_{ij}a_{ij^*}\boldsymbol{\xi}) \mid \mathbf{h}, \mathbf{a}\right] = \mathbf{h}_{ij}^{\top}\boldsymbol{\beta}$. With the expectation taken with respect to $p(Y \mid \mathbf{a}, \mathbf{h})$, we have

$$
\begin{aligned}
&U_k(\boldsymbol{\beta}_k, \boldsymbol{\Theta}_k) \\
&= \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} \int_y \hat{w}_{ijk}\mathbf{S}(\mathbf{h}_{ijk}, \mathbf{a}_{ik}) \cdot (Y_{ij} - \mathbf{h}_{ijk}^{\top}\boldsymbol{\beta}_k - \mathbf{h}_{ijk}^{\top}(a_{ijk}\boldsymbol{\psi}_k + a_{ij^*k}\boldsymbol{\eta}_k + a_{ijk}a_{ij^*k}\boldsymbol{\xi}_k)p(Y \mid \mathbf{a}, \mathbf{h})dY \\
&= \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} \hat{w}_{ijk}\mathbf{S}(\mathbf{h}_{ijk}, \mathbf{a}_{ik}) \cdot \int_y (Y_{ij} - \mathbf{h}_{ijk}^{\top}\boldsymbol{\beta}_k - \mathbf{h}_{ijk}^{\top}(a_{ijk}\boldsymbol{\psi}_k + a_{ij^*k}\boldsymbol{\eta}_k + a_{ijk}a_{ij^*k}\boldsymbol{\xi}_k)p(Y \mid \mathbf{a}, \mathbf{h})dY \\
&= \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} \hat{w}_{ijk}\mathbf{S}(\mathbf{h}_{ijk}, \mathbf{a}_{ik}) \cdot \mathbb{E}\left[(Y_{ij} - \mathbf{h}_{ijk}^{\top}\boldsymbol{\beta}_k - \mathbf{h}_{ijk}^{\top}(a_{ijk}\boldsymbol{\psi}_k + a_{ij^*k}\boldsymbol{\eta}_k + a_{ijk}a_{ij^*k}\boldsymbol{\xi}_k)\right]
\end{aligned}
$$

When the treatment-free model is correctly specified, then

$$
\begin{aligned}
\mathbb{E}\left[Y_{ij} - \mathbf{h}_{ijk}^{\top}\boldsymbol{\beta}_k - \mathbf{h}_{ijk}^{\top}(a_{ijk}\boldsymbol{\psi}_k + a_{ij^*k}\boldsymbol{\eta}_k + a_{ijk}a_{ij^*k}\boldsymbol{\xi}_k)\right] = \\
\mathbb{E}\left[Y_{ij} - \mathbf{h}_{ijk}^{\top}(a_{ijk}\boldsymbol{\psi}_k + a_{ij^*k}\boldsymbol{\eta}_k + a_{ijk}a_{ij^*k}\boldsymbol{\xi}_k)\right] - \mathbb{E}\left[\mathbf{h}_{ijk}^{\top}\boldsymbol{\beta}_k\right] = 0
\end{aligned}
$$

such that $U_k(\boldsymbol{\beta}_k, \boldsymbol{\Theta}_k)$ consistently estimates the blip parameters $\boldsymbol{\Theta}_k$ when the treatment-free model is correctly specified.

A.1.2 Treatment-free model misspecified, treatment models correctly specified

Let the treatment model be correctly specified i.e. $\mathbb{P}(\mathbf{A} = \mathbf{a}|\mathbf{h})$. We show that finding the root of the estimating function $U_k(\boldsymbol{\beta}_k, \boldsymbol{\Theta}_k)$ yields consistent estimators of $\boldsymbol{\Theta}$ if the weights satisfy the following balancing property.

Theorem A.1.1 (Balancing property)**.** *A weighted ordinary least-squares regression based*

on the model,

$$E[Y_j \mid \mathbf{h}_j, \mathbf{a}] = f(\mathbf{h}_j; \boldsymbol{\beta}) + \gamma(\mathbf{h}_j, \mathbf{a}; \boldsymbol{\Theta}),$$

will give consistent estimates for the dyadic-blip parameters $\boldsymbol{\Theta}$ *provided that the weights satisfy,*

$$\pi(\mathbf{a} \mid \mathbf{h}) w(\mathbf{a}, \mathbf{h}) = \pi(\mathbf{a}' \mid \mathbf{h}) w(\mathbf{a}', \mathbf{h}) \tag{A.1}$$

for any $\mathbf{a} \neq \mathbf{a}' \in \mathcal{A}^2$*, where* $\pi(\mathbf{a} \mid \mathbf{h}) = \mathbb{P}(\mathbf{A} = \mathbf{a} \mid \mathbf{h})$ *is the joint propensity score. In case of binary treatment, this condition translates to*

$$\pi(0,0,\mathbf{h})\ w(0,0,\mathbf{h}) = \pi(0,1,\mathbf{h})\ w(0,1,\mathbf{h}) = \pi(1,0,\mathbf{h})\ w(1,0,\mathbf{h}) = \pi(1,1,\mathbf{h})\ w(1,1,\mathbf{h}).$$

Proof. The estimates of $\boldsymbol{\Theta}$ in a ordinary least-squares regression will be biased unless we at least have $E(\mathbf{h} \mid \mathbf{a}) = E(\mathbf{h})$ for all $\mathbf{a}$ in $\mathcal{A}^2$. Hence, it is sufficient to find weights such that, in the weighted dataset, this condition holds. Specifically, using subscripts to denote the weighted data, we require weights such that $E_w(\mathbf{h} \mid \mathbf{a}) = E_w(\mathbf{h})$.

A sufficient condition to give us this result is that the conditional distribution of the dyad history in the weighted dataset is equal to the marginal, $f_w(\mathbf{h} \mid \mathbf{a}) = f_w(\mathbf{h})$. Expressing the left-hand side in terms of the joint weighted distribution, we find that this condition is equivalent to

$$\frac{p_w(\mathbf{a} \mid \mathbf{h})}{p_w(\mathbf{a})} f_w(\mathbf{h}) = f_w(\mathbf{h}).$$

It follows that we require weights such that,

$$\frac{p_w(\mathbf{a} = (0,0)|\mathbf{h})}{p_w(\mathbf{a} = (0,0))} = \frac{p_w(\mathbf{a} = (0,1)|\mathbf{h})}{p_w(\mathbf{a} = (0,1))} = \frac{p_w(\mathbf{a} = (1,0)|\mathbf{h})}{p_w(\mathbf{a} = (1,0))} = \frac{p_w(\mathbf{a} = (1,1)|\mathbf{h})}{p_w(\mathbf{a} = (1,1))} = 1. \tag{A.2}$$

Having the numerators and the denominators equal in Equation A.2 is sufficient for it to hold. Let $w(\mathbf{a}, x)$ denote the weight function. Beginning with the numerators, the

weighting scheming means we can write each numerator as,

$$p_w(\mathbf{a} = \mathbf{a}_0|\mathbf{h}) = \frac{p(\mathbf{a} = \mathbf{a}_0|\mathbf{h})w(\mathbf{a}_0, \mathbf{h})}{\sum_{\mathbf{u}\in\mathcal{A}^2} p(\mathbf{a} = \mathbf{u}|\mathbf{h})w(\mathbf{u}, \mathbf{h})},$$

where $a_0 \in \mathcal{A}^2$.

It follows immediately that when Equation A.1 holds, the numerators will all be equal.

Similarly, for any $\mathbf{a}_0$ in $\mathcal{A}^2$ the denominators we may written as,

$$\begin{aligned} p_w(\mathbf{a} = \mathbf{a}_0) &= \int \frac{p(\mathbf{a} = \mathbf{a}_0|\mathbf{h})w(\mathbf{a}_0, \mathbf{h})}{\sum_{\mathbf{u}\in\mathcal{A}^2} p(\mathbf{a} = \mathbf{u}|\mathbf{h})w(\mathbf{u}, \mathbf{h})} f(\mathbf{h})\, d\mathbf{h} \\ &= \int p(\mathbf{a} = \mathbf{a}_0|\mathbf{h})w(\mathbf{a}_0, \mathbf{h})k(\mathbf{h})^{-1} dP_h, \end{aligned}$$

where we have defined $k(\mathbf{h}) = \sum_{\mathbf{u}\in\mathcal{A}^2} p(\mathbf{a} = \mathbf{u}|\mathbf{h})w(\mathbf{u}, \mathbf{h})$. We can now see that when Equation A.1 holds, we will also have the numerators being equal. This concludes the proof. □

A.1.3 Interference models and weights estimation

We intend to discuss possible weight specifications satisfying the balancing conditions, and how they may simplify under different modelling casual graphs. Commonly, weights are a function of the joint propensity score, $P(\mathbf{a} \mid \mathbf{h})$, which is generally unknown and must be estimated from the data. We have called this model the treatment model. Under the causal framework explained in Section 3.2.2 the dyad-level treatment $\mathbf{a}$ can be considered as a categorical treatment which takes one of four values: $(0, 0), (0, 1), (1, 0)$, or $(1, 1)$, which we collect into the set S. For multiple nominal treatments, the generalized propensity scores are frequently modeled by a multinomial logistic regression. In the case of within-dyad interference, where each individual is exposed to a bivariate binary treatment, we can also implement a bivariate logistic model to estimate the weights. In another approach we can fit the marginal distributions $P(A_j \mid \mathbf{h}_j)$ and $P(A_{j*} \mid \mathbf{h}_{j^*})$, then model the association between the exposures via a Dale model [83].

The formulation of the dyad treatment as a categorical variable allows weights previous devised in the literature to be used in this dyadic-network context. The balancing weights

proposed for categorical treatment proposed in Li et al. [57], for example, will satisfy the balancing property. In their most general form these weights can be expressed as,

$$w(a, \mathbf{h}) = \frac{\frac{1}{P(a|\mathbf{h})}}{\sum_{s \in S} \frac{1}{P(s|\mathbf{h})}},$$

which can be interpreted as a Boltzmann distribution over the treatment space where the energy of each treatment is defined to be the negative log of the propensity score for that treatment.

Sometimes we estimate $\zeta(a_j \mid \mathbf{h}) = P(a_j \mid \mathbf{h})$ and $\nu(a_{j^*} \mid a_j, \mathbf{h}) = P(a_{j^*} \mid a_j, \mathbf{h})$ to form joint treatment model. Under this the balancing condition of A.2 can be written as,

$$\begin{aligned}
w(0,0,\mathbf{h}) =& \frac{\zeta(\mathbf{h})\nu(1,\mathbf{h})w(1,1,\mathbf{h})}{[1-\zeta(\mathbf{h})]\,[1-\nu(0,\mathbf{h})]}, \\
w(0,1,\mathbf{h}) =& \frac{\zeta(\mathbf{h})\nu(1,\mathbf{h})w(1,1,\mathbf{h})}{[1-\zeta(\mathbf{h})]\,\nu(0,\mathbf{h})}, \\
w(1,0,\mathbf{h}) =& \frac{\zeta(\mathbf{h})\nu(1,\mathbf{h})w(1,1,\mathbf{h})}{\zeta(\mathbf{h})\,[1-\nu(1,\mathbf{h})]},
\end{aligned}$$

such that $k = 4\zeta(\mathbf{h})\nu(1,\mathbf{h})w(1,1,\mathbf{h})$. Additionally, the weights can be written specifically in terms of this model formulation as,

$$w(a, \mathbf{h}) = \frac{\frac{1}{\zeta(a_j|\mathbf{h})\nu(a_{j^*}|a_j,\mathbf{h})}}{\sum_{s \in S} \frac{1}{\zeta(s_j|\mathbf{h})\nu(s_{j^*}|s_j,\mathbf{h})}},$$

As we have already mentioned, under simpler causal graphs than the most general case, as presented in Figure 3.1, the formulation of the treatment model, and hence the weights may be simplified. For example, under a model specified according to Figure 3.2, the treatment of each dyad member is independent of the other condition on their individual and shared covariates. In such a case, $P(a_{j^*} \mid a_j, \mathbf{h}) = P(a_{j^*} \mid \mathbf{h})$ and therefore $\nu(a_{j^*} \mid a_j, \mathbf{h}) = \nu(a_{j^*} \mid \mathbf{h})$. The balancing condition in Equation A.2 simplifies in this case

to the same condition for each set of marginal distributions, i.e.

$$\frac{p_w(a_i = a \mid \mathbf{h})}{p_w(a_i = a)} = 1, \quad \forall a \in \mathcal{A},$$

for $i = 1, 2$ indexing the dyad treatments. Therefore,

$$w(a, \mathbf{h}) = w(a_1, \mathbf{h}) w(a_2, \mathbf{h}).$$

It follows that that in such a case, any weights previously satisfying the balancing condition provided in, for example, Wallace and Moodie [9] can be used on each member of the dyad individually and multiplied together to form a dyad weight.

A.2 General Interference Model for a Two-stage DTR

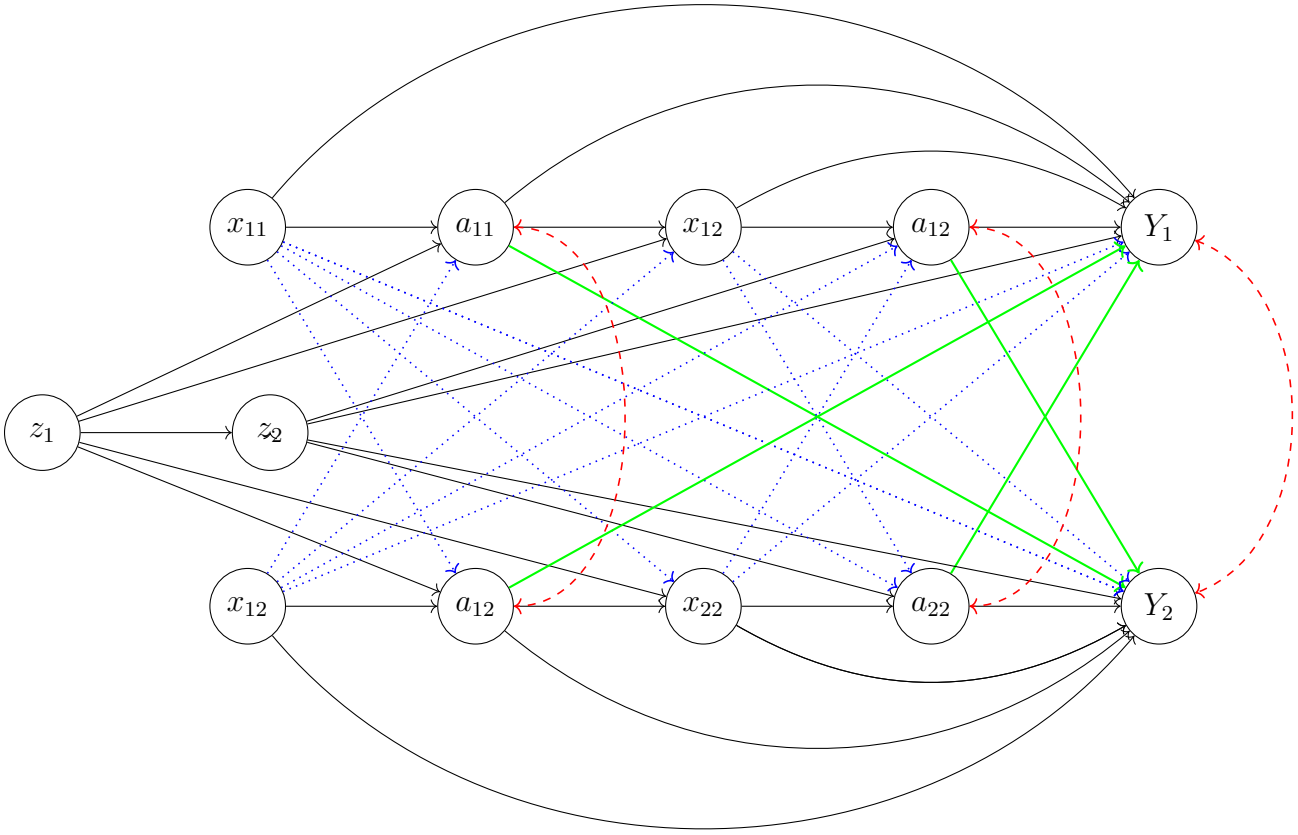


Figure A.1: General interference model within a given dyad for two stages. The treatments and individual characteristics of each member at each stage, can influence both the treatment and the outcome of the other member. A directed arrow signifies a potential causal effect of the first variable on the second, while a bidirectional arc between variables suggests the presence of unobserved factors that could be shared causes of the two variables.

Table A.1: Mean, standard error, bias, root mean squared error of the blip estimators with sample size n=1000 dyads across four scenarios for the single-stage simulation.

	Scenario	Mean	Standard Error	Bias	RMSE
ψ_0	1	2.402	0.021	0.402	0.555
	2	2.026	0.012	0.026	0.372
	3	2.011	0.011	0.011	0.346
	4	2.013	0.011	0.012	0.347
ψ_1	1	0.567	0.010	−0.433	0.541
	2	0.967	0.010	−0.032	0.322
	3	0.990	0.009	−0.010	0.286
	4	0.989	0.007	−0.011	0.291
η_0	1	-2.004	0.007	−0.004	0.225
	2	-2.005	0.007	−0.005	0.234
	3	-1.977	0.007	0.023	0.208
	4	-1.996	0.008	0.004	0.214
ξ_0	1	0.988	0.013	0.012	0.403
	2	0.988	0.012	0.012	0.392
	3	0.993	0.011	0.007	0.353
	4	0.989	0.011	0.011	0.361
ξ_1	1	-1.994	0.009	0.006	0.284
	2	-1.991	0.009	0.009	0.277
	3	-2.001	0.008	−0.001	0.242
	4	-1.997	0.008	0.003	0.252
True value of ψ_0 =2, ψ_1 =1 , η_0 =-2 , ξ_0=1 and ξ_1 =-2.					

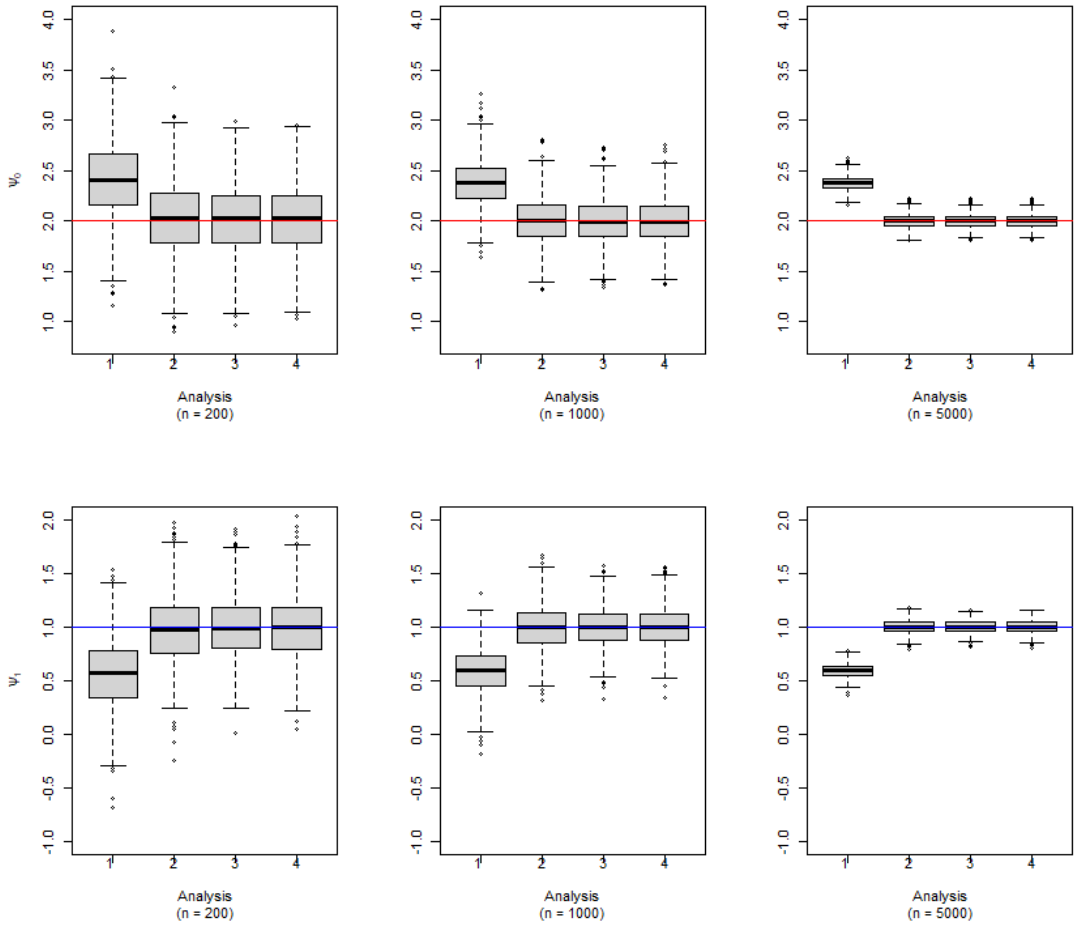


Figure A.2: Dyadic-blip function parameter estimates for 1000 simulated datasets via our dyadic-dWOLS approach when neither (Analysis 1), one (Analyses 2 and 3) or both (Analysis 4) treatment and treatment-free models are correct for single-stage simulation. n is the sample size of dyads.

A.3 Simulation

A.3.1 Single-stage regime

In this section we present additional results regarding the analysis presented in Section 3.4. These are provided in Figure(A.2) and Table(A.1).

A.3.2 Two-stage regime

Here we investigate the performance of the proposed model with two stages of treatment. For each dyad $i = 1, \ldots, n$ and each member $j = 1, 2$ we generate the patient information at stage one, X_{ij1}, according to a normal distribution with mean 2 and standard deviation 1. Patient information at the second stage, X_{ij2}, is generated according to a uniform distribution on the interval $[0, 2]$. The treatment model at stage k is defined as $P(A_{ijk} = 1 \mid X_{ijk} = x_{ijk}) = \text{expit}\big\{-0.25x_{ijk} + \sin(x_{ijk})\big\}$, for $k = 1, 2$. The outcome model is $Y_{ij} \mid (\mathbf{h}_{ij}, \mathbf{a}_i) = \mu(\mathbf{h}_{ij}, \mathbf{a}_i) + \varepsilon_{ij}$ where $\mu(\mathbf{h}_{ij}, \mathbf{a}_i) = \log(|x_{ij1}|) + \log(|x_{ij^*1}|) + \sin x_{ij2} + \log(|x_{ij^*2}|) + \sum_{k=1}^{2} \gamma_k$, where the dyadic-blip model at stage k is $\gamma_k = a_{ijk}(-2 + x_{ijk}) + a_{ij^*k}(-1 + 0.5x_{ijk}) + a_{ijk}a_{ij^*k}(1 + 0.5x_{ijk})$. Model misspecification for the treatment model, at each stage, is done by exclusion of the non-linear term. The treatment-free model at stage two is misspecified by eliminating $\log |x_{ij^*2}|$, and at stage one by excluding $\log |x_{ij^*1}|$.

Figure A.3 shows boxplots of the estimated dyadic-blip parameters for the first stage, while Figure A.4 plots the same for the second-stage estimates for the specific case of $n = 1000$ dyads. Blue horizontal lines have been drawn at the true value of each parameter. The label 1 denotes the case in which neither component of the model is correctly specified, 2 denotes the case in which only the treatment-free component is misspecified, 3 denotes the case where only the treatment model is misspecified, and 4 denotes the case in which everything is correctly specified. Consistent estimates of dyadic-blip parameters are achieved when at least one of the treatment model or treatment-free model is correctly specified.

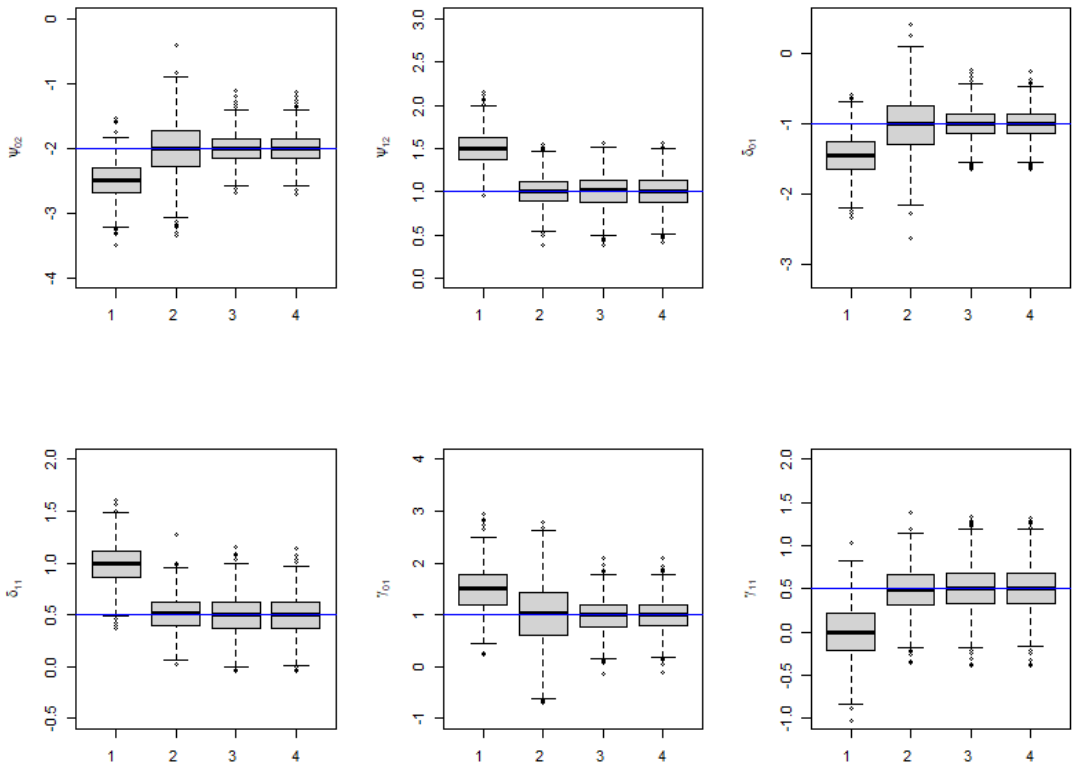


Figure A.3: First-stage dyadic-blip parameter estimates for 1000 simulated datasets via the dyadic-dWOLS approach for a two-stage DTR, when both (Analysis 1), one (Analyses 2 and 3) or neither (Analysis 4) treatment and treatment-free models are misspecified.

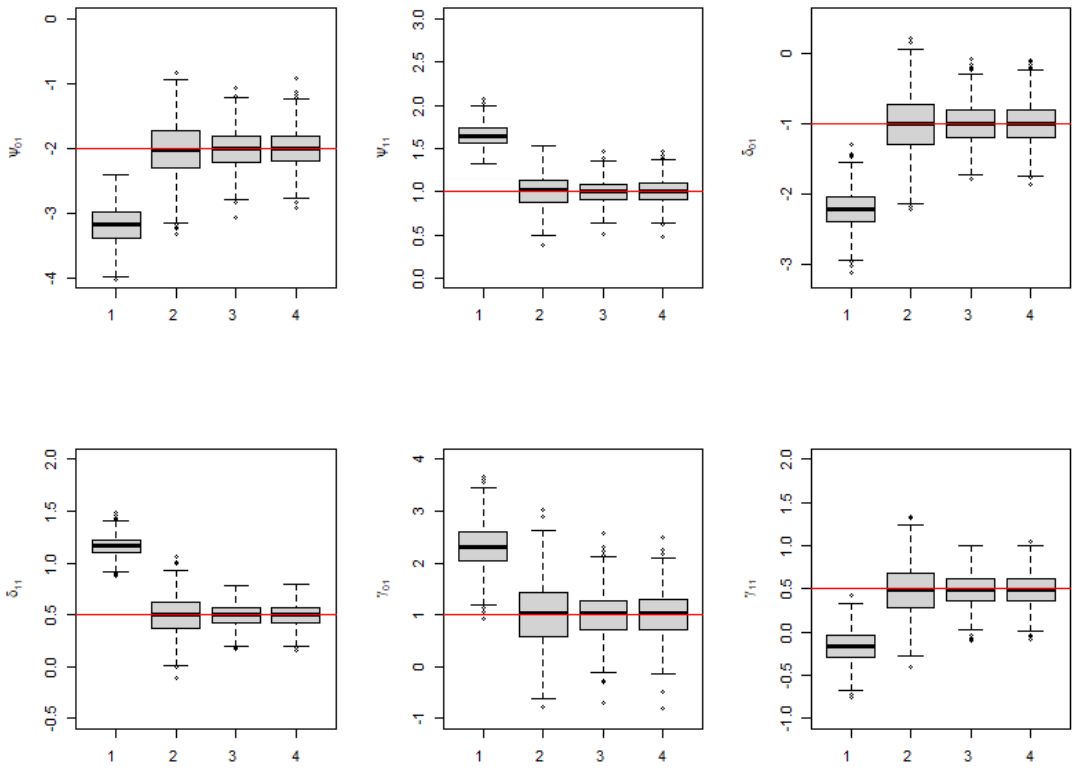


Figure A.4: Second-stage dyadic-blip parameter estimates for 1000 simulated datasets via the dyadic-dWOLS approach for a two-stage DTR, when both (Analysis 1), one (Analyses 2 and 3) or neither (Analysis 4) treatment and treatment-free models are misspecified.

Appendix B

Chapter 4 Appendix

B.1 Double Robustness Proof

Here we give the proof for Theorem 4.2.2.

Proof. This proof is similar to the proof of Theorem 3.3.1 from Chapter 4, which can be found in Appendix A. The difference is that under the assumption of a continuous treatment, Equation A.2, the balancing equation, is replaced with

$$\frac{\pi_w(a \mid \mathbf{h})}{\pi_w(a)} = 1,$$

where the ratio is now given in terms of probability density functions rather than probability mass functions. Again, a sufficient condition for this is to have equality in the denominators and numerators for each value of a. Starting with the numerators, if we write them explicitly as,

$$\pi_w(a \mid \mathbf{h}) = c(\mathbf{h})^{-1}\pi(a \mid \mathbf{h})w(a, \mathbf{h}),$$

where $c(\mathbf{h}) = \int_{\mathcal{A}} \pi(u \mid \mathbf{h})w(u, \mathbf{h})$, it can be seen immediately that if Equation 4.2 holds, then the denominators will be equal. With this result, if we similarly write the denominators

explicitly,

$$\pi_w(a) = \int c(\mathbf{h})^{-1} \pi(a \mid \mathbf{h}) w(a, \mathbf{h}) f(\mathbf{h}) d\mathbf{h}.$$

By our result regarding the numerators, the integrated will be the same for each value of a, hence we also have equality of the denominators when Equation 4.2 holds. This concludes the proof. □

B.2 Extra Simulation Results

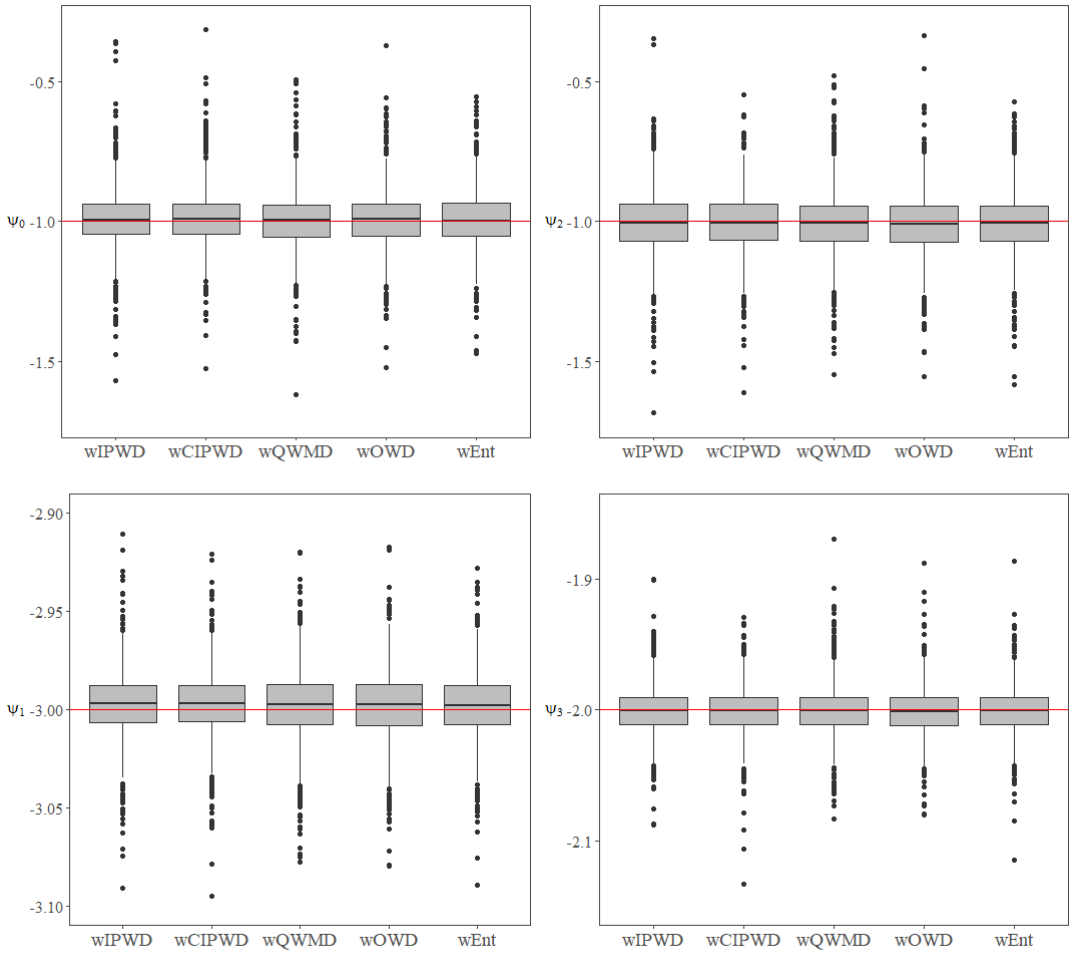


Figure B.1: Dyadic-blip parameter estimates for 1000 simulated datasets via the dyadic-dWOLS approach treatment-free models are misspecified and the treatments are normal.

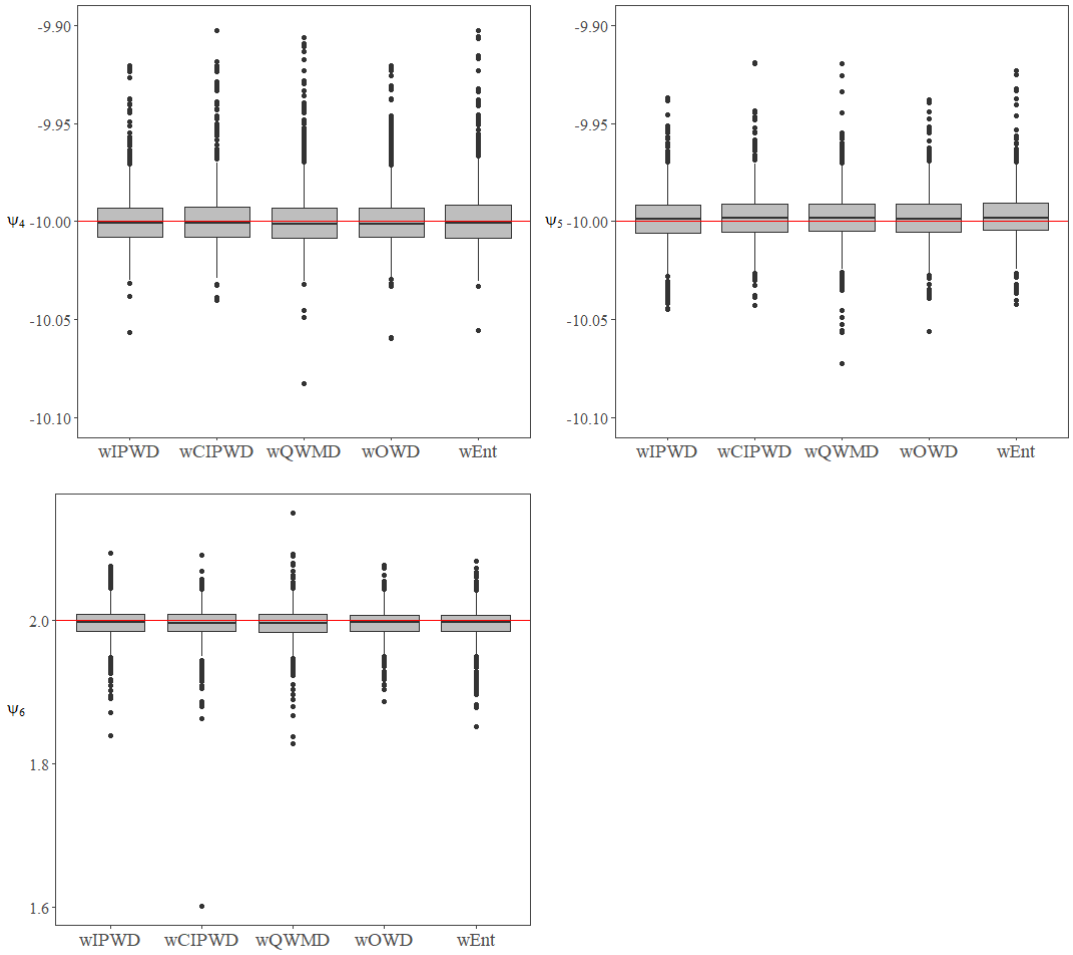


Figure B.2: Dyadic-blip parameter estimates for 1000 simulated datasets via the dyadic-dWOLS approach treatment-free models are misspecified and the treatments are normal.

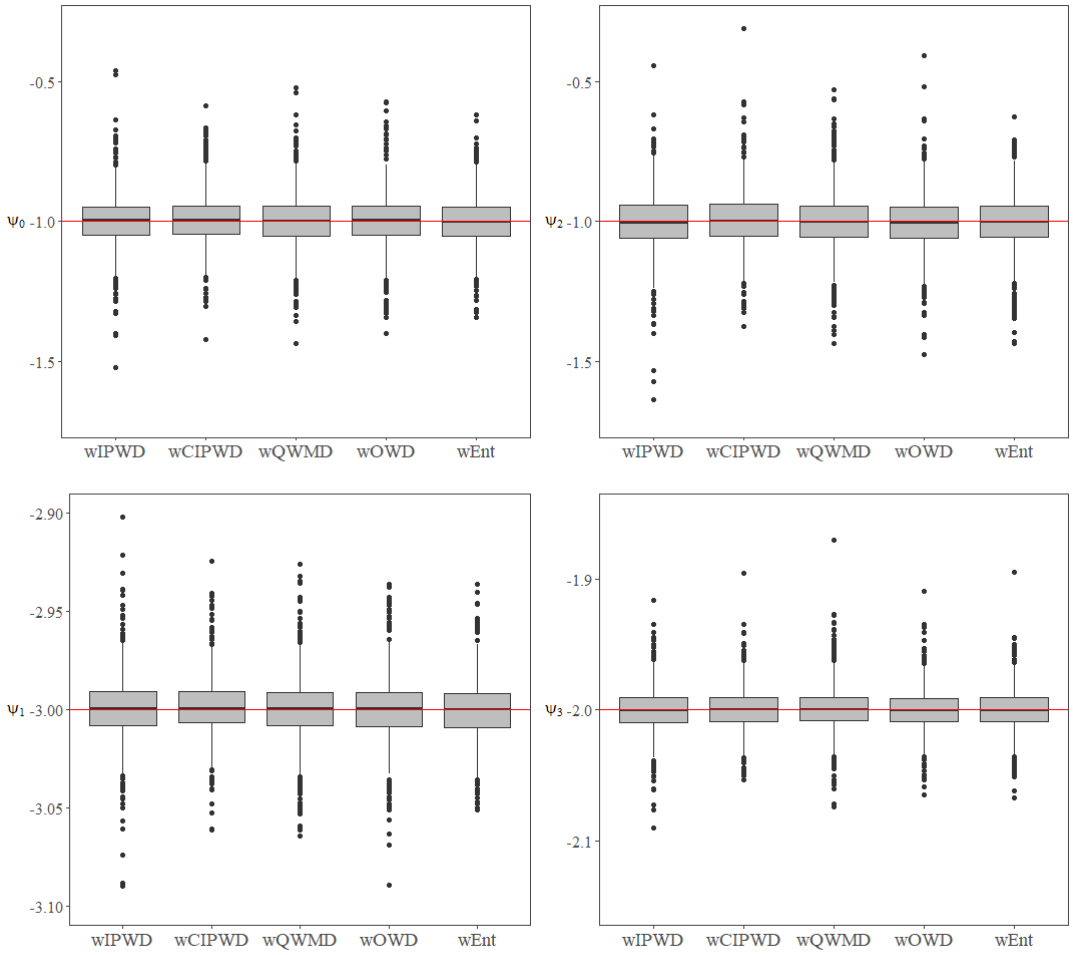


Figure B.3: Dyadic-blip parameter estimates for 1000 simulated datasets via the dyadic-dWOLS approach treatment models are misspecified and the treatments are normal.

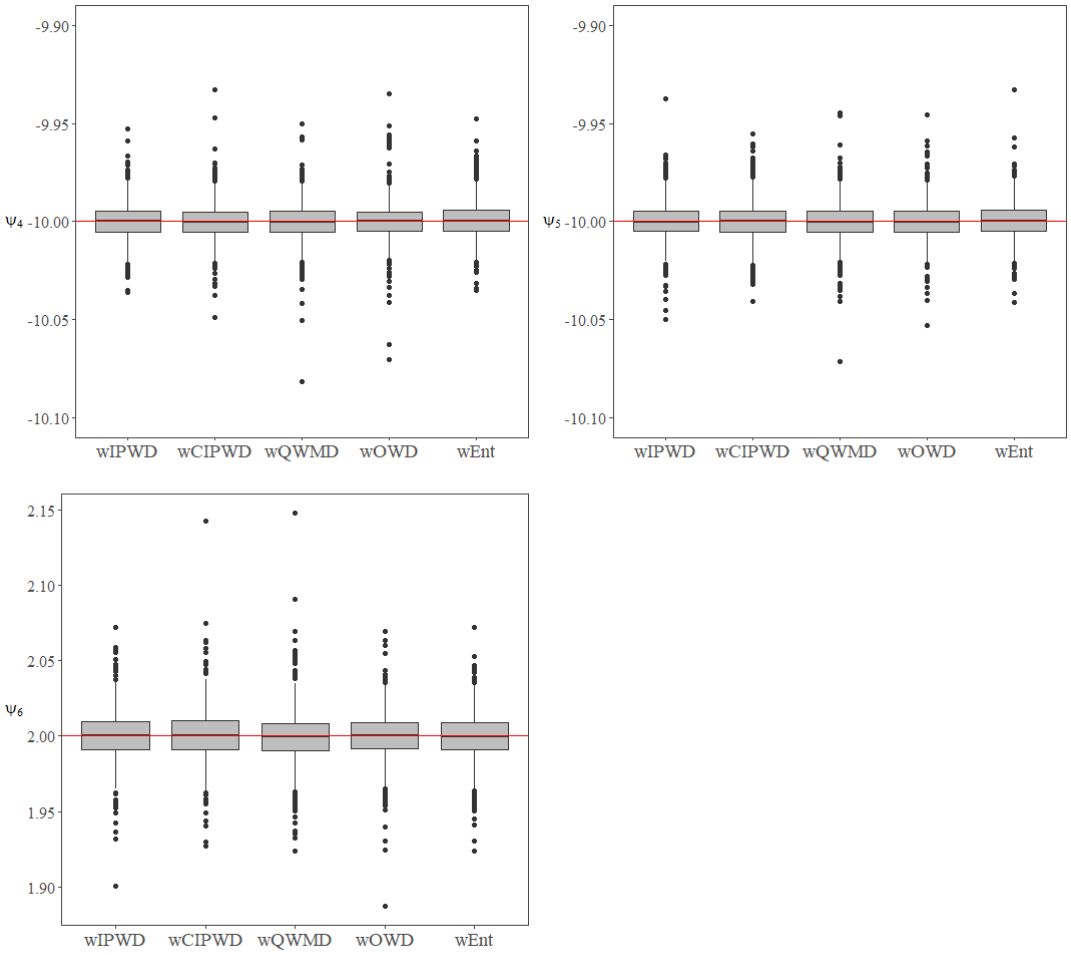


Figure B.4: Dyadic-blip parameter estimates for 1000 simulated datasets via the dyadic-dWOLS approach treatment models are misspecified and the treatments are normal.

	Dyadic-blip parameters						
	ψ_0	ψ_1	ψ_2	ψ_3	ψ_4	ψ_5	ψ_6
IPWD	0.437	0.035	0.419	0.034	0.033	0.035	0.028
CIPWD	0.176	0.017	0.169	0.016	0.018	0.018	0.015
WMD	0.195	0.030	0.156	0.015	0.036	0.018	0.017
OWD	0.188	0.029	0.154	0.015	0.034	0.017	0.016
EWD	0.153	0.015	0.159	0.016	0.017	0.018	0.015
mvGPS	0.194	0.022	0.207	0.020	0.028	0.023	0.022
STWD	0.162	0.014	0.159	0.016	0.015	0.018	0.015

Table B.1: RMSE: Normal treatment, treatment free model is misspecified. Sample size is 1000 dyads and the RSME are multiplied by 10^2

	Dyadic-blip parameters						
	ψ_0	ψ_1	ψ_2	ψ_3	ψ_4	ψ_5	ψ_6
IPWD	0.057	0.010	0.055	0.009	0.006	0.006	0.010
CIPWD	0.149	0.025	0.150	0.025	0.014	0.015	0.024
WMD	0.080	0.013	0.077	0.013	0.008	0.008	0.013
OWD	0.057	0.010	0.054	0.009	0.006	0.006	0.010
EWD	0.057	0.010	0.054	0.009	0.006	0.006	0.010
mvGPS	0.058	0.010	0.056	0.009	0.006	0.006	0.010
STWD	0.128	0.020	0.150	0.023	0.012	0.011	0.022

Table B.2: Log normal treatment. Treatment model is misspecified. Sample size is 1000 dyads and the RSME are multiplied by 10^2

	Dyadic-blip parameters						
	ψ_0	ψ_1	ψ_2	ψ_3	ψ_4	ψ_5	ψ_6
IWPD	0.117	0.019	0.118	0.019	0.016	0.013	0.022
CIPWD	0.125	0.020	0.123	0.020	0.015	0.013	0.024
WMD	0.126	0.022	0.120	0.020	0.020	0.014	0.027
OWD	0.116	0.020	0.123	0.021	0.017	0.014	0.025
EWD	0.124	0.020	0.120	0.021	0.017	0.014	0.022
mvGPS	0.109	0.019	0.121	0.020	0.017	0.014	0.024
STWD	0.119	0.019	0.114	0.019	0.016	0.013	0.024

Table B.3: RMSE: Log normal treatment, treatment free model is misspecified. Sample size is 1000 dyads and the RSME are multiplied by 10^2

	Dyadic-blip parameters						
	ψ_0	ψ_1	ψ_2	ψ_3	ψ_4	ψ_5	ψ_6
IPWD	0.100	0.016	0.106	0.017	0.009	0.009	0.015
CIPWD	0.101	0.016	0.109	0.017	0.009	0.010	0.016
WMD	0.096	0.016	0.101	0.016	0.009	0.010	0.017
OWD	0.099	0.017	0.109	0.018	0.010	0.010	0.017
EWD	0.097	0.016	0.100	0.016	0.010	0.010	0.016
mvGPS	0.090	0.015	0.105	0.017	0.010	0.010	0.016
STWD	0.099	0.016	0.097	0.016	0.009	0.010	0.016

Table B.4: Log Normal treatment, both treatment and treatment-free models are correct. Sample size is 1000 dyads and the RSME are multiplied by 10^2

	Dyadic-blip parameters						
	ψ_0	ψ_1	ψ_2	ψ_3	ψ_4	ψ_5	ψ_6
IPWD	0.204	0.036	0.222	0.036	0.038	0.025	0.044
CIPWD	0.235	0.039	0.218	0.037	0.036	0.026	0.046
WMD	0.254	0.045	0.230	0.038	0.044	0.028	0.052
OWD	0.212	0.038	0.214	0.037	0.039	0.026	0.046
EWD	0.410	0.062	0.408	0.065	0.037	0.027	0.042
mvGPS	0.210	0.038	0.210	0.036	0.039	0.026	0.046
STWD	0.242	0.041	0.218	0.036	0.036	0.026	0.048

Table B.5: RMSE: Log normal treatment, both models are misspecified. Sample size is 1000 dyads and the RSME are multiplied by 10^2.

B.3 Weights

- All weights except the entropy balancing weights are constructed by estimating the generalized propensity score [85–88].
- In entropy balancing, the balance is achieved by direct minimization of the loss function H(ω) subject to a balance and normalization constraints. q_i is a set of base weights chosen by the researcher.
- mvGPS weights are the generalization of SPW for multivariate normal treatments and modification of estimable exposure
- The balancing condition stated in Theorem 4.2.2 is sufficient condition but not necessary, meaning there are weights that don't satisfy the condition but still achieve balance. SPWD, EWD, and mvGPS are examples of such weights.

Table B.6: Weight functions implemented in the dyadic-dWOLS

Weight Function	**Description**
IPWD (inverse probability of treatment)	$\frac{1}{\pi(\mathbf{a}\|x)}$
CIPWD (capped inverse probability of treatment)	$\min\left\{\frac{1}{\pi(\mathbf{a}\|x)}, c\right\}$
STWD (stabilized inverse probability of treatment)	$\frac{\pi(\mathbf{a})}{\pi(\mathbf{a}\|x)}$
WMD (weights for categorized treatment POM)	$\frac{1-\pi(\mathbf{a}=\mathbf{k}\|x)}{\pi(\mathbf{a}=\mathbf{k}'\|x)}$
OWD (overlap weights for categorized treatment POM [57])	$\frac{\frac{1}{\pi(\mathbf{a}=\mathbf{k}\|x)}}{\sum_{j=1}^{m}\sum_{j^*=1}^{m}\frac{1}{\pi(\mathbf{a}=\mathbf{k}'\|x)}}$
EWD (entropy balancing with second moment) [76–78]	$\mathrm{H}(\omega_i) = \omega_i \ln(\frac{\omega_i}{q_i})$
mvGPS (multivariate generalized propensity score)[84]	$\frac{\pi(\mathbf{a})}{\pi(\mathbf{a}\|\boldsymbol{x})}$