INTERPRETABLE AND PERSONALIZED DECISION-MAKING IN HEALTHCARE

by

Qiaomei Li

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

(Industrial and Systems Engineering)

at the

UNIVERSITY OF WISCONSIN–MADISON

2024

Date of final oral examination: 5/16/2024

The dissertation is approved by the following members of the Final Oral Committee:
Yonatan Mintz, Assistant Professor, Industrial and Systems Engineering
Oguzhan Alagoz, Professor, Industrial and Systems Engineering
Justin J. Boutilier, Assistant Professor, Industrial and Systems Engineering
Corrine I. Voils, Professor, Surgery

To Mom & Dad

ACKNOWLEDGMENTS

First and foremost, I am immensely grateful to my advisor, Yonatan Mintz, without whom this thesis would not have been possible. He helped me navigate through the most challenging times and motivated me to reach my full potential. Not only did he teach me how to excel as a researcher, but he also inspired me to be a compassionate person. In addition, I would like to thank my past and current dissertation committee members - Oguzhan Alagoz, Justin J. Boutilier, Corrine I. Voils, and Nicole Werner - for their time and valuable feedback. Special thanks to Rachel Cummings and Corrine I. Voils for their consistent support and collaboration on multiple projects. Furthermore, I am grateful to my undergraduate mentors Joseph O'Rourke and Gwen Spencer for motivating me to pursue graduate school.

I've been fortunate to have met many talented individuals during my doctoral studies at the Georgia Institute of Technology and the University of Wisconsin - Madison. A heartfelt thank you to my labmates—Katherine Adams, Qinyang He, Vrishabh Patil, Eric Pulick, and Jinxin Tao—for fostering a productive research environment, engaging in insightful discussions during lab meetings, and participating in presentation practices. I also want to extend my gratitude to Valerie Odeh Couvertier and Xinyu Liu, for being my cheering squad and motivating me to become a better version of myself. Additionally, I would like to thank my fellow graduate students—Fernando A. Acosta-Pérez, Bainian Hao, Harshit Kothari, Vanessa Sawkmie, Eric Stratman, and Zach Zhou—for their support and kindness.

I want to express my gratitude to my friends and family for keeping me grounded throughout the past six years. A special thank you to Qinxuan Jin and Emma Ning, who have known me for over a decade, offering support through Zoom chats, endless texting, and so much more. Most importantly, I want to express my deepest appreciation to my parents, Zhonglin Li and Huixiang Peng, for their unconditional love and unwavering belief in me. Despite starting with little, they worked tirelessly to provide me the education they never had. My achievements would not have been possible without their sacrifices.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Personalized healthcare strives to offer customized prevention, diagnosis, and treatment plans for each individual, in contrast to the one-size-fits-all method, which offers standardized solutions for broad populations. By adapting treatments to individual variations such as disease progression and lifestyle factors, personalized healthcare aims to enhance efficient and effective patient-centered care. Additionally, it aids in reducing potential adverse effects and optimizing treatment efficacy and resource allocation. In this thesis, we focus on developing methodologies to anticipate the progression of diseases and forecast treatment outcomes for individuals. Furthermore, we study how to customize treatment strategies tailored to the unique responses from each individual.

1 INTRODUCTION

Personalized health planning is the cornerstone of personalized healthcare, focusing the unique health needs of each individual. Such a personalized plan holds the potential to enhance an individual's health trajectory while mitigating long-term adverse effects. This thesis aims to develop frameworks that facilitate the delivery of effective and tailored care, empowering individuals to adopt behaviors that manage disease progression and promote healthier lifestyles.

First, We introduce a novel way to construct optimal aggregate explainer models, aiding clinicians in comprehending predictions generated by machine learning and AI models. Our framework includes a local explainer aggregation method which uses non-convex optimization to select local explainers and an integer optimization framework for crafting a nearly global aggregate explainer. We assess the efficacy of our framework using the Parkinson's Progression Marker Initiative dataset and a geriatric mobility dataset, demonstrating that our approach outperforms five state-of-the-art explainer methods in terms of fidelity. Next, we propose a behavioral framework aimed at tackling the obstacle of distributing financial incentives to align with individual motivations and stimulate the highest participation in weight loss efforts. Our proposed framework consists of a behavioral model to capture the dynamics of one's motivational and physical states, a surrogate likelihood approach for estimating these unknown state parameters, and an algorithm to optimize the allocation of incentives to each individual. We provide both theoretical guarantees and experimental findings to demonstrate the cost-efficiency and efficacy of our personalized approach for weight loss interventions. Finally, we study the problem of identifying adherence patterns in a behavioral weight management intervention featuring financial incentives.We provide a general framework for designing incentive interventions through identifying and predicting adherence subtypes. We implement this framework in an ongoing randomized weight loss controlled trial. The results validate the framework's utility in identifying adherence subtypes within the intervention's scope, assisting researchers in effectively detecting key treatments and determining appropriate treatment levels for each participant.

2 OPTIMAL LOCAL EXPLAINER AGGREGATION FOR INTERPRETABLE PREDICTION

2.1 Introduction

When applying machine learning and AI models in high risk and sensitive settings, one of the biggest challenges for decision makers is to rationalize the insights provided by the model. In applications such as precision medicine, both prediction accuracy (e.g., anticipated efficacy of treatment) and transparency of how predictions are made are key for obtaining informed consent. However, the models that typically achieve the highest levels of accuracy also tend to be extremely complex, and even machine learning experts describe them as "black boxes" because it is difficult to explain why certain predictions are made (Breiman 2001b). One popular approach to resolve this trade off between explainability and accuracy is to extract simple *explainer* models from complex black box models. These models are intended to provide a simplified facsimile of the true model that is more useful for human interpretation of the generated predictions.

Two important widely-used metrics for evaluating explainer models are *fidelity* and *coverage*. Fidelity measures how well the explainer's predictions match the predictions of the original black box model, and coverage measure the fraction of the data universe that is reasonably explained by the explainer model. Explainer methods are generally classified as either *global* or *local*, based on how they trade off between these two quantities. Global explainers attempt to explain the full black box model across the entirety of the data. These models have a hard constraint to provide 100% coverage, often at the expense of fidelity. Local explainers, on the other hand, sacrifice coverage to potentially provide higher fidelity explanations in a smaller region of the data, usually centered around one single prediction.

Recent proposals suggest finding a middle ground between these two extremes by forming global (or near-global) explainers by aggregating local explainer models Ribeiro et al. (2016). This approach would allow the decision-maker to trade off

among coverage, fidelity, and explainability: including more local explainers in the aggregate model would improve coverage and fidelity, at the cost of a more complex—and hence less interpretable—aggregate model. However, the problem of computing the best subset of local explainers to explain a given black box model is combinatorial in nature, and hence computationally challenging to solve. All existing methods for building aggregate explainers use only heuristic approaches, and thus do not provide theoretical performance guarantees.

In this work, we present a novel way of constructing provably optimal aggregate explainer models from local explainers. We use an integer programming (IP) optimization framework that trades off between coverage of the aggregate model and fidelity of the local explainers that comprise the aggregate model. We also propose a local explainer methodology that uses an information filter for feature selection, and is designed for use in aggregation. We empirically validate the performance of this framework in two healthcare applications: Parkinson's Disease progression and geriatric mobility. These experimental results show that our model provides higher fidelity than existing methods. In this application, a clinician would use a black box model for their initial diagnosis of a patient, and then use that patient's data in the particular local explainer selected by our algorithm to understand why the black box model made its prediction.

Related Work

Our paper builds on previous work in the broader field of interpretable machine learning. The two primary types of interpretable learning include models that are interpretable by design (Aswani et al. 2019), and black box models that can be explained using global explainer (Wang and Rudin 2015, Lakkaraju et al. 2016, Ustun and Rudin 2016, Bastani et al. 2018) or local explainer (Ribeiro et al. 2016, 2018) methods.

Models that are interpretable by design are perhaps the gold standard for interpretable ML. However, these models often require significant domain knowledge to formulate and train, and are not suited for exploratory tasks such as the precision

healthcare applications we study in Section 2.4.

Global explainer methodology attempts to train an explainable model (e.g., a decision tree with minimal branching) to match the predictions of a black box model across the entirety of its feature space. While these models provide some understanding on the general behavior of the black box model, if the relationship between features and black-box predictions is too complex, then the global explainer may remove subtleties that are vital for explanation.

Local explainer methods attempt to train simpler models centered around the prediction for a single data point. The most commonly used local explainer methods are Local Interpretable Model-Agnostic Explanations (LIME) (Ribeiro et al. 2016) and anchors (Ribeiro et al. 2018). While local methods cannot validate the full black box model, they are useful for understanding the subtleties and justification for particular predictions. In recent literature several other local explainer methods have been proposed that draw inspiration from this stream (Rajapaksha et al. 2019, Sokol and Flach 2020, Plumb et al. 2018).

A third option which has been explored in recent literature, is that of aggregating several local explainer models to form a near-global explainer, as a method for improving the tradeoff between fidelity and coverage. Generally speaking, these methods have a budget for the maximum number of local explainers that can be incorporated into the aggregation and attempt to maximize possible coverage and fidelity within this budget. One method proposed to form such aggregate explainers is the submodular pick method (Ribeiro et al. 2016), which computes feature importance scores and greedily selects the features with highest importance. van der Linden et al. (2019) argue the Submodular Pick Algorithm has its limitations on predicting global behaviors from local explainers, and that the choice of aggregation function for local explainers is important for performance. They introduce the Global Aggregations of Local Explanations (GALE) method, which can be used to analyze how well the aggregation explains the model's global behavior. They compared the performance of global LIME aggregation with other global aggregation methods for binary and multi-class classification tasks, and found that different aggregation approaches performed best in binary and multi-class

settings. A recently proposed aggregation method (called GLocalX) hierarchically combines local explainers to from global explainers (GLocalX) Setzu et al. (2021). This methodology could be incorporated in to our optimization approach as pre-processing.

The methodology we propose in this paper builds on top of these existing explainer aggregation methods. In contrast to existing approaches which are heuristic in nature, we formulate the problem of choosing local explainers for the aggregate as an optimization problem. By doing so, our methods can produce explainer aggregates that provide both higher fidelity and higher coverage than existing approaches. In addition, our formulation includes parameters that allows for a direct tradeoff between coverage, fidelity, and interpretability. We believe this approach is especially appropriate for problems in explainable precision healthcare, where the relationship between diagnostic screening measures and the diagnosis is quite complex, and the model should incorporate the richness of this relationship in its predictions.

We propose a local explainer approach in Section 2.3 that includes a feature selection subroutine to improve explainability. Prior work on feature selection includes instance-wise feature selection (Chen et al. 2018) and Instance-wise Variable Selection using Neural Networks (INVASE) (Yoon et al. 2018). These approaches select the important features for each sample point using networks for classification with and without the features. Shapley values have also been used for complex model predictions, such as Shapley Sampling Values (Štrumbelj and Kononenko 2014, Aas et al. 2019) and Shapley Additive Explanations (SHAP) (Lundberg and Lee 2017, Lundberg et al. 2020), which computes Shapley values and presents the explanation as an additive feature attribution method. In contrast to these methods, our feature selection approach relies on a mutual information filter (Brown et al. 2012) to identify important features. While mutual information has been used in the past for feature selection, we introduce a computationally efficient way to compute this mutual information for the use of training local explainer models.

Our Contributions

In this paper, we formulate the problem of aggregating local explainers into an aggregate explainer algorithm as a non-convex optimization problem. In particular, we show that this aggregation problem can be written as an integer program (IP), that can be solved effectively using commercial solvers. This formulation is also helpful as it allows us to directly tradeoff coverage and fidelity of the resulting aggregation through parameters of the optimization problem. This approach provides flexibility to the practitioner to adapt the algorithm to the specific needs of her use case. Our approach easily handles multi-class prediction problems that arise in complex application domains such as precision healthcare, as well as the traditionally-studied binary classification.

Additionally, we design a new methodology for training local explainers for effective use in aggregation. Our local explainer algorithm directly computes locally significant features using an information filter, and we are the first to use information filters in local explainers. We introduce a novel computationally efficient algorithm for this filtering step, and our approach results in simpler (i.e., more interpretable) local explainers compared to prior work that used regularization for feature selection.

To validate our results, we compare our optimization based methodology against four other state of the art methods on two real world data sets. Both data sets come from the applications in the healthcare space. The first uses the Parkinson's Progression Marker Initiative PPMI (2019), where we create explainer methods for a model tasked with screening patients for Parkinson's Disease. The second uses a dataset of Geriatric activity, where we explain the predictions of a model that classifies the physical activity of geriatric patients to prevent falling. Our experiments show that our optimization method outperforms many of the existing explainer methods in terms of fidelity and coverage. In particular, when we examine cases of explaining multi-class model predictions, our explainer method can achieves 90% fidelity at 40-50% coverage, while existing global methods only achieved 70% fidelity, albeit at 100% coverage. Our results show that our approach on the Pareto frontier of the

fidelity and coverage tradeoff. Our IP framework outperforms existing aggregation methods in terms of both coverage and fidelity across all potential aggregation budgets (i.e., numbers of local explainers in the aggregate model).

2.2 Explainer Aggregation Methodology

Explainer models which can generalize to a large portion of the feature space are critical transparency. However, an explainer that is constrained to explain the entire feature space is likely have low fidelity since, by design, the explainer model is less complex than the black box model it is purported to explain. However, simpler models can achieve higher fidelity by attempting to explain the local behavior of the black box model at the cost of lower coverage.

One way to address the tradeoff between coverage and fidelity is to create a near-global aggregate explainer model by combining several local explainer models. Existing approaches have used this idea (Ribeiro et al. 2016) by formulating the construction of an aggregate explainer as an optimization problem: maximize coverage of the explainer subject to a constraint on the total number of local explainers included in the aggregate. Solving this optimization problem is conjectured to be computationally intractable (Ribeiro et al. 2016), and prior work has only attempted to solve it using heuristics.

In this section, we formulate the problem of constructing the aggregate explainer *from an arbitrary black-box model*, explicitly as an integer linear program that can be solved efficiently using commercial solvers, and allows us to directly trade off coverage and fidelity.

Mathematical Programming Formulation of Aggregation Problem

To formulate the optimization problem of constructing the aggregate explainer, we must first formally define the concepts of coverage and fidelity.

Let $\mathcal{X} \subset \mathbb{R}^m$ be the feature space that is modeled with a black box function, and let $f : \mathcal{X} \to \mathbb{Z}_+$ be the black box function of interest. Let $\mathcal{L} \subseteq \mathbb{Z}_+$ be the label space

in the image of f. We consider our explanation task over a dataset $\mathcal{D}$ containing n ordered pairs (x_i, y_i) for $i \in [n]$, where $x_i \in \mathcal{X}$ are the feature values and $y_i \in \mathcal{L}$ is the class label which has been generated by f. That is, $y_i = f(x_i)$.

Let $g_{i,r} : \mathcal{X} \to \mathcal{L}$ denote a local explainer model that explains the local behavior of the black box function f on inputs within a ball of radius $r \in \mathbb{R}_+$ centered around the point $x_i \in \mathcal{X}$. We use $\mathcal{X}_{i,r} := \{x \in \mathcal{X} : \|x - x_i\| \leqslant r\}$ to denote the region explained by $g_{i,r}$.

Define an aggregate explainer γ to be a set of local explainers centered around a subset of points in $\mathcal{D}$, where the local explainer for point $x_i \in \mathcal{D}$ has radius r_i.[1] We will refer to a generic local explainer $g \in \gamma$ and corresponding region of explanation $\mathcal{X}_g$.

Using these quantities we define the *coverage of aggregate explainer* γ on data set $\mathcal{D}$ as the total number of points in the data set that are covered by the explanation radius of at least one explainer contained in γ. We denote this as:

$$\text{Cov}(\gamma, \mathcal{D}) = \sum_{x \in \mathcal{D}} \max_{i \in \{i : g_{i,r} \in \gamma\}} \mathbb{1}[x \in \mathcal{X}_{i,r}]. \tag{2.1}$$

Next we note that the fidelity of a single local explainer is defined as the accuracy of that explainer with respect to the predicted labels of the black box model. We emphasize that fidelity captures the explainer's ability to replicate the predictions of the black-box model, and rather than ground truth predictive accuracy.

We define the *fidelity of aggregate explainer* γ on data set $\mathcal{D}$ as the minimum of the fidelity obtained by each individual local explainer in γ. We first need to define $\mathcal{D}_g$ as the number of points in the data set contained in the explanation region of g, i.e., $\mathcal{D}_g = \{x \in \mathcal{D} : x \in \mathcal{X}_g\}$. We denote this as:

$$\text{Fid}(\gamma, \mathcal{D}) = \min_{g \in \gamma} \frac{1}{|\mathcal{D}_g|} \sum_{x \in \mathcal{D}_g} \mathbb{1}[g(x) = f(x)]. \tag{2.2}$$

[1] More generally, any local explainers can be aggregated into γ. However, we assume the the explainer algorithm only has access to points in $\mathcal{D}$, so we restrict ourselves to only considering these points. It is assumed that the radii r_i are parameters of the problem and hence known to decision-maker.

While one could instead define the fidelity of γ as the average of the fidelities of its component explainers, our choice to use the minimum fidelity gives a stricter measure of how well the aggregate explainer captures the behavior of the black box model. This stricter measure is more appropriate for the healthcare applications we consider in Section 2.4, where a minimum standard of care is required. Note also that while we may be interested in the coverage and fidelity of γ across the entirety of $\mathcal{X}$, computing these quantities may be intractable or impossible in practice when $\mathcal{X}$ is not known a priori. Thus we consider these quantities only across an r-ball covering of our dataset.

Let K denote the budget of the maximum number of local explainers that can be contained in γ, and let φ be the minimum fidelity required for the aggregate explainer. Then the problem of computing an aggregate explainer can be formulated as the following optimization problem:

$$\max_{\gamma} \{\text{Cov}(\gamma, \mathcal{D}) : \text{Fid}(\gamma, \mathcal{D}) \geqslant \varphi, |\gamma| \leqslant K\}. \tag{2.3}$$

Reformulation as Integer Program (IP)

As written, optimization problem (2.3) is not trivial to solve, and could require enumerating all possible subsets γ of local explainers. To address this challenge, we propose reformulating problem (2.3) as an Integer Program (IP) that can be solved using commercial software. We first define three sets of binary variables w_i, y_j, z_{ij}. Let w_i be a binary variable that is equal to 1 if explainer $g_{i,r_i} \in \gamma$. That is, $w_i = \mathbb{1}[g_{i,r_i} \in \gamma]$. Let y_j be a binary variable that is equal to 1 if point j is covered by the aggregate explainer γ. That is $y_j = \mathbb{1}[x_j \in \cup_{g\in\gamma} \mathcal{X}_g]$. Finally, let z_{ij} be a binary variable that is equal to 1 if explainer $g_{i,r_i} \in \gamma$ covers point x_j. That is, $z_{ij} = \mathbb{1}[x_j \in \mathcal{X}_{i,r_i}]$. We now define the coverage and fidelity of aggregate explainer γ as IPs written in terms of these three sets of variables.

Proposition 1. *$\text{Cov}(\gamma, \mathcal{D})$, the coverage of aggregate explainer γ on dataset $\mathcal{D}$, can be*

expressed with the following set of integer variables and constraints:

$$\begin{aligned} Cov(\gamma, \mathcal{D}) = & \textstyle\sum_{j=1}^{n} y_j, \\ s.t. \quad & z_{ij} \leqslant w_i, \quad i, j \in [n], \\ & y_j \geqslant z_{ij}, \quad i, j \in [n], \\ & y_j \leqslant \textstyle\sum_{i=1}^{n} z_{ij}, \quad j \in [n], \\ & \|x_i - x_j\| z_{ij} \leqslant r_i, \quad i, j \in [n]. \end{aligned} \tag{2.4}$$

Proof. Recall the definition of $\mathrm{Cov}(\gamma, \mathcal{D})$ as given in Equation (2.1). We will directly reconstruct this definition using the binary variables defined above. First note that through a simple direct substitution we obtain $\mathrm{Cov}(\gamma, \mathcal{D}) = \sum_{x \in \mathcal{D}} \max_{i \in \{i : g_{i,r} \in \gamma\}} z_{ij}$. Since taking the maximum of binary variables is equivalent to the Boolean OR operator, we see that $y_j = \max_{i \in \{i : g_{i,r} \in \gamma\}} z_{ij}$, which provides us with the first equality. The next two inequalities directly capture that a local explainer g_{i,r_i} can only explain point x_j if g_{i,r_i} is included in γ, which is a standard way of modeling conditional logic in IP (Wolsey and Nemhauser 1999). The next two constraints come from modeling the Boolean OR operator using integer constraints (Wolsey and Nemhauser 1999). The final constraint ensures that a point x_j can only be covered by an explainer g_{i,r_i} if $x_j \in \mathcal{X}_{i,r_i}$, thus preserving the local region of the local explainer. □

Next we consider the minimum fidelity constraint.

Proposition 2. *The constraint* $Fid(\gamma, \mathcal{D}) \geqslant \varphi$ *can be modeled using the following set of integer linear constraints:*

$$\begin{aligned} & \|x_i - x_j\| z_{ij} \leqslant r_i, \quad i, j \in [n], \\ & z_{ij} \leqslant w_i, \quad i, j \in [n], \\ & \textstyle\sum_{j=1}^{n} (\mathbb{1}_{\{f(x_j) = g_{i,r_i}(x_j)\}} - \varphi) z_{ij} \geqslant 0, \quad i \in [n]. \end{aligned} \tag{2.5}$$

While the full proof of Proposition 2 is deferred to Appendix A.2, we note that the first two constraints ensure proper local behavior of the local explainer as in Proposition 1. The third constraint is derived by analysis of the definition of

$\text{Fid}(\gamma, \mathcal{D})$ in Equation (2.2), dis-aggregating the lower bound constraint across all $i \in [n]$, and re-writing the new lower-bound constraint to remove the min using properties of z_{ij}.

We can then use these expressions to for coverage and fidelity to re-write our optimization problem as an integer program that can then be solved using commercial solvers.

Proposition 3. *The optimization problem in* (2.3)*,*

$$\max_{\gamma} \{Cov(\gamma, \mathcal{D}) : Fid(\gamma, \mathcal{D}) \geqslant \varphi, |\gamma| \leqslant K\},$$

can be written as the following integer program:

$$\begin{aligned}
&\max \textstyle\sum_{j=1}^{n} y_j, \\
&\textit{s.t.} \quad z_{ij} \leqslant w_i, y_j, \quad i, j \in [n], \\
&\qquad y_j \leqslant \textstyle\sum_{i \in \mathcal{X}} z_{ij}, \quad j \in [n], \\
&\qquad \|x_i - x_j\| z_{ij} \leqslant r_i, \quad i, j \in [n], \\
&\qquad \textstyle\sum_{j=1}^{n} (\mathbb{1}_{\{f(x_j) = g_{i,r_i}(x_j)\}} - \varphi) z_{ij} \geqslant 0, \quad i \in [n], \\
&\qquad \textstyle\sum_{i \in \mathcal{X}} w_i \leqslant K, \\
&\qquad y_j, w_i, z_{ij} \in \{0, 1\} \quad i, j \in [n].
\end{aligned} \tag{2.6}$$

Proof. The objective function and first four constraints come directly from Propositions 1 and 2. The next constraint comes using the definition of w_i and direct substitution to obtain that $|\gamma| = \sum_{i \in =1}^{n} w_i$, which is then used to rewrite the budget constraint from (2.3). The final constraint ensures that our new variables are binary integers. □

2.3 Aggregate-Designed Efficient Local Explainer

While our main contribution in this paper is the local explainer aggregation methodology, we have additionally designed a new methodology for training local explain-

ers for effective use in aggregation. The key to our methodology is ensuring that local explainers only focus on the most relevant features in the particular region they are designed to explain. In contrast to previous methods that proposed the use of regularization to achieve this goal, we propose directly computing locally significant features using an information filter. Computing such filters are generally computationally expensive and requires the use of numerical integration; however, we introduce an efficient algorithm for filtering out less significant features. This methodology allows us to train local explainers that are significantly less complex than those that use regularization, with better fidelity for their specified region. In this section we present an overview of our methodology and highlight key results. Further details on the technical specifics of this methodology are deferred to the appendix.

Local Explainer Overview and Training Procedure

Our local explainer training methodology is formally presented in Algorithm 1. We give a brief overview of its operations here, and defer full details to Appendix A.4. The algorithm takes in hyper-parameters including number of points N to be sampled for training the explainer, a distance metric d, and a radius r around the point $\bar{x}$ being explained. First the algorithm samples N points uniformly from within a r radius of $\bar{x}$; we call this set of points $T(\bar{x})$. Depending on the distance metric being used this can often be done quite efficiently, especially if the features are binary valued or an ℓ^p metric is used (Barthe et al. 2005). Then using the sampled points, the algorithm uses the Fast Forward Feature Selection (FFFS) algorithm as a subroutine (discussed and formally presented in Appendices A.7), which uses a mutual-information-based information filter to remove unnecessary features and reduce the complexity of the explainer model. The FFFS algorithm uses an estimate of the joint empirical distribution of $(T(\bar{x}), f(T(\bar{x}))$ to select the most important features for explaining the model's predictions in the given neighborhood using tree traversal. We denote this set of features $\hat{\Phi}$. Then, using these features and the selected points, the local explainer model g is trained by minimizing an appropriate

loss function that attempts to match its predictions to those of the black box model. In principle, a regularization term can be added to the training loss of explainer g. However, in our empirical experiments (presented in Appendix A.5), we found that FFFS typically selected at most five features, so even the unregularized models where not overly complex.

Algorithm 1 Local Explainer Training Algorithm

Require: sampling radius r, number of sample points N, black box model f, data point to be explained $\bar{x}$, and loss function L for the explainer model $(\bar{x}, \bar{y})$

Initialize $T(\bar{x}) = \emptyset$
for $j = \{1, ..., N\}$ **do**
    Sample $x \sim U(\mathcal{B}(\bar{x}, r, d))$
    $T(\bar{x}) \leftarrow T(\bar{x}) \cup x$
end for
Obtain $\hat{\Phi}(\bar{x}) = \text{FFFS}(T(\bar{x}), \Phi, f)$
Train $g = _{\hat{g} \in \mathcal{G}} \{\sum_{x \in T(\bar{x})} L(f(x) - \hat{g}(x[\hat{\Phi}]))\}$
return g

2.4 Experimental Results

In this section we compare the performance of our IP method against five state-of-the-art explainer methods. We consider two local explainer aggregation methods—Submodular Pick and Anchor Points (Ribeiro et al. 2016, 2018)—and three global explainer methodologies—interpretable decision sets (Lakkaraju et al. 2016), active learning decision trees (Bastani et al. 2018), and naive decision tree global explainers (Friedman et al. 2001).

We compare these methods in both coverage and fidelity across two different datasets. These datasets are the Parkinson's Progression Marker Initiative (PPMI) data set, where we generate explainers for a black box model aimed at predicting Parkinson's Disease (PD) progression subtypes, and a Geriatric activity data set (Torres et al. 2013) where we generate explainers for a model that classifies the movement activities of geriatric patients based on wearable sensor data. We split

each dataset, using 80% for training and 20% as a holdout test set, and we apply 10-fold cross validation. One important feature of both these datasets is that they enable multi-class classification. Our experimental results show that our proposed optimization framework is better suited to these multi-class settings than existing state-of-the-art methods.

In addition to measuring the performance of our local aggregation methodology on different data sets and classification tasks, we also compare the performance of our information-filter-based decision-tree local explainer and LIME (Ribeiro et al. 2016) in the aggregation framework. We also measure performance for each of the aggregation-based methods under varying budgets of component local explainers. This budget is an informal measure of simplicity and interpretability, where aggregating fewer local explainers leads to a more interpretable aggregate explainer, but may sacrifice fidelity and/or coverage. Our results show that our methodology outperforms existing techniques in terms of fidelity and coverage, especially in the multi-class case.

PD Progression Cluster Classification

For our first set of experiments we used the PPMI data set to classify the disease progression of different patients into several subtypes based on screening measures. The PPMI study was a long run observational clinical study designed to verify progression markers for PD. To achieve this aim, the study collected data from multiple sites and includes lab test data, imaging data, and genetic data, among other potentially relevant features for tracking PD progression. The study includes measurements of all these features for the participants across 8 years at regularly scheduled follow up appointments. The complete data set contains information on 779 patients, and included 548 patients diagnosed with PD or some other kind of Parkinsonism and 231 healthy individuals as a control group. For our analysis we will focus on the first seven visits of this study which correspond to a span of approximately 21 study months, since these visits were conducted relatively close together temporally.

The classification task considered was the disease progression of the patients, and we performed a cluster analysis to generate labels, detailed in Appendix A.3. Our analysis identified four different subtypes of disease progression, corresponding to different trajectories of the diagnostic measurements' evolution over time. We also included one additional subtype corresponding to patients who did not have PD. Appendix A.3 presents a full description of these subtypes and their identification in the data.

As our black box model, we trained a random forest model to predict the progression subtype of a patient based on measurements taken during the baseline appointment and follow ups. We considered two different prediction tasks: (1) a binary prediction task to predict whether or not an individual has PD; (2) a multi-class prediction task to predict one of the five identified PD progression subtypes. Further details on the construction of the black box model and its performance on these tasks are given Appendix A.5.

We used each of the explainer methods presented above to explain the predictions made by these random forest models, and measured coverage and fidelity of these explainers. Coverage and fidelity for the binary prediction task are shown in Figure 2.1, and similar plots for the multi-class prediction task are shown in Figure 2.2.

Figures 2.1 and 2.2 show that for both prediction tasks, our optimization-based aggregation algorithm obtains a higher level of coverage than both Anchor points (Ribeiro et al. 2018) and Submodular Pick methods (Ribeiro et al. 2016) across all possible local explainer budgets. Note that when comparing coverage, global explainers are constrained to always achieve 100% coverage.

In terms of fidelity, Figure 2.1 shows that across fidelity lower bounds of 0.7 and 0.5, our methodology performs comparably with the other aggregate explainer methods and with the explainable decision set method. When increasing our fidelity lower bound to 0.9, our method significantly outperforms these methods. This shows that the fidelity lower bound parameter φ in our framework allows for higher fidelity explainers given proper tuning.

In the binary case our methodology does not outperform active learning and

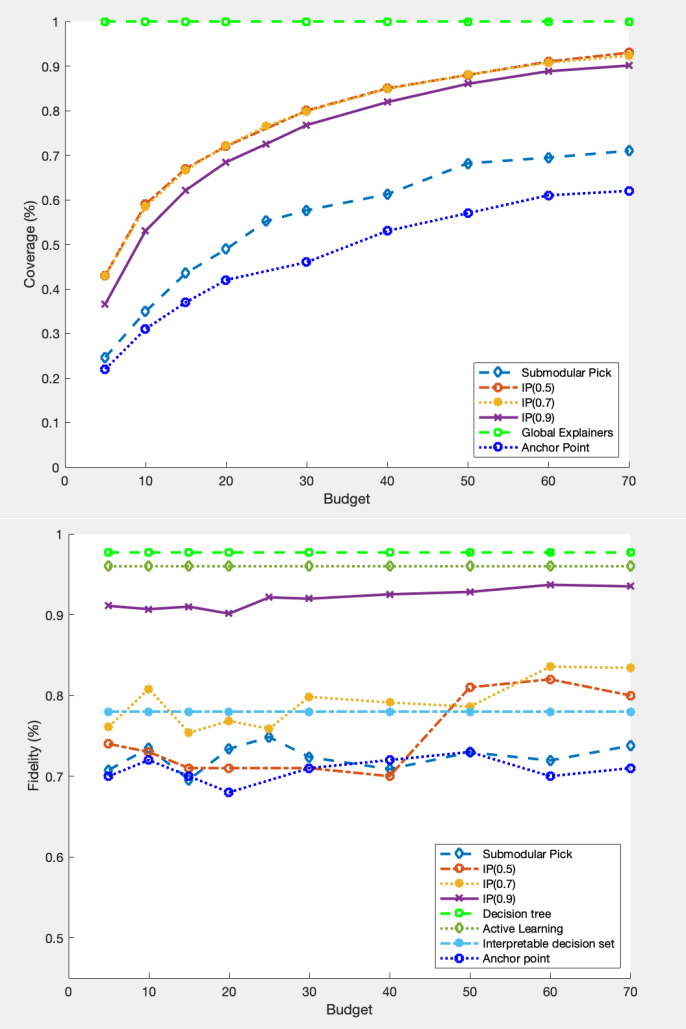


Figure 2.1: 2-class fidelity (bottom) and coverage (top) plots for various global explainers for a random forest model trained on the PPMI data set. The x-axis corresponds to the number of constituent local explainers that are used by the aggregation methods.

naive decision tree in terms of fidelity or coverage; however, when considering the multi-class setting of Figure 2.2, we see that our framework allows for significantly higher fidelity explanations. In particular, while active learning and naive decision trees achieve a fidelity of approximately 0.7 our optimization based global classifier with $\varphi = 0.9$ can achieve a fidelity of 0.9 in this case. While this is a significant increase, it does come with a cost for the coverage, as the explainer with this high

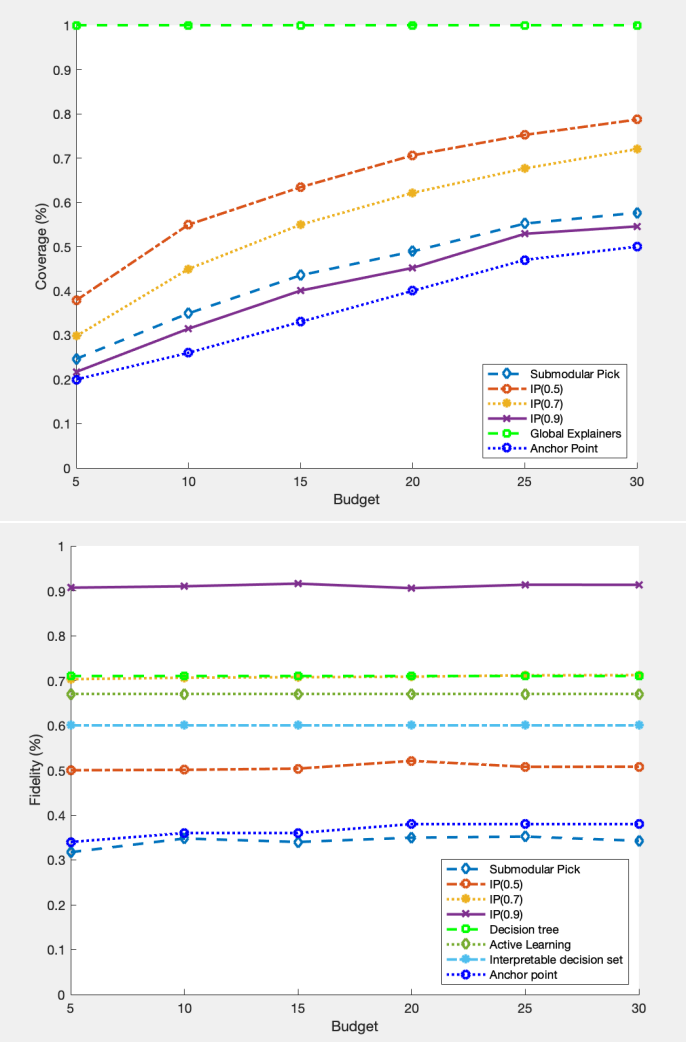


Figure 2.2: 5-class fidelity (bottom) and coverage (top) plots for various global explainers for a random forest model trained on the PPMI data set. The x-axis corresponds to the number of constituent local explainers that are used by the aggregation methods.

fidelity only covers 40–50% of the data, as compared to the global explainer methods of active learning and naive decision tree which cover 100% of the data.

Our methodology allows for greater flexibility in terms of trading off explainer coverage and fidelity, especially in this multi-class case. In contrast, the pure global explainer methods do not allow for this trade-off by ensuring a hard constraint of

100% coverage, which results in low fidelity explainers. Since our methodology outperforms existing aggregation methods, this indicates that using IP allows us to navigate the fidelity and coverage tradeoff more efficiently.

Empirical evaluation of our local explainer's performance compared with other local explainer methods, when used in the aggregate explainer are given in Appendix A.6. We find that our local explainer methodology outperforms LIME in both fidelity and coverage.

Figure 2.3 shows the Pareto frontier of the tradeoff between coverage and fidelity for the binary class prediction task. One advantage of our approach is that we allow a tunable tradeoff between the coverage and fidelity—corresponding to the three curves in the figure—while the other methods do not provide this option—corresponding to only a single point for the other methods. We see that our approach yields higher fidelity and higher coverage than most of the other local explainers, although there is less of a clear advantage of our proposed method compared to the global explainers. Figure 2.3 shows the tradeoff between the coverage and fidelity for the multi-class setting. In this setting we again see that our proposed local explainer provides better coverage and higher fidelity than other local explainers. In addition, our method also provides significantly higher fidelity than the global explainers at the expense of coverage.

Geriatric Activity Classification

For the second set of experiments we used a data set of Geriatric Activity based on the study conducted by (Torres et al. 2013). The main goal of this study was to provide ways of potentially reducing the likelihood of falls for geriatric individuals by classifying their activities when transferring beds. Generally, the highest risk for geriatric patients to fall is when getting out of bed so various sensors were deployed to detect whether an individual was attempting to leave their bed and detect other potentially risky activity. For this particular study, the authors used a novel wearable and environmental sensor which they validated with 14 individuals aged 66–86. The goal was to use this sensor data to classify between three different

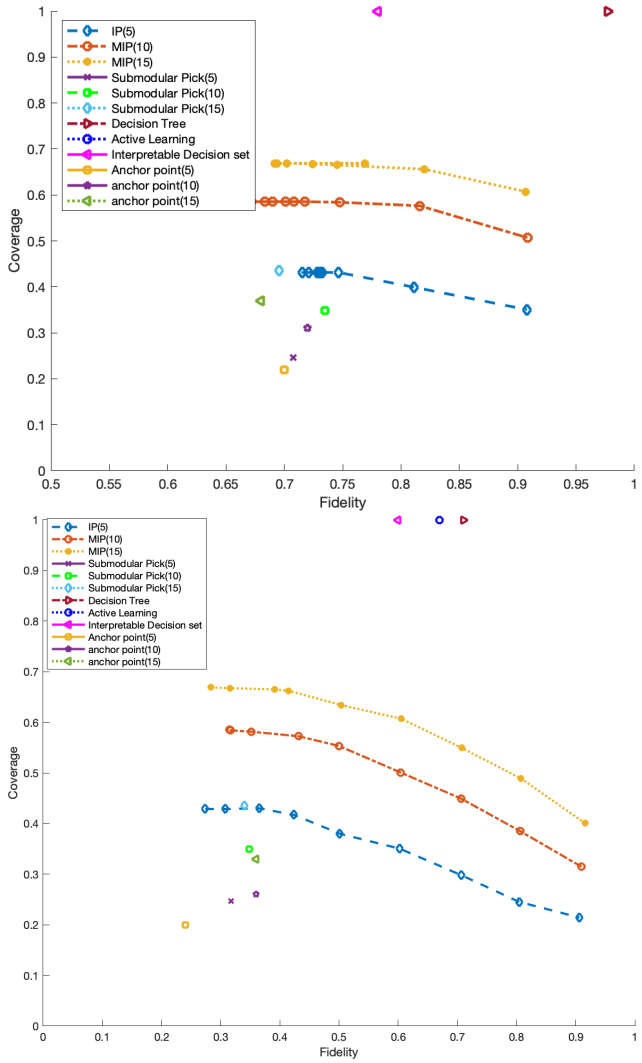


Figure 2.3: Pareto frontier for the tradeoff between fidelity and coverage on PPMI for binary (top) and multiclass (bottom, 5 class) classification task. The x-axis corresponds to fidelity and the y-axis corresponds to coverage.

activities, namely laying in bed, sitting in the bed, and getting out of the bed. To generate the data set, each of the participants was asked to perform a random set of five activities which ranged between the three potential activity classes.

Much like in the case of the PPMI data set, we trained a random forest model to classify between the various activity classes that we used to extract global explainers.

However, unlike the PPMI experiments, since there was no straight forward way to convert the multiclass classification task of detecting the different activities into a binary classification task we only performed the experiments for the multiclass case. The results for all explainer methods can be seen in Figure 2.5. Much like in the case for the PPMI data set, we note our methodology outperforms other aggregation based global explainers with respect to coverage across all budgets and fidelity lower bounds; however, it is still not obtaining 100% coverage like the pure global explainer methodologies. In terms of fidelity, much like in the multiclass case of the PPMI data, our methodology outperforms all other global explainers, with active learning being close to on par with our performance. This further suggests that using this form of optimization based local explainer aggregation is well suited to explaining multiclass predictions regardless of the underlying data set. Figure 2.4 shows the Pareto frontier of the tradeoff between coverage and fidelity across different explainers. Our methodology outperforms all local explainers in both metrics and all global explainers in terms of fidelity at the cost of decreased coverage.

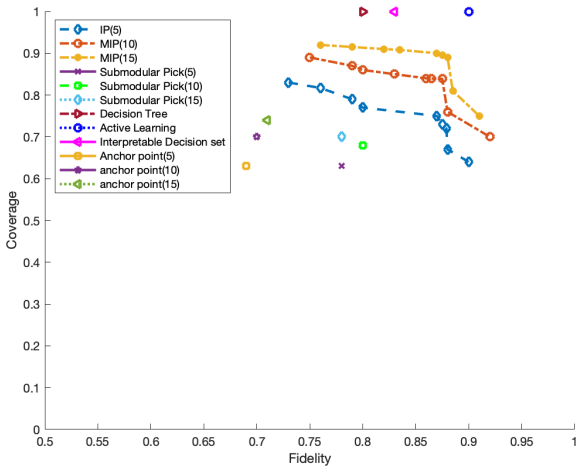


Figure 2.4: Pareto frontier for the tradeoff between fidelity and coverage on Geriatric Activity Dataset. The x-axis corresponds to fidelity and the y-axis corresponds to coverage.

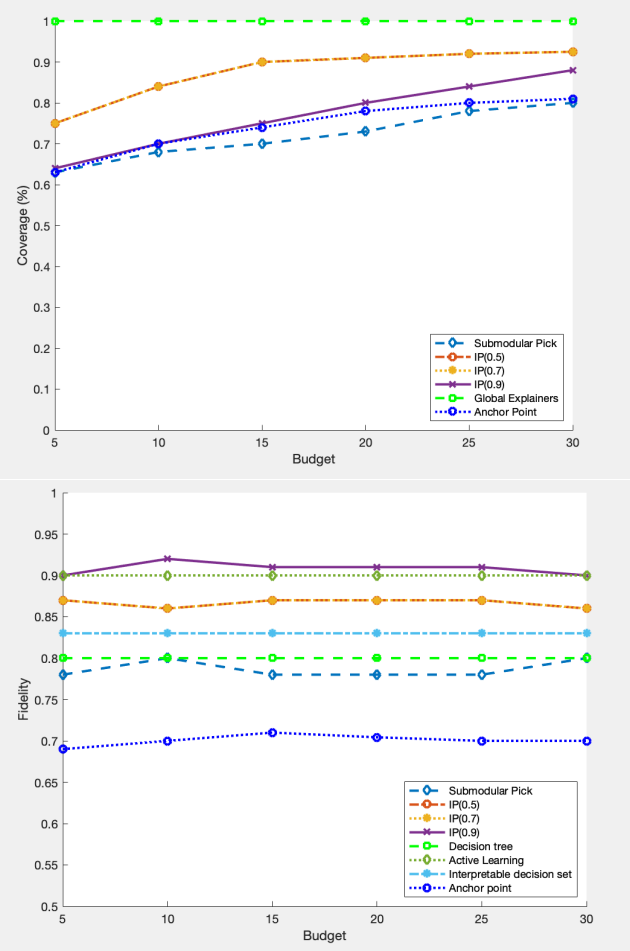


Figure 2.5: 5-class fidelity (bottom) and coverage (top) plots for various global explainers for a random forest model trained on the Geriatric Activity Dataset. The x-axis corresponds to the number of constituent local explainers used by the aggregation methods.

3 AN ADAPTIVE OPTIMIZATION APPROACH TO PERSONALIZED FINANCIAL INCENTIVES IN MOBILE BEHAVIORAL WEIGHT LOSS INTERVENTIONS

3.1 Introduction

The obesity epidemic is one of the most critical health issues facing the United States. According to the adult obesity data in 2017-2020 from the Center for Diseases and Prevention (CDC), the prevalence of obesity is 41.9% (Stierman et al. 2021). Obesity increases the risk of metabolic diseases such as type 2 diabetes and heart disease (Golay and Ybarra 2005) and has led to related medical costs of $173 billion in the United States in 2019 (Ward et al. 2021). If an individual with obesity is able to achieve a moderate reduction in weight (by 5%), they can mitigate many of these adverse effects (Wing et al. 1987, Krentz et al. 2016). Currently, the lowest risk treatments that have been found to be effective for treating obesity involve clinically monitored behavioral interventions (Grilo et al. 2011, Jakicic et al. 2016). Given advances in technology, recent generations of these interventions include a mobile application as a component in which individuals are asked to record their daily weight, exercise, and or daily calories consumed (Fukuoka et al. 2018). These applications can be used to provide participants with feedback and rewards to encourage behavioral change and weight loss. One key challenge in these interventions is that participant adherence decreases over time (Acharya et al. 2009, Lemstra et al. 2016).

Several studies have shown that financial incentives for weight loss could improve adherence and lead to clinically significant weight loss of at least 5% of baseline weight (John et al. 2011, Volpp et al. 2008). The primary objective of the intervention in these studies is to maximize the number of participants who achieve clinically significant weight loss at the end of the study (Wing et al. 1987). To achieve this goal, the interventionist can dispense financial incentives to each participant to encourage weight loss and calorie recording. Previous studies have

compared different reinforcement schedules, amounts, and targets in an attempt to determine the optimal structure on average of a financial incentives intervention. In these previous studies, incentives have followed a predetermined treatment schedule that does not adapt to participant data collected over the course of the intervention (Tsai and Liao 2020). In other words, all participants can receive the same amount of money for achieving the same criteria (e.g., weight loss, calorie recording). One key challenge for the interventionist in this setting is that each individual participant will have different levels of motivation stemming from financial incentives and internal desire to lose weight, leading to heterogeneity of response to financial incentives. Moreover, these individual motivations are unknown to the interventionist *a priori*, and must be inferred from participant data. A second challenge is that, to ensure the total intervention costs are manageable, the interventionist can only disburse incentives from some maximum total intervention budget. Typically, this budget is distributed evenly across all participants, such that each participant can earn a maximum amount. Accordingly, when scaling up the intervention to more participants, the cost increases linearly. A key operational challenge that must be addressed is how to design a modeling and optimization framework that can allow the interventionists to disburse costly incentives to match individual motivations and encourage the largest number of participants to lose weight.

In this paper, we propose a novel optimization framework that addresses these key challenges. While existing work in the operations literature has modeled how individuals respond to indirect motivational goals such as exercise goals and messaging (Aswani et al. 2019, Mintz et al. 2020), in this paper we focus on modeling how participants respond to direct monetary incentives for weight loss. Our proposed framework involves first providing a behavioral model for participants in the context of weekly financial incentives. This model is meant to capture the dynamics of how participant motivational states (i.e., intrinsic and extrinsic motivations) change over time and how they impact their choices with regards to calorie consumption and recording their physical state (i.e., their weight). We then propose a surrogate likelihood approach for estimating these unknown

participant state parameters and provide an approach to use these estimates to predict future participant response to potential interventions. The last step of our framework involves using this predictive model to optimize the incentives awarded to each participant. We provide an adaptive algorithm that can calculate an asymptotically optimal incentive policy while staying with in the financial resource constraints of the intervention. Our adaptive approach can be calculated weekly over the course of a weight loss intervention to improve its estimation using data obtained from participants currently participating in the trial, and compute new incentives that adapt to changing study conditions or participant response.

Clinical setting: the Log2Lose study

We developed our modeling and optimization methods using the data and structure from a study that investigated the impact of different financial incentive structures on weight loss called Log2Lose (Voils et al. 2018). The 24-week Log2Lose pilot trial was designed to evaluate the feasibility of delivering incentives in near real-time using data collected from cellular-connected scales and a mobile food and activity tracking. The goal of Log2Lose is to compare the efficacy of incentives for two different outcomes, either individually or combined: weekly calorie recording or weekly weight loss. Accordingly, participants were randomized to one of four arms: incentives for both calorie recording and weight loss (Arm A); incentives for calorie recording (Arm B); incentives for weight loss (Arm C); or no incentives (Arm D). The incentive schedule was based on psychological learning theory and involved the following principles: 1) It was fixed at $10 for the first four weeks to encourage learning of new behaviors, and 2) It varied between $0 and $30 per week thereafter. Thus, even if participants had the desired outcome, they did not receive a reward some weeks. The predetermined incentive scheduled applied to all participants. It was not known by participants *a priori* and thus appeared random.

All participants were invited to a biweekly group counseling session led by a registered dietitian that involved dietary education and behavioral skills such as regular self-weighing and calorie recording. Participants were encouraged to record

a minimum amount of calories (1,000 KCal for females and 1,200 KCal for males) in a mobile application for at least five days a week, with at least one of the days being a weekend day. Daily calories recorded were transmitted to the research team through an open application programming interface. If participants in Arm B met this goal, they were awarded a monetary payment between $0-$30. Additionally, all participants were given a cellular scale that transmitted their weight measures to the investigators whenever they weighed themselves. They were encouraged to weigh themselves at least two times each week. The difference between the first and last weight of the week was taken. Participants in Arm C received a payment between $0-$30 each week that the last weight was lower than the first weight. Arm A combined both weight loss and caloric recording incentives but reduced the reward range to $0-$15 for each to ensure the maximum amount that a participant could earn was $30 a week. The control arm did not receive any financial incentives. Recorded calories and weight were compiled and analyzed in a software application, and notice of incentives was provided using text messaging. For analysis in this paper, we used the cellular weight data, app data on calorie recording, the record of awarded incentives, as well as participant demographics for the purposes of validating our models and conducting our numerical studies. For the full demographic data and trial protocols please see (Voils et al. 2021, 2018). We note that, while our model is based on the structure of this particular intervention, we believe the approaches and techniques we develop will be widely applicable to other behavioral interventions developed in the future that may have direct financial incentive components.

Related literature

While our modeling is based in the clinical setting of behavioral interventions, through our modeling and optimization analysis we contribute to three streams of literature within the operations field. These include sequential decision making methods3.1, healthcare operations research 3.1, and predictive modeling for clinical weight loss 3.1.

Sequential decision making methods

Our setting of computing weekly individual level financial incentives for participants fits generally into the stream of sequential decision making methods with partial information. In particular we can think of our setting as that of a decision maker (the interventionist) taking sequential control actions (weekly incentives) with respect to a system state for which they have imperfect information (participant motivations and weight). Two of the main approaches for addressing the problem of making sequential decisions with imperfect information include partially observable Markov decision process (POMDP) (Yu and Bertsekas 2008, Ayer et al. 2012) and reinforcement learning (RL) (Sutton and Barto 2018). The key difference between these families of approaches is that, in the POMDP setting, the decision maker is assumed to have partial information of the system state while having information of the system dynamics; in contrast, in RL the decision maker is assumed to have full information of the system state while having partial information of the system dynamics.

The classical solution technique used for POMDP models involves reformulating the POMDP as what is known as a belief Markov Decision Process (MDP), by considering what is known as a belief state, a state that encodes the decision maker's belief they are in any of the POMDP states (Bertsekas 2012). In general, the belief state can be thought of as a distribution over the state space of the POMDP that reflects likelihood of a particular state being the true state of the system at any given point in time. However, solving the belief MDP in practice is quite challenging since even if the state space is finite, the belief state could be uncountably infinite. Therefore, in the POMDP literature different techniques such as approximate dynamic programming (ADP) (Yu and Bertsekas 2012, Dai and Shi 2019) and policy gradient (Zhang et al. 2021) have been used for approximating the optimal solutions. Our setting can be thought of as a particular instantiation of a POMDP, with specific model structure. Using our model structure, we develop an approximate solution method that is asymptotically optimal under a set of mild conditions.

Methods in the RL literature can be categorized into two broad families namely model-based RL (Zhou et al. 2018, Osband and Van Roy 2014), which use specific functional form (or parametric estimates) of the transition dynamics and value function, and model-free methods (Strehl et al. 2006, Akrour et al. 2016), which use stochastic approximations of the problem value functions and transition probabilities without explicit functional forms. In this paper, our proposed approach can be thought of as a form of model-based RL as we explicitly model system dynamics (e.g., dynamics of participant weight and motivations). Our modeling framework is related to existing model-based methods developed for behavioral weight loss interventions (Mintz et al. 2017, Zhou et al. 2018).

Healthcare operations research

Our setting is related to the large stream of existing work on applications of operations research to healthcare applications. In particular there has been a vast amount of work examining applications of sequential decision making in managing the operations of providing care (Ekici et al. 2014, Erdogan and Denton 2013, Childers et al. 2009), providing personalized treatment (Ayer et al. 2019, Bastani and Bayati 2020, Schell et al. 2016, Mintz et al. 2020, He and Mintz 2023), and intervention management (Deo et al. 2013, Lee et al. 2019). Our problem setting and methods contribute to these streams of literature, in particular to the work focused on personalized treatment and intervention management. Much like these settings, we consider a resource constrained problem, where decision makers must make costly decisions under uncertainty. One of the contributions of our work is in developing a framework that extends the existing work in these settings to behavioral interventions where a decision maker must disburse financial incentives to participants with imperfect information. In contrast to existing work that considers resource constraints on manpower or facilities, our work examines a constraint on the direct budget of the intervention and how it can be best disbursed amongst participants to motivate them to achieve weight loss.

Predictive models for weight loss

Our work is also related to a stream of literature that focuses on predicting an individual's weight loss success in the context of a clinically supervised intervention. Existing predictive models for this setting include differential equations (Thomas et al. 2011), Markov models (Bromberger et al. 2014), data mining methods (Batterham et al. 2017), and machine learning methods (Lee et al. 2020). In general, these methods were developed to perform a binary prediction task (i.e. whether or not a participant achieves clinically significant weight loss), making them challenging to use for optimization. In contrast, the behavioral framework we develop in this paper is capable of providing predictions for the full weight trajectory of study participants given a particular sequence of financial incentives. Our framework is also able to compute the likelihood of such a trajectory occurring and can thus also be used for binary prediction in addition to this regression task in a similar manner to (Aswani et al. 2019). However, our work differs from the predictive approach in (Aswani et al. 2019) in two key ways. First, we focus on a weight loss intervention with financial incentives instead of motivational goals, which are evaluated by participants in a slightly different manner and thus alter the model structure. In particular, the nature of the weekly financial incentives means participants value their actions over the course of several days (during the incentive evaluation period), unlike daily step goals that only impact participant behavior during a single day of the study. Second, our model incorporates both continuous and discrete measurements, making it challenging to use maximum likelihood estimation directly. We propose to solve this challenge using surrogate likelihood estimation, a more challenging method to analyze.

Contributions

In this paper, we develop a framework to design personalized financial incentives that encourage weight loss, while ensuring that intervention costs remain within a fixed budget. Through the development of this framework we make three key contributions:

1. We extend participant behavioral models in weight loss interventions to capture the effect of financial incentives on participant behavioral change. Our novel modeling additions include both medium and long term impacts of financial incentives, and capture how repeated use of financial incentives may not lead to meaningful long-run behavioral change. In particular, we are able to capture both short-term (in-week) participant decisions as well as long-term (between-weeks) participant behavioral change. Our model incorporates insights from self-determination theory (SDT), namely that it includes parameters both intrinsic and extrinsic motivation for weight loss. According to SDT, motivation ranges on a continuum from completely nonself-determined (lacking motivation) to self-determined (intrinsically motivated); in between are several levels of extrinsic motivation in which one's behavior can be completely or partially driven by external sources such as rewards and punishment (Deci and Ryan 2013).

2. We develop a novel inverse optimization approach for estimating unknown participant parameters and states that is statistically consistent. In contrast to existing literature which looked at inverse optimization for purely myopic participants (Mintz et al. 2017), our approach assumes participant's plan for the medium-term using dynamic programming, and uses the structure of the resulting optimal policy to construct a set of constraints for inverse optimization. We then use these constraints in a surrogate likelihood estimation model, which can be solved using commercial mixed integer programming solver. We further show the resulting estimates are statistically consistent, which, to our knowledge, is one of the first consistency guarantees shown for surrogate likelihood models trained with non-convex optimization. Furthermore we show how these estimated parameters can be used in an adaptive optimization framework to allocate incentives for weight loss given a budget that we call the Design of Incentives Algorithm (DIA). Through theoretical analysis, we show that the incentive policy output by DIA is asymptotically optimal.

3. We conduct a comprehensive set of numerical validation studies using real-world-data from the Log2Lose trial, which deployed financial incentives to help participants achieve clinically significant weight loss. Our experiments demonstrate that our proposed behavioral model is descriptive of participant behavior, and moreover is capable of better predictive performance than existing state-of-the-art machine learning approaches. Furthermore, through a simulation study we are able to show that our dynamic optimization framework is able to achieve improved clinical outcomes for less budget when compared to existing one-size-fits-all approaches, indicating that using our methods such interventions could be scaled to larger participant populations.

3.2 Participant behavioral model

Here, we present our model for participant behavior during a weight loss intervention. We use a utility maximization framework where participants are assumed to make weight loss-related decisions (namely how many calories to consume each day and whether or not to record their calories) based on individual utility functions that depend on their perceived health benefits, their responsiveness to financial incentives, and preferred level of caloric consumption. Our model consists of three key classes of variables we call physical system states, which are state variables that capture the physical aspects of weight loss (namely the participant's weight), motivational states that capture a participant's cognitive state and how much importance they place on different actions and health outcomes (i.e., intrinsic and extrinsic motivation for weight loss gained from financial incentives), and decision variables that represent a participant's actions that affect weight loss (i.e., daily caloric intake). A key feature of our model is that all physical and motivational states are modeled as individual specific, and thus will be different for each participant. Because of this, we focus our modeling discussion on modeling the behavior of a single participant.

To capture how participant behavior changes over time as a consequence of the intervention, we also define a set of dynamics that describe how the motivational

and physical states change over the course of the program as a consequence of the intervention treatment and individual participant decisions. Since financial incentives are administered to the participant based on their weight loss and calorie recording at the end of a study week our framework models the participant's decision-making process in two components: 1) A component that models the participant's daily actions over the course of a single week of a trial given their expectation for financial reward, we call this component the in-week decision model. 2) A component that models long term behavioral change by tracking how participant motivational states change as a consequence of the previous week's actions and financial incentives; we call this component the between-week decision model. Both of these time frames are fully integrated into a single participant model, which, as previously noted, is individual-specific and captures the unique way each participant will interact with the intervention. A key assumption to these models is that participants make decisions in a myopic utility-maximizing manner, that is, they only make decisions on calorie consumption during the course of a study week that will impact their financial incentive earned for that week and will not consider future incentives or long term health benefits. This behavior has been observed in the social science literature, and can be framed as participants making rational decisions with respect to high future discounting of health and monetary gain (Cawley 2004). Prior work has shown that models that incorporate this assumption still provide strong predictive and descriptive performance (Aswani et al. 2019, Mintz et al. 2017, Adams et al. 2023). We note that while existing myopic models consider participants that only consider single daily decisions, due to the structure of the financial incentives in our setting, the myopic assumption implies participants consider their decisions at the start of a week.

Participant in-week decision model

The first step of our framework is to describe the participant's daily decision making process during a single study week. Let t be the week index and $d \in \{0, ...6\}$ be the day index where each week starts on Monday ($d = 0$) and ends on Sunday ($d = 6$).

Let the physical system states $w_{t,d}, f_{t,d} \in \mathcal{W} \times \mathcal{F}$ be the participant's weight and caloric consumption on day d of week t, where $\mathcal{W}, \mathcal{F} \subset \mathbb{R}_+$ are closed intervals. Let the motivational states of the participant be given by $a_{1,t}, a_{2,t}, f_{b,t} \in \mathcal{A}^2 \times \mathcal{F}$ that represent the participant's internal motivation, that is a participant's personal motivation for weight loss, external motivation for weight loss, or how influential financial incentives are on the participant's motivation to lose weight, and the participant's preferred caloric consumption level on week t respectively. Here $\mathcal{A} \subset \mathbb{R}_+$ is assumed to be a closed interval. The participant's decisions in this model are denoted by $c_{t,d} \in \mathcal{F}$ that represent the participant's planned caloric intake on day d of week t. Note that unlike existing models that consider caloric consumption directly as a participant decision (Aswani et al. 2019, Mintz et al. 2017), a key feature of our model will be that, while participants are capable of planning to a particular value of caloric consumption, this may not equal the amount of calories they truly consume. This is a challenge in many calorie-recording based behavioral interventions since, even when trying to the best of their abilities, participants often cannot accurately estimate the amount of calories they consume with each meal (McKenzie et al. 2021). Furthermore, social desirability concerns may encourage under-reporting of caloric consumption. The final component of the in-week decision model is a motivational state that captures the participant's expectation for financial incentives at the end of the week. We denote the amount of financial incentive allocated by the interventionists for weight loss at the end of week t by $r_t^w \in \mathcal{R}$, where $\mathcal{R} \subset \mathbb{R}^+$ is a closed interval. However, since this amount is generated at the end of the week based on the participant's performance and the intervention is structured so that financial incentives seem randomly generated to the participant conditioned on meeting the goal, individuals cannot use the true value of the incentive for their decisions during week t. Instead participants form a belief on the financial reward they will potentially receive at the end of the week should they meet their weight loss goal based on their previous rewards received and knowledge of the intervention policies. We let $\hat{r}_t^w \in \mathcal{R}$ be a random variable that represents the participant's estimate of their potential financial reward for weight loss in week t that influences their decisions based on these beliefs.

Using these variables, we model the participant's in-week decision process for week t of the intervention as the following utility maximization problem,

$$\max_{\{c_{t,d}\}_{d=0}^{6}} \mathbb{E}\Big[-a_{1,t} \sum_{d=1}^{6} w_{t,d} + a_{2,t} \hat{r}_t^w \mathbb{1}\{w_{t,0} - w_{t,6} > 0\} - \sum_{d=0}^{6} (f_{t,d} - f_{b,t})^2 \Big] \quad (3.1a)$$

$$\text{subject to: } w_{t,d+1} = b w_{t,d} + c f_{t,d+1} + k, \quad d \in \{0, \cdots, 5\}, \quad (3.1b)$$

$$f_{t,d} = c_{t,d} + \xi_d, \quad d \in \{1, \cdots, 6\}, \quad (3.1c)$$

$$w_{t,d}, f_{t,d}, c_{t,d} \in \mathcal{W} \times \mathcal{F}^2, \quad d \in \{0, ..., 6\}. \quad (3.1d)$$

The interpretation of this model is that the participant's planned caloric intake at each day d of week t is given by the argmax of the above optimization problem where the objective given by (3.1a) represents the participant's utility function and (3.1b)-(3.1c) represent the dynamics of the participant's weight and caloric intake preferences. Note that (3.1a) contains three main components that impact the participant's decisions. The first term $-a_{1,t} \sum_{d=1}^{6} w_{t,d}$ indicates that the participant wants to reduce their future weight for each day of the week, and this is weighted by their motivation for weight loss $a_{1,t}$. The next term $a_{2,t} \hat{r}_t^w \mathbb{1}\{w_{t,0} - w_{t,6} > 0\}$ indicates participants would like to reduce their weight over the course of the week so that they can be eligible for the financial reward, and this is weighted by $a_{2,t}$ that expresses how motivated they are by financial rewards. The final term $-\sum_{d=0}^{6} (f_{t,d} - f_{b,t})^2$ indicates that participants want to choose foods with calories each day that are close to a certain caloric preference level $f_{b,t}$. This last component signifies that, without intervention, there exists some theoretical preferred amount participants would desire to eat that is not so little that they would feel hungry or so much that it would be physically impractical. Constraint (3.1b) represents the dynamics of weight loss using the Mifflin St. Jeor equation (Mifflin et al. 1990) where b, c are known constants and k is a constant computed from the participant's age, gender, and height. Constraint (3.1c) models that despite planning to consume $c_{t,d}$ participants may over or under eat since they cannot get an accurate estimate of their calories. This uncertainty is captured by i.i.d. disturbance variables $\xi_{t,d}$, that we assume are bounded such that $f_{t,d} \in \mathcal{F}$ with probability of one and $\mathbb{E}\xi_{t,d} = 0$.

Specifically we assume $\xi_{t,d} \sim U(-A, A)$, in other words that the deviation from the calorie plan is uniformly distributed within A calories. While other distributions could be used to model this uncertainty, we chose the uniform distribution for computational reasons to enable us to estimate the unknown model parameters using commercial mixed integer programming (MIP) solvers, this reformulation is described in detail in Section 3.3.

Participant between-week model

Next, we describe the model for how participant behavior evolves from week to week. While over the course of a single week participants do not respond directly to the financial incentives (since they are awarded at the end of the week) this model captures how weekly incentives change participant motivation over the course of the intervention. Therefore, unlike the in-week model, all components of this model describe the evolution of motivational states and not physical states or decisions.

Using the previous notation let $a_{1,t}, a_{2,t}$ describe the participant's internal motivation for weight loss and external motivation for weight loss from financial incentives on week t, and let $f_{b,t}$ represent the participant's preferred caloric intake on week t. Let g_t be an indicator variable that equals 1 when the participant successfully meets their calorie recording goal on week t. We model g_t as a Bernoulli random variable since different exogenous influences (such as participants not having time during the week or getting distracted) can influence whether or not they record calories (Raber et al. 2021). Let $p_t \in \mathcal{P} \subset (0, 1)$ represent the probability a participant will meet their calorie recording goal on week t (that is $p_t = \mathbb{E}g_t$).

We define the following set of dynamics that describes the transitions of motivational states $a_{1,t}, a_{2,t}, p_t, f_{b,t}, \hat{r}_t^w$:

$$a_{1,t+1} = \gamma_1(a_{1,t} - a_{1,b}) + a_{1,b} + k_1 \mathbb{1}\{(w_0 - w_6) > 0\} + r_t^c \mathbb{1}\{p_t - B \geqslant 0\}, \quad t \in \{0, \cdots, 23\}, \tag{3.2}$$

$$a_{2,t+1} = \gamma_2(a_{2,t} - a_{2,b}) + a_{2,b} + k_2 r_t^w \mathbb{1}\{(w_0 - w_6) > 0\}, \quad t \in \{0, \cdots, 23\}, \tag{3.3}$$

$$p_{t+1} = \gamma_p(p_t - p_b) + p_b + k_p g_t, \quad t \in \{0, \cdots, 23\}, \tag{3.4}$$

$$f_{b,t+1} = \gamma_f f_{b,t} + (1 - \gamma_f)\frac{1}{7}\sum_{d=0}^{6} f_{t,d}, \quad t \in \{0, \cdots, 23\}, \tag{3.5}$$

$$\hat{r}_{t+1}^w = \begin{cases} \frac{t}{t+1}\hat{r}_t^w + \frac{1}{t+1} r_t^w, & \text{if } w_0 - w_6 < 0, \\ \hat{r}_t^w, & \text{otherwise,} \end{cases} \quad t \in \{0, ..., 23\}. \tag{3.6}$$

The interpretation of these dynamics is that all motivational states have some baseline values that changed as participants interact with the intervention, but that, as time progresses, the impact of the intervention decays exponentially and the states tend to their baseline. Here, $a_{1,b}, a_{2,b}, p_{t,b}$ represent the baseline value of each motivational state, which can be interpreted as the motivational states of the participant without any interaction with the behavioral intervention, and $\gamma_1, \gamma_2, \gamma_p \in (0, 1)$ are the decay rates at which the states return to baseline. $k_1, k_2 \in \mathcal{K}$ represent the increase in motivational states when participants meet their weight loss goal and receive financial incentive respectively, where $\mathcal{K} \subset \mathbb{R}^+$ is a closed interval. $a_{1,w+1}$ increases by k_1 if the participant satisfies the weight loss requirements in previous week. This models that participants will be more motivated internally to meet the weight loss goal as they succeed initially in losing weight. Likewise $a_{2,t+1}$ increases by $k_1 r_t^w$ if the participant satisfies the weight loss requirements in week t and receives financial incentive r_t^w. This would indicate that if a constant positive reward

is given to the participant their motivation from financial incentives will increase rapidly. But in cases where $r_t^w = 0$ and the participant still manages to lose weight, only $a_{1,t}$ will increase while $a_{2,t}$ will return to baseline. This interaction in the dynamics ensures that, in order to impact long-term behavioral change and reduce dependence on incentives, effective policies should at some points provide zero reward even if a participant is likely to lose weight. This notion is well known in the behavioral literature and can be thought of as encoding the principle of intermittent reinforcement (Ferster and Skinner 1957). $B \in \mathcal{P}$ can be interpreted as the baseline probability of a participant satisfying the calorie recording requirements, and $a_{1,t+1}$ only increases when $p_t > B$, that is if the participant is motivated enough to record their calories that this would also reflect on their motivation to lose weight. k_p can be thought of as a parameter encoding the intrinsic motivation of the participant from calorie recording since p_{t+1} increases by k_p if the participant satisfies the calorie recording requirements in week t. Moreover r_t^c represents the amount of financial incentive awarded for meeting the calorie recording goal, and its inclusion in (3.2) signifies that if participants are rewarded for calorie recording this will increase their motivation for weight loss in the coming week.

There are two exceptions to these dynamics descriptions. The first is (3.5), which describes the long-term behavioral change of baseline caloric preference. Essentially, this equations states that future caloric consumption preference can be thought of as a geometric average of the previous caloric preference and the average caloric consumption in the previous week. Thus as participants modify their behavior and have lower weekly consumption this will result in a slow but long term change in the baseline caloric consumption preference of the participant. The second is (3.6), which indicates that participants estimate their expected reward as an arithmetic average of past rewards received so long as they've met the weight loss goal. In other words, if they did not meet the goal (and thus expected to receive zero reward) this belief does not update; however, if they do meet the goal but receive zero reward their reward belief decreases. This means that although providing a reward of zero would be beneficial for long run behavioral change, it could lead to a decrease in weight loss motivation in the short term presenting an

important trade-off to the decision maker.

3.3 Estimation and Prediction of Unknown Parameters

While the model described in Section 3.2 is able to capture mathematically the decision making process of participants and their interaction with the intervention, in practice most of the parameters in this model are not known *a priori* to the interventionist. In order to provide effective incentives to individuals so that they can lose weight, the interventionists must be able to estimate these individual level participant parameters using data collected through the intervention. This data comes in two main forms, observations of whether or not participants managed to meet their calorie recording goals on week t (g_t in the notation from Section 3.2) and noisy observations of weight at each day of the intervention that we call $\tilde{w}_{t,d}$. This estimation problem poses two main challenges, namely that the data is noisy, and there could be a significant amount of missing data. This second challenge is of particular interest to the Log2Lose case since in this intervention all weight is self generated through participants using a cellular scale. Depending on various factors (such as how busy their day was or if they are traveling) they may not weigh themselves every day throughout the intervention.

To address these challenges, we use a joint parameter estimation approach that formulates the estimation problem as a mixed integer program (MIP). Specifically, we consider an approach similar to (Aswani et al. 2019) by using a joint maximum likelihood estimation (MLE) approach. We assume that $\tilde{w}_{t,d} = w_{t,d} + \epsilon_{t,d}$, where $\epsilon_{t,d}$ are i.i.d. noise terms such that $\mathbb{E}\epsilon_{t,d} = 0$ and $\mathbb{E}\epsilon_{t,d}^2 = 0$. For our specific formulation we assume that $\epsilon_{t,d} \sim \text{Laplace}(0, \sigma)$, but note that our analysis could apply to all noise distributions that can be represented using a set of mixed integer linear constraints such as piece-wise linear distributions or the shifted exponential distribution.This is formalized by the following proposition.

Proposition 4. *The MLE problem can be formulated as the following constrained opti-*

mization problem:

$$\max_{\{w_{t,d},p_t,B,a_{1,t},a_{2,t},f_{b,t},f_{t,d},c_{t,d},\hat{r}_t^w\}_{t\in\mathcal{T},d\in\{0,\dots,6\}}} \sum_{t\in\mathcal{T},d\in\mathcal{D}_t} \log\mathbb{P}(\tilde{w}_{t,d}|w_{t,d}) + \sum_{t\in\mathcal{T}} \log\mathbb{P}(g_t|p_t), \tag{3.7a}$$

$$\text{subject to: } (3.1b), (3.1c), (3.2)-(3.6), \quad t\in\mathcal{T}, d\in\{0,\cdots,6\}, \tag{3.7b}$$

$$\{c_{t,d}\}_{d=0}^{6} \in \mathcal{C}(a_{1,t}, a_{2,t}, w_{t,0}, f_{b,t}, \hat{r}_t^w), \quad t\in\mathcal{T}, \tag{3.7c}$$

$$w_{t,d}\in\mathcal{W}, f_{t,d}, c_{t,d}\in\mathcal{F}, \quad t\in\mathcal{T}, d\in\{0,\dots,6\}, \tag{3.7d}$$

$$p_t, B\in\mathcal{P}, a_{1,t}, a_{2,t}\in\mathcal{A}, f_{b,t}\in\mathcal{F}, \hat{r}_t^w\in\mathcal{R}, \quad t\in\mathcal{T}. \tag{3.7e}$$

Full details of the formulation can be found in the appendix. Here $\mathcal{T}$ is the index set of all weeks in the study and $\mathcal{D}_t$ is the set of days during week $t\in\mathcal{T}$ that have weight observations. Note that $\mathcal{C}(a_1, a_2, w_{t,0}, f_{b,t}, \hat{r}_t^w)$ is the argmax set of (3.1), (i.e. is the set of decisions taken by the participant in the in-week maximization model). (3.1b) and (3.1c) are first introduced in the formulation of the participant in-week decision problem and they describe the daily transitions of variables like weight ($w_{t,d}$) and caloric intake variables ($f_{t,d}, f_{b,t}$). (3.2)- (3.5) correspond to the dynamics of states $a_{1,t}, a_{2,t}, p_t$, and $f_{b,t}$ between consecutive weeks.

Note that this formulation is non-linear and cannot be readily implemented using commercial solvers. This is due to non-linearity not only in the constraints but also in the objective function. While by assumption $\log\mathbb{P}(\tilde{w}_{t,d}|w_{t,d})$ can be expressed using linear constraints for any t, d, this is clearly not the case for $\log\mathbb{P}(g_t|p_t)$. Note that since $g_t \sim \text{Bernoulli}(p_t)$ we can express the p.d.f. of g_t as $\mathbb{P}(g_t|p_t) = p_t^{g_t}(1-p_t)^{1-g_t}$. So taking the log yields $\log\mathbb{P}(g_t|p_t) = g_t\log p_t + (1-g_t)\log(1-p_t)$, which is clearly non-linear and not readily expressible with mixed integer linear constraints. One approach for resolving this challenge is to use

a full descritization of the objective or use a piece-wise linear approximation of the natural log function (Wolsey and Nemhauser 1999). However, such approximations could be quite loose and have unfavorable statistical proprieties. Instead, we propose using a surrogate likelihood function (Bartlett et al. 2006, Nguyen et al. 2009, Goh and Rudin 2018, Awasthi et al. 2022), that is a function that can be more easily deployed in commercial solvers that will produce estimators with strong statistical properties such as consistency. In particular, we choose absolute error as our surrogate for this component of the likelihood function that is $\log \mathbb{P}(g_t|p_t) \approx |p_t - g_t|$. We will refer to the new minimization problem that only differs from (3.7) with the substitution of $\log \mathbb{P}(g_t|p_t)$ by the surrogate function as H_{SMLE}. Further we will refer to a specific problem instance with observations $\{\tilde{w}_{t,d}, g_t\}_{t\in\mathcal{T},d\in\mathcal{D}_t}$ and administered incentives $\{r_t^w, r_t^c\}_{t\in\mathcal{T}}$ as $H_{SMLE}(\{\tilde{w}_{t,d}, g_t, r_t^w, r_t^c\}_{t\in\mathcal{T},d\in\mathcal{D}_t})$.

While using the surrogate function allows us to linearize the objective there are still two key formulation challenges in the constraints that need to be addressed so that H_{SMLE} can be solved with commercial optimization software. First, H_{SMLE} is a bi-level optimization problem (Colson et al. 2007, Keshavarz et al. 2011, Bertsimas et al. 2015, Aswani et al. 2018), that is one of its constraints requires that variables be in the argmax set of a different optimization problem (usually referred to as the lower level problem). In this case, this lower level problem is a sequential decision making problem which means we will need to characterize the argmax set of an optimal policy. Second, we have nonlinear dynamics with bi-linear terms that need to be reformulated into a proper linearized form.

Characterizing participant decisions from the in-week model

To reformulate the bi-linear constraints, we will take a direct approach by showing that $\mathcal{C}(a_{1,t}, a_{2,t}, w_{t,0}, f_{b,t}, \hat{r}_t^w)$ can be characterized using a set of linear equations. To do this we will obtain a closed form solution of $c_{t,d}$ from the in-week decision model using dynamic programming. This solution is expressed in the following proposition.

Proposition 5. *The optimal solution* $\{c^*_{t,j}\}_{j=0}^{6}$ *of the in-week decision problem is* $c^*_{t,j} = f_{b,t} - \frac{a_{1,t} c \sum_{i=0}^{6-j} b^i}{2} - \frac{a_{2,t} \hat{r}^w_t c b^{6-j}}{4A}$ *for all* $j \in \{0, ...6\}$.

The complete proof can be found in Appendix B.1 here we provide a sketch. First, notice that once we take the expected value of (3.1a), then (3.1) consists of a quadratic objective function and a set of linear constraints, meaning that this problem should have a similar optimal policy structure to a linear quadratic regulator (LQR) (Bertsekas 2012). This implies the value function is quadratic, and there exists a unique optimal solution for $c_{t,d}$ that can be found using backward induction, and that this solution should be a linear function of the system states. Intuitively, this optimal solution of $c_{t,d}$ matches our expectations of how financial incentives and internal motivations impact participant weight loss. To interpret these, we can consider the solution in two parts: the internal motivation component $f_{b,t} - \frac{a_{1,t} c \sum_{i=0}^{6-j} b^i}{2}$ and external motivation component $-\frac{a_{2,t} \hat{r}^w_t c b^{6-j}}{4A}$. The internal motivation component implies caloric intake will decrease if the internal motivation $a_{1,t}$ increases, but will otherwise be close to the participant's preferred baseline if their internal motivation is low. The external motivation component shows that, so long as participants motivation from external compensation $a_{2,t}$ increases, so will their motivation for weight loss. Furthermore, as their expectation for future monetary incentives for weight loss increases, so too will they prefer to decrease their calories since they expect to receive a higher payoff. Interestingly, this component also shows that if participants have more uncertainty about their caloric consumption, that is A increases, the effects of the financial incentives decrease. This makes intuitive sense since the more uncertainty there is in true caloric consumption, the less control participants will be able to exert on their own behavior and so receiving the reward at the end of the week becomes less certain and less motivating.

Reformulation of bi-linear constraints

Next, we reformulate the bi-linear terms in (3.2)-(3.4) using a set of mixed integer linear constraints and variables. First, we define two sets of binary variables $l_{1,t}, l_{2,t}$ and three sets of continuous variables $z_{1,t}, z_{2,t}, z_{3,t}$. Let $l_{1,t} \in \mathbb{B}$ be equal to 1 if the

participant loses weight in week t, and let $l_{2,t} \in \mathbb{B}$ be equal to 1 if the probability of the participant satisfying calorie recording requirements is greater than B. Let $z_{1,t}$ be equal to $r_t^c \mathbb{1}\{p_t - B \geqslant 0\}$, $z_{2,t}$ be equal to $k_2 \mathbb{1}\{(w_{\underline{d}} - w_{\bar{d}}) > 0$, and $z_{3,t}$ be equal to $k_1 \mathbb{1}\{(w_{t,0} - w_{t,6}) > 0\}$. Using these quantities we will first consider the dynamics of $a_{1,t}$ from (3.2).

Proposition 6. (3.2) *can be expressed with the following set of integer variables and constraints:*

$$w_{t,0} - w_{t,6} \leqslant M_{1,t}(1 - l_{1,t}), \qquad t \in \{0, \cdots, 23\}, \tag{3.8}$$

$$p_t - B \leqslant M_{2,t} l_{2,t}, \qquad t \in \{0, \cdots, 23\}, \tag{3.9}$$

$$z_{1,t} \leqslant M_{z1} l_{2,t}, \qquad t \in \{0, \cdots, 23\}, \tag{3.10}$$

$$z_{1,t} \leqslant r_t^c, \qquad t \in \{0, \cdots, 23\}, \tag{3.11}$$

$$z_{1,t} \geqslant r_t^c - M_{z1}(1 - l_{2,t}), \qquad t \in \{0, \cdots, 23\}, \tag{3.12}$$

$$z_{3,t} \leqslant M_{z3} l_{1,t}, \qquad t \in \{0, \cdots, 23\}, \tag{3.13}$$

$$z_{3,t} \leqslant k_1, \qquad t \in \{0, \cdots, 23\}, \tag{3.14}$$

$$z_{3,t} \geqslant k_1 - M_{z3}(1 - l_{1,t}), \qquad t \in \{0, \cdots, 23\}, \tag{3.15}$$

$$z_{1,t}, z_{3,t} \geqslant 0, \qquad t \in \{0, \cdots, 23\}, \tag{3.16}$$

$$a_{1,t+1} = \gamma_1(a_{1,t} - a_{1,b}) + a_{1,b} + z_{1,t} + z_{3,t}, \qquad t \in \{0, \cdots, 23\}. \tag{3.17}$$

This reformulation can be done using big-M techniques for products of binary and continuous variables, as well as disjunctive constraints(Wolsey and Nemhauser 1999), the full details can be found in the appendix. Next we show that a similar approach can be used to reformulate the constraints that govern the dynamics of $a_{2,t}$.

Proposition 7. (3.3) *can be expressed with the following set of integer variables and*

constraints:

$$w_{t,0} - w_{t,6} \leqslant M_{1,t}(1 - l_{1,t}), \qquad t \in \{0, \cdots, 23\}, \tag{3.18}$$

$$z_{2,t} \leqslant M_{z2} l_{1,t}, \qquad t \in \{0, \cdots, 23\}, \tag{3.19}$$

$$z_{2,t} \leqslant k_2, \qquad t \in \{0, \cdots, 23\}, \tag{3.20}$$

$$z_{2,t} \geqslant k2 - M_{z2}(1 - l_{1,t}), \qquad t \in \{0, \cdots, 23\}, \tag{3.21}$$

$$z_{2,t} \geqslant 0, \qquad t \in \{0, \cdots, 23\}, \tag{3.22}$$

$$a_{2,t+1} = \gamma_2(a_{2,t} - a_{2,b}) + a_{2,b} + r_t^w z_{2,t}, \qquad t \in \{0, \cdots, 23\}. \tag{3.23}$$

The full MILP model that incorporates these constraints and the proper surrogate objective function can be found in the appendix.

Prediction and statistical consistency of surrogate likelihood estimation

While the surrogate likelihood estimation model can be thought of as descriptive, in practice clinicians are interested in predicting future participant behavior. Since we have a well defined likelihood model, we can use a Bayesian framework, similar to the one proposed in (Aswani et al. 2019), in order to predict future participant behavior using this model.

To simplify the notation for this conversion, let $\theta_t = \{a_{1,t}, a_{2,t}, p_t, B, f_{b,t}, \hat{r}_t^w, k_1, k_2, k_p\}$ be a shorthand for the full motivational state of the participant at week t, and let $\Theta = \mathcal{A}^2 \times \mathcal{P}^2 \times \mathcal{F} \times \mathcal{R}$ such that $\theta_t \in \Theta$. To convert the surrogate estimation problem into a Bayesian prediction problem we need to consider the posterior probability over the model parameters given observations $\{\tilde{w}_{t,d}, g_t\}_{t \in \mathcal{T}, d \in \{0,\ldots,6\}}$, namely $\mathbb{P}(\{\theta_t, w_{t,d}, c_{t,d}\}_{t \in \mathcal{T}, d \in \{0,\ldots,6\}} | \{\tilde{w}_{t,d}, g_t\}_{t \in \mathcal{T}, d \in \{0,\ldots,6\}})$. Using Bayes' Theorem we can write the posterior distribution in terms of the joint likelihood as follows:

$$\mathbb{P}(\{\theta_t, w_{t,d}, c_{t,d}\}_{t\in\mathcal{T},d\in\{0,\dots,6\}}|\{\tilde{w}_{t,d}, g_t, r_t^w, r_t^c\}_{t\in\mathcal{T},d\in\{0,\dots,6\}}) =$$
$$\frac{1}{Z}\mathbb{P}(\{\tilde{w}_{t,d}, g_t\}_{t\in\mathcal{T},d\in\mathcal{D}_t}|\{\theta_t, w_{t,d}, c_{t,d}, r_t^w, r_t^c\}_{(t,d)\in\mathcal{T}\times\{0,\dots,6\}})\mathbb{P}(\{\theta_t, w_{t,d}, c_{t,d}\}_{t\in\mathcal{T},d\in\{0,\dots,6\}}). \tag{3.24}$$

Here, Z is a normalization constant that ensures the posterior is a valid probability distribution and $\mathbb{P}(\{\theta_t, w_{t,d}, c_{t,d}\}_{t\in\mathcal{T},d\in\{0,\dots,6\}})$ is the prior probability distribution that reflects the clinician's initial beliefs over the values of $\{\theta_t, w_{t,d}, c_{t,d}\}_{t\in\mathcal{T},d\in\{0,\dots,6\}}$. Note that from the structure of the model, the participant's physical and behavioral state trajectory can be fully determined if the decision maker has knowledge of the initial values of the physical and motivational states (or equivalently their current value). Thus instead of considering joint posterior and prior distributions over all possible trajectories, we will focus our formulation on distributions for the initial participant physical and motivational states $\{\theta_0, w_{0,0}\}$. Note we do not need an explicit posterior or prior on $\{c_{0,d}\}_{d=0}^6$ since by Proposition 5 these values are fully determined by $\{\theta_0, w_{0,0}\}$. However, to obtain the posterior probability for some value of $\{\theta_0, w_{0,0}\}$ would still require us to integrate the joint posterior distribution over all possible trajectories with those initial conditions that could result in the observed data sequence $\{\tilde{w}_{t,d}, g_t\}_{t\in\mathcal{T},d\in\{0,\dots,6\}}$, which is numerically challenging to do. Instead, we will consider an approach similar to (Aswani et al. 2019) and use profile likelihood estimation to estimate the posterior distribution. For our analysis we make the following assumption on the prior distribution of $\{\theta_0, w_{0,0}\}$: For all $w_{0,0} \in \mathcal{W}, \theta_0 \in \Theta$, $\mathbb{P}(w_{0,0}, \theta_0) > 0$. Moreover, $\log \mathbb{P}(w_{0,0}, \theta_0)$ can be expressed as a set of mixed integer linear constraints and objective terms. The first part of the assumption is key for consistency and ensures that we consider every possible value of $w_{0,0}, \theta_0$ in our estimation. The second part is a relatively mild assumption that will allow us to pose the problem of obtaining our predictive estimates as a MILP. It is also satisfied by a variety of distributions such as the Laplace distribution and piece-wise linear distributions (such as those derived from histograms of previous

data), of note it is also satisfied by the uniform distribution. With this in mind, consider the following optimization problem:

$$\eta(\bar{w}_{0,0}, \bar{\theta}_0, \{r_t^w, r_t^c\}_{t\in\mathcal{T}}) =$$

$$\min_{\{w_{t,d},\theta_t,c_{t,d}\}_{t\in\mathcal{T},d\in\{0,\dots,6\}}} \sum_{t\in\mathcal{T},d\in\mathcal{D}_t} -\log\mathbb{P}(\tilde{w}_{t,d}|w_{t,d}) + \sum_{t\in\mathcal{T}} |g_t - p_t| - \log\mathbb{P}(w_{0,0},\theta_{0,0}) + \log Z, \tag{3.25a}$$

$$\text{subject to: } (3.1b), (3.1c), (3.2)-(3.6), \quad t\in\mathcal{T}, d\in\{0,\cdots,6\}, \tag{3.25b}$$

$$\{c_{t,d}\}_{d=0}^{6} \in \mathcal{C}(a_{1,t}, a_{2,t}, w_{t,0}, f_{b,t}, \hat{r}_t^w), \quad t\in\mathcal{T}, \tag{3.25c}$$

$$w_{0,0} = \bar{w}_{0,0}, \theta_{0,0} = \bar{\theta}_0, \tag{3.25d}$$

$$w_{t,d} \in \mathcal{W}, f_{t,d}, c_{t,d} \in \mathcal{F}, \quad t\in\mathcal{T}, d\in\{0,...,6\}, \tag{3.25e}$$

$$p_t, B \in \mathcal{P}, a_{1,t}, a_{2,t} \in \mathcal{A}, f_{b,t} \in \mathcal{F}, \hat{r}_t^w \in \mathcal{R}, \quad t\in\mathcal{T}. \tag{3.25f}$$

Note that (3.25) is essentially the same formulation as $H_{\mathbf{SMLE}}$ with the addition of the log prior and normalization terms to the objective and Constraint (3.25d) that sets the initial conditions. Problem (3.25) is in fact a feasibility problem, that when solved evaluates a function $\eta : \mathcal{W}\times\Theta\times\mathcal{R}^{2|\mathcal{T}|} \mapsto \mathbb{R}$, which is very similar to the log posterior distribution, but uses the surrogate likelihood instead of the true joint likelihood. By removing (3.25d) and the term $\log Z$ from the objective, we can transform (3.25) into a problem that calculates the surrogate maximum a posteriori estimate (MAP) for $w_{0,0}, \theta_0$, we will call these estimates $\hat{w}_{0,0}^{\text{MAP}}, \hat{\theta}_0^{\text{MAP}}$. One challenge with (3.25) is that the value of Z is not generally known and must be estimated by solving (3.25) at several initial conditions and then using numerical integration. Alternatively, we can estimate a surrogate posterior using the MAP estimates at a particular value of $\bar{w}_{0,0}, \bar{\theta}_0$ as follows:

$$\hat{\mathbb{P}}(\bar{w}_{0,0}, \bar{\theta}_0 | \{\tilde{w}_{t,d}, g_t, r_t^w, r_t^c\}_{t\in\mathcal{T},d\in\{0,\dots,6\}}) = \frac{\exp(-\eta(\bar{w}_{0,0}, \bar{\theta}_0, \{r_t^w, r_t^c\}_{t\in\mathcal{T}}))}{\exp(-\eta(\hat{w}_{0,0}^{\text{MAP}}, \hat{\theta}_0^{\text{MAP}}, \{r_t^w, r_t^c\}_{t\in\mathcal{T}}))} \tag{3.26}$$

Using this posterior distribution we can characterize the uncertainty around the

initial conditions and form predictions and scenarios for future participant behavior.

Consistency proof

We now proceed to prove that the estimates computed by H_{SMLE} and the predictive model are statistically consistent, that is, as more data is collected from the participant these estimates become closer to their ground truth value (or a value that is closest to the true distribution given the model definition). This condition is key for ensuring that any adaptive framework that is using a stream of participant data can provide effective incentives that are properly personalized to each participant. Moreover, this is a necessary condition to ensure such an adaptive policy is asymptotically optimal. In general, proving surrogate likelihood functions yield consistent estimates requires that the estimation problem have Lipschitz continuous objective function and constraints (and by extension a Lipschitz continuous value function) that allows using known asymptotic and finite time bounds (Bartlett et al. 2006, Nguyen et al. 2009). Since our estimation problem is a MILP, we do not necessarily satisfy this continuity condition. On the other hand, analysis of consistency of MILP based parameter estimates relies on an exact optimal solution of the optimization problem with respect to the true joint likelihood function of the problem (Mintz et al. 2017). Clearly, in the case of surrogate likelihood estimation this condition is not satisfied and so a different analysis is required. Our approach will extend the results for consistency of MILP estimates to the case of surrogate estimation, when the surrogate loss is within a multiplicative constant of the true likelihood. While we focus our analysis on the participant model in the context of weight loss interventions, the technique presented here can be generalized to any surrogate estimation using MIPs with a bounded likelihood function.

Let $\{\hat{w}_{0,0}, \hat{\theta}_0\} \in H_{SMLE}(\{\tilde{w}_{t,d}, g_t, r_t^w, r_t^c\}_{t\in\mathcal{T}, d\in\mathcal{D}_t})$ be the estimates calculated in the surrogate likelihood estimation problem, and let $w^*_{0,0}, \theta^*_0$ be the true value of these parameters for a particular participant. To show that $\hat{w}_{0,0}, \hat{\theta}_0 \xrightarrow{p} w^*_{0,0}, \theta^*_0$ we will first show that the surrogate posterior probability function defined in (3.26) is consistent in the Bayesian sense, which would then imply that $\hat{w}^{MAP}_{0,0}, \hat{\theta}^{MAP}_0$

are consistent estimators for any prior distribution that satisfies Assumption 3.3. Because the uniform distribution satisfies this assumption, and because under a uniform prior $\hat{w}_{0,0}^{MAP}, \hat{\theta}_0^{MAP} = \hat{w}_{0,}, \hat{\theta}_0$ this would mean that the H_{SMLE} estimates are also consistent. To formally conduct our analysis we will need the following definition for Bayesian consistency of a posterior distribution:

Definition 3.1. *For all* $(w_{0,0}^*, \theta_0^*) \in \mathcal{W} \times \Theta$ *and constants* $r, \delta > 0$*, we say the estimate of the posterior distribution* $\hat{\mathbb{P}}(\cdot|\{\tilde{w}_{t,d}, g_t, r_t^w, r_t^c\}_{t\in\mathcal{T}, d\in\{0,\dots,6\}})$ *is consistent if* $\mathbb{P}_{(w_{0,0}^*, g_0^*, \theta_0^*)}(\hat{\mathbb{P}}(S(\delta)|\{\tilde{w}_{t,d}, g_t, r_t^w, r_t^c\}_{t\in\mathcal{T}, d\in\{0,\dots,6\}}) \geqslant r) \to 0$ *as* $t \to \infty$*. Here* $\mathbb{P}_{(w_{0,0}^*, \theta_0^*)}$ *is the probability law where* $(w_{0,0}^*, \theta_0^*)$ *are the true initial conditions of the system, and where* $S(\delta) := \{(w_{0,0}, \theta_0) \notin \mathcal{B}((w_{0,0}^*, \theta_0^*), \delta)\}$*, where* $\mathcal{B}((w_{0,0}^*, \theta_0^*), \delta)$ *is an open ball with radius* δ *centered around* $(w_{0,0}^*, \theta_0^*)$*.*

The implication of Definition 3.1 is that if our posterior estimate is consistent, then as more data is collected it turns into a degenerate distribution at the true parameter values. While this is a stronger condition then parameter consistency we will show our estimate possesses this property and that this implies the point estimates are consistent as well. To proceed with the analysis we make the following technical assumption.

There exists $\epsilon > 0$ such that the set $\mathcal{P} := [\epsilon, 1 - \epsilon]$. In other words, for all $t \in \mathcal{T}$, $\epsilon \leqslant p_t \leqslant 1 - \epsilon$.

This assumption ensures that $\mathcal{P}$ is a compact set making it easily deployable with commercial optimization solvers. It also ensures that the value of p_t and by extension $\mathbb{P}(g_t|p_t)$ is bounded, which will be key in showing that surrogate posterior estimates are consistent. In practice this is a reasonable assumption since it guarantees that on any week in the trial a participant will have some positive probability of successfully completing their calorie recording goal or failing it. This is reflected in real-world interventions where no participant truly has an almost sure probability of failing to record or recording their calories. We will also require the following assumption on the history of the observations.

Let (w_0^*, θ_0^*) be the true initial conditions, the incentives $\{r_t^w, r_t^c\}_{t\in\mathcal{T}}$ are such that

for any $\delta > 0$,

$$\max_{S(\delta)} \lim_{|\mathcal{T}|\to\infty} \sum_{t\in\mathcal{T}, d\in\mathcal{D}_t} -\log \frac{\mathbb{P}(\tilde{w}_{t,d}|\bar{w}_{t,d})}{\mathbb{P}(\tilde{w}_{t,d}|w_{t,d})} + \sum_{t\in\mathcal{T}} -\log \frac{\mathbb{P}(g_t|\bar{p}_t)}{\mathbb{P}(g_t|p_t)} = -\infty. \tag{3.27}$$

where $\bar{w}_{t,d}, \bar{p}_t$ are the states and decisions under initial conditions $(w_{0,0}, p_0) \in S(\delta)$, and $w_{t,d}, p_t$ are the states and decisions under true initial conditions $(w^*_{0,0}, p^*_0)$. This assumption is known as a sufficient excitation condition and is a common assumption in the literature (Craig et al. 1987, Åström and Wittenmark 2013). Essentially, this assumption states that there is sufficient variance from the incentives administered so that it is possible for the clinician to identify the true states of the participants. In practice this assumption can be satisfied if there is sufficient process noise or by adding random perturbations to the incentives administered. Using this assumption, we can now proceed to prove the consistency of the posterior estimate. First, we prove a proposition on the structure of the surrogate likelihood function.

Proposition 8. *Given Assumptions 3.3– 3.3,* $|g_t - p_t|$ *can be bounded as:*

$$\frac{-\log(1-\epsilon)}{\epsilon}|g_t - p_t| \leqslant -\log(\mathbb{P}(g_t|p_t)) \leqslant \frac{-\log(\epsilon)}{1-\epsilon}|g_t - p_t|. \tag{3.28}$$

The complete proof can be found in the appendix, and here we present a brief sketch. Using Assumption 3.3, we consider two cases (one when $g_t = 0$ and one when $g_t = 1$) and use a calculus argument to show that the desired bounds hold. From (3.28), we see that so long as p_t is bounded then $\log(\mathbb{P}(g_t|p_t)) = \Theta(|g_t - p_t|)$, this will be key in showing convergence since it implies these expressions have similar asymptotic behavior. We note that the keys to this proposition are that the probability measure is log concave and bounded. Without these conditions, there could be edge-case observations that would make it difficult to distinguish between underlying values of p_t. With this structure we can now prove the main result on the posterior estimate.

Proposition 9. *Given Assumptions 3.3– 3.3, the surrogate posterior estimate* $\hat{\mathbb{P}}(w_{0,0}, \theta_0|\{\tilde{w}_{t,d}, g_t, r^w_t, r^c_t\}_{t\in\mathcal{T}, d\in\{0,\dots,6\}})$ *is consistent.*

The complete proof of this proposition can be found in the appendix here we present a sketch. The main arguments are first to use Proposition 8 to create a pointwise upper bound for the surrogate posterior function in terms of the true posterior function specified by the model. Then by Assumption 3.3 we show that this bound implies that for any initial conditions $(w_{0,0}, \theta_0) \neq (w^*_{0,0}, \theta^*_0)$ the posterior assigns zero probability in the limit. To complete the proof we use a volume bound to show that this condition holds uniformly over $\mathcal{W} \times \Theta$. This proposition shows that our posterior estimates satisfy Definition 3.1, and implies the following corrollary.

Corollary 3.2. *Given Assumptions 3.3– 3.3,* $(\hat{w}^{MAP}_{0,0}, \hat{\theta}^{MAP}_0) \xrightarrow{p} (w^*_{0,0}, \theta^*_0)$.

The complete proof of the corollary can be found in the appendix. Note that since Corollary 3.2 holds for surrogate maximum *a posteriori* estimates calculated with any prior distribution that satisfies Assumption 3.3, including the uniform distribution. However, if we use the uniform prior, then our predictive problem is exactly H_{SMLE} meaning that the estimators calculated from this problem are also consistent.

3.4 Financial Incentive Optimization

In this section, we show how the model and prediction techniques from Section 3.3 can be used to optimize personalized financial incentives for a cohort of participants in a weight loss trial. Recall that the goal of the interventionist is to administer financial incentives to each participant to maximize the number of participants that achieve clinically significant weight loss by the end of the trial while remaining within the intervention budget. Furthermore the incentive administered should seem random to the participant. To formally define our problem, let U be the set of participants. For this section, we will augment the notation from Sections 3.3 and 3.2 by including an additional index of $u \in U$ to indicate parameters specific to a trial participant u. So, for instance, the weight and motivational states of participant u at week t and day d will be given by $w_{u,t,d}, \theta_{u,t}$ respectively. Let $\mathcal{L} : \mathcal{W} \to \mathbb{R}$ be a loss function that captures if a participant is unable to lose a clinically significant

amount of weight. We leave this loss function in a general form since there are several ways of designing this incentive optimization problem depending on the interventionist's secondary outcomes, we present some illustrative examples of loss functions in Section 3.5. In week t, the clinician calculates a distribution $\pi_{u,t} \in \Delta_{\mathcal{R}^2}$ for each participant $u \in U$, where $\Delta_{\mathcal{R}^2}$ is the set of distribution with support over $\mathcal{R}^2$, and administers incentive $\{r^w_{u,t}, r^c_{u,t}\} \sim \pi_{u,t}$. Let G be the total intervention budget, that is, the clinician requires that with probability one $\sum_{u \in U, t \in \mathcal{T}} r^w_{u,t} + r^c_{u,t} \leqslant G$. The ultimate goal of the clinician is to find a sequence of distributions for all participants $\{\pi_{u,t}\}_{u \in U, t \in \mathcal{T}}$ such that $\mathbb{E} \sum_{u \in U} \mathcal{L}(w_{u,24,6}(\{\pi_{u,t}\}_{t \in \mathcal{T}}))$ is minimized and the budget constraint is not violated, where the expectation is taken over not only the uncertainty in the participant parameter values but also over the stochasticity of the incentive distribution.

As stated in this general form, this problem is challenging to solve due to the presence of a hard constraint, partially observed parameters, and randomized policy. Moreover, using standard techniques such as scenario generation (Kaut and Stein 2003) may not be tractable since different scenarios need to be created not only for each potential value of the unobserved states but also for each realization of the reward distribution and each participant. Instead we will consider a different approach that leverages the statistical properties of our posterior estimates from Section 3.3 and certainty equivalence to approximate a solution to the interventionist's problem. Specifically, we propose an adaptive approximation approach, where at each time t the interventionist will estimate the unknown participant parameters using H_{SMLE} for each participant and then calculate an incentive design based on these estimates. For our approximation approach, we restrict our policies to be only the set of deterministic policies over $\mathcal{R}$, equivalently distribution $\pi_{u,t}$ where for each $u \in U$ they assign a probability mass of 1 to a single element of $\mathcal{R}$. This restriction will simplify our formulation since we will not need to consider different realizations of the incentive distribution and can concentrate our efforts on the uncertainty in the unobserved participant parameters. Moreover, it will ensure that we can easily meet the budget constraint with probability one. In practice, financial incentives are rarely truly random in weight loss interventions but are in fact pre-

determined by interventionists to be perceived as random by participants (Leahey et al. 2015, Almeida et al. 2015). Since our adaptive approximation approach will be recomputing incentives at each time period, despite using a deterministic policy, since these rewards will be frequently changing, they should still be perceived as random by study participants making this approach suitable for our setting. In the remainder of this section, we will first present the details of our approximation algorithm and then provide guarantees that our method is asymptotically optimal over the class of deterministic policies. This guarantee ensures that under proper technical conditions the policy calculated by our method will converge to the best deterministic policy as more data is collected form the participants over the course of the intervention.

Approximation algorithm for personalized incentive design

To form our adaptive approach, we will consider a framework where interventionists minimize their loss with respect to their posterior information at each time step. To formalize this, suppose that it is currently the start of week T (where $1 < T < 24$) of the intervention, then let $\mathcal{F}_T = \{\tilde{w}_{u,t,d}, g_{u,t}, r^w_{u,t}, r^c_{u,t}\}_{u\in U, t\in\{0,\ldots,T\}, d\in\mathcal{D}_t}$. Our approach will solve the following deterministic policy problem formulation.

$$\min_{\{r^w_{u,i}, r^c_{u,i}\}_{i=T}^{24}\in\mathcal{R}^2} \{\mathbb{E}[\sum_{u\in U} \mathcal{L}(w_{u,24,6})|\mathcal{F}_T] : \sum_{u\in U, t\in\mathcal{T}} r^w_{u,t} + r^c_{u,t} \leqslant G\}. \tag{3.29}$$

From the modeling assumptions in Section 3.3 we note that knowledge of $(w_{u,0,0}, \theta_{u,0})$ for each participant are sufficient to determine the trajectory of all other parameters for participant u given a sequence of incentives. Thus by the smoothing theorem (Bickel and Doksum 2015), there exists some function $\phi : \mathcal{W}\times\Theta\times\mathcal{R}^2 \to \mathbb{R}$ such that (3.29) can be reformulated as:

$$\min_{\{r^w_{u,i}, r^c_{u,i}\}_{i=T}^{24}\in\mathcal{R}^2} \{\mathbb{E}[\sum_{u\in U} \phi(w_{u,0,0}, \theta_{u,0}, \{r^w_{u,t}, r^c_{u,t}\}_{t\in\mathcal{T}})|\mathcal{F}_T] : \sum_{u\in U, t\in\mathcal{T}} r^w_{u,t} + r^c_{u,t} \leqslant G\}. \tag{3.30}$$

In general the closed form of ϕ is difficult to obtain since it relies on the composition of the loss function and model dynamics; however, this reformulation illustrates that in order to approximate the expectation in the objective we would only need to consider an estimate of the posterior distribution for $(w_{u,0,0}, \theta_{u,0})$, such as the posterior estimate in (3.26). Thus one approach for solving (3.30) is using scenario generation and discretizing $\mathcal{W} \times \Theta$ into a grid of m scenarios. This would result in the following optimization problem:

$$\min_{\{r^w_{u,i}, r^c_{u,i}\}_{i=T}^{24}} \sum_{u \in U, k=\{0,\dots,m\}} \phi(w^k_{u,0,0}, \theta^k_{u,0}, \{r^w_{u,t}, r^c_{u,t}\}_{t\in\mathcal{T}}) \hat{\mathbb{P}}(w^k_{u,0,0}, \theta^k_{u,0} | \mathcal{F}_T), \tag{3.31a}$$

$$\text{subject to: } \sum_{u\in U, t\in\mathcal{T}} r^w_{u,t} + r^c_{u,t} \leqslant G, \tag{3.31b}$$

$$r^w_{u,t}, r^c_{u,t} \in \mathcal{R}^2. \tag{3.31c}$$

Solving this optimization problem is challenging first because the set Θ is high dimensional meaning that a large number of grid points may need to be selected in order to obtain a sufficiently close approximation to the distribution. Furthermore, recall that to compute $\hat{\mathbb{P}}(w^k_{u,0,0}, \theta^k_{u,0} | \{\tilde{w}_{u,t,d}, g_{u,t}, r^w_{u,t}, r^c_{u,t}\}_{t\in\mathcal{T}, d\in\{0,\dots,6\}})$ requires solving a MIP for each $k \in \{0, ..m\}$. Thus to form the objective would require solving $|U|m$ MIPs, which can be computationally expensive and would be challenging to scale to large weight loss interventions. Instead of using a full posterior approach, we instead propose to use the either the surrogate MAP or MLE estimates of $w_{u,0,0}, \theta_{u,0}$ with data up to time T, that we will denote as $\hat{w}^T_{u,0,0}, \hat{\theta}^T_{u,0}$, as single point estimates and optimizing future incentives with respect to these estimates.

We formalize this problem as follows:

$$\psi_T(\{\hat{w}^T_{u,0,0}, \hat{\theta}^T_{u,0}, \{\bar{r}^w_{u,t}, \bar{r}^c_{u,t}\}_{t=0}^T\}_{u\in U}) = \min \sum_{u\in U} \mathcal{L}(w_{u,24,6}), \tag{3.32a}$$

$$\text{subject to: } \sum_{u\in U}\sum_{t\in W} r^w_{u,t} + r^c_{u,t} \leqslant G, \tag{3.32b}$$

$$\text{Constraints (3.7b)-(3.7c)}, \quad \forall u \in U, \forall t \in \{0, ..., 24\}, \tag{3.32c}$$

$$w_{u,0,0} = \hat{w}^T_{u,0,0}, \hat{\theta}^T_{u,0} = \theta_{u,0}, \quad \forall u \in U, \tag{3.32d}$$

$$r^w_{u,t} = \bar{r}^w_{u,t}, r^c_{u,t} = \bar{r}^c_{u,t} \quad \forall u \in U, \forall t = \{0, ..., T\}, \tag{3.32e}$$

$$w_{u,t,d} \in \mathcal{W}, \theta_{u,t} \in \Theta, \quad \forall u \in U, t \in \{0, ..., 24\}. \tag{3.32f}$$

Here the values $\bar{r}^w_{u,t}, \bar{r}^c_{u,t}$ are the previously administered financial rewards from the beginning of the intervention up to the current time period T. Therefore ψ_T should be interpreted as the minimum possible value of the loss function if the true initial conditions of each participant u are the estimates $\{\hat{w}^T_{u,0,0}, \hat{\theta}^T_{u,0}\}$ and the rewards that they have received up to the current time period are fixed to their historical values.

Algorithm 2 Design of Incentives Algorithm (DIA)

Require: $\{\tilde{w}_{u,t,d}, g_{u,t}, r^w_{u,t}, r^c_{u,t}\}_{t\in\mathcal{T}, d\in\mathcal{D}_{u,t}}$ for all $u \in U$,
Compute $(\{\hat{w}^T_{u,0,0}, \hat{\theta}^T_{u,0}\}_{u\in U}) \in H_{SMLE}(\{\tilde{w}_{u,t,d}, g_{u,t}, r^w_{u,t}, r^c_{u,t}\}_{t\in\mathcal{T}, d\in\mathcal{D}_{u,t}})$,
Compute $\{r^w_{u,t}, r^c_{u,t}\}_{t\in\{T,...,24\}, u\in U} \in \{\psi_T(\{\hat{w}^T_{u,0,0}, \hat{\theta}^T_{u,0}, \{\bar{r}^w_{u,t}, \bar{r}^c_{u,t}\}_{t=0}^T\}_{u\in U}) | \{r^w_{u,t}, r^c_{u,t}\}_{t=T}^{24}\}$,
Apply $r^w_{u,T}, r^c_{u,T}$ back to $u \in U$.

Using this formulation we define our adaptive incentive calculation approach that we call the Design of Incentives Algorithm (DIA). The pseudocode of DIA is presented in Algorithm 2, and consists of three main steps. First, all data up to the current time step T is used to estimate model parameter for each participant using the SMLE model (H_{SMLE}) established in Section 3.3. Then, using the parameter

estimates of all participants and previously dispensed incentives as inputs, we solve (3.32) to compute a sequence of incentives from period T to 24. We then apply the incentive values for period T, $\{r^w_{u,T}, r^c_{u,T}\}$ to each participant $u \in U$ and collect new observations. These three steps are repeated for each week until we reach the end of the intervention. As new data is collected the parameter estimators are updated and new incentives are computed.

Asymptotic optimality

Here we show that the incentives output by DIA are asymptotically optimal with respect to the class of deterministic policies. This property ensures that as more data is collected from each participant over the course of the intervention, DIA produces incentives that approach the optimal incentives with respect to a full information problem, with policies restricted to the set of deterministic policies. In this section, we present sketches of proofs of each proposition and the detailed proofs can be found in the appendix.

Our proof approach will be similar to that proposed by (Mintz et al. 2017) with modification to our setting. In general, asymptotic optimality is not trivial to guarantee since it requires that the optima of an approximation problem converge in probability to the optima of the goal problem being approximated. Point-wise convergence of the value functions is usually insufficient to prove this property, and often it requires uniform convergence of the value function of the approximation problems to the objective of the goal problem. However, since our approximations are based on MIP formulations proving uniform convergence maybe difficult to guarantee. Thus, we will use a weaker condition known as epi-convergence (Lachout et al. 2005) that is sufficient to prove this result. This condition ensures that the epigraph of the value functions of the approximation problems converges stochastically to the epigraph of the target problem, and thus ensures convergence of the lower-level sets and minima.

For our analysis we will need to define the following value function:

$$\psi(\{\bar{w}_{u,0,0}, \bar{\theta}_{u,0}, \{\bar{r}^w_{u,t}, \bar{r}^c_{u,t}\}_{t=0}^{24}\}_{u\in U}) = \min \sum_{u\in U} \mathcal{L}(w_{u,24,6}), \tag{3.33a}$$

$$\text{subject to: } \sum_{u\in U}\sum_{t\in W} r^w_{u,t} + r^c_{u,t} \leqslant G, \tag{3.33b}$$

$$\text{Constraints (3.7b)-(3.7c)}, \quad \forall u \in U, \forall t \in \{0, ..., 24\}, \tag{3.33c}$$

$$w_{u,0,0} = \bar{w}_{u,0,0}, \bar{\theta}_{u,0} = \theta_{u,0}, \quad \forall u \in U, \tag{3.33d}$$

$$r^w_{u,t} = \bar{r}^w_{u,t}, r^c_{u,t} = \bar{r}^c_{u,t}, \quad \forall u \in U, \forall t = \{0, ..., 24\}, \tag{3.33e}$$

$$w_{u,t,d} \in \mathcal{W}, \theta_{u,t} \in \Theta, \quad \forall u \in U, t \in \{0, ..., 24\}. \tag{3.33f}$$

Note that $\psi(\{\bar{w}_{u,0,0}, \bar{\theta}_{u,0}, \{\bar{r}^w_{u,t}, \bar{r}^c_{u,t}\}_{t=0}^{24}\}_{u\in U})$ is the value function of a problem quite similar to (3.32). However, unlike (3.32), (3.33) is a feasibility problem where the incentive sequence is predefined for the entirety of the intervention and not only up to time T. Thus $\psi(\{\bar{w}_{u,0,0}, \bar{\theta}_{u,0}, \{\bar{r}^w_{u,t}, \bar{r}^c_{u,t}\}_{t=0}^{24}\}_{u\in U})$ can be interpreted as the minimum loss that would be expected by offering the predetermined sequence $\{\bar{r}^w_{u,t}, \bar{r}^c_{u,t}\}_{t=0}^{24}$ to each participant if their individual parameter values where truly equal to $\{\bar{w}_{u,0,0}, \bar{\theta}_{u,0}\}$.

To begin our analysis, we will show that $\psi(\{\bar{w}_{u,0,0}, \bar{\theta}_{u,0}, \{\bar{r}^w_{u,t}, \bar{r}^c_{u,t}\}_{t=0}^{24}\}_{u\in U})$ has structural properties that insure that given our estimation is consistent will result in asymptotically optimal incentives.

Proposition 10. *If Assumptions 3.3-3.3 hold, then the value function* $\psi(\{w_{u,0,0}, \theta_{u,0}, \{r^w_{u,t}, r^c_{u,t}\}_{t=0}^{T+n}\}_{u\in U})$ *is lower semi-continuous in each argument* $\{w_{u,0,0}\}_{u\in U}$, $\{\theta_{u,0}\}_{u\in U}$, *and* $\{\{r^w_{u,t}, r^c_{u,t}\}_{w=T+1}^{T+n}\}_{u\in U}$.

To prove this proposition we first show the problem can be reformulated as a parametric MILP with each of the parameter arguments as affine terms in the constraints, and then apply the results from (Hassanzadeh and Ralphs 2014). This

proposition ensures that the value function $\psi(\{w_{u,0,0}, \theta_{u,0}, \{r^w_{u,t}, r^c_{u,t}\}_{t=0}^{T+n}\}_{u\in U})$ has a closed epigraph and closed lower level sets, a key property for showing the convergence of minima.

For the remainder of the analysis, let $w^*_{u,0,0}, \theta^*_{u,0}$ be the true initial parameter values for each participant $u \in U$ and as before let $\{\hat{w}^T_{u,0,0}, \hat{\theta}^T_{u,0}\}$ be the estimates provided by H_{SMLE} estimates of these parameters at time T. Using the structure from Proposition 10 we analyze the manner by which $\psi(\{\hat{w}^T_{u,0,0}, \hat{\theta}^T_{u,0}, \{\bar{r}^w_{u,t}, \bar{r}^c_{u,t}\}_{t=0}^{24}\}_{u\in U})$ converges to $\psi(\{w^*_{u,0,0}, \theta^*_{u,0}, \{\bar{r}^w_{u,t}, \bar{r}^c_{u,t}\}_{t=0}^{24}\}_{u\in U})$. In particular we show the following convergence property.

Proposition 11. *If Assumptions 3.3-3.3 hold, then* $\psi(\{\hat{w}^T_{u,0,0}, \hat{\theta}^T_{u,0}, \{\bar{r}^w_{u,t}, \bar{r}^c_{u,t}\}_{t=0}^{24}\}_{u\in U}) \xrightarrow[\mathcal{R}^{2|U|})]{l-prob} \psi(\{w^*_{u,0,0}, \theta^*_{u,0}, \{\bar{r}^w_{u,t}, \bar{r}^c_{u,t}\}_{t=0}^{24}\}_{u\in U})$, *which means the function* $\psi(\{\hat{w}^T_{u,0,0}, \hat{\theta}^T_{u,0}, \{\bar{r}^w_{u,t}, \bar{r}^c_{u,t}\}_{t=0}^{24}\}_{u\in U})$ *is a lower semi-continuous approximation to the function* $\psi(\{w^*_{u,0,0}, \theta^*_{u,0}, \{\bar{r}^w_{u,t}, \bar{r}^c_{u,t}\}_{t=0}^{24}\}_{u\in U})$ (*Lachout et al. 2005*).

To prove this proposition we apply Proposition 9 and Proposition 10 in conjunction with results from (Lachout et al. 2005). This property ensures that any lower level set centered around some incentive sequence $\{\{\bar{r}^w_{u,t}, \bar{r}^c_{u,t}\}_{t=0}^{24}\}_{u\in U}$ for the value function evaluated at the estimates, will converge in probability to the lower level set of the corresponding problem with the initial parameters equal to their true values. Note that this is a stronger structural property then simple point-wise convergence since this condition must hold for any lower level set of ψ on the incentive space $\mathcal{R}^{2|U|}$. This property also essentially ensures that the value functions of the sequence of approximation problems that use the H_{SMLE} estimates will converge to the epigraphs of the value function of the problem with the true parameter values. This property leads us to the final result that shows the solution provided by DIA is asymptotically optimal for the participant's true initial conditions.

Theorem 3.3. *Denote the set of optimal deterministic financial incentives under the true initial conditions* $\{(w^*_{u,0,0}, \theta^*_{u,0})\}_{u\in U}$ *as*
$R^*(\{(w^*_{u,0,0}, \theta^*_{u,0})\}_{u\in U}) := \{\psi(\{w^*_{u,0,0}, \theta^*_{u,0}, \{\bar{r}^w_{u,t}, \bar{r}^c_{u,t}\}_{t=0}^{24}\}_{u\in U}) | \{\{\bar{r}^w_{u,t}, \bar{r}^c_{u,t}\}_{t=0}^{24}\}_{u\in U}\}$.

If Assumptions 3.3-3.3 hold and $\mathtt{dist}(x, Y) = \inf_{y \in Y} \|x - y\|$*, then*

$$\mathtt{dist}\big(\{r_{u,T}^{w,DIA}, r_{u,T}^{c,DIA}\}_{u \in U}, R^*(\{(w_{u,0,0}^*, \theta_{u,0}^*)\}_{u \in U})\}\big) \xrightarrow{p} 0 \tag{3.34}$$

for any $\{r_{u,T}^{w,DIA}, r_{u,T}^{c,DIA}\}_{u \in U}$ *returned by Algorithm DIA as* $T \to \infty$*.*

We prove the final results by combining the results of Proposition 9, 10, and 11. This result implies that as additional data is collected by the clinician on the participants, the recommended incentives calculated by DIA will approach the optimal deterministic incentives that should be allocated to each participant. The two keys to this result are that estimates computed from H_{SMLE} are consistent and that our problem structure results in lower semi-continuous value functions. Note that while this result shows asymptotic optimality with respect to the class of deterministic policies, it does not provide guarantees on how DIA would fair against the best stochastic policies, an analysis that is more complex to conduct analytically. In Section 3.5, we provide an empirical examination of several stochastic policies and compare their performance to DIA.

3.5 Numerical Studies

We conducted three sets of numerical studies using data from the Log2Lose trial (Voils et al. 2018). The first study analyzed the performance of our methodology for capturing a participant's true weight trajectory. For this study we fit the SMLE model to the weight records of participants with different weight trajectories for the entire 24 weeks of the trial and show how well our predicted trajectory fits this data. The second study examined the accuracy of our predictive method in predicting a participant's weight trajectory using weight and incentive data from a short time span. We compared the predictive performance of our behavioral model against three machine learning methods (logistic regression, linear support vector machine (SVM), and random forest).The third study examine how our DIA method performs in designing financial incentives to maximize clinical weight loss successes. For this study we compare the efficacy of different financial incentive policies (deterministic,

randomized, one-size-fits-all in Log2Lose) under different budget options using DIA in terms of number of participants able to achieve clinically significant weight loss and percentage of weight lost by the five participants who lost the least weight.

Our results show our approach is well-suited for capturing different weight trajectories and predicting the future trajectory. In terms of the financial incentives design, our results show that the deterministic and randomized policies, where the incentives are generated by DIA, are more effective for encouraging weight loss than the one-size-fits-all policy implemented in the original Log2Lose study. In addition, the results show the randomized policy is potentially better suited for weight loss intervention than the deterministic policy. We ran all the experiments in Python (Van Rossum and Drake 2009) and compute the optimization problems using Gurobi v9.1.1 (Gurobi Optimization, LLC 2022).

Describing different weight loss trajectories

In this study, we examine how well a behavioral model trained with H_{SMLE} is able to capture different weight loss trajectories using the entire 24 weeks of data from the Log2Lose trial. We found the weight loss trajectories fit three common patterns: 1) participants who lose weight initially but then later become resistant to the intervention, 2) participants who lose weight consistently over the course of the intervention, and 3) participants who are resistant to the intervention and do not lose much weight. We name these groups initial achievers, constant achievers, and intervention-resistant, respectively. and note that out of 67 participants they make up 10%, 73%, and 5% of study participants, respectively. The remaining 12% of study participants had too few weight and calorie records for this analysis and were thus excluded.

The results in Figure 3.1 show that using our behavioral model, the estimated trajectory is a good fit to the observed trajectory regardless of the missing weight records or weight loss pattern. Figure 3.1a shows how our model fits an early achiever, Figure 3.1b shows the fit to a constant achiever, and Figure 3.1c shows the fit to a trajectory of an intervention resistant participant. Figure 3.1d shows the pre-

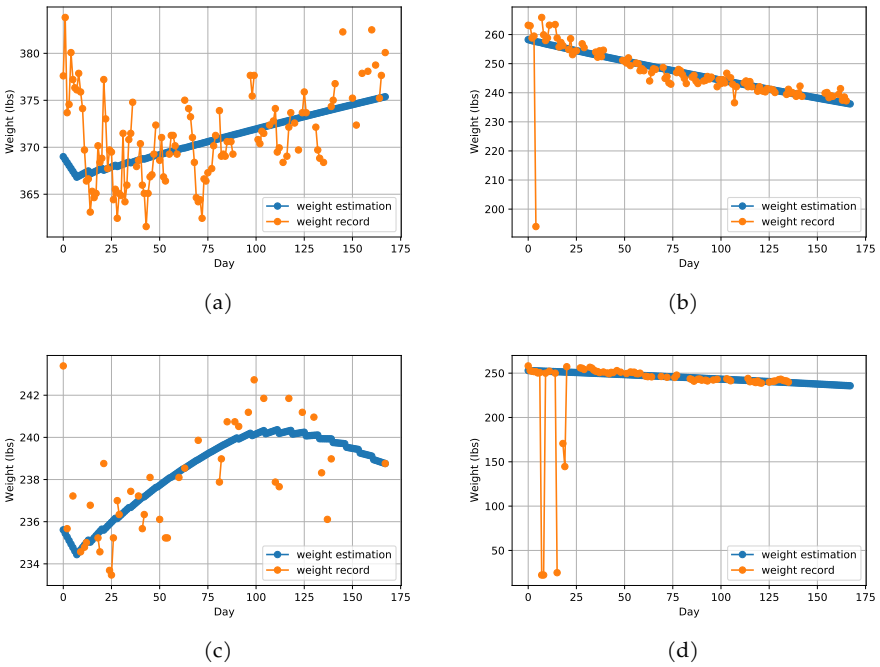


Figure 3.1: 4 examples of comparisons of true weight trajectory (orange) and the estimated fitting weight trajectory (blue) for week 0-24.

dicted weight trajectory remains accurate even when a participant's weight records have multiple missing consecutive measurements and anomalous measurements. These anomalous weight measurements could be caused by another household member of the study participant (or a pet) stepping onto the cellular scale. This result indicates that our proposed method is insensitive to these measures and can extract the underlying weight loss trajectory of the study participant.

Comparison of predictive performance

In this numerical study, we examined the performance of our behavioral model to predict whether or not a participant achieves clinically significant weight loss, defined as at least 5% weight loss, at the end the study (week 24). We compare our model against three common machine learning methods: linear SVM, logistic regression, and random forest (Breiman 2001a, Hastie et al. 2009). For this pre-

diction task we generated the labels by selecting either the weight at the end of the program or the last available weight record of week 24 (if the final weight is missing) as the true final weight of the participant and setting it to 1 if the final weight was no more than 95% of the initial weight and zero otherwise. For this study we included data from a total of 67 study participants who had at least 1 weight record in week 24.

Since our behavioral model performs a regression task, we used our posterior estimate from Section 3.3 to compute the probability the final weight would be below the clinically significant level using numerical integration (in a manner similar to (Aswani et al. 2019)). Then we varied a prediction threshold such that if this probability was larger than threshold our model would predict a label of 1. Our behavioral model only used daily weight and calorie measures and weekly incentive amounts as data for prediction.

All three machine learning methods were implemented using scikit-learn (Pedregosa et al. 2011a). For these models we used age, gender, height, body mass index, weekly average weight, two types of financial incentives, and weekly average caloric intake as training features. Since our data contained missing daily records, we could not directly use the data records. As an alternative option, we used weekly averages of caloric intake and weight since most weeks contained at least some measures of these features. We evaluated the predictive performance of these methods using five-fold cross validation, where in each fold 80% of the participants were used as a training set and 20% were used for validation. Within each fold we used another round of five-fold cross validation to optimize the hyperparameters of each of these ML methods.

To see how well each model is capable of using limited data and avoid over fitting we fit each model with feature sets that captured the first 4, 8, 12, 16, and 20 weeks. Note that, for each setting the models were tasked with predicting weight loss by week 24, meaning models trained on 4 weeks of data were predicting a measure 20 weeks in the future, models with 8 weeks of training data were predicting weight loss 16 weeks in the future, and so forth. We computed the false and true positive rates of each model and plotted them as ROC curves to analyze their predictive

performance. Figure 3.2 show the raw ROC curves for each time span. The figure shows that the performance of logistic regression and linear SVM does not improve as data from additional weeks is incorporated. Using random forest, we observe moderate improvement until the number of training weeks reaches 20. In contrast, our behavioral model improves consistently in its predictive capability as additional data is incorporated from the study participants. This suggests our proposed method is better suited for weight loss prediction even in the early weeks of a weight loss intervention. Our results also validate the consistency of the parameter estimates computed using H_{SMLE}. The results show our proposed behavioral model performs significantly better than the ML methods for longer training weeks and outperforms the ML methods for shorter training weeks with low false positive rate ($\text{FPR} \leqslant 0.4$). For instance, using the first 16 weeks of data as the training set, with a false positive rate at 0.43, the highest true positive rate achieved by the machine learning methods is 0.74 (random forest) while the true positive rate of our model is 0.92. We also note that with 20 weeks of data, the only competitive method to our behavioral model is the random forest predictor; however, the highest true positive it can achieve is 0.74 with a false positive between 0.21 and 0.78, and it is never able to achieve the 0.92 true positive rate which our model is able to achieve with a false positive of 0.17. This indicates that the random forest model is likely over fitting to data and may not be appropriate for incentive optimization in this setting, while our method is capable of leveraging the participant data effectively for prediction and optimization.

Simulation study of optimal incentive design

In the third study, we examine how well our adaptive methods perform in a simulated weight loss trial, and how deterministic policies compare to stochastic ones. We examine seven different incentive policies and examine their performance in terms of the number of participants able to achieve clinically significant weight loss, and the amount of weigh lost by the five participants who lost the least amount of weight. The policies we examined included six optimization based incentive

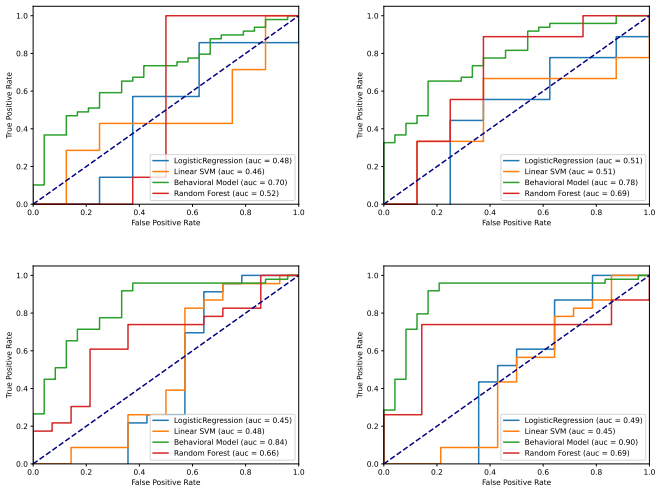


Figure 3.2: Raw ROC curves for various number of training weeks: (top: 4 weeks(left), 8 weeks(right); below: 16 weeks(left), 20 weeks(right).

policies that varied in whether they considered stochastic or deterministic incentives and what loss function they were optimizing. We also evaluated the incentive schedule implemented by the study investigators of Log2Lose. The deterministic optimization policies where based off of our proposed DIA method using a parameter estimate computed with H_{SMLE}, and distributed exactly the incentive amount computed by this method to each participant. To test stochastic policies that are still able to satisfy the budget constraint with probability one, we considered policies that at each week T would either provide a participant with their incentive amount as computed by DIA for some loss function or zero incentive with some non-zero probability $q \in \{0.25, 0.75\}$. For these optimization policies we considered two different loss functions: the indicator loss $(\min \sum_{u \in \mathcal{U}} \mathbb{1}\{w_{u,23,6} \leqslant 0.95 w_{u,0,0}\})$ and the hinge loss $(\min \sum_{u \in \mathcal{U}} (w_{u,23,6} - 0.95 w_{u,0,0})^{+})$. The indicator loss minimizes the number of participants who lose less than 5% of weight, and the hinge loss minimizes the gap between the final weight and the 5 % percent weight loss goal for participants who did not meet the weight goal.

We only included those study participants who had sufficient data and participated in the three treatment arms (A,B,C) of Log2Lose and were thus eligible for

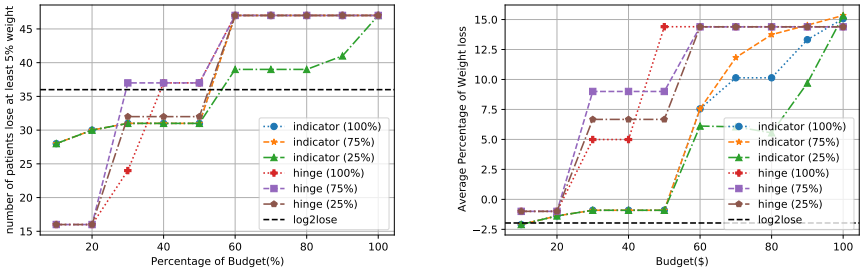
financial incentives. This resulted in the data of 47 participants being included in this study.

For each participant, we fit our behavioral model on their full study data (much like in Section 3.5) using H_{SMLE}, and used these fitted dynamics to simulate their behavior over the course of the trial. In each simulated week the particular incentive computation method would use available weight and recording goal measurements to compute a set of incentives for all 47 participants. Then the participants would receive this incentive and their dynamics would advance with the same functional and noise structure as detailed in Section 3.2. To simulate the noise over the trial we generate each new measurement of $g_{t,u}$ from a Bernoulli distribution with a mean equal to their respective $p_{t,u}$, set the value of $A = 500$ to reflect uncertainty in caloric intake of being within 500 calories, and set the variance of the Laplace noise of $w_{u,t}$ with a variance of 8 (parameter $b = 2$) derived from the empirical variance of weight measures observed in our data. Each simulated trial was run with 5 replicates. To ensure our estimation methods had sufficient observations to provide parameter estimates, we initialized each simulated trial with a two week run-in period where incentives were allocated at the same values they were disbursed in the Log2Lose trial. Thus from the second incentive given to each participant onwards, our optimization based methods began to differ from the incentives given by Log2Lose. To test how effective each method is with respect to intervention budget we ran simulated trials with 10 budget options in the range of \$520-\$5,857. The reason our range starts at \$520 is because this is the amount of money distributed to the group of participants in week 1 of the trial by Log2Lose (and thus during our simulated run-in period). We then constructed our range by increasing the budget by \$100 increments until we reached \$920. Since \$920 is approximately 15% of the total amount of incentives disbursed during Log2Lose, the remaining budgets we examined were at 20%, 40%, 60%, 80%, and 100% of the total amount spent which was \$5,857. When the budget was set to \$520, participants received no financial incentives after the first week regardless of the policy choice. We note that, because each participant received the same incentive as in the Log2Lose study the performance of the Log2Lose policy could not be evaluated at different budget

levels other then what was observed in the data.

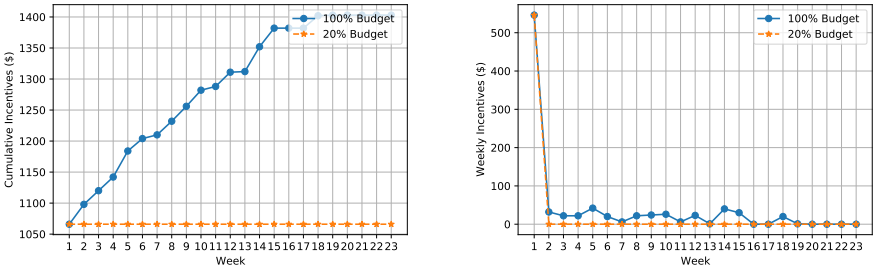
Figure 3.3a shows a comparison of the number of participants who achieved at least 5% weight loss with incentives provided by each of the different policies at different budget levels. Each optimization based approach is labeled as either indicator or hinge depending on the loss function used in its optimization; the percentage corresponds to the probability of the participant receiving the DIA incentive (with 100% corresponding to the DIA method). From this figure, we can see that our methods are able to achieve comparable performance to the Log2Lose policy with 20-60% of the budget spent during the Log2Lose trial. Moreover, from our simulation results, all optimization policies are able to assist nearly the whole participant cohort in achieving clinically significant weight loss when using 100% of the budget used by Log2Lose. This indicates that through our optimization-based approach, and predictive modeling, we are able to allocate incentives to participants when they are most likely to assist them in weight loss. Furthermore, since our approaches are personalized and not one-size-fits all, they are able to provide participants who are more externally motivated with greater incentives amounts to promote weight loss. This is in contrast to the one-size-fits-all approach, that is restricted in providing the same incentive schedule to all participants and thus spends some part of the budget on participants who may not need the added incentive to promote weight loss. Interestingly, the policy that is capable of achieving performance comparable to Log2Lose with the least amount of budget is a policy that only provides the DIA incentive with 75% probability and uses the hinge loss and not the deterministic DIA policy with the indicator loss. This indicates that by making the incentives intermittent-an approach consistent with psychological learning theory- weight loss behaviour can be promoted effectively and potentially more efficiently.

Figure 3.3b shows the average percentage of weight loss achieved by the five participants who lost the least percentage of weight over the 24 weeks. The results show both deterministic and randomized policies outperform the Log2Lose policy for a wide range of budgets, again reaffirming that, through personalization, resources can be spent on participants who are more likely to respond to finan-



(a) (b)

Figure 3.3: Number of participant achieving clinical weight loss success ($\geqslant$ 5% weight loss) (3.3a) and average percentage of weight loss across the bottom 5 participants who lose the least weight (3.3b) by the end of week 24 using 6 incentive policies and the randomized policy implemented in the Log2lose trial.



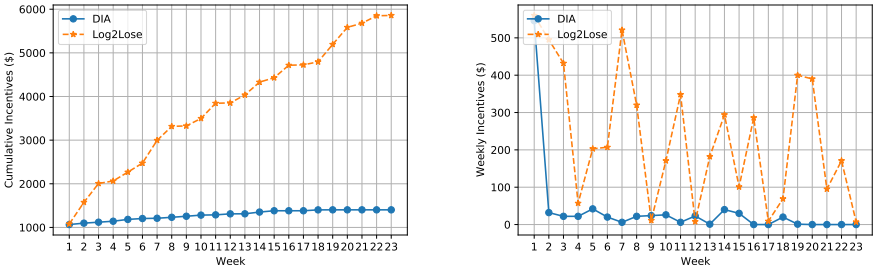
(a) (b)

Figure 3.4: Implementing the hinge loss function and deterministic incentive policy, Figure 3.4a shows the cumulative incentives distributed with 100% and 20% budget and Figure 3.4b shows the incentives distributed per week with 100% and 20% budget.

cial incentives and thus promote overall weight loss. Although the deterministic DIA policy guarantees each participant receives incentives and should prioritize weight loss by all participants with the hinge loss objective the randomized policies outperform the deterministic policies. Again, this reaffirms the effectiveness of intermittent incentives, and suggests that, in practice, a form of randomized policy could be effective in implementation.

Managerial insights

Our work provides several key insights to both the operators of Log2Lose and healthcare providers who would implement financial incentive-based interventions



(a) (b)

Figure 3.5: Figure 3.5a shows the cumulative incentives distributed using DIA with 100% budget and the cumulative incentives distributed in the original Log2lose study. Figure 3.5b shows the incentives distributed per week using DIA and the incentives distributed in the original Log2lose study.

for weight loss.

1. *Spending more of the budget on incentives early in the intervention improves outcomes.* The goal of providing financial incentives is to increase and maintain a participant's internal and external motivation for weight loss, and thus increase adherence to the intervention. Using our behavioral framework, we find that distributing larger incentives at the beginning helps increases a participant's external motivation, which increases a participant's weight loss for a moderate amount of time. Although such external motivation fades quickly, consistent weight loss success at the early stages can help increase the internal motivation, which has a long-lasting effect on increasing participant's adherence to the program. As a result, Figure 3.4a and 3.4b show that participants need significantly less incentives in the later weeks, and the surplus can be distributed to those who need extra incentive to increase their adherence to the program.

   Our findings are consistent with predictions from human behavioral literature, which suggests that larger incentives are more effective for individuals who are only starting to modify their behavior to induce behavioral changes and less effective for people who successfully incorporate the new behaviors into their lifestyle (Gneezy et al. 2011, Springer and Taylor 2016). The Log2Lose study attempted to provide greater rewards by providing participants with

$10 during each of the first four weeks, if they met incentive criteria (i.e., logged enough calories and/or lost weight, depending on randomization assignment). Our approach expands the Log2Lose approach by allowing greater amounts initially so as to increase extrinsic motivation, and thus weight loss.

2. *Moderate reward yielded large behavioral impact.* Once a significant amount of budget is distributed in the first weeks of the program, we find less incentives are required to keep participants losing weight. This eventually leads to moderate incentive reward on average across the entire intervention. The results in Figure 3.5a and 3.5b show $3-$4 on average per week per participant is sufficient for helping the entire group of participants to achieve 5% weight loss. Moderate rewards may be sufficient after initial weight loss as participants' motivation becomes more intrinsic as they succeed with weight loss. The greater intrinsic motivation may be sufficient to sustain their weight loss efforts throughout the trial.

3. *Intermittent rewards provide longer term benefit* (*i.e. implement a stochastic incentive policy over a deterministic one*).

   Our findings are consistent with research on reinforcement schedules and match the insights from the behavioral literature in that a random incentive scheme can induce more human efforts (Ederer et al. 2013). In our simulation study, random incentives lead to increasing numbers of participants achieving clinically significant weight loss success. The results in Figure 3.3a and Figure 3.3b show a randomized policy has the potential to outperform a deterministic policy. In particular, by implementing a randomized policy participants lose higher percentage of their initial weight.

3.6 Conclusion

In this paper, we develop a behavioral framework to design efficient and effective personalized financial incentives to help a large number of participants achieving

clinical weight loss success. This framework includes a behavioral model to describe the weekly decision process of a participant, a surrogate maximum likelihood estimation model for estimating model parameters, and an algorithm to optimize personalized financial incentives with limited budget. Under deterministic incentive policy, we show our estimated incentives converge to the optimal incentives which can be computed assuming we have full knowledge of each participant. Furthermore, we evaluate the performance of our personalized incentive design. The results show our approach outperforms existing machine learning methods in predicting weight loss success, and increases weight loss success with significantly less budget. In terms of healthcare practices, our framework can be applied to design personalized financial incentives, and it can be implemented with any deterministic or stochastic incentive policy for clinical weight loss programs.

4 APPLYING MACHINE LEARNING TO IDENTIFY PATTERNS OF ADHERENCE IN A BEHAVIORAL WEIGHT MANAGEMENT INTERVENTION WITH FINANCIAL INCENTIVES

4.1 Introduction

Obesity

Obesity, defined as a body mass index of at least 30 kg/m2, increases risk for health conditions such as type 2 diabetes, sleep apnea, arthritis, gallstones, and gallbladder disease (Bhaskaran et al. 2014, Ebbert et al. 2014). Data from the Centers for Disease Control and Prevention (CDC) indicates that the prevalence of obesity was 41.9% during the period of 2017-2020, with severe obesity (BMI of at least 35 kg/m^2) affecting 9.2% of the population (Stierman et al. 2021). The impact of obesity led to escalated costs across all categories of care, including inpatient, outpatient, and prescription drugs. In 2016, the cumulative obesity-related medical expenses among U.S. adults reached $260.6 billion (Cawley et al. 2021). Behavioral interventions apply behavior change techniques to help participants reduce caloric intake and increase physical activity. They usually start with a goal of achieving 5%-10% weight loss in 6 months (Kushner 2014). A behavioral intervention usually involves multiple behavior change techniques, including self-monitoring of weight, diet, and physical activity. Mobile-based applications can facilitate self-monitoring and have been found effective for stimulating behavior changes (Zhou et al. 2018).

Adherence Subtypes

Considering the availability, safety, and efficacy, behavioral intervention is the first-line treatment for obesity. Many studies have focused on increasing the effectiveness of such interventions, finding that adherence is strongly associated with participants' weight loss outcomes (Acharya et al. 2009). Most existing studies use a statistical approach to analyze the impact of different factors (e.g., social support)

on participants' adherence levels (Lemstra et al. 2016, Gibson and Sainsbury 2017). While these studies are capable of pinpointing certain crucial factors linked to adherence, the effects of these factors can greatly differ among individuals. Hence, we propose a data-driven, three-step framework to identify adherence subtypes based on individuals' responses and to predict adherence subtypes at the early stages of an intervention. This approach can help researchers identify participants who are more/less responsive to the intervention and adjust the intervention components, such as by adjusting sequential treatment options in a sequenced treatment to maximize the final weight outcomes.

Designing Sequenced Treatment for Weight Loss Intervention

Sequenced treatment is an approach in which multiple types of interventions or therapies are administered. Customizing treatment choices and levels for each individual is essential to optimize the potential effects of a sequenced treatment. We propose a framework to facilitate the design of sequenced treatments derived from an existing RCT for financial incentive-based weight loss programs. The approach consists of 3 steps. First, we define adherence based on participants' behavioral consistency. Then, after identifying the adherence subtypes, we proceed to analyze the correlation between adherence subtypes and the study's primary outcome (in most cases, the primary outcome would be weight loss). To predict adherence subtypes, potential features are chosen from treatment options (such as financial incentives) as well as demographic or clinical data provided by participants.

Financial Incentive-based Weight Loss Intervention

In this paper, we provide an application of our framework to the Log2Lose study, a randomized and single-blinded weight loss intervention that incorporates weekly financial incentives. The use of financial incentives has been shown to be effective for improving short-term weight outcomes (Volpp et al. 2008). The common approach is to distribute the same incentives to everyone who achieves the same criteria (e.g., lose weight by the end of some time period). We suggest that tailoring personalized

incentives based on individual differences will optimize weight loss outcomes while minimizing the intervention cost.

4.2 Methodology

In this section, we detail a three-step, data-driven framework for predicting an individual's long-term adherence to a financial incentives intervention for weight loss. Step 1: define weight loss intervention adherence; Step 2: identify adherence subtypes; and Step 3: predict each participant's adherence subtype. We explain each of these steps in Section 4.1 and then demonstrate how to apply them to the Log2Lose study in Section 4.2.

Overview

Since there are potentially multiple ways to define adherence to an intervention study, the first step is to provide a clear definition. In general, adherence can be defined as the behavioral consistency during the program, operationalized as the fraction of completed program components. In terms of the participant's behavioral consistency, we need to clarify the time interval and the behaviors that count.

Once we have clearly defined adherence to the study, the next step is to identify the adherence subtypes. We can achieve this by dividing participants into different clusters such that participants with similar adherence-related feature values will fall in the same cluster. Some of the common clustering methods include K-Means Clustering and Hierarchical Clustering (Burkardt 2009, Nielsen 2016). To determine the appropriate number of clusters, we can use the Elbow method for K-Means Clustering or plot the dendrogram for Hierarchical Clustering. However, these tests do not always guarantee the resulting clusters are the desired adherence subtypes. Further analysis may be necessary to assess whether there exist clear variances among adherence subtypes concerning financial incentives, response to incentives earned, and outcomes related to weight loss or weight maintenance.

The last step is to pick the best model for predicting an individual's adherence subtype. This selection process includes model training, feature selection, and evaluation of the model's performance with respect to some loss functions or performance metrics. Specifically, some of the interesting questions to look at during the evaluation include: Which subtypes are easier to predict? Which two subtypes are more likely to be mislabeled by each other? What's the tradeoff between the feature size and the prediction performance? Is there any correlation between the identified subtypes and individuals' weight outcomes?

Application: Log2Lose Case Study

Log2Lose is an ongoing randomized controlled trial at University of Wisconsin-Madison and Duke University. This single-blinded study involves 4 arms, where the difference across them is the presence and type of financial incentive. These 4 arms are labeled as Calorie Logging in which participants receive incentives for weekly calorie recording; Weight Loss in which participants receive incentives for weekly weight loss; Both in which participants receive incentives for self-monitory dietary and weight loss; and Neither in which participants receive no incentives. The duration of the study is 18 months, and it is divided into 3 phases. In Phase 1 (months 1-6), participants may receive weekly incentives if the last weight of the week is lower than the first weight of the week and/or if they record a sufficient number of calories on a sufficient number of days of the week (1000 for women, 1200 for men, at least 5 days of the week including one weekend day). In Phase 2 (months 7-12), participants may receive incentives if their weight by the end of each week is no more than 3 lbs compared to the weight achieved at the end of Phase I or if they log calories at least three days per week including a weekend day. In Phase III (months 13-18), participants cannot earn incentives for weight maintenance. For this case study, we used data from cohort 1 of Log2Lose. Cohort 1 involves 205 participants, and we only considered the records in Phase I and Phase II because participants can earn incentives (the variable in the 2x2 design) during these two phases. For the rest of this section, we explained how to apply

our proposed data-driven approach to identify and predict adherence subtypes in cohort 1 of Log2Lose.

First, we defined adherence as the number of weekly weight records and daily calorie logging records in Phase 1 and Phase 2. Next, we used the K-Means Clustering algorithm to identify the adherence subtypes. To do this, each time we take 1 data point (participant) as the holdout set, implement K-Means Clustering to classify the rest of participants, run the elbow test to determine the appropriate number of clusters, and then use the finalized cluster to find the subtype for the holdout participant. We repeat this process for all participants and check if there is any difference across the resulting clusters. If the resulting clusters are not the same, we may choose the clusters that are most frequently occurring as the most appropriate clustering. Otherwise, we can proceed to analyze the adherence subtypes using statistical tests.

To better understand the adherence subtypes, we first calculated the mean of each adherence-related feature (number of weekly weight records and daily calorie logging records) and plotted the trajectory of each feature for each of 4 adherence subtypes from week 0 to week 52. We compared the trajectories for weight records, calorie logging records, and financial incentives to identify the characteristics of each subtype. The plots can be found in Section 5.1. In addition, before analyzing the relationship between adherence subtypes and weight outcomes, we checked the percent of each arm included in each subtype to ensure there is no correlation between the subtypes and the study arms (Table 2). If there is any correlation between them, then the main difference across the resulting clusters from step 2 would be the study arms, and we could no longer establish the relationship between the predefined adherence and weight outcomes. If this happens, we need to go back to step 1 and update the definition for adherence. Next, we perform statistical analysis to determine the correlation between participants' adherence on weight loss and weight maintenance.

Lastly, we compared the performances of a set of Machine Learning models on predicting adherence subtypes. Our goal is to find the model and the set of features that provides the highest accuracy. We included a variety of linear, tree-

based, ensemble learning, and simple deep learning methods, including Logistic Regression, Random Forest, XGBoost, Linear SVM, and Neural Network. The set of feature candidates include study arms, weekly features (average weight per week, incentives received per week) and the following demographic features:

1. AGE: Patient Age
2. BP_MED: Do you have high blood pressure or are you on medications to control your blood pressure?
3. CPAP: Do you have sleep apnea or do you wear a CPAP at night to sleep?
4. DEPRESS: Have you ever been told by a healthcare provider that you have depression?
5. DIABETES: Do you have Type 2 Diabetes?
6. EDUCATION: What is the highest level of education that you have completed?
7. FINANCIAL: Which of the following categories best describes your financial situation?
8. RACE: Patient Race
9. SEX: Please set the sex listed on your original birth certificate.
10. TRY_LOSE_WEIGHT: Have you ever tried to lose weight before?
11. WEIGHT_ATTEMPT: How many serious attempts have you made?
12. WORK_STATUS: Which of the following best describes your current work status?

Next, we evaluated the performances of these ML algorithms using leave-one-out cross validation (LOOCV) and selected accuracy as the performance metric. For feature selection, our goal is to select a small set of features that are comparable to using the complete set of features. In addition, this process also helps remove

highly correlated features, which alleviate the concerns about correlation when selecting the set of feature candidates (e.g., WEIGHT_ATTEMPT is dependent on the condition of TRY_LOSE_WEIGHT). We started with training each type of ML model with the entire set of features and removed the one with least feature importance. Then, we trained each model again with the rest of features and removed the least important feature. We repeated these steps until removing any existing feature in the current set resulted in significant decrease in accuracy. The results can be found in Section 5.4.

4.3 Case Study Results

Adherence Subtypes

Following the procedures in Section 4.2, we found the same set of clusters for each training round, and the optimal number of adherence subtypes is k = 4. Next, we plotted the trajectories of the average number of calories, weight records per week, and the average amount of incentives received per week per cluster for Phase I and II (first 52 weeks) of the study. The trajectory of each cluster is colored in blue, orange, green, or red. Notice in Figure 1a there was a significant decline for one day, and when we examined it, it turned out to be the day after Thanksgiving day. This is potentially caused by the fact that most people chose not to restrict their caloric intake during the holiday. It will not affect the accuracy of our proposed adherence subtypes.

Figure 4.1a and 4.1c show the blue and orange subtypes are more consistent in submitting weight and calorie recording records while the green subtype are less adherent, and the red subtype barely adhere to the intervention. Next, we identified the difference between the blue and orange subtypes using financial incentives. In Figure 4.1b, we noticed the blue subtype received more incentives on average compared to the orange subtype, which rarely received incentives throughout the program. Therefore, we concluded both the blue and orange subtypes have relatively high adherence, but the blue subtype may be more motivated by

incentives, while the orange subtype is less so. Based on our observations, we named the orange subtype as the Self-motivated group and the blue one as the Reward-motivated group. On the other hand, the green subtype also received relatively high incentives, but the adherence level is significantly lower than the blue and orange subtypes, so we labeled the green subtype as the Low-adherent group. Lastly, since the red subtype failed to adhere to the intervention after the first 10 weeks and received minimal incentives across all weeks, we labeled it as the Non-adherent group.

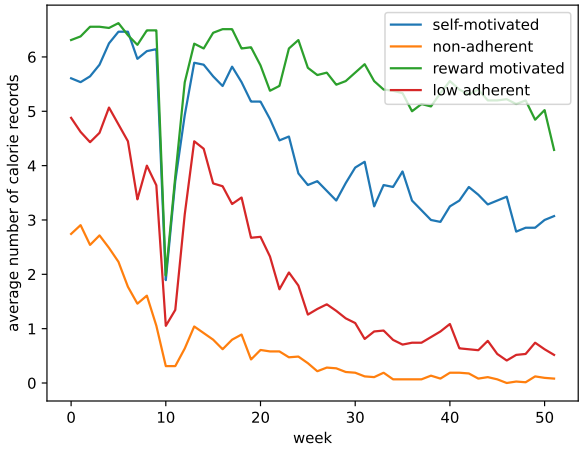
Table 4.1 shows the number of participants in each adherence group and the percent of participants in each group that belong to each of the four arms (Calorie Logging, Weight Loss, Both, Neither). Participants in reward-motivated and low-adherence groups qualified to receive at least 1 type of incentives; 27 out of 28 participants in self-motivated group received neither incentive; and the non-adherent group includes participants from all 4 arms.

	Neither	Calorie Logging	Weight Loss	Both	Total
Reward-motivated	0 (0%)	21 (46.7%)	20 (44.4%)	4 (8.9%)	45
Self-motivated	27 (96.4%)	0 (0%)	1 (3.6%)	0 (0%)	28
Low-adherent	0 (0%)	11 (19%)	16 (27.6%)	31 (53.5%)	58
Non-adherent	22 (29.7%)	19 (25.7%)	15 (20.3%)	18 (24.3%)	74
Total	49	51	52	53	205

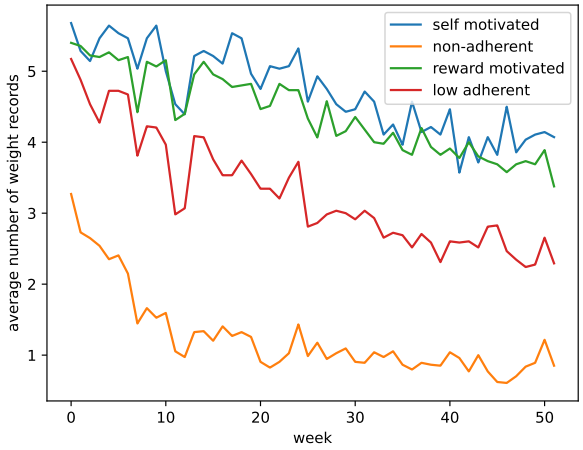
Table 4.1: Statistical Summary of Adherence Subtypes and Study Arms

Relationship between Adherence and Weight Outcomes

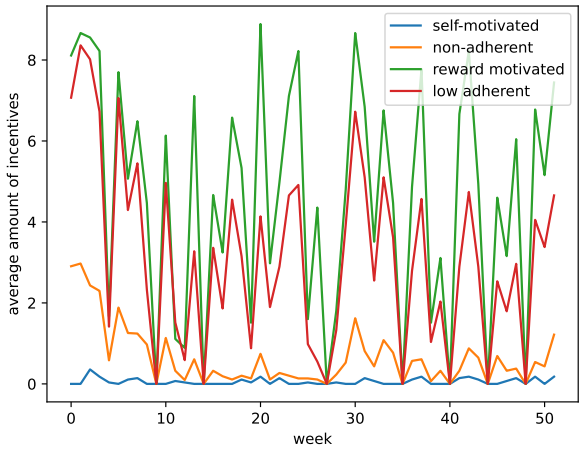
Next, we perform statistical analysis to determine the associations between participants' adherence on weight loss and weight maintenance. As shown in Table 4.2,



(a)



(b)



(c)

Figure 4.1: For the first 52 weeks: (a) Average number of calorie records per week per adherence subtype. (b) Average number of weight records per week per adherence subtype. (c) Average amount of incentives per week per adherence subtype.

there are significant differences in participants' weight loss and ability to maintain weight among the four adherence groups. Our results show that a larger percent of participants who are reward-motivated achieve clinically significant weight loss ($\geqslant$ 5%) at 6 months compared to the other adherence subtypes. As time progresses, we see an increasing number of participants achieve 5% weight loss in the self-motivated group, while around 9% of participants in the reward-motivated group who achieved 5% at 6 months fail to maintain it through 12 months. When comparing the average percent of weight loss achieved at the end of weight loss and maintenance periods, reward-motivated participants lose more weight on average compared to self-motivated participants.

Relationship between Adherence and Demographics

As shown in Table 4.3, we generally find no notable demographic differences among the adherence subtypes, except for AGE and DIABETES. However, there is a significant disparity in the number of serious weight loss attempts (WEIGHT_ATTEMPT) between the reward-motivated and self-motivated groups. The presence of significant differences in weight loss outcomes, coupled with the absence of distinct demographic variations across the adherence subtypes, suggests that one's adherence subtype is not solely associated with demographic factors.

	Reward-motivated	Self-motivated	Low-adherent	Non-adherent	p-value
AGE ($\mu \pm \sigma$)	$51.5 \pm 11.5^{w,x}$	$52.4 \pm 11.5^{y,z}$	$43.7 \pm 12.2^{w,y}$	$43.5 \pm 10.4^{x,z}$	< 0.01
BP_MED(%)	28.9%	28.6%	24.1%	29.7%	0.9
Yes	71.1%	71.4%	75.9%	70.3%	
No					
CPAP(%)	20%	25%	13.8%	23.0%	0.5
Yes	80%	75%	86.2%	75.7%	
No	0%	0%	0%	1.4%	
Other					

DEPRESS(%)	35.6%	25%	34.5%	32.4%	0.9
Yes	64.4%	75%	63.8%	63.5%	
No	0%	0%	1.7%	4.1%	
Other					
DIABETES(%)	2.2%	10.7%	6.9%	4.1%	<0.01
Yes	97.8%$^{\alpha}$	89.3%$^{\beta}$	93.1%$^{\phi}$	95.9%$^{\alpha,\beta,\phi}$	
No					
EDUCATION(%)	0%	7.1%	1.7%	2.8%	0.3
High school	20%	3.6%	17.3 %	13.5 %	
Trade/technical	31.1%	35.7%	29.3%	43.2%	
Bachelor	48.9%	53.6%	51.7%	40.5%	
Graduate					
FINANCIAL(%)	0%	0%	3.4%	2.8%	0.1
1-4:Worst-Best	2.3%	17.8%	12.1%	10.8%	
	24.4%	78.6%	82.8%	25.7%	
	73.3%	3.6%	1.7%	60.8%	
RACE(%)	84.4%	71.4%	72.4%	71.6%	0.4
White	11.1%	14.4%	13.8%	21.6%	
Black	0%	7.1%	6.9%	5.4%	
Asian	4.5%	7.1%	6.9%	1.4%	
Others					
SEX(%)	82.2%	78.6%	72.4%	85.1%	0.3
Female	17.8%	21.4%	27.6%	14.9%	
Male					
TRY_LOSE_	89.9%	96.4%	91.4%	89.2%	0.7
WEIGHT(%)	8.9%	3.6%	6.9%	5.4%	
Yes	2.2%	0%	1.7%	5.4%	
No					
Other					

WEIGHT_ ATTEMPT ($\mu \pm \sigma$)	$3.0 \pm 2.5^{\gamma}$	$4.6 \pm 2.4^{\gamma}$	5.4 ± 11.7	7.4 ± 17.3	0.28
WORK_	75.6%	71.4%	79.3%	75.4%	0.4
STATUS(%)	6.7%	3.6%	3.4%	10.8%	
Full-time	4.4%	10.7%	3.4%	4.1%	
Part-time	13.3%	14.3%	5.2%	4.1%	
Unemployed	0%	0%	0%	1.4%	
Retired	0%	0%	8.7%	2.8%	
Disabled	0%	0%	0%	1.4%	
Student					
Other					

Table 4.3: Statistical summary of demographic characteristics among adherence subtypes. Pairwise p-val: $\gamma = 0.01, w, x, y, z, \alpha, \beta, \phi < 0.01$.

Model Performance on Adherence Prediction

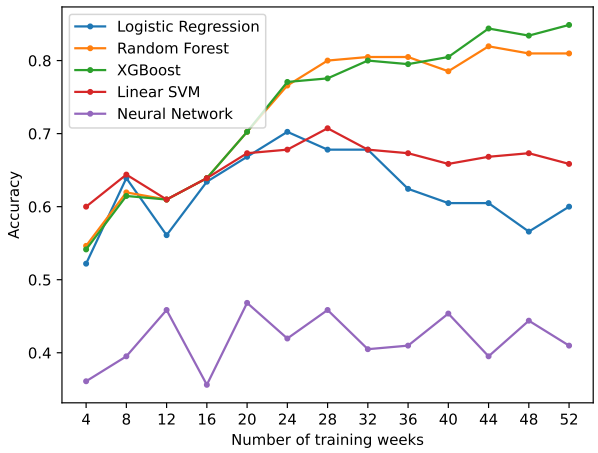
Figure 4.2a shows the accuracy of each ML model with the following features: average weight per week, incentives received per week, study arm, DEPRESS, CPAP, EDUCATION, WORK STATUS, and DIABETES. We found the Random Forest model to be the best option in this study, yielding 61.95% accuracy using the first 8 weeks of data, 70.24% accuracy using the first 20 weeks of data, and 80% accuracy using the first 28 weeks of data. This means we could predict each participant's long-term adherence with high accuracy by the end of Phase II using data from Phase I only.

When analyzing the average area under the curve (AUC), the Random Forest model demonstrates the following results: using the first 8 weeks of data, it yields a 95% confidence interval (CI) ranging from 0.76 to 0.89, with a mean of 0.83; using the first 20 weeks, the CI ranges from 0.88 to 0.95, with a mean of 0.91; and using

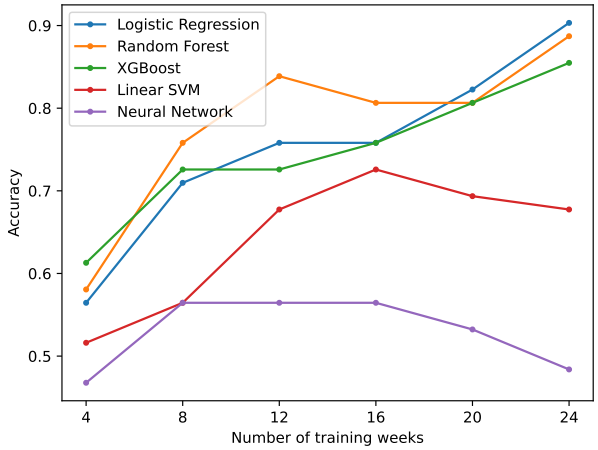
the first 28 weeks, the CI ranges from 0.91 to 0.97, with a mean of 0.94. Table 4.4 shows the one-vs-ALL AUC and average AUC for each ML model using the first 28 weeks of data as the training set. While the XGBoost and Random Forest models are similar in AUC performance, the Random Forest model achieves greater accuracy than the XGBoost model, with scores of 80% versus 77%, respectively.

Next, we assessed the predictive accuracy of machine learning models using the chosen features against those trained with the complete set of features plus two additional weekly variables (number of weight records per week, number of calorie records per week). The result can be found in Figure 4.2b. Since these two variables were also used to define adherence, we expected the performance of the models using these features to achieve the highest accuracy across all potential model candidates. Therefore, if the performance of our selected model is comparable to them, we further validated the prediction performance of our proposed model. Figure 4.2b shows the highest accuracy achieved is around 90% using the first 24 weeks of data. Our selected model achieves 76% accuracy using the same weeks of data. With a reduction of 14% in accuracy, the feature size is much smaller, and the model does not require the use of adherence features. In addition, by adding only 4 weeks of training data, our model's performance reached 80% accuracy, which is only 10% lower than the model with the complex feature set.

Last, but not least, we analyzed the prediction performance of our selected model for each adherence subtype. Tables 4.5-4.7 show the confusion matrices for 8, 20, and 28 training weeks. A confusion matrix is a performance measurement for binary or multi-classification problems. It is a 2x2 matrix for a binary classification problem, and a kxk matrix for a multi-classification problem where k is the number of classes. In our case study, the confusion matrix is a 4x4 matrix. The columns represent the predicted labels, and the rows represent the true labels. Each entry (i,j) represents the percent of total population in subtype i predicted as subtype j. The entries on the diagonal line of each matrix are the true positives. From Tables 4.5-4.7, we find the reward-motivated and non-adherent groups are easier to predict. However, the low-adherent and self-motivated groups are more likely to be mislabeled when the number of training weeks is small. It is important to correctly



(a)



(b)

Figure 4.2: (a) The accuracy rate of five ML models (Logistic Regression, Random Forest, XGBoost, Linear SVM, Neural Network) in predicting adherence subtypes using the selected features only. (b) The accuracy rate of five ML models (Logistic Regression, Random Forest, XGBoost, Linear SVM, Neural Network) in predicting adherence subtypes using the complete feature set plus the adherence-related variables.

identify participants in these subtypes because we want to avoid spending incentives on participants who are self-motivated and focus on distributing incentives to low-adherent participants who need significantly more incentives to achieve 5% weight loss. As we increase the number of training weeks to 28, the model achieves high accuracy in identifying the low-adherent and self-motivated groups. For example, for a participant whose true subtype is low-adherent, we can correctly identify the subtype with a probability of around 0.88.

4.4 Discussion

Empirical Results from the Case Study

Applying our general framework to the Log2Lose case study, we identified 4 adherence subtypes (reward-motivated, self-motivated, low-adherent, and non-adherent) that reflect participants' different responses to financial incentives. The reward-motivated group responded to financial incentives and achieved the highest adherence level; the self-motivated group did not require financial incentives and achieved the second highest adherence level; the low-adherent group also responded to financial incentives but had a low adherence level; and the non-adherent group did not respond to financial incentives and had the lowest adherence level. Our results are consistent with findings from existing studies that higher levels of adherence contribute to better weight outcomes. Our proposed ML model is able to predict participants' adherence subtypes with high accuracy. Specifically, the model excels at predicting reward-motivated and non-adherent groups. It is more challenging to distinguish self-motivated and low-adherent groups early in the program, but the model can predict these subtypes accurately with more training weeks. The correct identification of these subtypes could prevent researchers from distributing additional incentives to non-adherent participants who are not responsive to the intervention and distribute more incentives to reward-motivated participants to improve weight outcomes. Based on the predetermined intervention schedule, participants are eligible for a weekly incentive ranging from $0 to $10 if

they meet the criteria for dietary self-monitoring and/or weekly weight loss. By utilizing prediction results from the initial 8, 20, or 28 weeks of data, researchers could potentially save $782, $874, or $740 respectively across all participants in the study by withholding incentives from non-adherent participants identified as true positives. Alternatively, reallocating these funds to low-adherent participants could potentially enhance their weight loss outcomes with supplementary incentives. Moreover, this approach suggests that researchers could enroll more participants into the program within the existing budget. Through this case study, we demonstrated that the overarching framework effectively identifies and predicts adherence subtypes, offering valuable insights for designing cost-efficient interventions.

Implications for Designing Sequenced Treatments

To design an effective sequenced treatment, we propose to combine our framework with our incentive design algorithm developed for the Log2Lose pilot study Li et al. (2023). First, we proceed with the existing RCT and use the three-step framework to build the best ML model for predicting adherence subtypes. For the initial stage, participants can be randomly assigned to different arms. We propose an intervention that lasts eight weeks or longer. This extended time frame ensures the availability of adequate data for making reliable predictions regarding one's adherence subtype, allowing for appropriate treatment adjustments. It should be noted that the number of adherence subtypes may vary depending on the specific financial incentive-based intervention. Here, we elaborate on how to modify incentives based on the four subtypes identified in the Log2Lose study. If a participant is categorized as self-motivated, researchers should consider moving the participant to the control group. If a participant is non-adherent, researchers can reassign or randomize the participant to the control group or a group that receives different types of financial incentives. Participants, whether driven by rewards or exhibiting low adherence, should remain eligible for incentives. Our proposed incentive algorithm can then be used to customize the incentives for each individual accordingly. Using this approach offers the advantage of providing personalized incentives to

each participant. Additionally, the proposed incentive design can help participants achieve clinical weight loss success or improved weight outcomes, even with limited budget and resources.

Strengths and Limitations

Our approach can identify the adherence subtypes for any financial incentive based RCT, and it provides actionable results about how to adjust the financial incentives according to participants' adherence subtypes when designing sequenced treatments. For instance, one could assign self-motivated participants to the control group and non-adherent participants to a different treatment option. Another benefit of our approach is that it is easy to implement, and the prediction models can be built from existing classifier packages (e.g., scikit-learn). One limit of our approach is that it is not a causal framework. Additionally, our method depends on identifying adherence subtypes from at least one RCT, making it challenging to apply in the absence of prior RCTs. However, an alternative approach could involve leveraging existing RCTs with similar treatment and incentive structures. In such instances, we propose utilizing multiple RCTs to ensure the efficacy of the resulting subtypes and prediction model.

4.5 Conclusion

The study proposed a data-driven framework for predicting participants' adherence subtypes in financial incentive-based weight loss interventions. It revealed a strong correlation between adherence subtypes and participants' weight loss and maintenance outcomes, and it also demonstrated the proposed ML model's ability to predict adherence subtypes accurately. These findings validated the importance of identifying adherence subtypes for effective financial incentive interventions, and suggested that our approach is well-suited for designing sequenced treatments derived from weight loss RCTs. Our framework extends beyond weight loss interventions and can be applied to design sequenced treatments for various intervention

settings, such as Just-In-Time intervention (JITAI) for addictive behaviors and contingency management for substance use disorder (Yang et al. 2023, Petry 2011, Bolívar et al. 2021).Whether an intervention involves financial incentives or not, the framework can be utilized to identify the corresponding adherence subtypes within the context of the intervention and help researchers more effectively detect the key treatments and the appropriate levels of treatment for each participant.

	Adherence Subtype				χ^2/ANOVA
Weight Outcome	Reward-motivated	Self-motivated	Low-adherent	Non-adherent	p-value
Percent of participants who achieve 5% weight loss by the end of the 6-month weight loss period	$60\%^{a,b,c}$	$32.14\%^{a,d}$	$13.79\%^{b}$	$4.05\%^{c,d}$	< 0.01
Percent of participants who achieve 5% weight loss by the end of the 12-month weight maintenance period	$51.11\%^{e,f}$	$42.86\%^{g}$	$20.69\%^{e}$	$8.12\%^{f,g}$	< 0.01
Average percent of weight loss by the end of the 6-month weight loss period	$6.25\%^{h,i,j}$	$4.70\%^{h,k}$	$2.28\%^{i,l}$	$0.13\%^{j,k,l}$	< 0.01
Average percent of weight loss by the end of the 12-month weight maintenance period	$7.19\%^{m,n}$	$4.79\%^{o,p}$	$2.25\%^{m,o,q}$	$0.29\%^{n,p,q}$	< 0.01
Average number of weeks participants maintain their weight (weight achieved by the end of weight loss period) during the 6-month weight maintenance period (months 7-12)	$20.7^{r,s}$	$21.0^{t,u}$	$17.3^{r,t,v}$	$8.7^{s,u,v}$	< 0.01

Table 4.2: Statistical summary of the impacts of participants' adherence on weight loss and weight maintenance. Pairwise p-value: $a, i, m, q, r, t \leqslant 0.04$, $b - h, j - l, n - p, s, u, v < 0.01$.

	Reward-motivated	Self-motivated	Low-adherent	Non-adherent	Average AUC
Random Forest	0.97 [0.95, 0.99]	0.95 [0.92, 0.98]	0.90 [0.86, 0.94]	0.94 [0.91, 0.98]	0.94 [0.91, 0.97]
Logistic Regression	0.89 [0.84, 0.94]	0.86 [0.80, 0.93]	0.81 [0.74, 0.88]	0.78 [0.70, 0.85]	0.94 [0.77, 0.90]
XGBoost	0.95 [0.93, 0.98]	0.95 [0.92, 0.98]	0.91 [0.88, 0.95]	0.93 [0.89, 0.97]	0.94 [0.90, 0.97]
Linear SVM	0.85 [0.79, 0.92]	0.84 [0.77, 0.90]	0.72 [0.65, 0.79]	0.85 [0.79, 0.91]	0.82 [0.75, 0.88]
Neural Network	0.61 [0.52, 0.71]	0.62 [0.54, 0.70]	0.62 [0.54, 0.70]	0.75 [0.67, 0.83]	0.65 [0.57, 0.73]

Table 4.4: One-vs-All AUC for Random Forest, Logistic Regression, XGBoost, Linear SVM, and Neural Network models using the first 28 training weeks. The entries for each subtype contain the mean AUC and the 95% confidence interval of AUC.

Col: Predicted Labels Row: Ground Truth	Reward-motivated	Self-motivated	Low-adherent	Non-adherent
Reward-motivated	10.2%	0%	0%	3.4%
Self-motivated	0%	9.3%	10.2%	2.4%
Low-adherent	0%	6.3%	18.5%	3.4%
Non-adherent	4.4%	2.4%	5.4%	23.9%

Table 4.5: Confusion matrix of the Random Forest Model with 8 training weeks.

Col: Predicted Labels Row: Ground Truth	Reward-motivated	Self-motivated	Low-adherent	Non-adherent
Reward-motivated	10.7%	0%	0.5%	2.4%
Self-motivated	0%	13.7%	8.3%	0%
Low-adherent	0%	7.8%	18.0%	2.4%
Non-adherent	4.4%	0.5%	3.4%	27.8%

Table 4.6: Confusion matrix of the Random Forest Model with 20 training weeks.

Col: Predicted Labels Row: Ground Truth	Reward-motivated	Self-motivated	Low-adherent	Non-adherent
Reward-motivated	11.7%	0%	0.5%	1.5%
Self-motivated	0%	15.1%	6.8%	0%
Low-adherent	0%	3.4%	23.4%	0%
Non-adherent	4.4%	0.5%	1.5%	29.8%

Table 4.7: Confusion matrix of the Random Forest Model with 28 training weeks.

5 CONCLUSION

Patient-centered care, an approach aimed at delivering care that respects patient needs and values, is crucial for enhancing care quality and increasing patient satisfaction. Given the knowledge gap between patients and healthcare professionals, our work can enhance the comprehension of medical diagnoses and encourage patients' engagement.

We start with designing near-global aggregate explainers for complex diseases, which often go undiagnosed and untreated in developing countries. Our methodology helps improve diagnosis and awareness of such diseases, and it also allows direct trade-off among explainer coverage, fidelity, and interpretability. This trade-off can help health providers achieve a balance between transparency and the black box predictions. Then, we propose a behavioral framework which designs personalized financial incentives for enhancing clinical weight loss success among participants. The framework involves weekly decision-making, parameter estimation using a surrogate maximum likelihood estimation model, and an algorithm for incentive design. We demonstrate that our framework can contribute to customizing financial incentives with less budget, and it can accommodate different incentive policies. In particular, our work shows the potential of incorporating stochastic incentive policy to maximize individual's weight loss outcome. Next, we proceed to propose a framework to predict different adherence patterns among participants in incentive-based weight loss programs. We find a strong correlation between adherence subtypes and participants' success in losing and maintaining weight. We also suggest that by combining our proposed incentive algorithm and the identified adherence subtypes, healthcare researchers can design more efficient and effective sequenced treatments.

A ADDITIONAL PROOFS, EXPERIMENT RESULTS, AND ALGORITHM DETAILS IN CHAPTER TWO

A.1 Ethical Implications and Societal Impacts

Our aggregate explainer methodology provides explicit parameters that allow practitioners to clearly trade off among explainer coverage, fidelity, and interpretability. We note that in this trade off, low fidelity also results in low transparency, because the explanations offered by the explainer diverge significantly from the black box predictions that are meant to be explained. For example, explainers used for diagnostics might want to weigh more towards coverage, while explainers used for prediction transparency might want to weigh more towards fidelity. Navigating this tradeoff efficiently is critical to ensure that practitioners can correctly inform users or patients of the ML predictions. These contributions are particularly valuable in medical applications or other settings where informed consent is required.

Our methodology also contributes to providing better patient-centered care, which would be of particular value to low-income patients or patients in medically-underserved communities. Patient-centered care is an approach to provide care that is respectful of patient needs and values, and it has been found to be crucial for improving the quality of care and patient satisfaction. Since patients often have lower access to knowledge and understanding than their doctors, applying our work in the medical context will help them understand the complex medical diagnoses and take an active role in their health care. Complex diseases, such as Parkinson's disease and Alzheimer's disease, are often undiagnosed and untreated in developing countries. One major contributing factor is a lack of high quality diagnostic tests available for such diseases in low-income countries. Since these health problems usually require a multi-class classification setting, the fact that our methodology is well-suited for explaining multi-class predictions makes it valuable for improving the diagnosis and raising awareness among the general population.

A.2 Omitted Proofs

Proof of Prop 2. The first two constraints ensure proper local behavior of the local explainer as in Proposition 1. Thus we will focus the derivation of the final constraint. Using the definition of $\text{Fid}(\gamma, \mathcal{D})$ in Equation (2.2) and directly substituting variables for indicators, we can express the lower bound constraint as,

$$\min_{\{i: g_{i,r_i} \in \gamma\}} \frac{1}{|\mathcal{D}_{g_{i,r_i}}|} \sum_{x_j \in \mathcal{D}} \mathbb{1}[g_{i,r_i}(x_j) = f(x_j)] z_{ij} \geqslant \varphi.$$

Note that if the minimum over all explainers g_{i,r_i} must have fidelity of at least φ, then every local explainer must also have fidelity at least φ. This allows us to disaggregate this constraint across all $i \in [n]$. Consider the constraint for a single local explainer $g_{i,r_i} \in \gamma$. Using the definition of z_{ij}, note that $|\mathcal{X}_{i,r_i}| = \sum_{x_j \in \mathcal{X}} z_{ij}$. Thus the new lower bound fidelity constraint for a single explainer can be written as:

$$\frac{\sum_{j=1}^{n} \mathbb{1}[g_{i,r_i}(x_j) = f(x_j)] z_{ij}}{\sum_{j=1}^{n} z_{ij}} \geqslant \varphi. \tag{A.1}$$

Note that the denominator of the left hand side can only be zero when the numerator is also zero because $\sum_{j=1}^{n} z_{ij} \geqslant \sum_{j \in \mathcal{X}} \mathbb{1}[g_{i,r_i}(x_j) = f(x_j)] z_{ij}$. This means that we can multiply both sides of the inequality by the sum $\sum_{j=1}^{n} z_{ij}$ while still maintining its validity. Distributing φ and combining like terms gives us with the form of the constraint presented in the proposition statement. □

A.3 Clustering Methodology and PPMI Dataset

PD is a complex disorder, and is often expressed differently by different patients, which has motivated the need to create PD sub-types to better direct treatment. While many existing data-driven methods focus on clustering patients based on their baseline measurements (Fereshtehnejad and Postuma 2017), we propose clustering patients using the trajectory of how their symptoms progress.

We will use data collected in the PPMI study (Marek et al. 2011), which is a

long run observational clinical study designed to verify progression markers for PD. To achieve this aim, the study collected data from multiple sites and includes lab test data, imaging data, genetic data, among other potentially relevant features for tracking PD progression. The study includes measurements of all these various values for the participants across 8 years at regularly scheduled follow up appointments. The complete data set contains information on 779 patients, and included 548 patients diagnosed with PD or some other kind of Parkinsonism and 231 healthy individuals as a control group.

Determination of Criterion and Cluster Analysis

Since there is significant heterogeneity in how PD symptoms are expressed, there also is no agreement on a single severity score or measurement that can be used as a surrogate for PD progression. Thus instead of considering a single score, we will model the severity of the disease as a multivariate vector, and the disease progression as the trajectory of this vector through a multidimensional space. Using the PPMI data (Marek et al. 2011) and other previous literature on PD progression (Rao et al. 2006, Martinez-Martin et al. 2017, Bhat et al. 2018), we considered the following measures of severity to model disease progression:

- Unified Parkinson's Disease Rating Scale (UPDRS) II & III (Martínez-Martín et al. 1994): The UPDRS is a questionnaire assessment that is commonly used to track symptoms of PD by an observer. It consists of four major sections, each meant to measure a different aspect of the disease. These sections are: (I) Mentation Behavior and Mood, which includes questions related to depression and cognitive impairment; (II) Activities for Daily Living, which includes questions related to simple daily actions such as hygiene and using tools; (III) Motor Examination, which includes questions related to tremors and other physical ticks; and (IV) Complications of Therapy, which attempts to assess any adverse affects of receiving treatment. For our analysis we focused on the aggregate scores of sections II and III of the UPDRS to track physical symptoms of the disease.

- Montreal Cognitive Assessment (MoCA) (Nasreddine et al. 2005): Although not exclusively used for PD, the MoCA is a commonly used assessment for determining cognitive impairment and includes sections related to attention, executive functions, visual reasoning, and language. For our analysis, we used the MoCA scores of the individual patients as surrogates for their cognitive symptoms.

- Modified Schwab and England Activities of Daily Living Scale (MSES) (Siderowf 2010): The MSES is a metric used to measure the difficulties that individuals face when trying to complete daily chores due to motor deficiencies. This assessment is generally administered at the same time as the UPDRS and is often appended as a section V or VI. We used this score as a measure of how much autonomy the patients experience based on their symptoms.

We formed the empirical trajectory of these scores for each patient using the values measured during the patients' participation in the PPMI study (Marek et al. 2011). For our cluster analysis we used longitudinal measurements that were taken across the first seven visits of the study corresponding to a period of 21 months, where the first measurement formed the patient's baseline, and the next five measurements were taken at follow up visits at regular three month intervals; the final measurements were taken after six months. We chose this timeline for our analysis because participation was high among all participants in the study during this period, so we did not have to exclude any patients, and visits were more frequent to better capture disease progression over time. After these seven measurements, follow-up visits were scheduled too infrequently to provide useful trajectory modeling information.

We used these trajectories to cluster the patients together into progression sub-types. The main motivation for this approach is that if patients' severity scores progress in a similar way, then it may identify a useful sub-type for treatment design. Only patients diagnosed with PD were included in the cluster analysis, since we are interested in finding useful sub-types of disease progression. Each trajectory was then flattened out as a 28 dimensional vector, with the first four

entries corresponding the measurements at baseline, the next four for the 3 month follow up, and so on. Using scikit-learn and Python 3.7, we performed k-means clustering on these trajectories to define our sub-types (Pedregosa et al. 2011b, Friedman et al. 2001). Using cross validation and the elbow method (as seen in Figure A.8 in the appendix), we determined that there are four potential sub-types of disease progressions for the PPMI participants. We label these as: moderate physical symptoms cognitive decline cluster (Group 0), stagnant motor symptoms autonomy decline cluster (Group 1), motor symptom dominant cluster (Group 2), and moderate symptoms cluster (Group 3). The names we assigned to each individual cluster were given by the observed mean trajectories of the relevant scores for individuals that were classified into a particular cluster as shown in Figure A.1.

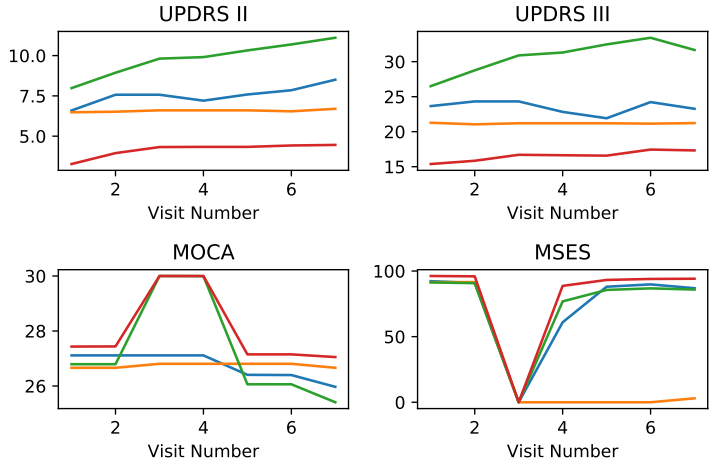


Figure A.1: Mean trajectory progression for given score by cluster. Blue corresponds to Group 0, orange corresponds to Group 1, green corresponds to Group 2, and red corresponds to Group 3. The y-axis of each plot the is numerical value of the corresponding disease severity measure.

In Figure A.2 we show two 2-dimensional projections of the different cluster groups. The plot on the top shows the projection onto the first two principal components of the data using PCA (Friedman et al. 2001); this projection method is meant to preserve linear relationships among data points as well as distances between data

points that are far apart. The projection shown in the bottom plot corresponds to the tSNE projection of the data onto a two-dimensional space (Maaten and Hinton 2008), this projection method was designed with manifolds in mind and is meant to preserve close distances (i.e., data points close in the tSNE projection should be also close in the higher dimensional space). Note that in both projections our resulting clusters are distinct and do not significantly overlap.

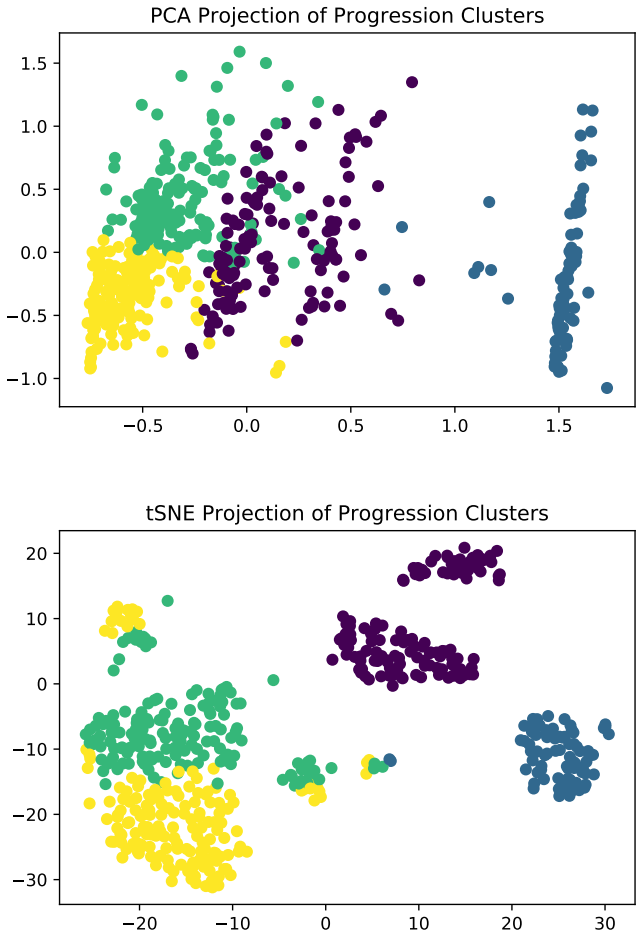


Figure A.2: PCA (top) and tSNE (bottom) 2-dimensional projections for visualizing trajectory clusters. Purple corresponds to Group 0, blue corresponds to Group 1, green corresponds to Group 2, and yellow corresponds to Group 3.

Validation of Clusters

To test whether these clustered sub-types provide additional insight into the health of the patients, we performed several statistical comparisons of each patients' characteristics at baseline across all four sub-types plus healthy patients, to determine if there were any statistically significant differences. The results and values of these comparisons are presented in Table A.1 below.

		Group 0	Group 1	Group 2	Group 3	Healthy	p value
Lymphocytes		1.643^{m}	1.749	1.642^{n}	1.704^{p}	1.850^{mnp}	0.01
REM Sleep Score		5.549^{de}	1.892^{dfgh}	5.969^{fij}	5.087^{gik}	3.247^{ehjk}	<0.001
UPDRS part II		6.594	6.482	7.981	3.272	N/A	<0.001
UPDRS part III		23.654	21.277	26.503	15.382	N/A	<0.001
Schwab & England Score		92.256	91.506	91.321	96.214	N/A	<0.001
Age		58.925^{a}	60.446	62.912^{abc}	58.387^{b}	59.571^{c}	0.02
	Anosmia	46	10	57	41	6	<0.001
Olfactory	Hyposmia	68	11	91	98	68	
	Normosmia	19	5	11	34	122	
Race White		95.49%	93.98%	94.34%	94.22%	94.37%	0.99
Gender Male		67.67%	57.83%	65.41%	63.01%	65.80%	0.63
Geriatric Depression Score		5.391	5.069	5.270	5.231	5.168	0.68

Table A.1: Comparison of baseline and screening measurements between clusters. p-values labeled in the table represent difference between all groups, and significant pairwise comparisons using a two sample T-test are marked by superscripts with p-values a-0.008; b-0.001; c-0.02;d,e,f,g,h,i,j,k-<0.001, m-0.003;n-0.004;p-0.04

As seen in Table A.1, many of the key screening measurements of the populations from the different clusters are significantly different, implying our clusters are informative about the health of individuals. In particular, we note that Group 0—which corresponds to moderate physical symptoms with cognitive decline—tends to be younger on average then the other groups, indicating this group may contain many more individuals with early onset PD. Moreover, the sub-types vary substantially in their sleep score and olfactory evaluation, which are both measures that have previously been shown to be strong indicators of PD (Rao et al. 2006) indicating that these progression sub-types are sensitive to these important predictors.

Overall, the comparisons shown in Table A.1 show that our data driven clusters are not only informative when comparing different forms of disease progression, but also correspond to variations in screening measurements. Based on this analysis, we believe that using screening data to predict these clusters could lead to clinically significant insights that can help with treatment.

A.4 Local Explainer Algorithm

After identifying the four disease progression sub-types, we would like to predict which kind of disease progression an individual might experience, given measurements collected during a screening visit. As we will show in our experiments in Section A.5, this task is best performed by complex black box models such as artificial neural networks (ANN) and bagged forests. This means that while the prediction may be accurate, it will not be easily explained, which make such models difficult to use for diagnosis recommendations. Our goal is to instead develop a method that trains simple auxiliary explainer models, and can still accurately describe the relationship between the data and the model output within a small region of a given prediction.

This methodology is known as *training local explainer models* and has been shown to be useful in understanding black box predictions (Ribeiro et al. 2016, 2018). One of the key tradeoffs in generating model explanations is that of *fidelity*—how well the explainer approximates the black box model—and *interpretability*—how easy it is for a practitioner to trace the predictions of the model. In contrast to previous literature which has proposed the use of regularization to achieve this goal, we propose directly computing locally significant features using an information filter. Generally, computing such filters can be computationally expensive and requires the use of numerical integration; however, one of our main contributions in this paper is to introduce an efficient algorithm for filtering out less significant features. This methodology will allow us to train local explainers that are significantly less complex then those that use regularization, with better fidelity.

Local Explainer Notation

Before proceeding to our discussion on the local explainer method, we will first establish some technical notation. We assume that for each patient $i = 1, ..., n$ we have an ordered pair (x_i, y_i), where $x_i \in \mathcal{X} \subseteq \mathbb{R}^m$ are the feature values of the patient and $y_i \in \mathcal{L} \subseteq \mathbb{Z}$ is the corresponding class label generated by a black box model f. Through our analysis we will also refer to this set of points through matrix notation where $X \in \mathcal{X}^n \subseteq \mathbb{R}^{m \times n}$ is the feature value matrix and $y \in \mathcal{L}^n \subseteq \mathbb{Z}^n$ is the vector of class labels, where each row in these matrices corresponds to a single patient's data. For our analysis we assume that $\mathcal{X}$ is a compact set. Let $\Phi = \{1, ..., m\}$ be the set of features, and it may also be used to denote the index set of the features. This set can be partitioned into two sets $\Phi_c, \Phi_b \subseteq \Phi$ that represent the set of continuous and binary features respectively.

Furthermore we define the set-valued function $\Phi^* : \mathcal{X} \to \Phi$ as the function which extracts the minimum set of necessary features to accurately predict the class of a point x. Namely,

$$\Phi^*(x) =_{\varphi \subseteq \Phi} \{|\varphi| : p(y|x) = p(y|x[\varphi])\}, \tag{A.2}$$

where $x[\varphi]$ is an indexing operation that maintains the values of x but only for the features in φ, and p is the conditional probability mass function of the labels y given the observation of some features. Specifically, if a feature index is not included $\Phi^*(x)$, then it is not required to understand the particular label of x. In addition, we will denote the ball around a point x of radius r with respect to a metric d as $\mathcal{B}(x, r, d)$.

Finally, a key feature of the explainer training method we propose includes the use of *mutual information*. In information theory, mutual information is a quantity that measures how correlated two random variables are with one another. If X, Y are two random variables with joint density p and marginal densities p_x, p_y, then

the mutual information between X and Y is denoted $I(X;Y)$ and calculated as:

$$I(x;y) = \mathbb{E}\log\frac{p(X,Y)}{p_x(X), p_y(Y)} = \int_x\int_y p(x,y)\log\frac{p(x,y)}{p_x(x), p_y(y)}dxdy. \tag{A.3}$$

If X and Y are independent then $I(X;Y) = 0$; otherwise $I(X;Y) > 0$, meaning that X contains some information about Y. A similar quantity can be computed using a conditional distribution on another random variable Z, known as the *conditional mutual information* and denoted $I(X;Y|Z)$.

Local Explainer Algorithm Description

Our main local explainer algorithm extends previous local explainer methods such as LIME (Ribeiro et al. 2016) by restricting the sampling region around the prediction, and including an information filter to ensure that fewer features are included in the final explainer mode.

Our general local explainer is formally presented in Algorithm 1, but we will give a brief overview of its operations here. The algorithm takes in hyper-parameters including number of points N to be sampled for training the explainer, a distance metric d, and a radius r around the point $\bar{x}$ being explained. First the algorithm samples N points uniformly from within a r radius of $\bar{x}$; we call this set of points $T(\bar{x})$. Depending on the distance metric being used this can often be done quite efficiently, especially if the features are binary valued or an ℓ^p metric is used (Barthe et al. 2005). Then using the sampled points, the algorithm uses the Fast Forward Feature Selection (FFFS) algorithm as a subroutine (formally presented in Section A.4 and Appendix A.7), which uses an information filter to remove unnecessary features and reduce the complexity of the explainer model. The FFFS algorithm uses an estimate of the joint empirical distribution of $(T(\bar{x}), f(T(\bar{x}))$ to select the most important features for explaining the model's predictions in the given neighborhood. We denote this set of features $\hat{\Phi}$. Then, using these features and the selected points, the explainer model g is trained by minimizing an appropriate loss function that attempts to match its predictions to those of the black box model. In principle

a regularization term can be added to the training loss of explainer g. However, through our empirical experiments in Section A.5 we found that FFFS typically selected at most five features, so even the unregularized models were not overly complex.

Fast Forward Selection Information Filter

A key step in our algorithm is the use of a mutual information filter to reduce the number of features that will be included in the training of the local explainer. Mutual information filters are commonly used in various signal processing and machine learning applications to assist in feature selection (Brown et al. 2012). However, these filters can be quite challenging to compute depending on the structure of the joint density function of the features and labels, and can require the use of (computationally expensive) numerical integration. We counteract this by considering an approximation of the density function, using histograms to calculate continuous features. When multiple combinations of features need to be considered as in our setting, the problem of finding the maximum-information minimum-sized feature set is known to be computationally infeasible (Brown et al. 2012). As such, our proposed method for computing the filter includes a common heuristic known as *forward selection*, which essentially chooses the next best feature to be included in the selected feature pool in a greedy manner. Using this method alone would still require recomputing the conditional distribution of the data based on previously selected features, which can result in long run times for large N. However, using some prepossessing techniques, we can show that these quantities can be stored efficiently using a tree structure, which allows quick computation of the filter.

The general idea of the FFFS algorithm is to consider the feature selection process as a tree construction. Part of this construction relies on an estimate of the empirical density of the features as a histogram with at most B bins and preprocessed summary tensor $M \in \{0,1\}^{B \times |\Phi| \times N}$ which indicates which bin of the histogram a feature value for a particular data point lays in. For each entry, $M[b, \varphi, x] = 1$ if the value of feature φ at point x falls in the bin b. Otherwise, $M[b, \varphi, x] = 0$. The depth of the

tree represents the number of selected features and each node of the tree is a subset of $T(\bar{x})$. For instance, at the beginning of the selection process, we have a tree with exactly one node R where $R = T(\bar{x})$. Assume binary feature φ_1 is selected in the first round. Then two nodes a, b are added under R, where $a = \{x_j : M[1, \varphi_1, j] = 1\}$ and $b = \{x_j : M(2, \varphi_1, j) = 1\}$. In the second round, we use the partition sets a, b to compute the mutual information instead of the complete set R. The set a is used for computing $\hat{p}(\varphi|\varphi_1 = 1)$, $\hat{p}(y|\varphi_1 = 1)$, and $\hat{p}(\varphi; y|\varphi_1 = 1)$, while b is used when the condition is $\varphi_1 = 2$. In each round the leaves $\mathcal{L}$ of the current tree represent the set of partition sets corresponding to all random permutation of selected features information. Therefore, $\mathcal{L}$ provides us sufficient information for calculating the desired mutual information. As shown in Algorithm 5, the algorithm only outputs the leaves $\mathcal{L}$, not the entire tree. The main algorithmic challenge is to efficiently calculate the marginal distributions $(\hat{p}(\varphi|S), \hat{p}(y|S)$ and joint distribution $\hat{p}(\varphi; y|S)$, which we are able to do using the tree structure.

The detailed structure of the FFFS algorithm used to compute the filtered feature set $\hat{\Phi}$ requires several subroutines, and the formal algorithmic construction for computing the filter is presented across Algorithms 3, 4, 5, and 6. The main FFFS algorithm is Algorithm 3, and it calls the subroutines for recursion (Algorithm 4), selecting features (Algorithm 5), and partitions (Algorithms 6). Formal presentation of these algorithms, as well as detailed descriptions, are given in Appendix A.7.

A.5 Experimental Validation of Local Explainer

In this section we empirically evaluate the quality of our local explainer methodology by first showing that accurate sub-type predictions of our PD sub-type clusters (as described in Section A.3) can be achieved using black-box methods applied to the data of individuals measured during the screening visit. We then apply our local explainer methodology developed in Section A.4 to explain the predictions given by these black-box models.

Our clusters were derived from longitudinal measurements of the four metrics

of disease severity described in Section A.3, measured across the first seven visits in the study over a period of 21 months. Treating these cluster (and the healthy patients) as our ground truth class labels, we first train black box machine learning models to predict which of these progression sub-types an individual will most likely experience given her screening data. This is meant to model the data available to a physician when she must make treatment decisions for a new patient. From screening data in the PPMI data set, we included the following 31 features: PTT, Lymphocytes, Hematocrit, Eyes, Psychiatric, Head-Neck-Lymphatic, Musculoskeletal, Sleep Score, Education Years, Geriatric Depression Score, Left Handed, Right Handed, Gender Male, Female Childbearing, Race White, Race Hispanic, Race American Indian, Race Asian, Race Black, Race PI, Anosmia, Hyponosmia, Normosmia, MRI Normal, MRI Abnormal Insignificant, MRI Abnormal Significant, BL/SC UPDRSII, BL/SC UPDRSIII, BL/SC MOCA, BL/SC MSES, and BL/SC Age. Among these 31 features, 20 features are binary variables and 11 features are continuous variables.

For accurate sub-type predictions using this data, in Section A.5 we trained three machine learning prediction models: one interpretable model (logistic regression) and two complex black box models (a feed forward ANN and a bagged forest). Our results indicate that the black box models outperform the simpler model, which necessitates the use of a local explainer method for this application to achieve both accurate classification and explainability.

In Section A.5 we computed local explanations based on the random forest model predictions (which was the model with the highest accuracy) using our proposed FFFS method with the information filter and a local explainer method. This is analogous to LIME (Ribeiro et al. 2016) which does not contain an information filter. Our results show that given a requirement of high explainer fidelity, the use of the information filter will result in less complex explainer models. All experiments described in this section were run on a laptop computer with a 1.2GHz Intel Core m3-7Y32 processor and MATLAB version R2019a with the machine learning and deep learning tool kits (MATLAB 2010).

Machine Learning Models for Cluster Prediction

We considered three different kinds of machine learning models for the task of predicting the progression cluster: logistic regression, feed forward ANN, and a bagged forest model. The patient data was split into training, validation, and testing sets with 70% of the data used for training, 15% for validation, and 15% for testing. Among 779 patients, 545 patients were selected for training, and 117 patients were selected for validation and testing.

Since bagged forests and ANNs are sensitive to hyperparamter settings, we used cross-validation to set their respective hyperparamters. Using cross validation and MATLAB's hyperparemeter optimization methods we found that the most effective ANN architecture for our task was with a single hidden layer containing one hundred hidden ReLu units. For the random forest model, we found that an ensemble of 50 bagged trees gave the best results compared to other forest sizes.

Figures A.3 and A.4 show the performance of the models on the same training, validation, and testing sets. In both figures, the classes 1-4 correspond to Groups 0-3, and class 5 corresponds to healthy patients (which we will also call Group 4). Figure A.3 contains the confusion matrix for each model. The rows of the matrix are the *output class*, which represents the predicted class, and the columns of the matrix are the *target class*, which is the true class. The cells on the diagonal of the matrix count accurate predictions. Each cell in the rightmost column has two values: the top number is the percentage of patients that are correctly predicted to each class, and the bottom number is the percentage of patients that are incorrectly predicted to each class. For each cell on the bottom row, the top number is the percentage of patients that belong to each class and is correctly predicted, and the bottom number is the percentage of patients that are incorrectly predicted. For the rest of cells in the matrix, the number in each cell counts for the number of patients that fall in this observation. The cell at the bottom right corner of each matrix shows the total percentage of patients that were correctly and incorrectly predicted.

As shown in Figures A.3 and A.4, the logistic regression model under-performs relative to the ANN and bagged forest models. Even though the bagged forest

model has a lower prediction rate for Group 0 compared to the ANN, it has equal or higher rates of accurate prediction for the other classes. Additionally, the bagged forest model consistently performed better than the ANN and logistic regression models in our experiments. We concluded from these results that the bagged forest classification model is the most effective for our prediction task, and we chose to consider its predictions when evaluating our local explainer method.

Local Explainer Validation

Since the main difference between our local explainer training algorithm and those in the literature is our use of the FFFS information filter, our experiments on the local explainer are focused on validating the effectiveness of using this information filter. We compare the performance of our local explainer training algorithm to a similar algorithm without a filtering step. We then compare the performance of these methods in terms of explainer complexity and fidelity, across different sampling radii and across all patients.

For the sampling parameters of our algorithm, we sampled $N = 10,000$ points centered around each patient within a radius r of either 3, 7, 11, or 15. The distance metric for computing this radius was a combination of the ℓ_∞ norm for the continuous features and the ℓ_1 norm for the binary features. The continuous value feature of each of the points was sampled uniformly using standard techniques (Barthe et al. (2005)). For binary valued features, we randomly chose at most r binary features and flipped their values. We first randomly generated an integer k between 0 and r, and randomly selected k binary features which we then flipped from their current value (that is, values of 1 were set to 0 and vice versa). To compute probability density estimates, we found that the method performed well with histograms with only three bins for continuous features and two bins for binary features. Intuitively three bins allows us to categorize feature values as low, medium, or high relative to their range.

For both training methods, we chose to train decision trees as our the local explainer class because these have been shown to be ergonomically suitable for

explaining black box models in healthcare contexts (Bastani et al. 2018). Then we computed the corresponding *fidelity score*, defined as the percentage of data where the prediction of the decision tree matched the prediction of the random forest model. We used the number of leaves on the decision tree as a measure of the explainer complexity.

In Figure A.5, we compare the explainer complexity and fidelity level of the explainers generated by the two different training methodology across the four different tested sampling radii. Unsurprisingly, when the sampling radius is small (i.e., $r = 3$), there is not much advantage to using the information filter in terms of reducing model complexity for a given fidelity level. Since all points are sampled so closely together, the relevant features are easily learned in explainer training. Conversely, when the sampling radius is large ($r = 15$), the addition of the information filter only helps slightly. With such a large radius, sampling feature values that are far from the point that is meant to be explained may not give useful information for that prediction. However, when considering the medium radius ranges, for high levels of fidelity, the inclusion of the information filter provides simpler models across the board. In particular, consider the plots corresponding for local explainer radius of $r = 7$ and $r = 11$ in Figure A.5. Note that in both of these figures, when considering high fidelity explainers generated by both methods (fidelity $\geqslant 0.6$), the explainers generated by the information filter method are less complex then those generated without the filter. This would indicate that using our information filter, we can obtain high fidelity local explainers that are on average less complex then those generated without this filter. When considering low fidelity explainers, the no filter method creates less complex models then the filter method. This is because our filter method is better equipped to find relevant features even in more complex regions of the black box model, while the no filter method is unable to learn these regions effectively with a fixed sample size. This is significant since this would indicate that our proposed methodology is able to explain a larger portion of the feature space using less complex models while still finding meaningful features for explanations, relative to existing methodologies.

Overall, the plots in Figure A.5 show that incorporating an information filter

into local explainer training can be more effective in extracting relevant features then using regularization, and can generate less complex models with high fidelity. In addition, these results indicate that using an information filter allows for local explainers with information filters to obtain higher fidelity over a larger radius with relatively less complex models. This is particularly significant since less complex models can be me more easily interpreted by domain experts, making it easier for them to translate the clinical significance of the black box model outputs. while larger explanation radii are useful for model validation and generalization of explanations. Moreover, even in complex decision regions generated by the black box model, using an information filter in conjunction with local explainers is better at extracting relevant features for predictions which again can be useful for model validation and providing clinical insights.

A.6 Comparison of Local Explainer Performance in Aggregate

To evaluate the performance of our proposed local explainer methodology in the context of explainer aggregation, we considered the impact on aggregate fidelity and coverage of our aggregate explainer using different base local explainers. For this experiment we used our IP methodology as the mode of aggregation and evaluated the difference between using our proposed information filter based local explainer (labeled in the plots as “filtered”) and LIME (labeled in the plots as “unfiltered”) as the base local explainers to be aggregated. For these experiments, we fixed the lower bound on fidelity of the IP at 70% and plotted both the coverage and fidelity of the aggregate with different explainer budgets for both binary prediction and multi-class prediction.

Figures A.6 shows the coverage and fidelity comparisons for the binary prediction class. We see that the use of our information-filter-based local explainer provides a better coverage and roughly 4% higher fidelity score then those obtained by our aggregation method in conjunction with LIME across all budget

levels. These results indicate that our proposed local explainer methodology leads to aggregate explainers that include both simpler component explainers, and can achieve improved coverage and fidelity in the binary classification case.

Figure A.7 shows the coverage and fidelity comparisons for the multi class prediction task. In this setting we again see that our proposed local explainer provides improved coverage and fidelity across all potential aggregate budgets. The advantage in the multi-class setting is less pronounced than in the binary prediction case, but our method still provides on average 5% improvement in coverage over LIME for the resulting aggregate explainer.

A.7 FFFS Algorithmic Details

In this appendix, we present and discuss the FFFS algorithm used in our local explainer method. The main algorithm is presented in Algorithm 3, and the required subroutines are presented in Algorithms 4, 5, and 6.

Since the main structure of the algorithm requires a recursive tree traversal, Algorithm 3 includes a general prepossessing wrapper algorithm that starts the recursion. In this part of the algorithm, the sampled data points are used to compute the empirical densities of their feature values. These densities are approximated using histograms which can vary in the number of bins. For simplicity of presentation, we assume each histogram has the same bin size, but of course this detail can be modified in implementation. The key addition here is the computation of tensor M, which tracks the inclusion of each data point's features into their respective histogram bin.

Algorithm 4 contains the main recursion of the filter computation, and it outputs the selected features when it terminates. The recursion of Algorithm 4 requires a set of selected features S, a set of unselected features U, the binary tensor M, the black box model predictions Y, and $\mathcal{L}$, which is a set of partition sets of points in $T(\bar{x})$. Since no features are selected prior to the first call to Algorithm 4, we initialize the inputs $S = \emptyset$, $U = \Phi$, $Y = f(T(\bar{x}))$ and $\mathcal{L} = T(x_i)$ when it is first called in Algorithm 3. The recursion terminates and outputs the current set of selected features when

either all features are selected or $\mathcal{L}$ becomes empty. If the termination condition is not met, Algorithm 4 calls Algorithm 5, which updates $S, U,$ and $\mathcal{L}$ using a bin expansion. Then Algorithm 4 makes a recursive call with updated inputs and repeat the previous steps.

Algorithm 5 is used to select one feature from the set of unselected features that maximizes the mutual information $I(\varphi; Y|U)$, and to update $\mathcal{L}$ given the current selected feature. We apply forward selection in Algorithm 5. In order to find $\varphi^* =_{\varphi \in U} I(\varphi; Y|U)$, we compute $I(\varphi; y|S)$ for each unselected feature φ. The approximated mutual information $I(\varphi; y|S)$ is computed using the following equation (Brown et al. 2012):

$$I(\varphi; y|S) \approx \hat{I}(\varphi; y|S) = \frac{1}{|T(\bar{x})|} \sum_{i=1}^{N} \log \frac{\hat{p}(\varphi; y|S)}{\hat{p}(\varphi|S)\hat{p}(\varphi|S)}.$$

If $I(\varphi^*; y|S)$ is not positive, then we do not select any new features. If no new feature is selected, we terminate the process by setting $U = \emptyset$, which satisfies the termination condition of Algorithm 4, and the feature selection process will be complete. If $I(\varphi^*; y|S) > 0$, then we can obtain additional information on the prediction by adding φ^* to the set of selected features S and removing it from the set of unselected features U. Algorithm 5 then calls Algorithm 6 to update $\mathcal{L}$ to $\mathcal{L}'$. Algorithm 6 is used to partition each set in $\mathcal{L}$ given current selected feature φ^*. Using the binary tensor M, we can collect the set of bins for φ^*. As an illustrative example of this process, let $B_{\varphi^*} = \{b_1, b_2\}$ and $\mathcal{L} = T(\bar{x}) = \{x_1, x_2, ..., x_p\}$. Assume $x_i^{\varphi^*} \in b_1$ for $i < 5$ and $x_i^{\varphi^*} \in b_2$ otherwise. Then we can partition the set $\{x_1, x_2, ..., x_p\}$ into 2 sets ℓ_1, ℓ_2 s.t. $\ell_1 = \{x_1, ..., x_4\}$ and $\ell_2 = \{x_5, ..., x_p\}$. Next we add sets ℓ_1, ℓ_2 to $\mathcal{L}'$. Since $\mathcal{L}$ contains exactly one set, we finish the partition process, and Algorithm 6 outputs $\mathcal{L}' = \{\{x_1, ..., x_4\}, \{x_5, ..., x_p\}\}$.

Proposition 12. *The time complexity of the FFFS algorithm for a fixed maximum discretization bin size is* $\mathcal{O}(N|\Phi|)$.

Proof. Note that the size of the generated points is given by the input parameter N, and the set of all features is denoted by Φ. First, since the bin sized is fixed as

Algorithm 3 Fast Forward Feature Selection (FFFS)

Require: $T(\bar{x}), \Phi, f$
for $\varphi \in \Phi_c$ **do**
  Form histogram with bin set B_φ and frequencies $\hat{p}_\varphi$
end for
set $M \in |B_\varphi| \times |\Phi| \times N$ as a zero tensor
for $x \in T(\bar{x})$ **do**
  for $\varphi \in \Phi$ **do**
    for $b \in B_\varphi$ **do**
      if $x[\varphi] \in b$ **then**
        Set $M[b, \varphi, x] = 1$
      end if
    end for
  end for
end for
return RecursionFFS$(\emptyset, \Phi, M, f(T(\bar{x})), T(\bar{x}))$

Algorithm 4 Recursion FFS

Require: $S, U, M, Y, \mathcal{L}$
if $U = \emptyset$ or $\mathcal{L} = \emptyset$ **then**
  return S
else
  $[S', U', \mathcal{L}'] =$ SelectFeature$(S, U, M, Y, \mathcal{L})$
  return RecursionFFS$(S', U', M, Y, \mathcal{L}')$
end if

a constant, and the preprocessing step requires a nested **for** loop, the total time complexity of the preprocessing is $\mathcal{O}(N|\Phi|)$. The FFFS algorithm operates as a tree traversal, where the depth of the tree at the final stage corresponds to the number of selected features. In each level of the tree, the mutual information of all points is evaluated using Algorithm 5 and the sets of generated points are partitioned into smaller sets using Algorithm 6, which combined require $\mathcal{O}(N)$ operations. Next, since in the worst case, all features contain positive mutual information on the prediction value of the black box model, the maximum possible tree depth is given by $|\Phi|$. Combining these two facts gives the desired result. □

Algorithm 5 Select Feature

Require: $S, U, M, Y, \mathcal{L}$

$f^* =_{f \in U} I(f; Y|U)$
if $I(f^*; Y|U) > 0$ **then**
    $U = U \setminus f^*$
    $S = S \cup f^*$
    $\mathcal{L}'$= BinPartition $(M, \mathcal{L}, f^*)$
    return $S, U, \mathcal{L}'$
else
    $U = \emptyset$
    return $S, U, \mathcal{L}$
end if

Algorithm 6 Bin Partition

Require: $M, \mathcal{L}, f^*$

Use M to find B_{f^*} s.t. $B_{f^*} = \{b_1, b_2, ..., b_k\}$ is the set of bins for feature f^*
$\mathcal{L}' = \emptyset$
for $\ell \in \mathcal{L}$ **do**
    Partition ℓ into smaller sets $\{\ell_1, \ell_2, ...\ell_k\}$ w.r.t B_{f^*}: $\ell_i = \{t \in l : t^{f^*} \in b_i\} \forall i \in \{1, ..., k\}$
    $\mathcal{L}' = \mathcal{L}' \cup \{l_1, ..., l_k\}$
end for
return $\mathcal{L}'$

A.8 Additional figures

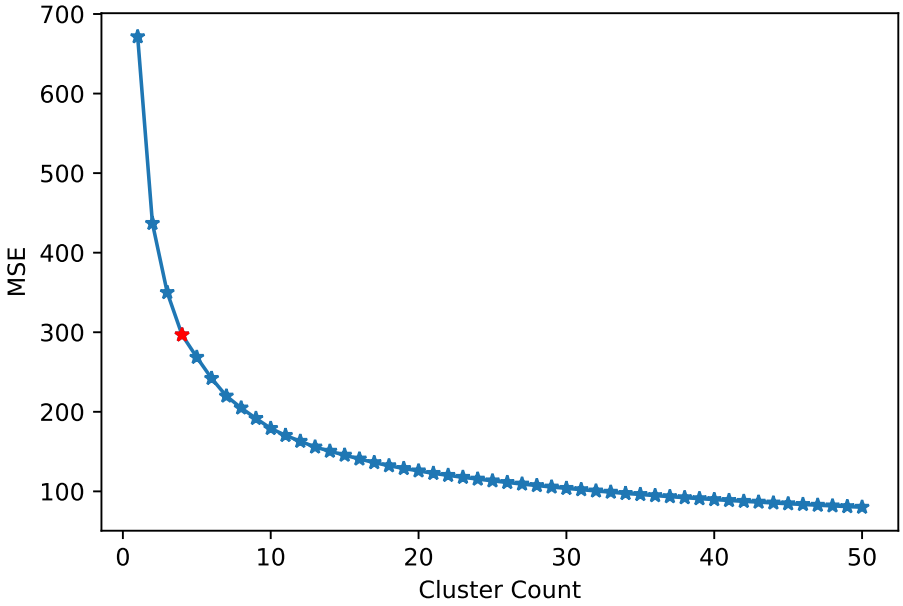
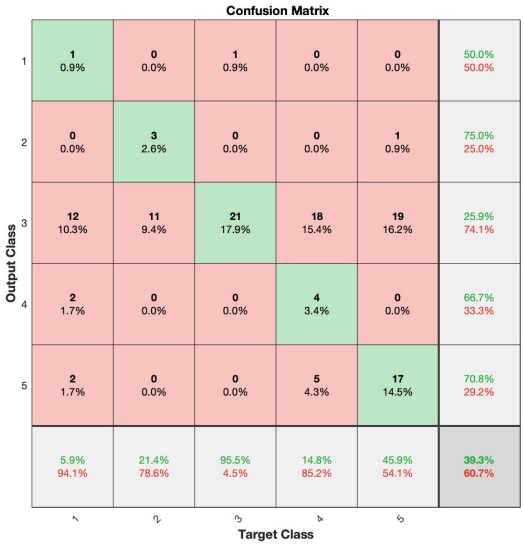
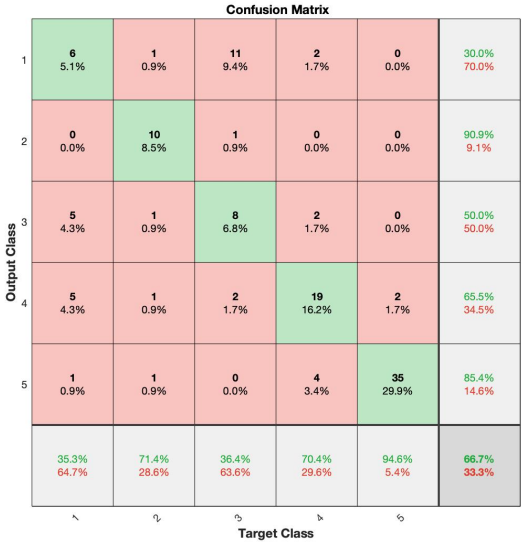
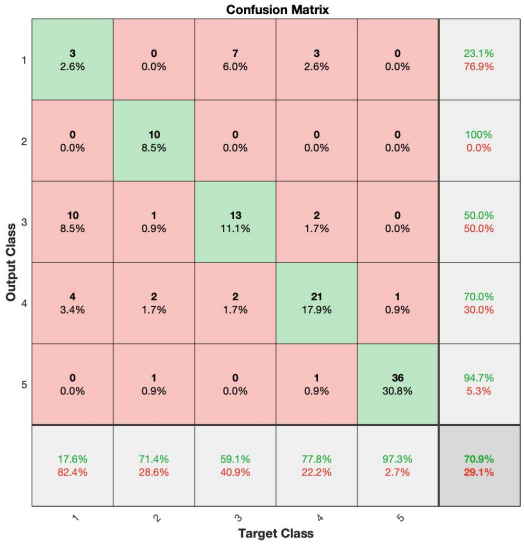
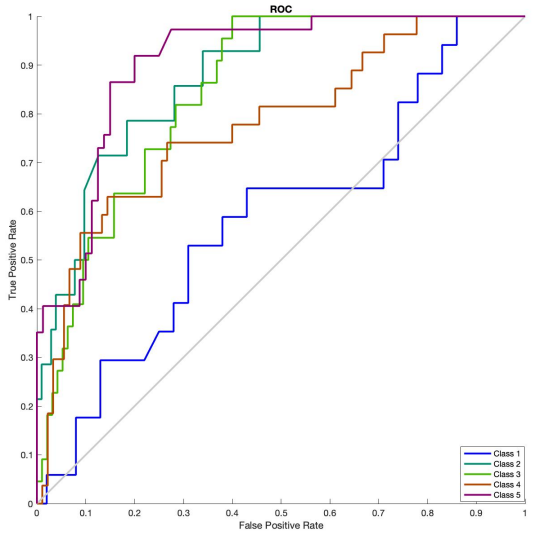


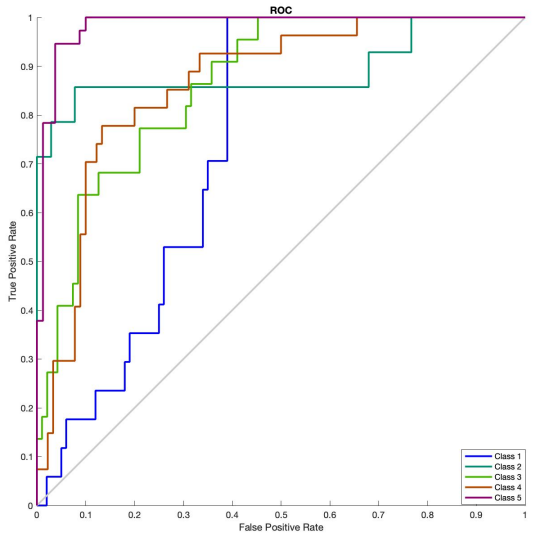
Figure A.8: Elbow plot for determining number of clusters to use for k-means clustering. Red marked value is located at 4 clusters and roughly corresponds to the bend in the elbow. The x-axis describes the total number of clusters used in k-means clustering, and the y-axis represents the MSE loss associated with the resulting clusters.


Confusion Matrix
Output Class
Target Class
1 0.9% 0 0.0% 1 0.9% 0 0.0% 0 0.0% 50.0% 50.0%
0 0.0% 3 2.6% 0 0.0% 0 0.0% 1 0.9% 75.0% 25.0%
12 10.3% 11 9.4% 21 17.9% 18 15.4% 19 16.2% 25.9% 74.1%
2 1.7% 0 0.0% 0 0.0% 4 3.4% 0 0.0% 66.7% 33.3%
2 1.7% 0 0.0% 0 0.0% 5 4.3% 17 14.5% 70.8% 29.2%
5.9% 94.1% 21.4% 78.6% 95.5% 4.5% 14.8% 85.2% 45.9% 54.1% 39.3% 60.7%


Confusion Matrix
Output Class
Target Class
6 5.1% 1 0.9% 11 9.4% 2 1.7% 0 0.0% 30.0% 70.0%
0 0.0% 10 8.5% 1 0.9% 0 0.0% 0 0.0% 90.9% 9.1%
5 4.3% 1 0.9% 8 6.8% 2 1.7% 0 0.0% 50.0% 50.0%
5 4.3% 1 0.9% 2 1.7% 19 16.2% 2 1.7% 65.5% 34.5%
1 0.9% 1 0.9% 0 0.0% 4 3.4% 35 29.9% 85.4% 14.6%
35.3% 64.7% 71.4% 28.6% 36.4% 63.6% 70.4% 29.6% 94.6% 5.4% 66.7% 33.3%


Confusion Matrix
Output Class
Target Class
3 2.6% 0 0.0% 7 6.0% 3 2.6% 0 0.0% 23.1% 76.9%
0 0.0% 10 8.5% 0 0.0% 0 0.0% 0 0.0% 100% 0.0%
10 8.5% 1 0.9% 13 11.1% 2 1.7% 0 0.0% 50.0% 50.0%
4 3.4% 2 1.7% 2 1.7% 21 17.9% 1 0.9% 70.0% 30.0%
0 0.0% 1 0.9% 0 0.0% 1 0.9% 36 30.8% 94.7% 5.3%
17.6% 82.4% 71.4% 28.6% 59.1% 40.9% 77.8% 22.2% 97.3% 2.7% 70.9% 29.1%





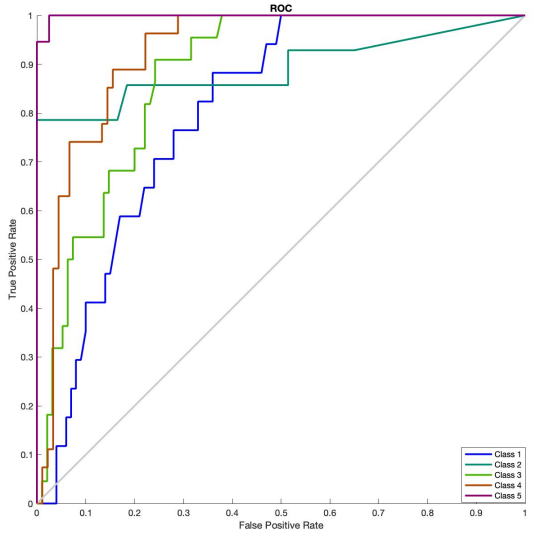


Figure A.4: ROC Curves: Logistics Regression (top), Neural Network (center) and Random Forest (bottom)

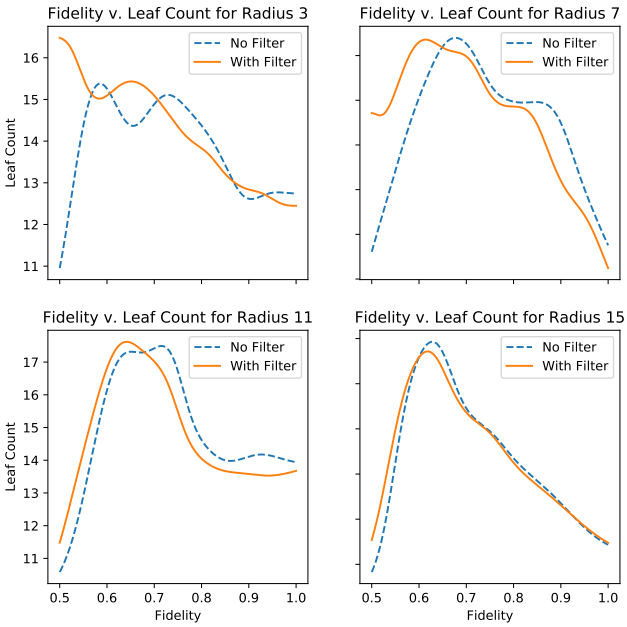


Figure A.5: Comparison of local explainer algorithm with the information filter (solid line) and without the the information filter (dashed line) for various different radius settings for the algorithms. The x-axis corresponds to the given fidelity score of the model and the y-axis measures the complexity of the decision tree explainer by the number of leaves. For a small radius ($r = 3$) and large radius ($r = 15$), the addition of an information filter does not lead to a significant difference in model complexity across all levels of fidelity. However, using the information filter in explainer training for moderate sized radii ($r = 7$ and $r = 11$) results in less complex models at higher levels of fidelity (> 0.6).

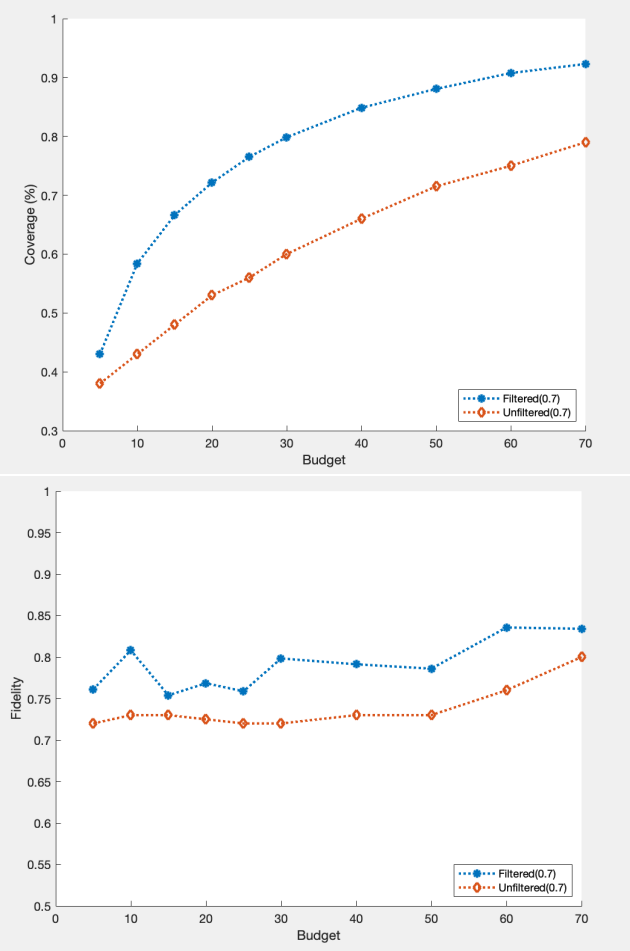


Figure A.6: 2-class fidelity (bottom) and coverage (top) plots for an IP based explainer aggregate using both an information filter based local explainer (labeled filtered) and LIME type local explainer (labeled unfiltered). The x-axis corresponds to the number of constituent local explainers that are used by the aggregation methods.

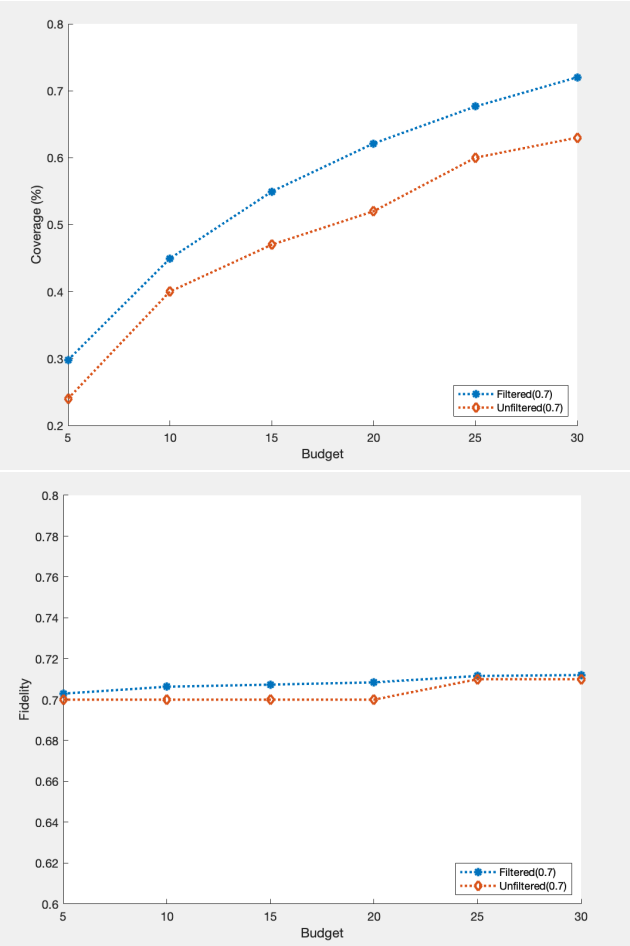


Figure A.7: Fidelity and coverage plots for an IP based explainer aggregate using both an information filter based local explainer (labeled filtered) and LIME type local explainer (labeled unfiltered). These plots are for a multiclass classification task. The x-axis corresponds to the number of constituent local explainers that are used by the aggregation methods.

B.1 Proofs of propositions in text

Proof of Proposition 4

To formulate this MLE problem recall that $\tilde{w}_{t,d} = w_{t,d} + \epsilon_{t,d}$, where $\epsilon_{t,d}$ are i.i.d. $\epsilon_{t,d} \sim \text{Laplace}(0, \sigma)$. Recalling from Section 3.2 that $g_t \sim \text{Bernoulli}(p_t)$, and letting $\mathcal{T}$ be the index set of all weeks in the study and $\mathcal{D}_t$ be the set of days during week $t \in \mathcal{T}$ that have weight observations, we can expand the joint likelihood function as follows:

$$\begin{aligned}
&\mathbb{P}(\{\tilde{w}_{t,d}, g_t\}_{t\in\mathcal{T}, d\in\mathcal{D}_t} | \{w_{t,d}, c_{t,d}, a_{1,t}, a_{2,t}, p_t, f_{b,t}, \hat{r}_t^w, B, k_1, k_2, k_p, r_t^w, r_t^c\}_{(t,d)\in\mathcal{T}\times[0,\dots,6]}) = \\
&\prod_{t\in\mathcal{T}} \big(\mathbb{P}(g_t|p_t)\mathbb{P}(f_{b,t}|f_{b,t-1}, \{c_{t,d}\}_{d\in D})\mathbb{P}(a_{1,t}|a_{1,t-1}, a_{1,b}, k_1, r_t^c, p_t, \{w_{t,d}\}_{d=0}^6, B) \\
&\mathbb{P}(a_{2,t}|a_{2.t-1}, a_{2,b}, k_2, r_t^w, \{w_{t,d}\}_{d=0}^6)\mathbb{P}(p_t|p_{t-1}, p_b, k_p, g_{t-1}) \\
&\mathbb{P}(\tilde{w}_{t,0}|w_{t,0})\mathbb{P}(w_{t,0}|w_{t-1,6}, c_{t,0})\mathbb{P}(c_{t,0}|w_{t-1,6}, a_{1,t}, a_{2,t}, f_{b,t}, \hat{r}_t^w)\big) \\
&\prod_{t\in\mathcal{T}, d\in\mathcal{D}_t} \mathbb{P}(\tilde{w}_{t,d}|w_{t,d}) \prod_{(t,d)\in\mathcal{T}\times[0,\dots,6]} \big(\mathbb{P}(w_{t,d}|w_{t,d-1}, c_{t,d})\mathbb{P}(c_{t,d}|w_{t,d-1}, a_{1,t}, a_{2,t}, f_{b,t}, \hat{r}_t^w)\big)
\end{aligned} \tag{B.1}$$

Note that many of the terms in the joint likelihood function are in fact degenerate distributions by the assumptions of the model in Section 3.2. Thus by taking the log of the above expression and expressing degenerate distributions as deterministic constraints we get the desired formulation.

Proof of Proposition 5

To prove the proposition we will solve the in week problem explicitly with dynamic programming. Let $V_{t,6}(w_{t,j})$ be the value function of a sub-problem maximizing the utility function from day $j \in \{0, .., 5\}$ to the end of day 6 (Sunday) of week t. We

want to show that:

$$V_{t,6}(w_{t,j}) = \max_{c_{t,j}} -a_{1,t}\left((\sum_{i=0}^{6-j} b^{i+1})w_{t,j-1} + (\sum_{d=j}^{6}(\sum_{i=0}^{6-d} cb^i)c_{t,j}) + (\sum_{d=j}^{6}(\sum_{i=0}^{6-d} b^i)k)\right)$$
$$+ a_{2,t}\hat{r}_t^w \frac{w_{t,\bar{d}} - b^{5-j}w_{t,j-1} - \sum_{d=j}^{6}(cb^{6-d}c_{t,d}) - \sum_{d=j}^{6}(b^{6-d}k) + A}{2A}$$
$$- \sum_{d=j}^{6}(c_{t,d}^2 - 2c_{t,d}f_{b,t} + \mathbb{E}[\xi_{t,d}^2] + f_{b,t}^2) \tag{B.2}$$

Because, if (B.2) is the correct structure, then the sub problems of the in-week model can be written as a sequence of convex optimization problems. First consider the base case $j = 5$:

$$\begin{aligned} V_{t,6}(w_{t,5}) &= \max_{c_{t,6}} \mathbb{E}[-a_{1,t}(bw_{t,5} + c(c_{t,6} + \xi_{t,5}) + k) + a_{2,t}\hat{r}_t^w \mathbb{1}\{w_{t,0} - w_{t,6} > 0\} - (c_{t,6} + \xi_{t,6} - (f_{b,t})^2] \\ &= \max_{c_{t,6}} -a_{1,t}(bw_{t,5} + cc_{t,6} + ck) + a_{2,t}\hat{r}_t^w \mathbb{P}(w_{t,0} - w_{t,6} > 0) \\ &\quad - (c_{t,6}^2 - 2c_{t,6}f_{b,t} + \mathbb{E}[\xi_{t,6}^2] + f_{b,t}^2) \end{aligned} \tag{B.3}$$

Since $\mathbb{P}(w_{t,0} - w_{t,6} > 0) = \mathbb{P}(\xi_{t,6} \leqslant \frac{w_{t,0} - bw_{t,5} - cc_{t,6} - k}{c})$ and $\xi_{t,6} \sim U(-A, A)$, $\mathbb{P}(\xi_{t,6} \leqslant \frac{w_{t,0} - bw_{t,5} - cc_{t,6} - k}{c}) = \frac{w_{t,0} - bw_{t,5} - cc_6 - k + A}{2A}$. Substituting this into (B.3):

$$\begin{aligned} V_{t,6}(w_{t,5}) = \max_{c_{t,6}} &- a_{1,t}(bw_{t,5} + cc_{t,6} + ck) + a_{2,t}\hat{r}_t^w \frac{w_{t,0} - bw_{t,5} - cc_6 - k + A}{2A} \\ &- (c_{t,6}^2 - 2c_{t,6}f_{b,t} + \mathbb{E}[\xi_{t,6}^2] + f_{b,t}^2) \end{aligned} \tag{B.4}$$

This proves the base case since (B.4) follows the desired form. Note, that this is a concave quadratic optimization problem, so a stationary point will be a global optimal solution. Next we take the derivative of the equation with respect to $c_{t,6}$ and set it equal to 0, we find the optimal solution $c_{t,6}^* = f_{b,t} - \frac{a_{1,t}c}{2} - \frac{a_{2,t}\hat{r}_t^w c}{4A}$.

Next we make the following inductive hypothesis for some $0 \leqslant j < 6$:

$$
\begin{aligned}
V_{t,6}(w_{t,j}) = \max_{c_{t,j+1}} & -a_{1,t}\left(\left(\sum_{i=0}^{6-j-1} b^{i+1}\right)w_{t,j} + \left(\sum_{d=j+1}^{6}\left(\sum_{i=0}^{6-d} cb^{i}\right)c_{t,j+1}\right) + \left(\sum_{d=j+1}^{6}\left(\sum_{i=0}^{6-d} b^{i}\right)k\right)\right) \\
& + a_{2,t}\hat{r}_t^{w}\frac{w_{t,0} - b^{5-j-1}w_{t,j} - \sum_{d=j+1}^{6}(cb^{6-d}c_{t,d}) - \sum_{d=j+1}^{6}(b^{6-d}k) + A}{2A} \\
& - \sum_{d=j+1}^{6}(c_{t,d}^2 - 2c_{t,d}f_{b,t} + \mathbb{E}[\xi_{t,d}^2] + f_{b,t}^2),
\end{aligned}
\tag{B.5}
$$

Then the $V_{t,6}(w_{t,j-1})$ can be computed as:

$$
\begin{aligned}
V_{t,6}(w_{t,j-1}) &= \max_{c_{t,j}} -a_{1,t}w_{t,j} - (c_{t,j} + \xi_j - f_{b,t})^2 + V_{w,6}(w_{t,j}) \\
&= \max_{c_{t,j}} -a_{1,t}(bw_{t,j-1} + cc_{t,j} + k) - (c_{t,j}^2 - 2c_{t,j}f_{b,t} + \xi_j^2 + f_{b,t}^2) \\
&\qquad - a_{1,t}\left(\left(\sum_{i=0}^{6-j-1} b^{i+1}\right)w_{t,j} + \left(\sum_{d=j+1}^{6}\left(\sum_{i=0}^{6-d} cb^{i}\right)c_{t,j+1}\right) + \left(\sum_{d=j+1}^{6}\left(\sum_{i=0}^{6-d} b^{i}\right)k\right)\right) \\
&\qquad + a_{2,t}\hat{r}_t^{w}\frac{w_{t,0} - b^{5-j-1}(bw_{t,j-1} + cc_{t,j} + k) - \sum_{d=j+1}^{6}(cb^{6-d}c_{t,d}) - \sum_{d=j+1}^{6}(b^{6-d}k}{2A} \\
&\qquad - \sum_{d=j+1}^{6}(c_{t,d}^2 - 2c_{t,d}f_{b,t} + \xi_{t,d}^2 + f_{b,t}^2) \\
&= \max_{c_{w,j}} -a_{1,t}\left(\left(\sum_{i=0}^{6-j} b^{i+1}\right)w_{t,j-1} + \left(\sum_{d=j}^{6}\left(\sum_{i=0}^{6-d} cb^{i}\right)c_{t,j}\right) + \left(\sum_{d=j}^{6}\left(\sum_{i=0}^{6-d} b^{i}\right)k\right)\right) \\
&\qquad + a_{2,t}\hat{r}_t^{w}\frac{w_{t,0} - b^{5-j}w_{t,j-1} - \sum_{d=j}^{6}(cb^{6-d}c_{t,d}) - \sum_{d=j}^{6}(b^{6-d}k) + A}{2A} \\
&\qquad - \sum_{d=j}^{6}(c_{t,d}^2 - 2c_{t,d}f_{b,t} + \xi_{t,d}^2 + f_{b,t}^2)
\end{aligned}
\tag{B.6}
$$

Which proves our claim that the structure of (B.2) holds for all days of the week as desired. To complete the proof and show that $c_{i,j}^*$ has the desired form, we can take the derivative of (B.6) with respect to $c_{t,j}$ and set it equal to 0, which yields

$c^*_{t,j} = f_{b,t} - \frac{a_{1,t} c \sum_{i=6-j}^{6} b^i}{2} - \frac{a_{2,t} \hat{r}^w_t c b^{6-j}}{4A}$ as desired.

Proof of Proposition 6

First we define two sets of binary variables $\{l_{1,t}\}_{t=0}^{23}$ and $\{l_{2,t}\}_{t=0}^{23}$. Using the Big-M technique Wolsey and Nemhauser (1999), let $l_{1,t} = 1$ if $w_{t,6} < w_{t,0}$ and $l_{1,t} = 0$ if $w_{t,6} \geqslant w_{t,0}$. Similarly, let $l_{2,t} = 1$ if $p_t \geqslant B$ and $l_{2,t} = 0$ if $p_t < B$. Constraint 3.8 enforces $l_{1,t} = 1$ if $w_{t,0} < w_{t,6}$ and $l_{1,t} = 0$ if $w_{t,0} \geqslant w_{t,6}$. Similarly, Constraint 3.9 enforces $l_{2,t} = 1$ if $p_t \geqslant B$ and $l_{2,t} = 0$ if $p_t < B$. Constraint 3.10-3.12 is the reformulation of $r^c_t \mathbb{1}\{p_w - B \geqslant 0\}$. If $l_{2,t} = 0$, then $z_{1,t} = 0$. If $l_{2,t} = 1$, $z_{1,t} = r^c_t$ since Constraint 3.11 is a tighter upper bound for $z_{1,t}$ than Constraint 3.10, and Constraint 3.12 ensures $z_{1,t}$ must be greater than or equal to r^c_t. Similarly, Constraint 3.13-3.16 ensures $z_{3,t} = 0$ if $l_{1,t} = 0$ and $z_{3,t} = k_1$ if $l_{1,t} = 1$. Then in Constraint 3.17 we replace the nonlinear terms with $z_{1,t}$ and $z_{3,t}$.

Proof of Proposition 7

Similar to the proof of Proposition 6, we use the Big-M technique Wolsey and Nemhauser (1999) to reformulate nonlinear constraints as linear ones. First we introduce the binary variables $l_{1,t}$, $l_{2,t}$, $z_{1,t}$, and $z_{2,t}$. Constraint (3.18) enforces $l_{1,t} = 1$ if $w_{t,6} < w_{t,0}$ and $l_{1,t} = 0$ if $w_{t,6} \geqslant w_{t,0}$. Constraint (3.19)-(3.22) are the reformulation of $k_2 \mathbb{1}\{(w_{t,0} - w_{t,6}) > 0\}$, which indicates $z_{2,t} = 0$ if $l_{1,t} = 0$ and $z_{2,t} = k_2$ if $l_{1,t} = 1$. Constraint 3.19-3.20 ensures $z_{2,t} \leqslant k_2$, and Constraint 3.21 ensures $z_{1,t}$ must be greater than or equal to r^c_t. Lastly, in Constraint 3.23 the product of the binary and the continuous variables is replaced with $z_{2,t}$.

Proof of Proposition 8

We prove the inequalities hold for $g_t = 0$ and $g_t = 1$ separately.

If $g_t = 0$, then the log-likelihood function $-\log(\mathbb{P}(g_t = 0|p_t)) = -\log(p_t^{g_t}(1 - p_t)^{1-g_t}) = -g_t \log(p_t) - (1-g_t)\log(1-p_t) = -\log(1-p_t)$ and $|g_t - p_t| = |0 - p_t| = p_t$. Let $f : [\epsilon, 1-\epsilon] \mapsto \mathbb{R}$ be: $f(x) = \frac{-\log(1-x)}{x} p_t + \log(1-p_t)$. Note, $f(p_t) = 0$.

Computing the first derivative of f yields $\frac{df}{dx} = \frac{\frac{x}{1-x}+\log(1-x)}{x^2}$. Note that $\frac{df}{dx} > 0$, meaning f is monotonically increasing in x, and thus $f(\epsilon) \leqslant f(p_t) \leqslant f(1-\epsilon)$ which gives the desired inequalities.

If $g_t = 1$, then $\log(\mathbb{P}(g_t = 1|p_t)) = -\log(p_t)$ and $|g_t - p_t| = 1 - p_t$. Define $h : [\epsilon, 1-\epsilon] \mapsto \mathbb{R}$ as $h(x) = \frac{-\log(x)}{1-x}(1-p_t) + \log(p_t)$, note $h(p_t) = 0$. We can compute the first derivative of h as $\frac{dh}{dx} = \frac{x+x(-\log(x))-1}{x(1-x)^2}$, and note that $\frac{dh}{dx} < 0$ meaning h is monotonically decreasing. Therefore, $h(1-\epsilon) \leqslant h(p_t) \leqslant h(\epsilon)$ which provides the desired result.

Proof of Proposition 9

Let $(w^*_{0,0}, \theta^*_0)$ be the true initial conditions. Then for any possible initial conditions $(w_{0,0}, \theta_0) \neq (w^*_{0,0}, \theta^*_0)$ we can express the surrogate posterior as follows:

$$\log(\hat{\mathbb{P}}(w_{0,0}, \theta_0|\{\tilde{w}_{t,d}, g_t, r^w_t, r^c_t\})) = \log(\hat{\mathbb{P}}(w^*_{0,0}, \theta^*_0|\{\tilde{w}_{t,d}, g_t, r^w_t, r^c_t\})) + \sum_{t\in\mathcal{T}, d\in\mathcal{D}_t} \log\frac{\mathbb{P}(\bar{w}_{t,d} - \tilde{w}_{t,d})}{\mathbb{P}(w_{t,d} - \tilde{w}_{t,d})} + \sum_{t\in\mathcal{T}}(|g_t - \bar{p}_t| - |g_t - p^*_t|) - \log\frac{\mathbb{P}(w_{0,0}, \theta_0)}{\mathbb{P}(w^*_{0,0}, \theta^*_0)} \tag{B.7}$$

Using the results of Proposition 8, we can bound $\sum_{t\in\mathcal{T}}(|g_t - p^*_t| - |g_t - \bar{p}_t|) \leqslant \epsilon_{max}\sum_{t\in\mathcal{T}}\frac{\log(\mathbb{P}(g_t|p^*_t))}{\log(\mathbb{P}(g_t|\bar{p}_t))}$, where $\epsilon_{max} = \max\{\frac{\epsilon}{\log(1-\epsilon)}, \frac{1-\epsilon}{\log(\epsilon)}\}$. Thus we see:

$$\text{(B.7)} \leqslant \log(\hat{\mathbb{P}}(w^*_{0,0}, \theta^*_0|\{\tilde{w}_{t,d}, g_t, r^w_t, r^c_t\})) + \sum_{t\in\mathcal{T}, d\in\mathcal{D}_t} \log\frac{\mathbb{P}(\bar{w}_{t,d} - \tilde{w}_{t,d})}{\mathbb{P}(w_{t,d} - \tilde{w}_{t,d})} + \epsilon_{max}\sum_{t\in\mathcal{T}}\frac{\log(\mathbb{P}(g_t|p^*_t))}{\log(\mathbb{P}(g_t|\bar{p}_t))} - \log\frac{\mathbb{P}(w_{0,0}, \theta_0)}{\mathbb{P}(w^*_{0,0}, \theta^*_0)} \tag{B.8}$$

Since $\frac{\mathbb{P}(w_{0,0}, \theta_0)}{\mathbb{P}(w^*_{0,0}, \theta^*_0)}$ is a constant and $\log(\hat{\mathbb{P}}(w^*_{0,0}, \theta^*_0|\{\tilde{w}_{t,d}, g_t, r^w_t, r^c_t\})) \in [0, 1]$ by definition, then combined with Assumption **??** this implies $\max_{S(\delta)} \log(\hat{\mathbb{P}}(w_{0,0}, \theta_0|\{\tilde{w}_{t,d}, g_t, r^w_t, r^c_t\})) \to -\infty\ \forall\delta > 0$. This implies $\max_{S(\delta)} \hat{\mathbb{P}}(w_{0,0}, \theta_0|\{\tilde{w}_{t,d}, g_t, r^w_t, r^c_t\}) \to 0$.

To complete the proof consider the probability mass placed on $S(\delta)$ given by $\hat{\mathbb{P}}(S(\delta)|\{\tilde{w}_{t,d}, g_t, r^w_t, r^c_t\}) = \int_{S(\delta)} \hat{\mathbb{P}}(w_{0,0}, \theta_0|\{\tilde{w}_{t,d}, g_t, r^w_t, r^c_t\})dw_{0,0}d\theta_0 \leqslant \text{Vol}(\mathcal{W} \times$

$\subseteq) \max_{S(\delta)} \hat{\mathbb{P}}(w_{0,0}, \theta_0 | \{\tilde{w}_{t,d}, g_t, r_t^w, r_t^c\}) \to 0$. Thus our surrogate posterior meets the definition as desired.

Proof of Corollary 3.2

Since the event $\left\{(\hat{w}_{0,0}^{MAP}, \hat{\theta}_0^{MAP}) \notin \mathcal{B}((w_{0,0}^*, \theta_0^*), \delta))\right\}$ is a subset of the event $\left\{ \max_{S(\delta)} \hat{\mathbb{P}}(w_{0,0}, \theta_0 | \{\tilde{w}_{t,d}, g_t, r_t^w, r_t^c\}) \geqslant \max_{w_{0,0}, \theta_0 \in \mathcal{B}((w_{0,0}^*, \theta_0^*), \delta)} \hat{\mathbb{P}}(w_{0,0}, \theta_0 | \{\tilde{w}_{t,d}, g_t, r_t^w, r_t^c\})\right\}$, which implies $\mathbb{P}((\hat{w}_{0,0}^{MAP}, \hat{\theta}_0^{MAP}) \notin \mathcal{B}((w_{0,0}^*, \theta_0^*), \delta)) \leqslant \mathbb{P}(\max_{S(\delta)} \hat{\mathbb{P}}(w_{0,0}, \theta_0 | \{\tilde{w}_{t,d}, g_t, r_t^w, r_t^c\}) \geqslant \max_{w_{0,0}, \theta_0 \in \mathcal{B}((w_{0,0}^*, \theta_0^*), \delta)} \hat{\mathbb{P}}(w_{0,0}, \theta_0 | \{\tilde{w}_{t,d}, g_t, r_t^w, r_t^c\}))$. By Proposition 9, $\mathbb{P}(\max_{S(\delta)} \hat{\mathbb{P}}(w_{0,0}, \theta_0 | \{\tilde{w}_{t,d}, g_t, r_t^w, r_t^c\})) \to 0$ as $\mathcal{T} \to \infty$ and hence the result of the corollary follows.

Proof of Proposition 10

Propositions 6 and 7 indicate the problem described in (3.33) can be reformulated with a set of linear constraints which are affine in $(w_{u,0,0}, \theta_{u,0} \{r_{u,t}^w, r_{u,t}^c\}_{t=0}^T) \forall u \in \mathcal{U}$. This implies the function $\psi(\{w_{u,0,0}, \theta_{u,0}, \{r_{u,t}^w, r_{u,t}^c\}_{t=0}^{T+n}\}_{u \in U})$ is lower semi-continuous to each argument by applying results from Hassanzadeh and Ralphs (2014).

Proof of Proposition 11

Corollary 3.2 implies the surrogate posterior estimates $\hat{\mathbb{P}}(w_{u,0,0}, \theta_{u,0} | \{\tilde{w}_{u,t,d}, g_{u,t}, r_t^w, r_t^c\}_{t=0}^T)$ are statistically consistent and Proposition 10 implies $\psi(\{w_{u,0,0}, \theta_{u,0}, \{r_{u,t}^w, r_{u,t}^c\}_{t=0}^{T+n}\}_{u \in U})$ is lower semi-continuous in all of its arguments. Hence by Proposition 2.1.ii of Lachout et al. (2005) $\psi(\{\hat{w}_{u,0,0}^T, \hat{\theta}_{u,0}^T, \{\bar{r}_{u,t}^w, \bar{r}_{u,t}^c\}_{t=0}^{24}\}_{u \in U})$ is a lower semi-continuous approximation of the function $\psi(\{w_{u,0,0}^*, \theta_{u,0}^*, \{\bar{r}_{u,t}^w, \bar{r}_{u,t}^c\}_{t=0}^{24}\}_{u \in U})$ with respect to the true initial conditions.

Proof of Theorem 3.3

Since Corollary 3.2 implies $(\hat{w}_{0,0}^{MAP}, \hat{\theta}_0^{MAP}) \xrightarrow{p} (w_{0,0}^*, \theta_0^*)$ and Proposition 11 implies $\psi(\hat{w}_{u,0,0}^T, \hat{\theta}_{u,0}^T, \{\bar{r}_{u,t}^w, \bar{r}_{u,t}^c\}_{t=0}^{24}\}_{u \in U})$ is a lower semi-continuous approximation to the

function $\psi(\{w^*_{u,0,0}, \theta^*_{u,0}, \{\bar{r}^w_{u,t}, \bar{r}^c_{u,t}\}_{t=0}^{24}\}_{u\in U})$, the result follows Theorem 4.3 of (Lachout et al. (2005)) which implies any solution $\{r^{w,DIA}_{u,T}, r^{c,DIA}_{u,T}\}_{u\in U}$ returned by Algorithm 2 are asymptotically optimal.

Rellich's identity

Standard developments of Pohozaev's identity used an identity by Rellich **?**, reproduced here.

Lemma B.1 (Rellich)**.** *Given* L *in divergence form and* a, d *defined above,* $u \in C^2(\Omega)$*, we have*

$$\int_\Omega (-Lu)\nabla u \cdot (x - \bar{x})\, dx = (1 - \frac{n}{2}) \int_\Omega a(\nabla u, \nabla u)\, dx - \frac{1}{2}\int_\Omega d(\nabla u, \nabla u)\, dx \quad \text{(B.9)}$$

$$+\frac{1}{2}\int_{\partial\Omega} a(\nabla u, \nabla u)(x - \bar{x}) \cdot \nu\, dS - \int_{\partial\Omega} a(\nabla u, \nu)\nabla u \cdot (x - \bar{x})\, dS.$$

Proof:
There is no loss in generality to take $\bar{x} = 0$. First rewrite L:

$$Lu = \frac{1}{2}\left[\sum_i \sum_j \frac{\partial}{\partial x_i}\left(a_{ij}\frac{\partial u}{\partial x_j}\right) + \sum_i \sum_j \frac{\partial}{\partial x_i}\left(a_{ij}\frac{\partial u}{\partial x_j}\right)\right]$$

Switching the order of summation on the second term and relabeling subscripts, $j \to i$ and $i \to j$, then using the fact that $a_{ij}(x)$ is a symmetric matrix, gives the symmetric form needed to derive Rellich's identity.

$$Lu = \frac{1}{2}\sum_{i,j}\left[\frac{\partial}{\partial x_i}\left(a_{ij}\frac{\partial u}{\partial x_j}\right) + \frac{\partial}{\partial x_j}\left(a_{ij}\frac{\partial u}{\partial x_i}\right)\right]. \quad \text{(B.10)}$$

Multiplying $-Lu$ by $\frac{\partial u}{\partial x_k}x_k$ and integrating over Ω, yields

$$\int_\Omega (-Lu)\frac{\partial u}{\partial x_k}x_k\, dx = -\frac{1}{2}\int_\Omega \sum_{i,j}\left[\frac{\partial}{\partial x_i}\left(a_{ij}\frac{\partial u}{\partial x_j}\right) + \frac{\partial}{\partial x_j}\left(a_{ij}\frac{\partial u}{\partial x_i}\right)\right]\frac{\partial u}{\partial x_k}x_k\, dx$$

Integrating by parts gives

$$= \frac{1}{2}\int_\Omega \sum_{i,j} a_{ij}\left[\frac{\partial u}{\partial x_j}\frac{\partial^2 u}{\partial x_k \partial x_i} + \frac{\partial u}{\partial x_i}\frac{\partial^2 u}{\partial x_k \partial x_j}\right] x_k \, dx$$

$$+\frac{1}{2}\int_\Omega \sum_{i,j} a_{ij}\left[\frac{\partial u}{\partial x_j}\delta_{ik} + \frac{\partial u}{\partial x_i}\delta_{jk}\right]\frac{\partial u}{\partial x_k}\, dx$$

$$-\frac{1}{2}\int_{\partial\Omega} \sum_{i,j} a_{ij}\left[\frac{\partial u}{\partial x_j}\nu_i + \frac{\partial u}{\partial x_i}\nu_j\right]\frac{\partial u}{\partial x_k}x_k \, dx$$

$= I_1 + I_2 + I_3$, where the unit normal vector is ν. One may rewrite I_1 as

$$I_1 = \frac{1}{2}\int_\Omega \sum_{i,j} a_{ij}\frac{\partial}{\partial x_k}\left(\frac{\partial u}{\partial x_i}\frac{\partial u}{\partial x_j}\right) x_k \, dx$$

Integrating the first term by parts again yields

$$I_1 = -\frac{1}{2}\int_\Omega \sum_{i,j} a_{ij}\left(\frac{\partial u}{\partial x_i}\frac{\partial u}{\partial x_j}\right)\, dx + \frac{1}{2}\int_{\partial\Omega}\sum_{i,j} a_{ij}\left(\frac{\partial u}{\partial x_i}\frac{\partial u}{\partial x_j}\right) x_k \nu_k \, dS$$

$$-\frac{1}{2}\int_\Omega \sum_{i,j}\left(\frac{\partial u}{\partial x_i}\frac{\partial u}{\partial x_j}\right) x_k \frac{\partial a_{ij}}{\partial x_k}\, dx.$$

Summing over k gives

$$\int_\Omega (-Lu)(\nabla u \cdot x)\, dx = -\frac{n}{2}\int_\Omega \sum_{i,j} a_{ij}\left(\frac{\partial u}{\partial x_i}\frac{\partial u}{\partial x_j}\right)\, dx$$

$$+\frac{1}{2}\int_{\partial\Omega}\sum_{i,j} a_{ij}\left(\frac{\partial u}{\partial x_i}\frac{\partial u}{\partial x_j}\right)(x\cdot\nu)\, dS - \frac{1}{2}\int_\Omega \sum_{i,j}\left(\frac{\partial u}{\partial x_i}\frac{\partial u}{\partial x_j}\right)(x\cdot\nabla a_{ij})\, dx$$

$$+\frac{1}{2}\int_\Omega \sum_{i,j,k} a_{ij}\left[\frac{\partial u}{\partial x_j}\frac{\partial u}{\partial x_k}\delta_{ik} + \frac{\partial u}{\partial x_i}\frac{\partial u}{\partial x_k}\delta_{jk}\right]\, dx$$

$$-\frac{1}{2}\int_{\partial\Omega}\sum_{i,j}a_{ij}\left[\frac{\partial u}{\partial x_j}\nu_i+\frac{\partial u}{\partial x_i}\nu_j\right](\nabla u\cdot x)\,dS.$$

Combining the first and fourth term on the right-hand side simplifies the expression

$$\int_\Omega(-Lu)(\nabla u\cdot x)\,dx=(1-\frac{n}{2})\int_\Omega\sum_{i,j}a_{ij}\left(\frac{\partial u}{\partial x_i}\frac{\partial u}{\partial x_j}\right)\,dx$$

$$+\frac{1}{2}\int_{\partial\Omega}\sum_{i,j}a_{ij}\left(\frac{\partial u}{\partial x_i}\frac{\partial u}{\partial x_j}\right)(x\cdot\nu)\,dS-\frac{1}{2}\int_\Omega\sum_{i,j}\left(\frac{\partial u}{\partial x_i}\frac{\partial u}{\partial x_j}\right)(x\cdot\nabla a_{ij})\,dx$$

$$-\frac{1}{2}\int_{\partial\Omega}\sum_{i,j}a_{ij}\left[\frac{\partial u}{\partial x_j}\nu_i+\frac{\partial u}{\partial x_i}\nu_j\right](\nabla u\cdot x)\,dS.$$

Using the notation defined above, the result follows.

B.2 Complete MILP formulation of SMLE problem

$$
\begin{aligned}
\min \quad & \alpha \sum_{t\in\mathcal{T},d\in\mathcal{D}_t} |w_{t,d} - \tilde{w}_{t,d}| + \beta \sum_{t\in\mathcal{T}} |p_t - g_t|. \\
\text{s.t.} \quad & w_{t,d+1} = bw_{t,d} + cf_{t,d+1} + k, && t \in \{0,\cdots,23\}, \quad d \in \{0,\cdots,6\} \\
& f_{t,d} = dc_{t,d} + \xi_d, && t \in \{0,\cdots,23\}, \quad d \in \{0,\cdots,6\} \\
& f_{b,t+1} = \gamma_f f_{b,t} + (1-\gamma_f)\sum_{d=0}^{6} f_{t,d}, && t \in \{0,\cdots,23\} \\
& c_{t,d} = f_{b,t} - \frac{a_{1,t} c \sum_{i=6-d}^{6} b^i}{2} - \frac{a_{2,t}\hat{r}_t^w c b^{6-d}}{4A} && t \in \{0,\cdots,23\}, \quad d \in \{0,\cdots,6\} \\
& p_{t+1} = \gamma_p(p_t - p_b) + p_b + k_p g_t, && t \in \{0,\cdots,23\} \\
& w_{t,0} - w_{t,6} \leqslant M_{1,t}(1-l_{1,t}), && t \in \{0,\cdots,23\} \\
& p_t - B \leqslant M_{2,t} l_{2,t}, && t \in \{0,\cdots,23\} \\
& z_{1,t} \leqslant M_{z1} l_{2,t}, && t \in \{0,\cdots,23\} \\
& z_{1,t} \leqslant r_t^c, && t \in \{0,\cdots,23\} \\
& z_{1,t} \geqslant r_t^c - M_{z1}(1-l_{2,t}), && t \in \{0,\cdots,23\} \\
& z_{1,t} \geqslant 0, && t \in \{0,\cdots,23\} \\
& z_{3,t} \leqslant M_{z3} l_{1,t}, && t \in \{0,\cdots,23\} \\
& z_{3,t} \leqslant k_1, && t \in \{0,\cdots,23\} \\
& z_{3,t} \geqslant k_1 - M_{z3}(1-l_{1,t}), && t \in \{0,\cdots,23\} \\
& z_{3,t} \geqslant 0, && t \in \{0,\cdots,23\} \\
& a_{1,t+1} = \gamma_1(a_{1,t} - a_{1,b}) + a_{1,b} + z_{1,t} + z_{3,t}, && t \in \{0,\cdots,23\} \\
& z_{2,t} \leqslant M_{z2} l_{1,t}, && t \in \{0,\cdots,23\} \\
& z_{2,t} \leqslant k_2, && t \in \{0,\cdots,23\} \\
& z_{2,t} \geqslant k2 - M_{z2}(1-l_{1,t}), && t \in \{0,\cdots,23\} \\
& z_{2,t} \geqslant 0, && t \in \{0,\cdots,23\} \\
& a_{2,t+1} = \gamma_2(a_{2,t} - a_{2,b}) + a_{2,b} + r_t^w z_{2,t}, && t \in \{0,\cdots,23\} \\
& l_{1,t}, l_{2,t} \in \{0,1\} && t \in \{0,\cdots,23\}
\end{aligned}
\tag{B.11}
$$

REFERENCES

Aas, Kjersti, Martin Jullum, and Anders Løland. 2019. Explaining individual predictions when features are dependent: More accurate approximations to shapley values. *arXiv preprint arXiv:1903.10464*.

Acharya, Sushama D, Okan U Elci, Susan M Sereika, Edvin Music, Mindi A Styn, Melanie Warziski Turk, and Lora E Burke. 2009. Adherence to a behavioral weight loss treatment program enhances weight loss and improvements in biomarkers. *Patient preference and adherence* 3:151.

Adams, Katherine B, Justin J Boutilier, Sarang Deo, and Yonatan Mintz. 2023. Planning a community approach to diabetes care in low-and middle-income countries using optimization. *arXiv preprint arXiv:2305.06426*.

Akrour, Riad, Gerhard Neumann, Hany Abdulsamad, and Abbas Abdolmaleki. 2016. Model-free trajectory optimization for reinforcement learning. In *Icml*, 2961–2970. PMLR.

Almeida, Fabio A, Wen You, Samantha M Harden, Kacie CA Blackman, Brenda M Davy, Russell E Glasgow, Jennie L Hill, Laura A Linnan, Sarah S Wall, Jackie Yenerall, et al. 2015. Effectiveness of a worksite-based weight loss randomized controlled trial: the worksite study. *Obesity* 23(4):737–745.

Åström, Karl J, and Björn Wittenmark. 2013. *Adaptive control*. Courier Corporation.

Aswani, Anil, Philip Kaminsky, Yonatan Mintz, Elena Flowers, and Yoshimi Fukuoka. 2019. Behavioral modeling in weight loss interventions. *EJOR* 272(3): 1058–1072.

Aswani, Anil, Zuo-Jun Shen, and Auyon Siddiq. 2018. Inverse optimization with noisy data. *OR* 66(3):870–892.

Awasthi, Pranjal, Anqi Mao, Mehryar Mohri, and Yutao Zhong. 2022. H-consistency bounds for surrogate loss minimizers. In *Icml*, 1117–1174. PMLR.

Ayer, Turgay, Oguzhan Alagoz, and Natasha K Stout. 2012. Or forum—a pomdp approach to personalize mammography screening decisions. *OR* 60(5):1019–1034.

Ayer, Turgay, Can Zhang, Anthony Bonifonte, Anne C Spaulding, and Jagpreet Chhatwal. 2019. Prioritizing hepatitis c treatment in us prisons. *OR* 67(3):853–873.

Barthe, Franck, Olivier Guédon, Shahar Mendelson, and Assaf Naor. 2005. A probabilistic approach to the geometry of the ℓ_p^n-ball. *The Annals of Probability* 33(2):480–513.

Bartlett, Peter L, Michael I Jordan, and Jon D McAuliffe. 2006. Convexity, classification, and risk bounds. *JASA* 101(473):138–156.

Bastani, Hamsa, Osbert Bastani, and Carolyn Kim. 2018. Interpreting predictive models for human-in-the-loop analytics. *arXiv preprint 1705.08504*.

Bastani, Hamsa, and Mohsen Bayati. 2020. Online decision making with high-dimensional covariates. *OR* 68(1):276–294.

Batterham, Marijka, L Tapsell, K Charlton, J O'shea, and R Thorne. 2017. Using data mining to predict success in a weight loss trial. *Journal of Human Nutrition and Dietetics* 30(4):471–478.

Bertsekas, Dimitri. 2012. *Dynamic programming and optimal control: Volume i*, vol. 1. Athena scientific.

Bertsimas, Dimitris, Vishal Gupta, and Ioannis Ch Paschalidis. 2015. Data-driven estimation in equilibrium using inverse optimization. *MPA* 153:595–633.

Bhaskaran, Krishnan, Ian Douglas, Harriet Forbes, Isabel dos Santos-Silva, David A Leon, and Liam Smeeth. 2014. Body-mass index and risk of 22 specific cancers: a population-based cohort study of 5· 24 million uk adults. *The Lancet* 384(9945): 755–765.

Bhat, Shreya, U. Rajendra Acharya, Yuki Hagiwara, Nahid Dadmehr, and Hojjat Adeli. 2018. Parkinson's disease: Cause factors, measurable indicators, and early diagnosis. *Computers in Biology and Medicine* 102:234–241.

Bickel, Peter J, and Kjell A Doksum. 2015. *Mathematical statistics: basic ideas and selected topics, volumes i-ii package*. Chapman and Hall/CRC.

Bolívar, Hypatia A, Elias M Klemperer, Sulamunn RM Coleman, Michael DeSarno, Joan M Skelly, and Stephen T Higgins. 2021. Contingency management for patients receiving medication for opioid use disorder: a systematic review and meta-analysis. *JAMA psychiatry* 78(10):1092–1102.

Breiman, Leo. 2001a. Random forests. *ML* 45:5–32.

———. 2001b. Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical Science* 16(3):199–231.

Bromberger, Bianca, Paige Porrett, Rashikh Choudhury, Kristoffel Dumon, and Kenric M Murayama. 2014. Weight loss interventions for morbidly obese patients with compensated cirrhosis: a markov decision analysis model. *Journal of Gastrointestinal Surgery* 18(2):321–327.

Brown, Gavin, Adam Pocock, Ming-Jie Zhao, and Mikel Luján. 2012. Conditional likelihood maximisation: A unifying framework for information theoretic feature selection. *Journal of Machine Learning Research* 13(Jan):27–66.

Burkardt, John. 2009. K-means clustering. *Virginia Tech, Advanced Research Computing, Interdisciplinary Center for Applied Mathematics*.

Cawley, John. 2004. An economic framework for understanding physical activity and eating behaviors. *American Journal of Preventive Medicine* 27(3):117–125.

Cawley, John, Adam Biener, Chad Meyerhoefer, Yuchen Ding, Tracy Zvenyach, B Gabriel Smolarz, and Abhilasha Ramasamy. 2021. Direct medical costs of obesity in the united states and the most populous states. *Journal of managed care & specialty pharmacy* 27(3):354–366.

Chen, Jianbo, Le Song, Martin J Wainwright, and Michael I Jordan. 2018. Learning to explain: An information-theoretic perspective on model interpretation. *arXiv preprint arXiv:1802.07814*.

Childers, Ashley Kay, Gurucharann Visagamurthy, and Kevin Taaffe. 2009. Prioritizing patients for evacuation from a health-care facility. *Transportation research record* 2137(1):38–45.

Colson, Benoît, Patrice Marcotte, and Gilles Savard. 2007. An overview of bilevel optimization. *Annals of OR* 153:235–256.

Craig, John J, Ping Hsu, and S Shankar Sastry. 1987. Adaptive control of mechanical manipulators. *The International Journal of Robotics Research* 6(2):16–28.

Dai, Jim G, and Pengyi Shi. 2019. Inpatient overflow: An approximate dynamic programming approach. *M&SOM* 21(4):894–911.

Deci, Edward L, and Richard M Ryan. 2013. *Intrinsic motivation and self-determination in human behavior*. Springer Science & Business Media.

Deo, Sarang, Seyed Iravani, Tingting Jiang, Karen Smilowitz, and Stephen Samuelson. 2013. Improving health outcomes through better capacity allocation in a community-based chronic care model. *OR* 61(6):1277–1294.

Ebbert, Jon O, Muhamad Y Elrashidi, and Michael D Jensen. 2014. Managing overweight and obesity in adults to reduce cardiovascular disease risk. *Current atherosclerosis reports* 16:1–7.

Ederer, Florian, Richard Holden, Margaret A Meyer, et al. 2013. *Gaming and strategic ambiguity in incentive provision*. Centre for Economic Policy Research.

Ekici, Ali, Pınar Keskinocak, and Julie L Swann. 2014. Modeling influenza pandemic and planning food distribution. *M&SOM* 16(1):11–27.

Erdogan, S Ayca, and Brian Denton. 2013. Dynamic appointment scheduling of a stochastic server with uncertain demand. *IJoC* 25(1):116–132.

Fereshtehnejad, Seyed-Mohammad, and Ronald B. Postuma. 2017. Subtypes of Parkinson's disease: What do they tell us about disease progression? *Current neurology and neuroscience reports* 17(4):34.

Ferster, Charles B, and Burrhus Frederic Skinner. 1957. Mixed schedules.

Friedman, Jerome, Trevor Hastie, and Robert Tibshirani. 2001. *The elements of statistical learning: Data mining, inference, and prediction*. Springer Series in Statistics, Springer.

Fukuoka, Yoshimi, Teri G Lindgren, Yonatan Dov Mintz, Julie Hooper, Anil Aswani, et al. 2018. Applying natural language processing to understand motivational profiles for maintaining physical activity after a mobile app and accelerometer-based intervention: the mped randomized controlled trial. *JMIR mHealth and uHealth* 6(6):e10042.

Gibson, Alice A, and Amanda Sainsbury. 2017. Strategies to improve adherence to dietary weight loss interventions in research and real-world settings. *Behavioral Sciences* 7(3):44.

Gneezy, Uri, Stephan Meier, and Pedro Rey-Biel. 2011. When and why incentives (don't) work to modify behavior. *Journal of economic perspectives* 25(4):191–210.

Goh, Siong Thye, and Cynthia Rudin. 2018. A minimax surrogate loss approach to conditional difference estimation. *arXiv preprint arXiv:1803.03769*.

Golay, Alain, and Juan Ybarra. 2005. Link between obesity and type 2 diabetes. *Best practice & research Clinical endocrinology & metabolism* 19(4):649–663.

Grilo, Carlos M, Robin M Masheb, G Terence Wilson, Ralitza Gueorguieva, and Marney A White. 2011. Cognitive–behavioral therapy, behavioral weight loss, and sequential treatment for obese patients with binge-eating disorder: A randomized controlled trial. *Journal of consulting and clinical psychology* 79(5):675.

Gurobi Optimization, LLC. 2022. Gurobi Optimizer Reference Manual.

Hassanzadeh, Anahita, and Ted K Ralphs. 2014. A generalization of benders' algorithm for twostage stochastic optimization problems with mixed integer recourse. In *Technical report*.

Hastie, Trevor, Robert Tibshirani, Jerome H Friedman, and Jerome H Friedman. 2009. *The elements of statistical learning: data mining, inference, and prediction*, vol. 2. Springer.

He, Qinyang, and Yonatan Mintz. 2023. Model based reinforcement learning for personalized heparin dosing. *arXiv preprint arXiv:2304.10000*.

Jakicic, John M, Kelliann K Davis, Renee J Rogers, Wendy C King, Marsha D Marcus, Diane Helsel, Amy D Rickman, Abdus S Wahed, and Steven H Belle. 2016. Effect of wearable technology combined with a lifestyle intervention on long-term weight loss: the idea randomized clinical trial. *JAMA* 316(11):1161–1171.

John, Leslie K, George Loewenstein, Andrea B Troxel, Laurie Norton, Jennifer E Fassbender, and Kevin G Volpp. 2011. Financial incentives for extended weight loss: a randomized, controlled trial. *Journal of general internal medicine* 26(6): 621–626.

Kaut, Michal, and W Stein. 2003. *Evaluation of scenario-generation methods for stochastic programming*. Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät ....

Keshavarz, Arezou, Yang Wang, and Stephen Boyd. 2011. Imputing a convex objective function. In *2011 ieee international symposium on intelligent control*, 613–619. IEEE.

Krentz, AJ, K Fujioka, and M Hompesch. 2016. Evolution of pharmacological obesity treatments: focus on adverse side-effect profiles. *Diabetes, Obesity and Metabolism* 18(6):558–570.

Kushner, Robert F. 2014. Weight loss strategies for treatment of obesity. *Progress in cardiovascular diseases* 56(4):465–472.

Lachout, Petr, Eckhard Liebscher, and Silvia Vogel. 2005. Strong convergence of estimators as ε n-minimisers of optimisation problemsof optimisation problems. *Annals of the Institute of Statistical Mathematics* 57(2):291–313.

Lakkaraju, Himabindu, Stephen H Bach, and Jure Leskovec. 2016. Interpretable decision sets: A joint framework for description and prediction. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 1675–1684. KDD '16.

Leahey, Tricia M, Leslee L Subak, Joseph Fava, Michael Schembri, Graham Thomas, Xiaomeng Xu, Katie Krupel, Kimberly Kent, Katherine Boguszewski, Rajiv Kumar, et al. 2015. Benefits of adding small financial incentives or optional group meetings to a web-based statewide obesity initiative. *Obesity* 23(1):70–76.

Lee, Elliot, Mariel S Lavieri, and Michael Volk. 2019. Optimal screening for hepatocellular carcinoma: A restless bandit model. *M&SOM* 21(1):198–212.

Lee, Sang Ho, Peijin Han, Russell K Hales, K Ranh Voong, Kazumasa Noro, Shinya Sugiyama, John W Haller, Todd R McNutt, and Junghoon Lee. 2020. Multi-view radiomics and dosiomics analysis with machine learning for predicting acute-phase weight loss in lung cancer patients treated with radiotherapy. *Physics in Medicine & Biology* 65(19):195015.

Lemstra, Mark, Yelena Bird, Chijioke Nwankwo, Marla Rogers, and John Moraros. 2016. Weight loss intervention adherence and factors promoting adherence: a meta-analysis. *Patient preference and adherence* 10:1547.

Li, Qiaomei, Yonatan Mintz, Kara Gavin, and Corrine Voils. 2023. An adaptive optimization approach to personalized financial incentives in mobile behavioral weight loss interventions. *arXiv preprint arXiv:2307.00444*.

van der Linden, Ilse, Hinda Haned, and Evangelos Kanoulas. 2019. Global aggregations of local explanations for black box models. *arXiv preprint arXiv:1907.03039*.

Lundberg, Scott M, Gabriel Erion, Hugh Chen, Alex DeGrave, Jordan M Prutkin, Bala Nair, Ronit Katz, Jonathan Himmelfarb, Nisha Bansal, and Su-In Lee. 2020. From local explanations to global understanding with explainable ai for trees. *Nature machine intelligence* 2(1):56–67.

Lundberg, Scott M, and Su-In Lee. 2017. A unified approach to interpreting model predictions. In *Advances in neural information processing systems*, 4765–4774.

Maaten, Laurens van der, and Geoffrey Hinton. 2008. Visualizing data using t-SNE. *Journal of Machine Learning Research* 9(Nov):2579–2605.

Marek, Kenneth, Danna Jennings, Shirley Lasch, Andrew Siderowf, Caroline Tanner, Tanya Simuni, Chris Coffey, Karl Kieburtz, Emily Flagg, Sohini Chowdhury, et al. 2011. The parkinson progression marker initiative (ppmi). *Progress in neurobiology* 95(4):629–635.

Martinez-Martin, P, C Rodriguez-Blazquez, and M João Forjaz. 2017. *Rating scales in movement disorders*. Elsevier.

Martínez-Martín, P1, A Gil-Nagel, L Morlán Gracia, J Balseiro Gómez, J Martinez-Sarries, F Bermejo, and Cooperative Multicentric Group. 1994. Unified Parkinson's disease rating scale characteristics and structure. *Movement Disorders* 9(1):76–83.

MATLAB. 2010. *version 7.10.0 (r2010a)*. Natick, Massachusetts: The MathWorks Inc.

McKenzie, Briar L, Daisy H Coyle, Joseph Alvin Santos, Tracy Burrows, Emalie Rosewarne, Sanne AE Peters, Cheryl Carcel, Lindsay M Jaacks, Robyn Norton, Clare E Collins, et al. 2021. Investigating sex differences in the accuracy of dietary assessment methods to measure energy intake in adults: a systematic review and meta-analysis. *The American Journal of Clinical Nutrition* 113(5):1241–1255.

Mifflin, Mark D, Sachiko T St Jeor, Lisa A Hill, Barbara J Scott, Sandra A Daugherty, and Young O Koh. 1990. A new predictive equation for resting energy expenditure in healthy individuals. *The American journal of clinical nutrition* 51(2):241–247.

Mintz, Yonatan, Anil Aswani, Philip Kaminsky, Elena Flowers, and Yoshimi Fukuoka. 2017. Behavioral analytics for myopic agents. *arXiv preprint arXiv:1702.05496*.

———. 2020. Nonstationary bandits with habituation and recovery dynamics. *OR* 68(5):1493–1516.

Nasreddine, Ziad S, Natalie A Phillips, Valérie Bédirian, Simon Charbonneau, Victor Whitehead, Isabelle Collin, Jeffrey L Cummings, and Howard Chertkow. 2005. The montreal cognitive assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society* 53(4):695–699.

Nguyen, XuanLong, Martin J Wainwright, and Michael I Jordan. 2009. On surrogate loss functions and f-divergences.

Nielsen, Frank. 2016. *Introduction to hpc with mpi for data science*. Springer.

Osband, Ian, and Benjamin Van Roy. 2014. Model-based reinforcement learning and the eluder dimension. *NeurIPS* 27.

Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. 2011a. Scikit-learn: Machine learning in Python. *JMLR* 12:2825–2830.

Pedregosa, Fabian, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, and Vincent Dubourg. 2011b. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research* 12(Oct):2825–2830.

Petry, Nancy M. 2011. Contingency management: what it is and why psychiatrists should want to use it. *The psychiatrist* 35(5):161–163.

Plumb, Gregory, Denali Molitor, and Ameet S Talwalkar. 2018. Model agnostic supervised local explanations. In *Advances in neural information processing systems*, 2515–2524.

PPMI. 2019. Parkinson's progression markers initiative.

Raber, Margaret, Yue Liao, Anne Rara, Susan M Schembre, Kate J Krause, Larkin Strong, Carrie Daniel-MacDougall, and Karen Basen-Engquist. 2021. A systematic review of the use of dietary self-monitoring in behavioural weight loss interventions: delivery, intensity and effectiveness. *Public health nutrition* 24(17):5885–5913.

Rajapaksha, Dilini, Christoph Bergmeir, and Wray Buntine. 2019. Lormika: Local rule-based model interpretability with k-optimal associations. *arXiv preprint arXiv:1908.03840*.

Rao, Shobha S, Laura A Hofmann, and Amer Shakil. 2006. Parkinson's disease: Diagnosis and treatment. *American Family Physician* 74(12):2046–2054.

Ribeiro, Marco Tulio, Sameer Singh, and Carlos Guestrin. 2016. "Why should I trust you?": Explaining the predictions of any classifier. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 1135–1144. KDD '16.

———. 2018. Anchors: High-precision model-agnostic explanations. In *Proceedings of the 32nd aaai conference on artificial intelligence*, 1527–1535. AAAI '18.

Schell, Greggory J, Wesley J Marrero, Mariel S Lavieri, Jeremy B Sussman, and Rodney A Hayward. 2016. Data-driven markov decision process approximations for personalized hypertension treatment planning. *MDM policy & practice* 1(1): 2381468316674214.

Setzu, Mattia, Riccardo Guidotti, Anna Monreale, Franco Turini, Dino Pedreschi, and Fosca Giannotti. 2021. Glocalx-from local to global explanations of black box ai models. *Artificial Intelligence* 294:103457.

Siderowf, A. 2010. *Schwab and England activities of daily living scale*, 99–100. Elsevier.

Sokol, Kacper, and Peter Flach. 2020. Limetree: Interactively customisable explanations based on local surrogate multi-output regression trees. *arXiv preprint arXiv:2005.01427*.

Springer, Matthew G, and Lori L Taylor. 2016. Designing incentives for public school teachers: Evidence from a texas incentive pay program. *Journal of Education Finance* 344–381.

Stierman, B, J Afful, MD Carroll, TC Chen, O Davy, S Fink, CD Fryar, Q Gu, CM Hales, JP Hughes, et al. 2021. National health and nutrition examination survey 2017–march 2020 prepandemic data files-development of files and prevalence estimates for selected health outcomes. *National Health Statistics Reports*.

Strehl, Alexander L, Lihong Li, Eric Wiewiora, John Langford, and Michael L Littman. 2006. Pac model-free reinforcement learning. In *Icml*, 881–888.

Štrumbelj, Erik, and Igor Kononenko. 2014. Explaining prediction models and individual predictions with feature contributions. *Knowledge and information systems* 41(3):647–665.

Sutton, Richard S, and Andrew G Barto. 2018. *Reinforcement learning: An introduction*. MIT press.

Thomas, Diana M, Corby K Martin, Steven Heymsfield, Leanne M Redman, Dale A Schoeller, and James A Levine. 2011. A simple model predicting individual weight change in humans. *Journal of biological dynamics* 5(6):579–599.

Torres, Roberto L Shinmoto, Damith C Ranasinghe, Qinfeng Shi, and Alanson P Sample. 2013. Sensor enabled wearable rfid technology for mitigating the risk of falls near beds. In *2013 ieee international conference on rfid (rfid)*, 191–198. IEEE.

Tsai, Wen-Hsuan, and Xingmiu Liao. 2020. Mobilizing cadre incentives in policy implementation: Poverty alleviation in a chinese county. *China Information* 34(1): 45–67.

Ustun, Berk, and Cynthia Rudin. 2016. Supersparse linear integer models for optimized medical scoring systems. *Machine Learning* 102(3):349–391.

Van Rossum, Guido, and Fred L. Drake. 2009. *Python 3 reference manual*. Scotts Valley, CA: CreateSpace.

Voils, Corrine I, Erica Levine, Jennifer M Gierisch, Jane Pendergast, Sarah L Hale, Megan A McVay, Shelby D Reed, William S Yancy Jr, Gary Bennett, Elizabeth M Strawbridge, et al. 2018. Study protocol for log2lose: A feasibility randomized controlled trial to evaluate financial incentives for dietary self-monitoring and interim weight loss in adults with obesity. *Contemporary clinical trials* 65:116–122.

Voils, Corrine I, Jane Pendergast, Sarah L Hale, Jennifer M Gierisch, Elizabeth M Strawbridge, Erica Levine, Megan A McVay, Shelby D Reed, William S Yancy Jr, and Ryan J Shaw. 2021. A randomized feasibility pilot trial of a financial incentives intervention for dietary self-monitoring and weight loss in adults with obesity. *Translational Behavioral Medicine* 11(4):954–969.

Volpp, Kevin G, Leslie K John, Andrea B Troxel, Laurie Norton, Jennifer Fassbender, and George Loewenstein. 2008. Financial incentive–based approaches for weight loss: a randomized trial. *JAMA* 300(22):2631–2637.

Wang, Fulton, and Cynthia Rudin. 2015. Falling rule lists. In *Proceedings of the 18th international conference on artificial intelligence and statistics*, 1013–1022. AISTATS '15.

Ward, Zachary J, Sara N Bleich, Michael W Long, and Steven L Gortmaker. 2021. Association of body mass index with health care expenditures in the united states by age and sex. *PloS one* 16(3):e0247307.

Wing, Rena R, Randi Koeske, Leonard H Epstein, Mary Patricia Nowalk, William Gooding, and Dorothy Becker. 1987. Long-term effects of modest weight loss in type ii diabetic patients. *Archives of internal medicine* 147(10):1749–1753.

Wolsey, Laurence A, and George L Nemhauser. 1999. *Integer and combinatorial optimization*, vol. 55. John Wiley & Sons.

Yang, Min-Jeong, Steven K Sutton, Laura M Hernandez, Sarah R Jones, David W Wetter, Santosh Kumar, and Christine Vinci. 2023. A just-in-time adaptive intervention (jitai) for smoking cessation: Feasibility and acceptability findings. *Addictive behaviors* 136:107467.

Yoon, Jinsung, James Jordon, and Mihaela van der Schaar. 2018. Invase: Instance-wise variable selection using neural networks.

Yu, Huizhen, and Dimitri Bertsekas. 2012. Discretized approximations for pomdp with average cost. *arXiv preprint arXiv:1207.4154*.

Yu, Huizhen, and Dimitri P Bertsekas. 2008. On near optimality of the set of finite-state controllers for average cost pomdp. *Math of OR* 33(1):1–11.

Zhang, Junyu, Chengzhuo Ni, Csaba Szepesvari, Mengdi Wang, et al. 2021. On the convergence and sample efficiency of variance-reduced policy gradient method. *NeurIPS* 34:2228–2240.

Zhou, Mo, Yonatan Mintz, Yoshimi Fukuoka, Ken Goldberg, Elena Flowers, Philip Kaminsky, Alejandro Castillejo, and Anil Aswani. 2018. Personalizing mobile fitness apps using reinforcement learning. In *Ceur*, vol. 2068. NIH Public Access.