Behavioral Analytics for Myopic Agents

Yonatan Mintz, Anil Aswani, Philip Kaminsky

Department of Industrial Engineering and Operations Research, University of California, Berkeley, CA 94720, {aaswani,kaminsky,ymintz}@berkeley.edu

Elena Flowers

Department of Physiological Nursing, School of Nursing, University of California, San Francisco, CA 94143, elena.flowers@ucsf.edu

Yoshimi Fukuoka

Department of Physiological Nursing & Institute for Health & Aging, School of Nursing, University of California, San Francisco, CA 94143, Yoshimi.Fukuoka@ucsf.edu

Many multi-agent systems have the structure of a single coordinator providing behavioral or financial incentives to a large number of agents. In such settings, two challenges faced by the coordinator are a finite budget from which to allocate incentives, and an initial lack of knowledge about the utility function of the agents. Here, we present a behavioral analytics approach to solve the coordinator's problem when the agents make decisions by maximizing utility functions that depend on prior system states, inputs, and other parameters that are initially unknown and subject to partially unknown temporal dynamics. Our behavioral analytics framework involves three steps: first, we develop a behavioral model that describes the decision-making process of an agent; second, we use data to estimate behavioral model parameters for each agent and then use these estimates to predict future decisions of each agent; and third, we use the estimated behavioral model parameters to optimize a set of costly incentives to provide to each agent. In this paper, we describe a specific set of tools, models, and approaches that fit into this framework, and that adapt models and incentives as new information is collected by repeating the second and third steps of this framework. Furthermore, we prove that the incentives computed by this adaptive approach are asymptotically optimal with respect to a given loss function that describes the coordinator's objective. We optimize incentives utilizing a decomposition scheme, where each sub-problem solves the coordinator's problem for a single agent, and the master problem is a pure integer program. We conclude with a simulation study to evaluate the effectiveness of our behavioral analytics approach in designing personalized treatment plans for a weight loss program. The results show that our approach maintains efficacy of the program while reducing its costs by up to 60%, while adaptive heuristics provide substantially less savings.

Key words: health care, optimization, statistical inference

1. Introduction

The increasing availability of data presents an opportunity to transform the design of incentives (i.e., costly inputs that are provided to agents to modify their behavior and decisions) from a single analysis into an adaptive and dynamic process whereby the incentive design is optimized as new data becomes available. Historically, this adaptive setting has been studied under the framework of repeated games (Radner 1985, Fudenberg et al. 1994, Laffont and Martimort 2002), where

researchers have focused on the analysis and identification of structural properties of effective policies, and on equilibria. In contrast, continuing advances in optimization software and statistical estimation tools, utilized with the vast amount of data now available in many settings, enable a new approach that in many circumstances has the potential to lead to practical tools for designing effective incentives in real-world settings. This approach, which we call *behavioral analytics*, is built around a three step framework: first, we develop a behavioral model that describes the decisionmaking process of an agent; next, we iterate repeatedly over two steps as new information is collected. In the second step, we use data to estimate behavioral model parameters for each agent and then use these estimates to predict future decisions of each agent; and in the third, we use the estimated behavioral model parameters to optimize a set of costly incentives to provide to each agent. In this paper, we describe a specific set of tools, models, and approaches that fit into this framework, and that adapt models and incentives as new information is collected while the second and third steps of the framework are repeated.

Specifically, we consider the following discrete-time setting: There is a large pool of agents each with a set of utility function parameters (which we will refer to as the motivational state) and system state at time t, and each agent makes a decision at t by maximizing a myopic utility function. A single coordinator makes noisy observations at t of the system states and decisions of each agent, and then assigns behavioral or financial incentives (e.g., bonuses, payments, behavioral goals, counseling sessions) at t to a subset of agents. The incentives change the motivational and system states of assigned agents at time t + 1, while the motivational and system states of non-assigned agents evolve at t + 1 according to some dynamics. This process repeats, and time t advances to wards infinity in unit increments. Here, the coordinator's problem is to decide what incentives to provide to which agents in order to minimize the coordinator's loss function, a function that depends on the system states and decisions of all agents. This problem is challenging because the motivational states of agents are neither known nor measured by the coordinator, because agents make decisions by maximizing an unknown utility function, because measurements are noisy, and because the coordinator has a fixed budget (over a specified time horizon) from which to allocate incentives.

1.1. Potential Applications for Behavioral Analytics

The setting described above is found in many domains, including personalized healthcare, demand response programs, and franchise logistics. Below, we elaborate on these potential applications of our framework. The first application is the design of a weight loss program. The coordinator is a clinician and the agents are individuals trying to lose body weight. The next application is the design of a demand response program in which the coordinator is an electric utility company and the agents are homeowners (since they consume electricity). The final application is in franchise logistics, where a parent company is a coordinator for a group of franchisees (who are the agents).

1.1.1. Weight Loss Programs In a clinically-supervised weight loss program, a clinician provides two types of behavioral incentives to a group of individuals who are trying to lose body weight. The first type of incentive is behavioral goals provided to each individual by the clinician, and it is costless when communication costs are negligible, as is the case with mobile phone-delivered programs (Fukuoka et al. 2015). The second type of incentive is that the clinician can provide a limited amount of counseling to individuals, but this is costly and the clinician must decide how to allocate a limited budget of counseling sessions to the entire pool of individuals. For example, the Diabetes Prevention Program Research Group (2002, 2003, 2009) has shown that such programs lead to a clinically significant loss of 5-7% body weight on average, which can prevent or delay the onset of type 2 diabetes with few side effects. However, these programs are difficult to design because variations in individual motivational states mean there is not just one set of optimal behavioral goals and assignment of counseling sessions, but rather that behavioral goals and the number/timing of counseling needs to be personalized to individuals' motivational states to maximize weight loss.

Personalizing the behavioral incentives for each individual can improve efficacy of weight loss programs and reduce the associated program costs through a reduction in the average amount of counseling for each individual. Mobile phone technology is one promising avenue for implementing such personalization, due to its relatively low cost and pervasiveness among diverse communities (Lopez et al. 2013). Mobile phones allow clinicians to collect real time health data through use of personal logs and devices such as accelerometers, which provides noisy measurements of the health state and decisions of each individual. Randomized controlled trials (RCT's) have found that the use of mobile phones can reduce the cost of implementing weight loss programs with maintaining efficacy (Fukuoka et al. 2015); however, little research to date has explored how to use the data generated by mobile phones and digital accelerometers in order to personalize behavioral incentives (Fukuoka et al. 2011, O'Reilly and Spruijt-Metz 2013, Azar et al. 2013, Pagoto et al. 2013).

Several adaptive methods have been proposed for designing personalized healthcare treatments, including: multi-armed bandits (Negoescu et al. 2014, Deo et al. 2013, Bastani and Bayati 2015), robust optimization (Bertsimas and O'Hair 2013), and dynamic programming (Engineer et al. 2009). One common approach for optimal treatment design and clinical appointment scheduling has been Markov decision process (MDP) models (Ayer et al. 2015, Mason et al. 2013, Deo et al. 2013, Kucukyazici et al. 2011, Leff et al. 1986, Liu et al. 2010, Wang and Gupta 2011, Gupta and Wang 2008, Savelsbergh and Smilowitz 2016). These methods are designed for situations with

infrequent data collection (e.g., only during clinical visits), whereas in weight loss programs the data is collected daily (or more often) using mobile devices. Our work develops an approach that can leverage this increased data availability to better design incentives. Moreover, existing approaches focus either on motivational states characterizing adherence (Mason et al. 2013) or health states describing prognosis (Ayer et al. 2012, Deo et al. 2013, Helm et al. 2015, Wu et al. 2013, Negoescu et al. 2014, Engineer et al. 2009). In contrast, we seek to combine motivational and health states into a single predictive model that is used for personalizing the weight loss program.

1.1.2. Demand Response (DR) DR programs are used by electric utilities to alter homeowners' electricity usage to better match electricity generation and reduce peak electricity demand. Utilities incentivize homeowners to shift or reduce electricity consumption using price-based programs (e.g., time-of-day electricity rates). Utilities also incentivize reduced electricity consumption through exchange programs in which a homeowner's inefficient appliances are replaced (for free by the electric utility) with efficient appliances (Palensky and Dietrich 2011, Deng et al. 2015). Implementations of such DR programs have decreased peak electricity demand by almost 10% and have improved the balance between electricity supply and demand (Lee et al. 2014). However, adverse selection is a major issue in these programs because incentives are often provided to homeowners who already have low electricity consumption or already had plans to replace inefficient appliances.

Better targeting in a DR program may lead to improved efficacy with lower associated costs. For example, electric utilities have the capability to send an auditor to homes to assess what appliance upgrades are needed (PG&E 2016). Consequently, an electric utility would be interested in finding the most effective way to schedule its auditors and set its rebates and tariffs. Homeowner electricity usage data can be collected by the utility in real time using smart electricity meters, and the adoption rate of these smart electricity meters is increasing in the US (Lee et al. 2014). Moreover, the two way communication capabilities of smart electricity meters and mobile phones can be used to communicate billing and incentive information to homeowners (Darby 2010), which opens the possibility for better targeting of price-based programs and appliance-replacement programs.

DR programs are often designed using game-theoretic approaches (Saghezchi et al. 2015, Samadi et al. 2010, 2012), multi-armed bandits (Wijaya et al. 2013), convex optimization (Li et al. 2011, Mohsenian-Rad and Leon-Garcia 2010, Ratliff et al. 2014), dynamic programming (Jiang and Low 2011, Costa and Kariniotakis 2007, Molderink et al. 2010), and MDP's (O'Neill et al. 2010, Kim and Poor 2011). These approaches commonly assume the electric utility has perfect information on the motivational state of each homeowner, and that the uncertainty is primarily in electricity generation and pricing. In contrast, our proposed methodology has the ability to estimate the motivational state of each homeowner to better design DR programs through improved targeting of

price-based and appliance-replacement incentives. (Existing work also does not consider the option of the power company to provide rebates for upgrading inefficient equipment, while our framework can incorporate this scheduling problem.)

1.1.3. Franchise Logistics A franchise contract governs the relationship between a parent company (with a developed brand/products/services) and a separate franchisee company that retails this brand for a limited amount of time (Rubin 1978). Franchisees enter these contracts because they lack the brand recognition, economies of scale, or expertise needed. Each franchisee desires to maximize their individual profit while the parent company works to ensure brand quality and profitability are maintained. Towards this end, franchise contracts have several common clauses that help the parent company ensure that franchisees operate within quality standards: The parent company provides managerial assistance, training, and funding for advertising campaigns (Rubin 1978), and the parent company may also monitor a specific franchisee by assigning an employee to supervise the franchisee's daily operations (Gal-Or 1995).

A common feature of such franchise contracts is the requirement for the franchisee to provide sales figures, customer satisfaction metrics, and expense reports to the parent company; and new online survey tools give the capability to monitor customer satisfaction and sales at a finer time-scale than previously possible. Such data lends itself towards approaches for the improved allocation of parent company resources (e.g., managerial assistance, training, funding for advertising campaigns, and monitoring of specific franchisees), which can improve the performance of both the parent company and the franchisee. Restated, these new data streams can be used by the parent company to develop a better understanding of the competencies and weaknesses of each franchisee, and to then use this to better allocate resources and set contract terms.

Characteristics of optimal franchise contracts have been studied extensively in the operations research literature (Gal-Or 1995, Lal 1990, Li 1997, Xie et al. 2016). This stream of literature primarily extends earlier approaches used to solve contracting problems in which it is assumed that data acquisition and monitoring is expensive, making it difficult to estimate the utility function of the franchisee (Lal 1990, Gal-Or 1995). Additionally, much of this work is focused on modeling the interaction and motivations of the franchisees and parent company and not on developing algorithms for optimal contract calculation. In contrast, our behavioral analytics framework for adaptively designing incentives – which consists of repeatedly estimating utility functions and then refining the incentives using optimization modeling – applies to settings where data (e.g. sales, customer satisfaction, etc.) is readily available to the parent company.

1.2. Literature Review

The behavioral analytics framework we develop in this paper builds upon existing literature on data driven and adaptive methods for stochastic optimization. Ban and Rudin (2016) and Vahn (2015) consider how predictive and data driven models can be incorporated into inventory management problems, and both parametric and nonparametric predictive models are used by a decision maker to estimate demand and compute an optimal reorder policy. These models are constructed to estimate demand through i.i.d observations; this differs from the setup in our paper where the observations are generated by temporal dynamics and are thus not i.i.d. A more general set of approaches are reinforcement learning and Bayesian optimization (Aswani et al. 2013, Frazier and Wang 2016, Osband and Roy 2015, Osband et al. 2016), which leverage statistical estimation to compute asymptotically optimal control inputs for systems with appropriate model structures. However, the relationship between the computed control inputs and the estimated model is often difficult to interpret because of the nonparametric nature of the estimation (Breiman et al. 2001). Our approach offers improved interpretability of the incentives computed by our framework because we simultaneously generate estimates of the parameters of the utility function (i.e., motivational states) for each agent. These estimates provide insights into the resulting incentive allocations computed by our framework because these parameters usually have behavioral or financial interpretations (e.g., responsiveness to incentives, production efficiency, level of risk aversion).

Our behavioral analytics framework is also related to research that explores stochastic control of multi-agent systems. Related methods include decentralized control (Li et al. 2012), approximate dynamic programming (Boukhtouta et al. 2011, George and Powell 2007), game-theoretic approaches (Adlakha and Johari 2013, Iyer et al. 2011, 2014, Zhou et al. 2016), and robust optimization (Blanchet et al. 2013, Bertsimas and Goyal 2012, Lorca and Sun 2015). In general, these models consider very different settings from the ones we consider in this paper. This body of work studies settings where the agents can strategically interact with other agents (without the presence of a coordinator) and where the agents are able to consider long time horizons when making decisions. Our setting differs in that we have a single coordinator that provides incentives to a group of agents that do not interact strategically with other agents, and where the agents are myopic (meaning they make decisions based on short time horizons). The three examples of weight loss programs, demand response programs, and franchise logistics more closely match the setting we consider in this paper.

1.3. Contributions

Our overall goal in this paper is to provide tools and approaches that form a specific implementation of the three steps of a behavioral analytics framework, and our secondary goal is to give an example that demonstrates how our implementation of behavioral analytics can be applied to a real-world engineering problem. Recall that these three steps involve designing a behavioral model, and then repeatedly estimating the parameters of this behavioral model, and using the estimated parameters to optimize the incentives provided to each agent. To do this, we first need to identify a general (and practically useful) class of models that describe agent behavior and can be incorporated into optimization models for incentive design. This is non-obvious because incentive design in principle requires solving bilevel programs, precluding the straightforward use of commercial optimization software packages. We address this by abstracting and generalizing our earlier work on the development of predictive models of the behavior of individuals participating in a weight loss program (Aswani et al. 2016). Given these behavioral models, we design an optimization approach that, rather than directly solving the relevant bilevel program, is built around formulations that incorporate the individual behavior model into mathematical programs that can be solved in a straightforward way with commercial solvers, and that lead to incentives that are asymptotically optimal as more data is collected. Below, we describe these contributions in further detail:

First, we develop and analyze an abstract model of agent behavior. This model consists of a myopic utility function (meaning the agent makes decisions based on a utility function that depends on states only one time period into the future) and temporal dynamics on the system states and on the parameters of the utility function. It abstracts and generalizes a predictive model we created in our prior work on behavioral modeling for weight loss (Aswani et al. 2016). In addition, we explore (for the first time) theoretical questions related to statistical consistency of utility function parameter estimates. Such consistency is important because in order to design optimal incentives we need to be able to correctly estimate the parameters of the utility functions of each agent, and it was recently shown that not all approaches that have been proposed for estimating parameters of utility functions are statistically consistent (Aswani et al. 2015). Here, we provide mixed integer linear programming (MILP) formulations for estimating the parameters of the utility functions, and we prove these formulations generate estimates that are statistically consistent.

We also develop novel mathematical programs for incentive design that incorporate our model for agent behavior, and we prove that the incentives are asymptotically optimal (in time). Incentive design in principle requires solving a bilevel program, and the situation is complicated in our setting because the mathematical structure of our abstract model for agent behavior leads to incentive design problems that consist of bilevel mixed integer programs (BMIP's). BMIP's are computationally difficult to solve (Ralphs and Hassanzadeh 2014, DeNegre and Ralphs 2009, James and Bard 1990, Moore and Bard 1992) since solution techniques for continuous bilevel programming (Ahuja and Orlin 2001, Aswani et al. 2015, Aswani and Ouattara 2016, Dempe 2002, Heuberger 2004) cannot be used. Consequently, we develop an adaptive two-stage decomposition algorithm. In the first stage, we solve the coordinator's problem for each agent considered individually by estimating the utility function parameters of an agent by solving a single MILP and then solving a series of MILP sub-problems. The second stage consists of an integer linear program (ILP) master problem that aggregates the sub-problem solutions and solves the coordinator's problem for all agents considered jointly. We prove this asymptotically designs the optimal incentives.

To evaluate the efficacy of the specific behavioral models, parameter estimation techniques, and optimization models in our instantiation of a behavioral analytics framework, we perform computational experiments in the context of goal-setting and clinical appointment scheduling for individuals participating in a clinically-supervised weight loss program. The first step of our behavioral analytics approach involves constructing a model that describes individuals' decisions on how much to eat and how much physical activity (in terms of daily steps) to do – subject to a utility function that captures the tradeoffs inherent in achieving one-day-ahead weight loss with reducing dietary consumption and increasing physical activity. The second step of our behavioral analytics approach uses past data for each individual in order to quantify (for each individual) the tradeoffs captured by the utility function, as well as estimate the responsiveness of each individual to the incentives of providing physical activity goals and providing counseling sessions, and the third step of our behavioral analytics approach uses the behavioral model and estimated parameters to determine what physical activity goals to provide to each individual and to determine how to allocate a fixed number of counseling sessions to a pool of individuals participating in the program. These second and third steps are repeated as more data is collected from each individual. Through a simulation study, we compare personalized treatment plans computed by our approach with treatment plans computed by an adaptive heuristic, and we find that our approach performs substantially better than the heuristic. Common heuristics implicitly assume monotonicity in individuals' behaviors with respect to the treatment plan values, while actual behavior (captured by our predictive models) displays substantial non-monotonicity: For example, losing weight causes individuals to eat more and exercise less, so the speed of weight loss can impact the final weight loss outcomes.

1.4. Outline

Section 2 describes the first step of our behavioral analytics framework – the development of the behavioral model. The model consists of a utility function – describing how an agent makes decisions – and temporal dynamics on the system states and parameters of the utility function. We refer jointly to both components of this abstract model as the behavioral model. Section 3 presents approaches for estimating parameters of this behavioral model using MILP formulations to solve the problems of maximum likelihood estimation (MLE) and Bayesian inference. We prove that solutions of our MILP formulations provide consistent estimates of the agent's parameters. In Section 4, we present algorithms for optimizing the incentives provided to agents by the coordinator. We first present an algorithm based on solving two MILP's that allows the coordinator to allocate incentives in the situation where there is only a single agent with unknown-to-the-coordinator parameters, and we prove that this algorithm computes incentives that are asymptotically optimal (in the sense of minimizing the coordinator's loss function) as time t goes to infinity. Next, we develop a two-stage decomposition algorithm (building on the single-agent formulation) to solve the coordinator's problem in a multi-agent setting, and we generalize our proof of asymptotic optimality to this setting Finally in Section 5, we study (via simulation) the effectiveness of our algorithms for designing personalized weight loss treatment plans. Our results show that treatment plans computed by our behavioral analytics approach could potentially reduce the cost of running such weight loss programs by as much as 60% without affecting the efficacy of these programs.

2. Predictive Modeling of a Single Myopic Agent

In this section, we present our behavioral model for a single myopic agent. This forms the first step of our specific implementation of a behavioral analytics framework, and the key design problem is formulating a predictive model that is amenable to performing the second and third steps of our behavioral analytics framework of parameter estimation and incentive optimization. This model is an extension and abstraction of a behavioral model that was validated in our past work on behavioral modeling for weight loss (Aswani et al. 2016), in which we used cross-validation (i.e., out-of-sample comparisons) to perform a data-based validation of the predictive accuracy of our behavioral model by comparison to a standard machine learning algorithm for prediction.

Let $\mathcal{X}, \mathcal{U}, \Pi, \Theta$ be compact finite-dimensional sets with $\mathcal{X}, \mathcal{U}, \Theta$ convex. We will refer to the agent's system states $x_t \in \mathcal{X}$, motivational states (or type) $\theta_t \in \Theta$, and decisions $u_t \in \mathcal{U}$ at time t. The coordinator provides an incentive (or input) $\pi_t \in \Pi$ to the agent at time t, and we assume that the motivational states are unknown to the coordinator but known to the agent. In our behavioral model, the system and motivational states are subject to temporal dynamics:

$$x_{t+1} = h(x_t, u_t),$$

$$\theta_{t+1} = g(x_t, u_t, \theta_t, \pi_t).$$
(1)

The intuition of the above dynamics is that future system states x_{t+1} depend on the current system states x_t and decision u_t , while future motivational states θ_{t+1} depend on the current system states x_t , decision u_t , motivational states θ_t , and incentives π_t .

The agents are modeled to be myopic in the sense that agents make decisions at time t by considering only their present utility function. We assume the agent's utility function belongs to a parametrized class of functions $\mathcal{F} := \{(x, u) \mapsto f(x, u, \theta, \pi) : \theta \in \Theta, \pi \in \Pi\}$; and the agent's utility function at time t is $f(\cdot, \cdot, \theta_t, \pi_t)$. Thus at time t the agent's decisions are

$$u_t \in \arg\max\left\{f(x_{t+1}, u, \theta_t, \pi_t) \mid x_{t+1} = h(x_t, u), \ u \in \mathcal{U}\right\},$$
(2)

which means we are assuming the agent has perfect knowledge of x_t, θ_t, π_t . This model says that the agent's decisions depend on the current system states, motivational states, and incentives.

Though the coordinator also has perfect knowledge of the incentives π_t , the coordinator can only make noisy observations of past system states and agent decisions:

$$\widetilde{x}_{t_i} = Dx_{t_i} + \nu_{t_i} \quad \forall i = 0, \dots, n_x,
\widetilde{y}_{\tau_i} = Cu_{\tau_i} + \omega_{\tau_i} \quad \forall i = 0, \dots, n_u,$$
(3)

where C, D are known output matrices, and x_{t_i}, u_{τ_i} are the systems states and agent decisions generated by (1) and (2) with initial conditions (x_0, θ_0) and incentives (π_1, π_2, \ldots) .

For our subsequent optimization modeling and theoretical analysis, we make the following assumptions about this behavioral model:

ASSUMPTION 1. The sets $\mathcal{X}, \mathcal{U}, \Pi, \Theta$ are bounded and finite-dimensional. Moreover, the sets $\mathcal{X}, \mathcal{U}, \Theta$ are convex polyhedra described by a finite number of linear inequalities, and Π can be described by a finite number of mixed integer linear constraints.

This mild assumption ensures that states, decisions, and inputs are bounded; that the range of possible values for states and inputs are polytopes; and that the set of possible incentives is representable by mixed integer linear constraints.

ASSUMPTION 2. The function $f: \mathcal{X} \times \mathcal{U} \times \Theta \times \Pi \to \mathbb{R}$ is deterministic, concave in x, strictly concave in u, and concave in θ ; moreover, f can be expressed as

$$f(x, u, \theta, \pi) = -(x; u)^T \cdot Q \cdot (x; u) + (\theta; \pi)^T \cdot H \cdot (x; u) + \sum_{i=1}^K \min_{j \in J_i} \left\{ F_{i,j} \cdot (x; u; \theta; \pi) + \zeta_{i,j} \right\}, \quad (4)$$

where Q is a positive semidefinite matrix, the $F_{i,j}$, H are matrices of appropriate dimension, the $\zeta_{i,j}$ are scalars, and the J_i are sets of indices.

Strict concavity in u ensures $\arg \max_{u \in \mathcal{U}} f(x, u, \theta, \pi)$ is singleton for all $(x, \theta, \pi) \in \mathcal{X} \times \Theta \times \Pi$, and the concavity assumptions also model diminishing returns and ensure u_t is polynomial-time computable by the agent (Brock and Wartman 1990, Gafni 1990, Cawley 2004).

ASSUMPTION 3. The functions $h: \mathcal{X} \times \mathcal{U} \times \Theta \times \Pi \to \mathcal{X}$ and $g: \mathcal{X} \times \mathcal{U} \times \Theta \times \Pi \to \Theta$ are deterministic surjective functions of the form

$$h(x,u) = Ax + Bu + k$$

$$g(x,u,\theta,\pi) = G_i \cdot (x;u;\theta;\pi) + \chi_i \text{ when } B_i \cdot (x;u;\theta;\pi) \le \psi_i$$
(5)

where A, B, G_i, B_i are matrices; γ_i, ψ_i, k are vectors; χ_i are scalars; and the interiors of the polytopes $B_i \cdot (x; u; \theta; \pi) \leq \psi_i$ are disjoint.

This condition on h, g allows us to formulate problems of statistical estimation as a MILP.

ASSUMPTION 4. The $\{\nu_{t_i}\}_{i=0}^{n_x}$ and $\{\omega_{\tau_i}\}_{i=0}^{n_u}$ from the measurement noise model (3) are sequences of *i.i.d* random vectors with *i.i.d* components with zero mean and (known) finite variance. Moreover, the logarithm of their probability density functions can be expressed using integer linear constraints.

This means $\mathbb{E}\omega_{\tau_i} = \mathbb{E}\nu_{t_i} = 0$ and $\mathbb{E}(\nu_{t_i})_j^2 = \sigma_{\nu}^2 < \infty$ and $\mathbb{E}(\omega_{\tau_i})_j^2 = \sigma_{\omega}^2 < \infty$ with known $\sigma_{\nu}^2, \sigma_{\omega}^2$. Examples of noise distributions satisfying the integer linear representability assumption include the Laplace distribution, the shifted exponential distribution, and piecewise linear distributions. This assumption can be relaxed to requiring integer quadratic representability (such as is the case for Gaussian distributed noise), and the subsequent results change in that the optimization formulations become MIQP's, rather than the MILP's that occur with the above assumption.

ASSUMPTION 5. The discrete-time system with temporal dynamics (1) and (2) and measurement model (3) is observable (i.e., there exists a T and sequence π_t such that (x_0, θ_0) can be exactly computed if the measurements from $0 \le t \le T$ are noiseless).

This last assumption is an identifiability condition (Bickel and Doksum 2006), meaning that different initial conditions (x_0, θ_0) on the agent's system and motivational states produce different sequences of measurements and states, and this is a common assumption for control systems (Callier and Desoer 1994). The second assumption is common for utility functions (Brock and Wartman 1990, Gafni 1990, Cawley 2004). The third assumption says the system state dynamics are linear, and that the motivational state dynamics are piecewise affine, which are common models for control systems (Callier and Desoer 1994, Mignone et al. 2000, Aswani and Tomlin 2009). We believe all five assumptions are satisfied by agents in the three examples of weight loss programs, demand response programs, and franchise logistics. Section 5 provides a behavioral model for agents in a weight loss program that satisfies our above assumptions, and we conclude that section with a computational study where we solve the coordinator's problem for a weight loss program.

3. Estimating Model Parameters

In this section, we explore how the coordinator can estimate the agent's initial parameters (x_0, θ_0) , and predict the agent's future behavior for a fixed policy π . This forms the second step of our implementation of a behavioral analytics framework, and we leverage the mathematical structure of the behavioral model described in Section 2 to construct techniques and methods for estimation and prediction. This second step is important because the estimated parameters of the behavioral model and subsequent predictions of future agent behavior are used to optimize incentives in the third step of our behavioral analytics approach. We will assume the coordinator makes noisy and partial observations – according to the measurement model (3) – of the agent's state and decisions for ntime periods (with some missing observations). In Section 3.1, we present a Maximum Likelihood Estimation approach to estimate the agent's initial system states and motivational states. In Section 3.2, we consider a setting in which the coordinator has some prior knowledge about the possible values of the motivational states, and consider a Bayesian setting.

3.1. Maximum Likelihood Estimation

Let $\{\tilde{x}_{t_i}\}_{i=0}^{n_x}$ denote the process of the state observations, and let $\{\tilde{y}_{\tau_i}\}_{i=0}^{n_u}$ denote the process of the behavior observations.

Our approach to estimating the agent's initial parameters will be to compute estimates $(\hat{x}_0, \hat{\theta}_0) \in \arg\min_{(x_0,\theta_0)\in\mathcal{X}\times\Theta} \mathcal{L}(x_0,\theta_0,\{\tilde{x}_{t_i}\}_{i=0}^{n_x},\{\tilde{y}_{\tau_i}\}_{i=0}^{n_u};\pi)$ by minimizing an appropriately chosen loss function \mathcal{L} . More specifically, we use the approach of Maximum Likelihood Estimation (MLE). Let p_{ν}, p_{ω} be the density functions of $\nu_{t_i}, \omega_{\tau_i}$; then the joint likelihood function of (θ_0, x_0) for a fixed π is

$$\mathcal{L}(x_0, \theta_0, \{\tilde{x}_{t_i}\}_{i=0}^{n_x}, \{\tilde{y}_{\tau_i}\}_{i=0}^{n_u}, \pi) = p(\{\tilde{x}_{t_i}\}_{i=0}^{n_x}, \{\tilde{y}_{\tau_i}\}_{i=0}^{n_u} | \theta_0, x_0, \pi)$$

$$= \prod_{i=0}^{n_x} p_{\nu}(\tilde{x}_{t_i} - Dx_{t_i}) \prod_{j=0}^{n_u} p_{\omega}(\tilde{y}_{\tau_j} - Cu_{\tau_j})$$
(6)

Thus the coordinator's estimation problem is given by the following:

$$(\hat{x}_{0}, \hat{\theta}_{0}) \in \underset{(x_{t}, \theta_{t}, u_{t})}{\operatorname{arg\,max}} \sum_{i=0}^{n_{x}} \log p_{\nu} (\tilde{x}_{t_{i}} - Dx_{t_{i}}) + \sum_{j=0}^{n_{u}} \log p_{\omega} (\tilde{y}_{\tau_{j}} - Cu_{\tau_{j}})$$
s.t. $u_{t} \in \arg \max \left\{ f(x_{t+1}, u, \theta_{t}, \pi_{t}) \mid x_{t+1} = h(x_{t}, u), \ u \in \mathcal{U} \right\} \quad \forall t,$

$$x_{t+1} = h(x_{t}, u_{t}) \qquad \forall t,$$

$$\theta_{t+1} = g(x_{t}, u_{t}, \theta_{t}, \pi_{t}) \qquad \forall t,$$

$$(7)$$

$$x_t \in \mathcal{X}, \theta_t \in \Theta \qquad \qquad \forall t.$$

Problem (7) is a bilevel optimization problem because the u_t are minimizers of $f(x_{t+1}, \cdot, \theta_t, \pi_t)$, and such bilevel problems frequently arise in the context of estimating utility functions (Keshavarz et al. 2011, Bertsimas et al. 2014, Aswani et al. 2015). For the setting we consider in this paper, we show that the bilevel program for MLE (7) can be exactly reformulated as a MILP.

PROPOSITION 1. If Assumptions 1–5 hold; then the feasible region of (7) can be formulated as a set of mixed integer linear constraints with respect to $(x_t, u_t, \theta_t, \pi_t)$.

Proof: The constraints $\theta_{t+1} = g(x_t, u_t, \theta_t, \pi_t)$ can be reformulated using Assumption 3 as $\theta_{t+1} \leq G_i \cdot (x_t; u_t; \theta_t; \pi_t) + \xi_i + (1 - \iota_i) \cdot M$ $\theta_{t+1} \geq G_i \cdot (x_t; u_t; \theta_t; \pi_t) + \xi_i - (1 - \iota_i) \cdot M$ $B_i \cdot (x_t; u_t; \theta_t; \pi_t) \leq \psi_i + (1 - \iota_i) \cdot M$ $\iota_i \in \{0, 1\}$

where M > 0 is a large-enough constant. Such a finite M exists because $\mathcal{X}, \mathcal{U}, \Pi, \Theta$ are compact. Hence it suffices to show $u_t \in \arg \max \{f(x_{t+1}, u, \theta_t, \pi_t) \mid x_{t+1} = h(x_t, u), u \in \mathcal{U}\}$ can be represented (by its optimality condition) using a finite number of mixed integer linear constraints. Suppose $\mathcal{U} = \{u : \Xi u \leq \kappa\}$, where Ξ is a matrix and κ is a vector. Recall Assumption 2

$$f(x, u, \theta, \pi) = -(x; u)^T \cdot Q \cdot (x; u) + (\theta; \pi)^T \cdot H \cdot (x; u) + \sum_{i=1}^K \min_{j \in J_i} \left\{ F_{i,j} \cdot (x; u; \theta; \pi) + \zeta_{i,j} \right\}.$$
(9)

We cannot characterize optimality by differentiating f because it is generally not differentiable, but we can reformulate the maximization of (9) as the following convex quadratic program:

$$u_{t} \in \arg \max - (x_{t+1}; u)^{T} \cdot Q \cdot (x_{t+1}; u) + (\theta_{t}; \pi_{t})^{T} \cdot H \cdot (x_{t+1}; u) + \sum_{i=1}^{K} w_{i}$$

s.t. $w_{i} \leq F_{i,j} \cdot (x_{t+1}; u; \theta_{t}; \pi_{t}) + \zeta_{i,j}$ for all i, j
 $\Xi u \leq \kappa$ (10)

Using Assumption 3, we can rewrite the above as

$$u_{t} \in \arg \max - u^{T} \cdot (B; \mathbb{I})^{T} \cdot Q \cdot (B; \mathbb{I}) \cdot u + ((\theta_{t}; \pi_{t})^{T} H - 2(Ax + k; 0)^{T} Q) \cdot (B; \mathbb{I}) \cdot u + \sum_{i=1}^{K} w_{i}$$

s.t. $w_{i} \leq F_{i,j} \cdot (B; \mathbb{I}; 0; 0) \cdot u + F_{i,j} \cdot (Ax_{t} + k; 0; \theta_{t}; \pi_{t}) + \zeta_{i,j}$ for all i, j
 $\Xi u \leq \kappa$ (11)

where we have eliminated the constant $(\theta_t; \pi_t)^T \cdot H \cdot (Ax_t + k; 0)$ since x_t, θ_t, π_t are known to the agent. The above optimization problem is strictly convex by assumption, and all constraints are linear for fixed π_t . Hence the optimality conditions for (11) can be characterized using the KKT conditions (Dempe 2002, Boyd and Vandenberghe 2004). Let $\lambda_{i,j}$ and μ be the Lagrange Multipliers for the first and second set of constraints given in (11), and note the KKT conditions are

$$2(B;\mathbb{I})^{T} \cdot Q \cdot (B;\mathbb{I}) \cdot u + (B;\mathbb{I})^{T} \cdot (2Q \cdot (Ax+k;0) - H^{T} \cdot (\theta_{t};\pi_{t})) + \Xi^{T} \cdot \mu = \sum_{i=1}^{K} \sum_{j \in J_{i}} \lambda_{i,j} \cdot (B;\mathbb{I};0;0)^{T} \cdot F_{i,j}^{T}$$

$$\sum_{i=1}^{K} \sum_{j \in J_{i}} \lambda_{i,j} \cdot (B;\mathbb{I};0;0)^{T} \cdot F_{i,j}^{T}$$

$$\sum_{i=1}^{K} \sum_{j \in J_{i}} \lambda_{i,j} \cdot (B;\mathbb{I};0;0)^{T} \cdot F_{i,j}^{T}$$

$$\sum_{i=1}^{K} \sum_{j \in J_{i}} \lambda_{i,j} \cdot (B;\mathbb{I};0;0)^{T} \cdot F_{i,j}^{T}$$

$$\sum_{i=1}^{K} \sum_{j \in J_{i}} \lambda_{i,j} \cdot (B;\mathbb{I};0;0) \cdot u + F_{i,j} \cdot (Ax_{t}+k;0;\theta_{t};\pi_{t}) + \zeta_{i,j}) = 0 \text{ for all } i,j$$

$$w_{i} \leq F_{i,j} \cdot (B;\mathbb{I};0;0) \cdot u + F_{i,j} \cdot (Ax_{t}+k;0;\theta_{t};\pi_{t}) + \zeta_{i,j} \text{ for all } i,j$$

$$\Xi u \leq \kappa \text{ and } \lambda_{i,j} \geq 0 \text{ for all } i,j$$

$$(12)$$

(8)

Note that the only nonlinear conditions are those which represent complimentary slackness. However, these conditions can be reformulated as integer linear conditions by posing them as disjunctive constraints (Wolsey and Nemhauser 1999): For sufficiently large M – which exists because of the compactness of $\mathcal{X}, \mathcal{U}, \Pi, \Theta$ – the complimentary slackness conditions are

$$\lambda_{i,j} \leq M \iota_{i,j} \text{ for all } i, j$$

$$F_{i,j} \cdot (B; \mathbb{I}; 0; 0) \cdot u + F_{i,j} \cdot (Ax_t + k; 0; \theta_t; \pi_t) + \zeta_{i,j} \leq w_i + M \cdot (1 - \iota_{i,j}) \text{ for all } i, j$$

$$\iota_{i,j} \in \{0, 1\} \text{ for all } i, j$$

$$(13)$$

This shows that feasible region of (7) of can be represented using a finite number of mixed integer linear constraints. \Box

An important consequence of this proposition is that it is possible to compute the global solution of the MLE problem (7) using standard optimization software.

COROLLARY 1. If Assumptions 1–5 hold, then the MLE problem (7) can be expressed as a MILP. REMARK 1. If the logarithm of the noise densities can be expressed using integer quadratic constraints (e.g., Gaussian distributions), then the MLE problem (7) can be expressed as a MIQP.

3.2. Bayesian Estimation

Solving the MLE problem (7) gives an estimate of the agent's initial system states and motivational states, which completely characterize the agent. However, the coordinator often has some prior knowledge about the possible values of the motivational states. In such a case, a Bayesian framework is a natural setting for making predictions of the agent's future system states.

Suppose the coordinator has interacted with the agent over T time periods, has measured $\{\tilde{x}_{t_i}\}_{i=0}^{n_x}, \{\tilde{y}_{\tau_i}\}_{i=0}^{n_u}$ with $T_{n_x}, T_{n_u} \leq T$, and wants to predict the agent's future states and decisions $\{x_i, \theta_i, u_i\}_{i=T}^{T+n}$ for some n > 0 time steps into the future. In principle, this means the coordinator wants to calculate the posterior distribution of $\{x_i, \theta_i, u_i\}_{i=T}^{T+n}$. But (x_0, θ_0) completely characterize the agent in our model (recall Assumption 5 states that distinct initial conditions produce different state and decision trajectories), and so we can predict the agent's future states and decisions using the posterior distribution of (x_0, θ_0) . Hence we focus on computing the posterior of (x_0, θ_0) . A direct application of Bayes's Theorem (Bickel and Doksum 2006) gives

$$p(x_0,\theta_0|\{\tilde{x}_{t_i}\}_{i=0}^{n_x},\{\tilde{y}_{\tau_i}\}_{i=0}^{n_u},\{\pi_i\}_{i=0}^{T+n}) = Z^{-1} \times p(\{\tilde{x}_{t_i}\}_{i=0}^{n_x},\{\tilde{y}_{\tau_i}\}_{i=0}^{n_u}|x_0,\theta_0,\{\pi_i\}_{i=0}^{T+n}) \times p(x_0,\theta_0).$$
(14)

Here Z is a normalization constant that ensures the right hand side is a probability distribution, and $p(x_0, \theta_0)$ reflects the coordinator's prior beliefs. We begin with an assumption on $p(x_0, \theta_0)$.

ASSUMPTION 6. The function $\log p(x_0, \theta_0)$ can be expressed using a finite number of mixed integer linear constraints, and $p(x_0, \theta_0) > 0$ for all $(x_0, \theta_0) \in \mathcal{X} \times \Theta$.

This is a mild assumption because it holds for the Laplace distribution, the shifted exponential distribution, and piecewise linear distributions. Significantly, it is true when the prior distribution $p(x_0, \theta_0)$ is an empirical histogram with data in each histogram bin (Aswani et al. 2016).

Next, we describe an optimization approach to computing the posterior distribution of (x_0, θ_0) . Consider the following feasibility problem:

$$\begin{split} \psi_{T}(\bar{x}_{0},\bar{\theta}_{0}) &= \log p\left(x_{0} = \bar{x}_{0}, \theta_{0} = \bar{\theta}_{0} \middle| \{\tilde{x}_{t_{i}}\}_{i=0}^{n_{x}}, \{\tilde{y}_{\tau_{i}}\}_{i=0}^{n_{u}}, \{\pi_{i}\}_{i=0}^{T}\right) + \log Z \\ &= \max_{(x_{t},\theta_{t},u_{t})} \sum_{i=0}^{n_{x}} \log p_{\nu}(\tilde{x}_{t_{i}} - Dx_{t_{i}}) + \sum_{j=0}^{n_{u}} \log p_{\omega}(\tilde{y}_{\tau_{j}} - Cu_{\tau_{j}}) + \log p(x_{0},\theta_{0}) \\ &\text{s.t.} \quad u_{t} \in \arg \max \left\{ f(x_{t+1}, u, \theta_{t}, \pi_{t}) \middle| x_{t+1} = h(x_{t}, u), \ u \in \mathcal{U} \right\} \quad \forall t, \\ &x_{t+1} = h(x_{t}, u_{t}) \quad \forall t, \\ &\theta_{t+1} = g(x_{t}, u_{t}, \theta_{t}, \pi_{t}) \quad \forall t, \\ &x_{0} = \bar{x}_{0}, \theta_{0} = \bar{\theta}_{0}, \end{split}$$

$$x_t \in \mathcal{X}, \theta_t \in \Theta \qquad \qquad \forall t.$$

The above problem is almost the same as the MLE problem (7), with the only differences that the above has additional constraints $x_0 = \bar{x}_0, \theta_0 = \bar{\theta}_0$ and an additional term in the objective $\log p(x_0, \theta_0)$. Thus we have that the above problem (15) can be expressed as a MILP or MIQP.

COROLLARY 2. If Assumptions 1-6 hold, then (15) can be formulated as a MILP.

REMARK 2. Under appropriate relaxed representability conditions on the noise distributions and the prior distribution, the problem (15) can be formulated as a MIQP.

Solving (15) does not directly provide the posterior distribution of (x_0, θ_0) because Z is not known a priori, though it can be computed using numerical integration. (See for instance the approach by Aswani et al. (2016).) But since Z only scales the posterior estimate, we instead propose a simpler scaling. Let $(\hat{x}_{0,T}, \hat{\theta}_{0,T}) \in \arg \max_{(x_0,\theta_0)} \psi_T(x_0,\theta_0)$ be the maximum a posteriori (MAP) estimates of the initial conditions, and note that the above corollaries apply to the computation of the MAP because the corresponding optimization problem for computing the MAP is simply (15) but with the constraints $x_0 = \bar{x}_0, \theta_0 = \bar{\theta}_0$ removed. We propose using

$$\hat{p}(x_0, \theta_0 | \{\tilde{x}_{t_i}\}_{i=0}^{n_x}, \{\tilde{y}_{\tau_i}\}_{i=0}^{n_u}, \{\pi_i\}_{i=0}^T) = \frac{\exp(\psi_T(x_0, \theta_0))}{\exp(\psi_T(\hat{x}_{0,T}, \hat{\theta}_{0,T}))}$$
(16)

as an estimate of the posterior distribution of (x_0, θ_0) . Two useful properties of our estimate are that $\hat{p}(x_0, \theta_0 | \{\tilde{x}_{t_i}\}_{i=0}^{n_x}, \{\tilde{y}_{\tau_i}\}_{i=0}^{n_u}, \{\pi_i\}_{i=0}^T) \in [0, 1]$ by construction, and that $\hat{p}(\hat{x}_{0,T}, \hat{\theta}_{0,T} | \{\tilde{x}_{t_i}\}_{i=0}^{n_x}, \{\tilde{y}_{\tau_i}\}_{i=0}^{n_u}, \{\pi_i\}_{i=0}^T) = 1$ by construction. We will show this estimate is statistically consistent in a Bayesian sense (Bickel and Doksum 2006): DEFINITION 1. The posterior estimate (16) is consistent if for all $(x_0^*, \theta_0^*) \in \mathcal{X} \times \Theta$ and $\epsilon, \delta > 0$ we have $p_{(x_0^*, \theta_0^*)}(\hat{p}(\mathcal{E}(\delta) | \{\tilde{x}_{t_i}\}_{i=0}^{n_x}, \{\tilde{y}_{\tau_i}\}_{i=0}^{n_u}, \{\pi_i\}_{i=0}^T) \ge \epsilon) \to 0$ as $T \to \infty$, where $p_{(x_0^*, \theta_0^*)}$ is the probability law under (x_0^*, θ_0^*) , $\mathcal{E}(\delta) = \{(x_0, \theta_0) \notin \mathcal{B}(x_0^*, \theta_0^*, \delta)\}$, and $\mathcal{B}(x_0^*, \theta_0^*, \delta)$ is an open δ ball around (x_0^*, θ_0^*) .

The meaning of this definition is that if (x_0^*, θ_0^*) are the true initial conditions of the agent, then a consistent posterior estimate is such that it collapses until all probability mass is on the true initial conditions. Statistical consistency of (16) also needs an additional technical assumption:

ASSUMPTION 7. Let (x_0^*, θ_0^*) be the agent's true initial conditions. The incentives π_t are such that

$$\max_{\mathcal{E}(\delta)} \lim_{T \to \infty} \sum_{i=0}^{n_x} \log \frac{p_\nu(\tilde{x}_{t_i} - D\overline{x}_{t_i})}{p_\nu(\tilde{x}_{t_i} - Dx_{t_i})} + \sum_{j=0}^{n_u} \log \frac{p_\omega(\tilde{y}_{\tau_j} - C\overline{u}_{\tau_j})}{p_\omega(\tilde{y}_{\tau_j} - Cu_{\tau_j})} = -\infty$$
(17)

for any $\delta > 0$, almost surely, where x_t, u_t are the states and decisions under initial conditions (x_0^*, θ_0^*) , and $\overline{x}_t, \overline{u}_t$ are the states and decisions under initial conditions (x_0, θ_0) .

This type of assumption is common in the adaptive control literature (Craig et al. 1987, Astrom and Wittenmark 1995), and is known as a *sufficient excitation* or a *sufficient richness* condition. It is a mild condition because there are multiple ways of ensuring this condition holds (Bitmead 1984, Craig et al. 1987, Astrom and Wittenmark 1995). One simple approach (Bitmead 1984) is to compute an input π_t and then add a small amount of random noise (whose value is known since it is generated by the coordinator) to the input before applying the input to the agent.

PROPOSITION 2. If Assumptions 1–7 hold, then the estimated posterior distribution denoted by $\hat{p}(x_0, \theta_0 | \{\tilde{x}_{t_i}\}_{i=0}^{n_x}, \{\tilde{y}_{\tau_i}\}_{i=0}^{n_u}, \{\pi_i\}_{i=0}^T)$ and given in (16) is consistent.

Proof: Let (x_0^*, θ_0^*) be the agent's true initial conditions, and observe that

$$\log \hat{p}(x_{0},\theta_{0}|\{\tilde{x}_{t_{i}}\}_{i=0}^{n_{x}},\{\tilde{y}_{\tau_{i}}\}_{i=0}^{n_{u}},\{\pi_{i}\}_{i=0}^{T}) = \log \hat{p}(x_{0}^{*},\theta_{0}^{*}|\{\tilde{x}_{t_{i}}\}_{i=0}^{n_{x}},\{\tilde{y}_{\tau_{i}}\}_{i=0}^{n_{u}},\{\pi_{i}\}_{i=0}^{T}) + \sum_{i=0}^{n_{x}} \log \frac{p_{\nu}(\tilde{x}_{t_{i}}-D\overline{x}_{t_{i}})}{p_{\nu}(\tilde{x}_{t_{i}}-Dx_{t_{i}})} + \sum_{j=0}^{n_{u}} \log \frac{p_{\omega}(\tilde{y}_{\tau_{j}}-C\overline{u}_{\tau_{j}})}{p_{\omega}(\tilde{y}_{\tau_{j}}-Cu_{\tau_{j}})} + \log \frac{p(x_{0},\theta_{0})}{p(x_{0}^{*},\theta_{0}^{*})}, \quad (18)$$

where x_t, u_t are the states and decisions under initial conditions (x_0^*, θ_0^*) , and $\overline{x}_t, \overline{u}_t$ are the states and decisions under initial conditions (x_0, θ_0) . But $\log \frac{p(x_0, \theta_0)}{p(x_0^*, \theta_0^*)}$ is a constant by assumption, and $\log \hat{p}(x_0^*, \theta_0^* | \{\tilde{x}_{t_i}\}_{i=0}^{n_x}, \{\tilde{y}_{\tau_i}\}_{i=0}^{n_u}, \{\pi_i\}_{i=0}^T) \leq 0$ since $\hat{p}(x_0, \theta_0 | \{\tilde{x}_{t_i}\}_{i=0}^{n_x}, \{\tilde{y}_{\tau_i}\}_{i=0}^{n_u}, \{\pi_i\}_{i=0}^T) \in [0, 1]$ by construction. So using Assumption 6 gives $\max_{\mathcal{E}(\delta)} \log \hat{p}(x_0, \theta_0 | \{\tilde{x}_{t_i}\}_{i=0}^{n_x}, \{\tilde{y}_{\tau_i}\}_{i=0}^{n_u}, \{\pi_i\}_{i=0}^T) \to -\infty$ for any $\delta > 0$ almost surely. Equivalently, $\max_{\mathcal{E}(\delta)} \hat{p}(x_0, \theta_0 | \{\tilde{x}_{t_i}\}_{i=0}^{n_x}, \{\tilde{y}_{\tau_i}\}_{i=0}^{n_u}, \{\pi_i\}_{i=0}^T) \to 0$ for any $\delta > 0$ almost surely. Thus for any $\delta > 0$ we have that

$$\hat{p}(\mathcal{E}(\delta)|\{\tilde{x}_{t_{i}}\}_{i=0}^{n_{x}},\{\tilde{y}_{\tau_{i}}\}_{i=0}^{n_{u}},\{\pi_{i}\}_{i=0}^{T}) = \int_{\mathcal{E}(\delta)} \hat{p}(x_{0},\theta_{0}|\{\tilde{x}_{t_{i}}\}_{i=0}^{n_{x}},\{\tilde{y}_{\tau_{i}}\}_{i=0}^{n_{u}},\{\pi_{i}\}_{i=0}^{T}) \times dx_{0} \times d\theta_{0} \leq$$

$$\text{volume}(\mathcal{X} \times \Theta) \cdot \max_{\mathcal{E}(\delta)} \hat{p}(x_{0},\theta_{0}|\{\tilde{x}_{t_{i}}\}_{i=0}^{n_{x}},\{\tilde{y}_{\tau_{i}}\}_{i=0}^{n_{u}},\{\pi_{i}\}_{i=0}^{T}) \to 0 \quad (19)$$

almost surely. This proves the result since (19) holds almost surely for any $\delta > 0$. \Box

COROLLARY 3. If Assumptions 1–7 hold, then $(\hat{x}_{0,T}, \hat{\theta}_{0,T}) \xrightarrow{p} (x_0^*, \theta_0^*)$ as $T \to \infty$.

The above two results imply that future agent behavior can be reasonably predicted using the MAP parameters. Recall that calculating the MAP can be formulated as a MILP or MIQP, since the corresponding optimization problem is (15) with the constraints $x_0 = \bar{x}_0, \theta_0 = \bar{\theta}_0$ removed.

4. Optimizing Incentives

The final step of our behavioral analytics framework involves using estimates of behavioral model parameters for each agent to optimize the design of costly incentives provided to the agents by the coordinator. In Section 4.1, we develop an algorithm for the single agent case. In Section 4.2, we use this single-agent algorithm as a sub-problem in the multi-agent case. In both cases, we show that our algorithms are asymptotically optimal (as time continues and more data is collected) with respect to the coordinator's loss function when the agents behave according to the model constructed in Section 2. The two algorithms we present in fact combine the second and third steps of our framework by first applying the parameter estimation algorithms (described in Section 3) that comprise the second step, and then optimizing incentives. The benefit of combining the second and third steps into a single algorithm is that this makes it easier to recompute the incentives as more data is collected over time from each agent.

4.1. Optimizing Incentives for a Single Agent

Consider the problem of designing optimal incentives for a single agent at time T by choosing $\{\pi_i\}_{i=T+1}^{T+n} \in \Pi^n$ to minimize a bounded loss function $\ell : \mathcal{X}^n \times \mathcal{U}^n \to \mathbb{R}$ of the agent's system states and decisions over the next n time periods. We consider losses of a fairly general form:

Assumption 8. The loss function ℓ can be described by mixed integer linear constraints.

Since the coordinator only has noisy and incomplete observations of the agent's system states and decisions $\{\tilde{x}_{t_i}\}_{i=0}^{n_x}, \{\tilde{y}_{\tau_i}\}_{i=0}^{n_u}$, one design approach is to minimize the expected posterior loss

$$\min\left\{ \mathbb{E}\left[\ell(\{x_t, u_t\}_{t=T+1}^{T+n}) \middle| \{\tilde{x}_{t_i}\}_{i=0}^{n_x}, \{\tilde{y}_{\tau_i}\}_{i=0}^{n_u}, \{\pi_i\}_{i=0}^{T}\right] \middle| \{\pi_i\}_{i=T+1}^{T+n} \in \Pi^n \right\}.$$
(20)

However, recalling our previous discussion, the agent's behavior is completely characterized by the initial conditions (x_0, θ_0) , and so by the sufficiency and the smoothing theorem (Bickel and Doksum 2006), there exists $\varphi : \mathcal{X} \times \Theta \times \Pi^n \mapsto \mathbb{R}$ such that the design problem can be exactly reformulated as

$$\min\left\{\mathbb{E}\left[\varphi(x_0,\theta_0,\{\pi_i\}_{i=0}^{T+n})\big|\{\tilde{x}_{t_i}\}_{i=0}^{n_x},\{\tilde{y}_{\tau_i}\}_{i=0}^{n_u},\{\pi_i\}_{i=0}^{T}\right] \mid \{\pi_i\}_{i=T+1}^{T+n} \in \Pi^n\right\}.$$
(21)

Calculating this expectation is difficult because the posterior distribution of (x_0, θ_0) does not generally have a closed form expression. In principle, discretization approaches from scenario generation (Kaut and Wallace 2003) could be used to approximate the design problem as

$$\min\left\{\frac{\sum_{i=1}^{M}\varphi(x_{i,0},\theta_{i,0},\pi)\exp(\psi_{T}(x_{i,0},\theta_{i,0};\pi))}{\sum_{i=1}^{M}\exp(\psi_{T}(x_{i,0},\theta_{i,0};\pi))} \mid \{\pi_{i}\}_{i=T+1}^{T+n} \in \Pi^{n}\right\}.$$
(22)

where $(x_{i,0}, \theta_{i,0})$ is an exhaustive enumeration of $\mathcal{X} \times \Theta$. This approximation (22) is still challenging to solve because the objective has a fractional, nonconvex form, and ψ_T is defined as the value function of a MILP, meaning that it does not have an easily computable closed form expression (Ralphs and Hassanzadeh 2014). This means (22) is a BMIP with lower level problems that are MILP's. This is a complex class of optimization problems for which existing algorithms can only solve small problem instances (James and Bard 1990, Moore and Bard 1992, DeNegre and Ralphs 2009).

In this section, we develop a practical algorithm for optimizing incentives for a single agent. We first summarize our algorithm, and show it only requires solving two MILP's. Next we prove this algorithm can be interpreted as solving an approximation of solving either (20) or the optimal incentive design problem under perfect noiseless information. More substantially, we also show that our algorithm provides a set of incentives that are asymptotically optimal as time advances.

4.1.1. Two Stage Adaptive Algorithm (2SSA) Algorithm 1 summarizes our two stage adaptive approach (2SSA) for designing optimal incentives for a single agent. The idea of the algorithm is to first compute a MAP estimate of the agent's initial conditions, use the MAP estimate as data for the first two arguments of φ , and then minimize φ . In fact, we can solve this minimization problem without having to explicitly compute φ . Because φ is defined as the composition of the agent's dynamics with initial conditions (x_0, θ_0) and the coordinator's loss function ℓ , it can be written as the value function of a feasibility problem:

$$\varphi(\overline{x}_{0}, \overline{\theta}_{0}, \{\overline{\pi}_{i}\}_{i=0}^{T+n}) = \min_{x_{t}, u_{t}, \theta_{t}, \pi_{t}} \ell(\{x_{t}, u_{t}\}_{t=T+1}^{T+n})$$
s.t. $u_{t} \in \arg \max \{f(x_{t+1}, u, \theta_{t}, \pi_{t}) \mid x_{t+1} = h(x_{t}, u), u \in \mathcal{U}\} \quad \forall t,$

$$x_{t+1} = h(x_{t}, u_{t}) \qquad \forall t,$$
(23)

$$\theta_{t+1} = g(x_t, u_t, \theta_t, \pi_t) \qquad \forall t,$$

$$x_t \in \mathcal{X}, \theta_t \in \Theta, \pi_t \in \Pi \qquad \qquad \forall t$$

$$x_0 = \overline{x}_0, \ \theta_0 = \overline{\theta}_0, \\ \pi_t = \overline{\pi}_t \qquad \forall t.$$

More importantly, the problem of minimizing this φ can be formulated as a MILP.

COROLLARY 4. If Assumptions 1-8 hold, then $\varphi(x_0, \theta_0, \{\pi_i\}_{i=0}^{T+n})$ is lower semicontinuous in $x_0, \theta_0, \{\pi_i\}_{i=T+1}^{T+n}$, and the optimization problem given by $\min\{\varphi(x_0, \theta_0, \{\pi_i\}_{i=0}^{T+n}) \mid \{\pi_i\}_{i=T+1}^{T+n} \in \Pi^n\}$ can be formulated as a MILP for all fixed values of $(x_0, \theta_0, \{\pi_i\}_{i=0}^T) \in \mathcal{X} \times \Theta \times \Pi^{T+1}$.

Algorithm 1 Two Stage Single Agent Algorithm (2SSA)
Require: $\{\tilde{x}_{t_i}\}_{i=0}^{n_x}, \{\tilde{u}_{\tau_i}\}_{i=0}^{n_u}, \{\pi_i\}_{i=0}^T$
1: compute $(\hat{x}_{0,T}, \hat{\theta}_{0,T}) \in \arg \max_{(x_0,\theta_0)} \psi_T(x_0, \theta_0)$
2: return $\pi_{2SSA}(T) \in \arg\min\{\varphi(\hat{x}_{0,T}, \hat{\theta}_{0,T}, \{\pi_i\}_{i=0}^{T+n}) \mid \{\pi_i\}_{i=T+1}^{T+n} \in \Pi^n\}$

Proof: For the first result, note Proposition 1 implies the feasible region of (23) can be expressed as mixed integer linear constraints with respect to $(x_t, u_t, \theta_t, \pi_t)$. Thus $\varphi(\overline{x}_0, \overline{\theta}_0, \{\overline{\pi}_i\}_{i=0}^{T+n})$ is the value function of a MILP in which $x_0, \theta_0, \overline{\pi}_t$ belong to an affine term. Standard results (Ralphs and Hassanzadeh 2014) imply the value function is lower semicontinuous with respect to $x_0, \theta_0, \{\pi_i\}_{i=T+1}^{T+n}$.

To show the second result, note that the problem of $\min\{\varphi(x_0, \theta_0, \{\pi_i\}_{i=0}^{T+n}) \mid \{\pi_i\}_{i=T+1}^{T+n} \in \Pi^n\}$ is equivalent to (23) but with removal of the constraints $\pi_t = \overline{\pi}_t$ for $t = T + 1, \ldots, T + n$. And so the result follows by Proposition 1 and by recalling the assumptions on Π and ℓ . \Box

4.1.2. Asymptotic Optimality of 2SSA The next result provides the underlying intuition of 2SSA. In particular, we are approximating $\mathbb{E}[\varphi(x_0, \theta_0, \{\pi_i\}_{t=0}^{T+n}) | \{\tilde{x}_{t_i}\}_{i=0}^{n_x}, \{\tilde{y}_{\tau_i}\}_{i=0}^{n_u}, \{\pi_i\}_{i=0}^{T}]$ using $\varphi(\hat{x}_{0,T}, \hat{\theta}_{0,T}, \{\pi_i\}_{i=0}^{T+n})$, and both these functions are converging to $\varphi(x_0^*, \theta_0^*, \{\pi_i\}_{t=0}^{T+n})$.

PROPOSITION 3. Suppose that Assumptions 1-8 hold. Then as $T \to \infty$ we have that: $\mathbb{E}[\varphi(x_0, \theta_0, \{\pi_i\}_{t=0}^{T+n}) | \{\tilde{x}_{t_i}\}_{i=0}^{n_x}, \{\tilde{y}_{\tau_i}\}_{i=0}^{n_u}, \{\pi_i\}_{i=0}^{T}] \xrightarrow{p} \varphi(x_0^*, \theta_0^*, \{\pi_i\}_{t=0}^{T+n}) \text{ for all fixed } \{\pi_i\}_{i=0}^{T+n}; \text{ and } \varphi(\hat{x}_{0,T}, \hat{\theta}_{0,T}, \{\pi_i\}_{t=0}^{T+n}) \xrightarrow{l-prob} \varphi(x_0^*, \theta_0^*, \{\pi_i\}_{t=0}^{T+n}). \text{ Here, } \Lambda_n \xrightarrow{l-prob}_{\mathcal{X}} \Lambda \text{ means random function } \Lambda_n : \mathcal{X} \to \mathbb{R} \text{ is a lower semicontinuous approximation to function } \Lambda : \mathcal{X} \to \mathbb{R} \text{ (Vogel and Lachout 2003a).}$

Proof: We begin by proving the first result. By definition we have that

$$\mathbb{E}\left[\varphi(x_{0},\theta_{0},\{\pi_{i}\}_{i=0}^{T+n})\big|\{\tilde{x}_{t_{i}}\}_{i=0}^{n_{x}},\{\tilde{y}_{\tau_{i}}\}_{i=0}^{n_{u}},\{\pi_{i}\}_{i=0}^{T}\right] = \int_{\mathcal{X}\times\Theta}\varphi(x_{0},\theta_{0},\{\pi_{i}\}_{i=0}^{T+n})p(x_{0},\theta_{0}|\{\tilde{x}_{t_{i}}\}_{i=0}^{n_{x}},\{\tilde{y}_{\tau_{i}}\}_{i=0}^{n_{u}},\pi)\times dx_{0}\times d\theta_{0}. \quad (24)$$

Also, Proposition 2 implies the posterior $p(x_0, \theta_0 | \{\tilde{x}_{t_i}\}_{i=0}^{n_x}, \{\tilde{y}_{\tau_i}\}_{i=0}^{n_u}, \{\pi_i\}_{i=0}^T)$ is consistent, and thus becomes degenerate at (x_0^*, θ_0^*) in the limit. Hence the Dominated Convergence Theorem gives

$$(24) \xrightarrow{p} \int_{\mathcal{X} \times \Theta} \varphi(x_0, \theta_0, \pi) \times \delta(x_0 - x_0^*) \times \delta(\theta_0 - \theta_0^*) \times dx_0 \times d\theta_0 = \varphi(x_0^*, \theta_0^*, \pi),$$
(25)

where in the equation above $\delta(\cdot)$ is the Dirac delta function.

For the second result, recall that Corollary 3 implies $(\hat{x}_{0,T}, \hat{\theta}_{0,T}) \xrightarrow{p} (x_0^*, \theta_0^*)$. And Corollary 4 gives that $\varphi(x_0, \theta_0, \{\pi_i\}_{i=0}^{T+n})$ is lower semicontinuous in $x_0, \theta_0, \{\pi_i\}_{i=T+1}^{T+n}$. The result then follows by direct application of Proposition 2.1.ii of (Vogel and Lachout 2003b). \Box

If the coordinator had perfect knowledge of the agent's true initial conditions (x_0^*, θ_0^*) , then the optimal incentives are $\arg\min\{\varphi(x_0^*, \theta_0^*, \{\pi_i\}_{i=0}^{T+n}) \mid \{\pi_i\}_{i=T+1}^{T+n} \in \Pi^n\}$. But since we do not know the initial conditions, the above result shows that both (20) and $\arg\min\{\varphi(\hat{x}_{0,T}, \hat{\theta}_{0,T}, \{\pi_i\}_{i=0}^{T+n}) \mid \{\pi_i\}_{i=T+1}^{T+n} \in \Pi^n\}$ are reasonable approximations. In fact, we can show a stronger result for the solution generated by 2SSA.

THEOREM 1. Note that $\arg\min\{\varphi(x_0^*, \theta_0^*, \{\pi_i\}_{i=0}^{T+n}) | \{\pi_i\}_{i=T+1}^{T+n} \in \Pi^n\}$ is the set of optimal solutions under the agent's true initial conditions (x_0^*, θ_0^*) . If Assumptions 1–8 hold, then we have that

dist
$$\left(\pi_{2SSA}(T), \arg\min\{\varphi(x_0^*, \theta_0^*, \{\pi_i\}_{i=0}^{T+n}) \mid \{\pi_i\}_{i=T+1}^{T+n} \in \Pi^n\}\right) \xrightarrow{p} 0$$
 (26)

as $T \to \infty$, for any $\pi_{2SSA}(T)$ returned by 2SSA. Note $\operatorname{dist}(x, B) = \inf_{y \in B} ||x - y||$.

Proof: The result follows by combining the second part of our Proposition 3 with Theorem 4.3 from (Vogel and Lachout 2003a). \Box

This result suggests that any solution returned by 2SSA is asymptotically included within the set of optimal incentives computed for the agent's true initial conditions. Restated, the result says 2SSA provides a set of incentives that are asymptotically optimal. This is a non-obvious result because in general pointwise-convergence of a sequence of stochastic optimization problems is not sufficient to ensure convergence of the minimizers of the sequence of optimization problems to the minimizer of the limiting optimization problem. Rockafellar and Wets (2009) provide an example that demonstrates this possible lack of convergence of minimizers.

4.2. Policy Calculation With Multiple Agents

We next study the general setting where the coordinator designs incentives for a large group of agents. We let \mathcal{A} be the set of agents, and the quantities corresponding to a specific agent $a \in \mathcal{A}$ are denoted using subscript a. Now suppose that at time T the coordinator measures $\{\tilde{x}_{t_i}^a\}_{i=0}^{n_x^a}, \{\tilde{y}_{\tau_i}^a\}_{i=0}^{n_u^a}$ for all agents $a \in \mathcal{A}$. One approach to designing incentives is by solving:

$$\min\left\{ \mathbb{E}\left[\Phi(x_{0}^{a},\theta_{0}^{a},\{\pi_{i}^{a}\}_{i=0}^{T+n} \text{ for } a \in \mathcal{A}) | \{\tilde{x}_{t_{i}}^{a}\}_{i=0}^{T_{x}^{a}}, \{\tilde{y}_{\tau_{i}}^{a}\}_{i=0}^{T_{u}^{a}}, \{\pi_{i}^{a}\}_{i=0}^{T} \text{ for } a \in \mathcal{A}\right] \right| \\ \left\{\{\pi_{i}^{a}\}_{i=T+1}^{T+n} \text{ for } a \in \mathcal{A}\} \in \Omega\right\}$$
(27)

Here, $\Phi: \mathcal{X}^{\#\mathcal{A}} \times \theta^{\#\mathcal{A}} \times \Omega \to \mathbb{R}$ is a joint loss function that depends on the behavior all agents. For the settings we are interested in, this loss function has a separable structure.

ASSUMPTION 9. Loss Φ is additively $\Phi(x_0^a, \theta_0^a, \{\pi_i^a\}_{i=0}^{T+n} \text{ for } a \in \mathcal{A}) = \sum_{a \in \mathcal{A}} \varphi^a(x_0^a, \theta_0^a, \{\pi_i^a\}_{i=0}^{T+n})$ or multiplicatively separable $\Phi(x_0^a, \theta_0^a, \{\pi_i^a\}_{i=0}^{T+n})$ for $a \in \mathcal{A}) = \prod_{a \in \mathcal{A}} \varphi^a(x_0^a, \theta_0^a, \{\pi_i^a\}_{i=0}^{T+n})$.

Without loss of generality, we assume Φ is additively separable since we can obtain similar results for the case of multiplicative separability by taking the logarithm of Φ . We also make an assumption that states Ω is decomposable in a simple way. ASSUMPTION 10. There exist a finite set $V = \{v_1, v_2, \ldots\}$ with vector-valued, sets $S_v \subseteq \Pi^n$ for $v \in V$, and a vector-valued constant β such that

$$\Omega = \left\{ \{\pi_i^a\}_{i=T+1}^{T+n} \text{ for } a \in \mathcal{A} : y_v^a \in \{0,1\}, \sum_{v \in V} y_v^a = 1 \text{ for } a \in \mathcal{A}, \sum_{a \in \mathcal{A}} \sum_{v \in V} v \cdot y_v^a \leq \beta \\ \{\pi_i^a\}_{i=T+1}^{T+n} \in S_v \text{ if } y_v^a = 1 \right\}.$$
(28)

Moreover, the sets S_v are compact sets that are representable by a finite number of mixed integer linear constraints.

The underlying idea of this assumption is that the set V describes a vector of discrete elements that can be used as incentives, and β is the vector-valued budget on the discrete incentives. When the discrete incentives are fixed at v, the set S_v keeps the discrete incentives fixed and describes the feasible set of continuous incentives.

Even with these assumptions on separability and decomposibility, solving (27) is difficult because it is a BMIP with $\#\mathcal{A}$ MILP's in the lower level. Thus, we develop an adaptive algorithm (based on the 2SSA algorithm) for optimizing incentives for multiple agents. We first summarize our algorithm, and demonstrate that it only requires solving a small number of computable MILP's. Next we prove this algorithm provides a set of incentives that are asymptotically optimal as time advances.

4.2.1. Adaptive Algorithm for Multiple Agents We design incentives for multiple agents with the Adaptive Behavioral Multi-Agent Algorithm (ABMA) presented in Algorithm 2. The main idea behind this method is to use the assumptions on Φ and Ω to decompose the initial problem into $\#\mathcal{A}$ sub-problems that solve a single agent problem, and a single master problem that combines these solutions into a global optimum across all agents. Because of the assumptions on Ω , each sub-problem can be further decomposed into #V sub-problems. For each sub-problem, we use the 2SSA algorithm to solve the $\#\mathcal{A} \cdot \#V$ sub-problems; however, we do not explicitly call the 2SSA algorithm because it is more efficient to solve the MAP estimator once and then solve the incentive design problem for each single agent. Our first result concerns the computability of this algorithm.

PROPOSITION 4. If Assumptions 1–10 hold, then the main computational steps of the ABMA algorithm involve solving a total of $\#A \cdot (\#V + 1)$ MILP's and 1 ILP.

Proof: Step 2 of ABMA is a MAP estimate, which can be computed by solving a single MILP by Corollary 2. A similar argument used to prove Corollary 4 shows that Steps 4 and 5 can be computed by solving a single MILP. Step 8 can be seen to be an ILP by construction. The remaining steps of ABMA are assignment steps and do not require solving any optimization problems. \Box

Algorithm 2 Adaptive Behavioral Multi-Agent Algorithm (ABMA)

Require: $\{\tilde{x}^a_{t_i}\}^{n^a_x}_{i=0}, \{\tilde{u}^a_{\tau_i}\}^{n^a_u}_{i=0}, \{\pi^a_i\}^T_{i=0}$ for $a \in \mathcal{A}$ 1: for all $a \in \mathcal{A}$ do compute $(\hat{x}_{0,T}^a, \hat{\theta}_{0,T}^a) = \arg \max_{(x_0,\theta_0)} \psi_T(x_0, \theta_0)$ 2: for all $v \in V$ do 3: set $\pi_v^a \in \arg\min\{\varphi^a(\hat{x}_{0,T}, \hat{\theta}_{0,T}, \{\pi_i\}_{i=0}^{T+n}) \mid \{\pi_i\}_{i=T+1}^{T+n} \in S_v\}$ 4: set $\phi_v^a = \varphi^a(\hat{x}_0^a, \hat{\theta}_0^a, \pi_v^a)$ 5:end for 6: 7: end for 8: compute: $y \in \arg\min \sum_{a \in A} \sum_{v \in V} \phi_v^a \cdot y_v^a$ s.t. $\sum_{a \in \mathcal{A}} \sum_{v \in V} v \cdot y_v^a \leq \beta$ $\sum_{v \in V} y_v^a = 1$ for $a \in \mathcal{A}$ $y_v^a \in \{0, 1\}$ 9: for all $a \in \mathcal{A}$ and $v \in V$ do set $\pi^a_{ABMA}(T) = \pi^a_v$ if $y^a_v = 1$ 10: 11: end for 12: return $\pi^a_{ABMA}(T)$ for $a \in \mathcal{A}$

This means the ABMA algorithm performs incentive design for the multi-agent case by solving $#\mathcal{A} \cdot (\#V+1) + 1$ MILP's, which is significantly less challenging than solving a BMIP with $#\mathcal{A}$ MILP's in the lower level as would be required to solve (27).

The ABMA algorithm also has an alternative interpretation, and to better understand this consider the following feasibility problem:

$$\begin{split} \Phi(\overline{x}_{0}^{a},\overline{\theta}_{0}^{a},\{\overline{\pi}_{i}^{a}\}_{i=0}^{T+n} \text{ for } a \in \mathcal{A}) = \\ & \underset{x_{t}^{a},u_{t}^{a},\theta_{t}^{a},\pi_{t}^{a}}{\min} \Phi(x_{0}^{a},\theta_{0}^{a},\{\pi_{i}^{a}\}_{i=0}^{T+n} \text{ for } a \in \mathcal{A}) \\ & \text{ s.t. } u_{t}^{a} \in \arg\max\left\{f(x_{t+1}^{a},u,\theta_{t}^{a},\pi_{t}^{a}) \mid x_{t+1}^{a} = h(x_{t}^{a},u), \ u \in \mathcal{U}\right\} \quad \forall a,t, \\ & x_{t+1}^{a} = h(x_{t}^{a},u_{t}^{a}) \qquad \qquad \forall a,t, \\ & \theta_{t+1}^{a} = q(x_{t}^{a},u_{t}^{a},\theta_{t}^{a},\pi_{t}^{a}) \qquad \qquad \forall a,t, \end{split}$$

$$\sum_{t=1}^{a} \sum_{t=1}^{a} \sum_{$$

$$x_t^a \in \mathcal{X}^a, \theta_t^a \in \Theta, \pi_t^a \in \Pi \qquad \forall a, t,$$

$$x_0^a = \overline{x}_0^a, \theta_0^a = \overline{\theta}_0^a, \{\pi_t^a\}_{t=0}^{T+n} = \{\overline{\pi}_t^a\}_{t=0}^{T+n} \qquad \forall a, t \in \mathbb{R}$$

$$\{\{\pi_t^a\}_{t=T+1}^{T+n} \text{ for } a \in \mathcal{A}\} \in \Omega.$$

Our first result concerns regularity properties of the above written feasibility problem.

PROPOSITION 5. If Assumptions 1–10 hold, then $\Phi(x_0^a, \theta_0^a, \{\pi_i^a\}_{i=0}^{T+n} \text{ for } a \in \mathcal{A})$ is lower semicontinuous in its arguments, and $\min\{\Phi(x_0^a, \theta_0^a, \{\pi_i^a\}_{i=0}^{T+n} \text{ for } a \in \mathcal{A}) \mid \{\{\pi_t^a\}_{t=T+1}^{T+n} \text{ for } a \in \mathcal{A}\} \in \Omega\}$ can be formulated as a MILP for all fixed values of $(x_0^a, \theta_0^a, \{\pi_i^a\}_{i=0}^T) \in \mathcal{X} \times \Theta \times \Pi^{T+1}$ for $a \in \mathcal{A}$.

Proof: Using the assumptions on separability of the joint loss function Φ (Assumption 9) and decomposibility on the incentive set Ω (Assumption 10), we have that (29) can be reformulated as

$$\min_{y_v^a} \sum_{a \in \mathcal{A}} \sum_{v \in V} \phi_v^a \cdot y_v^a$$
s.t.
$$\min\{\varphi^a(x_{0,T}, \theta_{0,T}, \{\pi_i\}_{i=0}^{T+n}) \mid \{\pi_i\}_{i=T+1}^{T+n} \in S_v\} \le \phi_v^a$$

$$x_0^a = \overline{x}_0^a, \theta_0^a = \overline{\theta}_0^a, \{\pi_t^a\}_{t=0}^{T+n} = \{\overline{\pi}_t^a\}_{t=0}^{T+n} \text{ for all } a, t$$

$$\sum_{a \in \mathcal{A}} \sum_{v \in V} v \cdot y_v^a \le \beta$$

$$\sum_{v \in V} y_v^a = 1 \text{ for } a \in \mathcal{A}$$

$$y_v^a \in \{0, 1\}$$
(30)

Since $\phi_v^a \cdot y_v^a$ is the product of a continuous and binary decision variable, standard integer programming reformulation techniques allow us to reformulate the above as

$$\min_{y_v^a} \sum_{a \in \mathcal{A}} \sum_{v \in V} z_v^a$$
s.t. $\varphi^a(x_{0,T}, \theta_{0,T}, \{\pi_i\}_{i=0}^{T+n}) \leq \phi_v^a$
 $\{\pi_i\}_{i=T+1}^{T+n} \in S_v$
 $x_0^a = \overline{x}_0^a, \theta_0^a = \overline{\theta}_0^a, \{\pi_t^a\}_{t=0}^{T+n} = \{\overline{\pi}_t^a\}_{t=0}^{T+n} \text{ for all } a, t$
 $\sum_{a \in \mathcal{A}} \sum_{v \in V} v \cdot y_v^a \leq \beta$
 $\sum_{v \in V} y_v^a = 1 \text{ for } a \in \mathcal{A}$ (31)
 $y_v^a \in \{0, 1\}$
 $z_v^a \geq \phi_v^a - M \cdot (1 - y_v^a)$
 $z_v^a \leq \phi_v^a + M \cdot (1 - y_v^a)$
 $z_v^a \leq M \cdot y_v^a$
 $z_v^a \geq -M \cdot y_v^a$

where M > 0 is a large-enough constant. Such a finite M exists because $\mathcal{X}, \mathcal{U}, \Pi, \Theta$ are compact, and because ℓ^a is representable by a finite number of mixed integer linear constraints. Since Corollary 4 (and its proof) implies we can represent $\varphi^a(x_{0,T}, \theta_{0,T}, \{\pi_i\}_{i=0}^{T+n}) \leq \phi_v^a$ and $\{\pi_i\}_{i=T+1}^{T+n} \in S_v$ by mixed integer linear constraints, this means we can reformulate (29) as a MILP with linear constraints that are affine in $(\overline{x}_0^a, \overline{\theta}_0^a, \{\overline{\pi}_i^a\}_{i=0}^{T+n}$ for $a \in \mathcal{A}$). And so standard results (Ralphs and Hassanzadeh 2014) imply its value function is lower semicontinuous with respect to these variables, which is our first result. The second result follows by noting $\min\{\Phi(x_0^a, \theta_0^a, \{\pi_i^a\}_{i=0}^{T+n} \text{ for } a \in \mathcal{A}) \mid \{\{\pi_t^a\}_{t=T+1}^{T+n} \text{ for } a \in \mathcal{A}\} \in \Omega\}$ is equivalent to (29) but with removal of the constraints $\pi_t^a = \overline{\pi}_t^a$ for $t = T+1, \ldots, T+n$. \Box The optimization problem (29) and the above result provide an alternative interpretation of the ABMA algorithm, which is formalized by the next corollary.

COROLLARY 5. If Assumptions 1–10 hold, then the solution of $\min\{\Phi(x_0^a, \theta_0^a, \{\pi_i^a\}_{i=0}^{T+n} \text{ for } a \in \mathcal{A}\} \mid \{\{\pi_i^a\}_{t=T+1}^{T+n} \text{ for } a \in \mathcal{A}\} \in \Omega\}$, is given by the ABMA algorithm but with Step 2 replaced with the step: set $(\hat{x}_{0,T}^a, \hat{\theta}_{0,T}^a) = (x_0^a, \theta_0^a)$.

Proof: This is straightforward from the reformulation shown in (30). \Box

Thus, though (29) is a large MILP, the assumptions we have made allow us to decompose the solution of this problem into a series of substantially smaller MILP's.

4.2.2. Asymptotic Optimality of ABMA The optimization problem in (29) is a useful construction because it can also be used to compute the optimal set of incentives. If each agent's true initial conditions $(x_0^{*,a}, \theta_0^{*,a})$ were known, then an optimal solution belongs to $\arg\min\{\Phi(x_0^{*,a}, \theta_0^{*,a}, \{\pi_i^a\}_{i=0}^{T+n} \text{ for } a \in \mathcal{A}) \mid \{\{\pi_i^a\}_{t=T+1}^{T+n} \text{ for } a \in \mathcal{A}\} \in \Omega\}$. More importantly, we have the following relationship to the solutions of the ABMA algorithm:

THEOREM 2. Note that $\arg\min\{\Phi(x_0^{*,a}, \theta_0^{*,a}, \{\pi_i^{*,a}\}_{i=0}^{T+n} \text{ for } a \in \mathcal{A}) \mid \{\{\pi_t^a\}_{t=T+1}^{T+n} \text{ for } a \in \mathcal{A}\} \in \Omega\}$ is the set of optimal solutions under the agents' true initial conditions $(x_0^{*,a}, \theta_0^{*,a})$. If Assumptions 1–8 hold, then we have that

$$\operatorname{dist}\left(\{\pi_{ABMA}^{a}(T) \text{ for } a \in \mathcal{A}\}, \\ \operatorname{arg\,min}\left\{\Phi(x_{0}^{*,a}, \theta_{0}^{*,a}, \{\pi_{i}^{a}\}_{i=0}^{T+n} \text{ for } a \in \mathcal{A}\} \mid \{\{\pi_{t}^{a}\}_{t=T+1}^{T+n} \text{ for } a \in \mathcal{A}\} \in \Omega\}\right) \xrightarrow{p} 0 \quad (32)$$

as $T \to \infty$, for any $\pi^a_{ABMA}(T)$ returned by ABMA. Recall that $\operatorname{dist}(x, B) = \inf_{y \in B} ||x - y||$.

Proof: Corollary 3 implies $(\hat{x}_{0,T}^a, \hat{\theta}_{0,T}^a) \xrightarrow{p} (x_0^{*,a}, \theta_0^{*,a})$, and Corollary 4 states Φ is lower semicontinuous in its arguments. This means $\Phi(\hat{x}_0^a, \hat{\theta}_0^a, \{\pi_i^a\}_{i=0}^{T+n}$ for $a \in \mathcal{A})$ is a lower semicontinuous approximation to $\Phi(x_0^{*,a}, \theta_0^{*,a}, \{\pi_i^a\}_{i=0}^{T+n}$ for $a \in \mathcal{A})$ by Proposition 2.1.ii of (Vogel and Lachout 2003b). But Corollary 5 shows that

$$\{\pi^a_{ABMA}(T) \text{ for } a \in \mathcal{A}\} \in \arg\min\{\Phi(\hat{x}^a_0, \hat{\theta}^a_0, \{\pi^a_i\}_{i=0}^{T+n} \text{ for } a \in \mathcal{A}) \mid \{\{\pi^a_t\}_{t=T+1}^{T+n} \text{ for } a \in \mathcal{A}\} \in \Omega\}.$$
(33)

This means that the result follows by applying Theorem 4.3 from (Vogel and Lachout 2003a). \Box

Thus any solution returned by ABMA is asymptotically included within the set of optimal incentives computed for the agents' true initial conditions. Restated, the above result says ABMA provides a set of incentives that are asymptotically optimal. This is a non-obvious result because in general pointwise-convergence of a sequence of stochastic optimization problems is not sufficient to ensure convergence of the minimizers of the sequence of optimization problems to the minimizer of the limiting optimization problem. Rockafellar and Wets (2009) provide an example that demonstrates this possible lack of convergence of minimizers.

5. Computational Experiments: Weight Loss Program Design

We have completed computational experiments applying the tools and techniques developed in this paper that form a specific implementation of a behavioral analytics framework. We compare several approaches, including ours, for designing incentives for multiple myopic agents to the problem described in Section 1.1.1 of designing personalized behavioral incentives for a clinically-supervised weight loss program. The first step of our behavioral analytics approach is to construct a behavioral model of individuals in weight loss programs. We describe the data source used for the simulations, and then summarize our behavioral model (Aswani et al. 2016) for individuals participating in such loss programs. To demonstrate the second and third steps of our framework, we simulate a setting in which behavioral incentives chosen using our ABMA algorithm are evaluated against behavioral incentives computed by (intuitively-designed) adaptive heuristics. Both our implementation of a behavioral analytics framework and the heuristic provide adaptation by recomputing the incentives at regular intervals as more data is collected from each individual. Our metric for comparison is the number of individuals who achieve clinically significant weight loss (i.e., a 5% reduction in body weight) at the end of the program. We also compare the percentage of weight loss for individuals who do not achieve clinically significant weight loss in order to better understand how clinical visits are allocated by the different methods. We conclude by performing a sensitivity analysis of design choices for the second and third steps of our behavioral analytics framework.

5.1. mDPP Program Trial Data Source

Our computational experiments used data from the mDPP trial (Fukuoka et al. 2015). This was a randomized control trial (RCT) that was conducted to evaluate the efficacy of a 5 month mobile phone based weight loss program among overweight and obese adults at risk for developing type 2 diabetes. This program was adapted from the Diabetes Prevention Program (DPP) (Diabetes Prevention Program Research Group 2002, 2009), but the number of in person clinical visits was reduced from 16 to 6 per person, and group exercise sessions were replaced with a home based exercise program to reduce costs. Sixty one overweight adults were randomized into an active control group which only received an accelerometer (n=31) or a treatment group which receive the mDPP mobile app plus the accelerometer and clinical office visits (n=30). Changes in primary and secondary outcomes for the trial were promising. The treatment group lost an average of 6.2 ± 5.9 kg ($-6.8\% \pm 5.7\%$) between baseline and the 5 month follow up while the control group gained 0.3 ± 3.0 kg ($0.3\% \pm 5.7\%$) (p < 0.001). The treatment group's steps per day increased by 2551 ± 4712 compared to the control group's decrease of 734 ± 3308 steps per day (p < 0.001). Additional details on demographics and other treatment parameters are available in (Fukuoka et al. 2015).

The data available from the mDPP trial includes step data (from accelerometer measurements), body weight data (which was measured at least twice a week every week and recorded in the mobile app by individuals in the treatment group, as well as measured three times in a clinical setting at baseline, 3 month, and 5 month), and demographic data (i.e., age, gender, and height of each individual). We note that this data matches the assumptions in Section 2.

5.2. Summary of Behavioral Model

We construct a behavioral model for each individual participating in the weight loss program. Using the terminology and notation of Section 2, the system state of each individual x_t is their body weight on day t, and their decisions $u_t = (f_t, s_t)$ on day t are how many calories they consume f_t and how many steps they walk s_t . The behavioral incentives $\pi_t = (g_t, d_t)$ provided to an individual on day t consist of (numeric) step goals g_t and an indicator d_t equal to one if a clinical visit was scheduled for that day. The motivational state (or type) is $\theta_t := (u_b, f_{b,t}, F_{b,t}, p_t, \mu, \delta, \beta)$. The state u_b captures the individual's baseline preference for number of steps taken each day, while $f_{b,t}, F_{b,t}$ capture the individual's short term and long term caloric intake preference, respectively. The variable p_t captures the disutility an individual experiences from not meeting a step goal. The last set of motivational states describe the individual's response to behavioral incentives. The states β, δ describe the amount of change in the individual's caloric consumption and physical activity preferences, respectively, after undergoing a single clinical visit. The state μ describes the self efficacy effect (Bandura 1998, Conner and Norman 1996) from meeting exercise goals.

The utility function of an individual on day t is given by $f(w_{t+1}, f_t, u_t; \theta_t, \pi_t) = -w_{t+1}^2 - r(u - u_b)^2 - r(f - f_{b,t})^2 + p_t(u_t - g_t)^-$. In our past modeling work (Aswani et al. 2016), we found that the predictions of this behavioral model were relatively insensitive to the value of r. And our numerical experiments in (Aswani et al. 2016) found that fixing the value of r to be the same for each individual provided accurate predictions. And so we assume that r is a fixed and known constant in our numerical experiments here. The temporal dynamics for an individual's system and motivational states are the dynamics of an individual's type by:

$$w_{t+1} = a \cdot w_t + b \cdot u_t + f_t + k \tag{34}$$

$$F_{b,t+1} = (1-\alpha) \cdot F_{b,t} + \alpha \cdot f_{b,t} \tag{35}$$

$$f_{b,t+1} = \gamma \cdot (f_{b,t} - F_{b,t}) + F_{b,t} - \beta \cdot d_n \tag{36}$$

$$p_{t+1} = \gamma \cdot p_t + \delta \cdot d_t + \mu \cdot \mathbb{1}(u_t \ge g_t).$$
(37)

Equation (34) is a "calories in minus calories out" description of weight change, and a standard physiological formula (Mifflin et al. 1990) is used to compute the values of a, b, k based on the

demographics of the individual. Equation (35) models the long term caloric intake preference as an exponential moving average of the short term caloric intake preferences. We found that the predictions for different individuals were relatively insensitive to the value of α , and so in our numerical experiments we assume α is known and fixed to a value satisfying $\alpha < 1$. In (36), we model the dynamics of baseline food consumption as always tending towards their initial value unless perturbed by a clinical visit. In (37) we model the tendency for meeting the step goal as tending towards zero unless there is a clinical visit or the individual has met the previous exercise goal, which increases their self efficacy and makes the individual more likely to meet their step goal in the future. In both (36) and (37), $\gamma < 0$ is assumed to be a known decay factor since we found that predictions were relatively insensitive to the value of γ (Aswani et al. 2016). Note that these temporal dynamics and utility functions satisfy the assumptions in Section 2.

For the MLE and MAP calculations, we assumed that step and weight data were measured with zero-mean noise distributed according to a Laplace distribution with known variance. We found that predictions of the behavioral model estimated when assuming the noise had a Gaussian distribution were of the same quality as those estimated when assuming Laplace noise, and so we assume Laplace noise so that the MAP and MLE problems can be formulated as MILP's (as shown in Section 2). Furthermore, the prior distribution used for the MAP calculation for each individual was a histogram of the MLE estimates of all the other individual's parameters. Note that this form for a prior distribution can be expressed using integer linear constraints (Aswani et al. 2016). The complete MILP formulations for MAP and MLE are provided in the appendix.

5.3. Weight Loss Program Design

Since the majority of implementation costs for weight loss programs are due to clinical visits, the clinician's design problem is to maximize the expected number of individuals who reduce their weight by a clinically significant amount (i.e., 5% reduction in body weight). The clinician is able to personalize the step goals for each individual, and can change the number and timing of clinical visits for each individual. However, there is a budget constraint on the total number of visits that can be scheduled across all individuals. This constraint captures the costliness of clinical visits.

We optimize the weight loss program using our ABMA algorithm to implement the second and third steps of our behavioral analytics framework. This requires choosing a loss function for each individual, and Figure 1 shows three choices that we considered. These three losses make varying tradeoffs between achieving the primary health outcome of number of individuals with clinically significant weight loss (i.e., 5% weight loss) at the end of the program versus the secondary health outcome of maximizing weight loss of individuals who were not able to achieve 5% weight loss. The first choice of a loss function is the *step loss*, which is given by

$$\varphi = \begin{cases} -1, & \text{if } x_{T+n} \le 0.95 \cdot \tilde{x}_0 \\ 0, & \text{otherwise} \end{cases}$$
(38)

This discontinuous choice of a loss function gives minimal loss to 5% or more reduction in body weight and maximal loss to less than 5% reduction in body weight. The second choice of a loss function is the *hinge loss*, which is given by

$$\varphi = \begin{cases} -1, & \text{if } x_{T+n} < 0.95 \cdot \tilde{x}_0 \\ -0.2 \cdot (x_{T+n}/\tilde{x}_0 - 1), & \text{if } 0.95 \cdot \tilde{x}_0 \le x_{T+n} \le \tilde{x}_0 \\ 0, & \text{if } x_{T+n} > \tilde{x}_0 \end{cases}$$
(39)

This continuous choice of a loss function gives minimal loss to 5% or more reduction in body weight, maximal loss to less than 0% reduction in body weight, and an intermediate loss for intermediate reductions in body weight. The third choice of a loss function is the *time-varying hinge loss*, which is given by

$$\varphi = \begin{cases} -1, & \text{if } \frac{x_{T+n}}{\tilde{x}_0} < 0.95 - \frac{0.05}{\log T} \\ 10(\log T)(\frac{x_{T+n}}{\tilde{x}_0} - (0.95 + \frac{0.05}{\log T})), & \text{if } 0.95 - \frac{0.05}{\log T} \le \frac{x_{T+n}}{\tilde{x}_0} \le 0.95 + \frac{0.05}{\log T} \\ 0, & \text{if } \frac{x_{T+n}}{\tilde{x}_0} > 0.95 + \frac{0.05}{\log T} \end{cases}$$
(40)

Much like the hinge loss (39), it promotes intermediate amounts of weight loss that might not meet the 5% threshold of clinically significant weight loss. However, as more data is collected it approaches the step loss (38) to reflect a higher degree of confidence in the estimated parameters. Thus, this choice of loss function can be considered an intermediate between the hinge (39) and step (38) losses. There is one computational note. Since these losses are non-decreasing, we can modify Step 4 of the ABMA algorithm to instead minimize the body weight of each individual and then compose the body weight with the loss function.

For the purpose of comparing various program designs through simulations, we considered three additional designs for the weight loss program. We used an adaptive heuristic to design the weight loss program: Clinical visits were scheduled towards the end of the treatment at least one week apart, with more visits given to individuals who were closer to meeting the weight loss goal of a 5% weight reduction based on their latest observed weight, and step goals were set to be a 10% increase over a linear moving average of the individual's observed step count over the prior week. The second design was a "do nothing" plan where individuals were given exactly one clinical visit after two weeks, and their step goals were a constant 10,000 steps each day. The third design was the original design of the mDPP trial: Clinical visits were scheduled on predetermined days during the treatment after 2, 4, 6, 19, 14, 18, and 20 weeks of the trials. The first two weeks of this design



Figure 1 The left plot shows the step loss function (38), and the right plot shows the hinge loss function (39). The x-axis on both plots is $100 \cdot (x_{T+n}/\tilde{x}_0 - 1)$, which is percent reduction in body weight.

did not contain any clinical visits or exercise goals but instead served as an initialization period. After the first two weeks, exercise goals increased 20% each week, starting with a 20% increase over the average number of steps taken by individuals during the the two week initialization period. The exercise goals were capped at a maximum of 12,000 steps a day. Since the adaptive heuristic and ABMA are both adaptive, we recalculated both at the beginning of each month of the treatment and allowed both adaptive methods a 2 week initialization time similar to the mDPP trial.

5.4. Simulation Comparison

We compared the six different program designs using simulations of a weight loss program with a five month duration and with 30 individuals participating. Each individual in the simulation followed our behavioral model, and the parameters corresponding to the behavioral model for each individual were chosen to be those estimated by computing the MLE using the data from the mDPP trial. Since we also wanted to test how these different designs account for missing data, we assumed that the data available to each algorithm would be limited to days of the mDPP study where a particular individual reported their weight and steps. Since the adaptive heuristic and our behavioral analytics framework are both adaptive, we recalculated the program design at the beginning of each month of the program (by re-runnning the heuristic calculations and rerunning the ABMA algorithm) and allowed both adaptive methods a two-week initialization time similar to the design of the program in the mDPP trial. All simulations were run using MATLAB on a laptop computer with a 2.4GHz processor and 16GB RAM. The Gurobi solver (Gurobi Optimization 2015) in conjunction with the CVX toolbox for MATLAB (Grant and Boyd 2014) were used to perform the initial estimation of the individual parameters, compute designs for the weight loss program, and perform simulations of each design.

Figure 2 compares the primary outcome of interest to clinicians, which is the number of individuals that achieve clinically significant weight loss (i.e., 5% or more reduction in body weight). We



Figure 2 Comparison of different program design methods with respect to number of successful individuals (i.e., lost 5% or more body weight)

repeated the simulations for our behavioral analytics framework and the adaptive heuristic under different constraints on the total number of clinical visits that could be allocated to individuals. The x-axis of Figure 2 is the average number of clinical visits provided to individuals. The horizontal line at 18 is the number of individuals who achieved 5% weight loss in the actual mDPP trial, in which each individual received 7 clinical visits. Figure 2 shows that all forms of behavioral analytics program and adaptive heuristic program designs outperform the "do nothing" policy. Furthermore, our behavioral analytics approach and the adaptive heuristic achieve results comparable to the original mDPP program design but with significantly less resources (i.e., less clinical visits). The simulations predict that using our behavioral analytics approach in which ABMA has a step (38) or time-varying hinge loss (40) to design the weight loss program can provide health outcomes comparable to current clinical practice while using only 40-60% of the resources (i.e., clinical visits) of current practice. In contrast, the adaptive heuristic would require 80-95% of resources (i.e., clinical visits) to attain health outcomes comparable to current clinical practice. This suggests that appropriate choice of the loss function for our ABMA algorithm, as part of behavioral analytics approach, to personalize the design of a weight loss program could increase capacity (in terms of the number of individuals participating in the program for a fixed cost) by up to 60%, while achieving comparable health outcomes.

Figure 3 compares the different program designs using a secondary outcome of interest to clinicians of the average amount of weight loss of individuals who did not successfully achieve 5% weight



Figure 3 Comparison of different program design methods with respect to the percent weight lost by unsuccessful individuals (i.e., lost less than 5% body weight)

loss. The original treatment plan of mDPP and the "do nothing" treatment plan slightly outperform the adaptive program designs at certain clinical visit budgets. This effect however is mainly due to these static plans not identifying individuals who are on the cusp of achieving 5% weight loss but might still achieve around 3-4% weight loss, while both adaptive program designs allocate clinical visits to these individuals and ensure they reach the weight loss goal of 5% weight loss. Restated, the lower weight loss of unsuccessful individuals under the behavioral analytics treatment plans is an artifact of the improved success rate of the behavioral analytics plans in helping individuals achieve 5% weight loss. This effect is further exemplified in the region of between an average of 2.8-4.2 visits per individual, where we see that individuals who were not successful in achieving 5% weight loss in the behavioral analytics treatment plans on average lose more weight then those under the heuristic policy, while in the region of an average of 5.6-7 visits per individual we see that individuals under the heuristic treatment lose more weight. The effect in this last region is mainly due to the behavioral model used in our behavioral analytics framework, which is more effective at identifying the individuals who would most benefit from additional clinical visits. Therefore, more resources are spent on individuals who could potentially reach their 5% weight loss goal, while the adaptive heuristic uses these resources in a less effective manner.

Figures 2 and 3 demonstrate a tradeoff between the primary and secondary outcomes, and the various loss functions provide different tradeoffs. Note the line for the step loss (38) is the first to achieve a primary outcome comparable to that of the original mDPP trial while fluctuating

relatively little in terms of the secondary outcome. This matches intuition that the step loss function (38) is focused on ensuring individuals achieve 5% weight loss while not being concerned with their final weight. On the other hand, the line for the hinge loss (39) lags behind the other behavioral analytics policies in achieving comparable primary outcomes to the mDPP trial while having an extremely effective secondary outcome. These results follow our intuition that this loss favors intermediate weight loss over achieving clinically significant weight loss. Finally, the line corresponding to the time-varying hinge loss function (40) has a clear transition at an average of 5 visits per individual from favoring the primary outcome to the secondary outcome. This behavior indicates that using such time scaling leads to interventions that focus on primary outcomes when resources are constrained but also accounts for secondary outcomes when resources are less scarce. Such behavior may be useful for implementing a behavioral analytics approach when the relative abundance of resources is not known *a priori*.

5.5. Computational Performance and Sensitivity Analysis

The simulation experiments assumed that treatment plans were updated at the beginning of each month by re-running the second and third steps of our behavioral analytics approach (through applying the ABMA algorithm), and so we conducted a sensitivity analysis to examine the effect of updating the program design more or less frequently. Figures 4 and 5 compare the health outcomes of using a program designed by our behavioral analytics framework with a time-varying hinge loss (40), where the treatment was recalculated once every two weeks, once a month, and once every two months. These results show that recomputing the treatment plan with lesser or higher frequency does not significantly impact the efficacy of the resulting treatment. This indicates that for practical implementation, the statistical convergence rate of estimated parameters in our behavioral model is sufficiently fast that it would suffice to rerun the weight loss program design algorithm at most once a month.

We also conducted time-benchmarks for the sub-problems involved in computing 2SSA, MAP, and ABMA, which are the algorithms comprising the second and third steps of our implementation o a behavioral analytics framework. The results of the time-benchmarks are summarized in Tables 1,2, and 3 in the appendix. On average, solving all sub-problems took 17s per individual. This is promising for practical implementation, particularly because each sub-problem calculation can be performed in parallel for each size constraint of the clinical visit schedule (from 1 to 7 visits). The results show that computation time increases with respect to the number of available visits and data available in the treatment plan calculation. However, the calculation times still remain below 30s on average per individual for each step of the program calculation. This would imply that our methodology for weight loss program calculation is suitable for large scale program design since the program design would be updated at most once every month.



Figure 4 Comparison of calculation schedules and their effects on the number of successful individuals (i.e., lost 5% or more body weight)



Figure 5 Comparison of calculation schedules and their effects on the weight lost by unsuccessful individuals (i.e., lost less than 5% body weight)

6. Conclusion

In this paper, we develop a *behavioral analytics* framework for multi-agent systems in which a single coordinator provides behavioral or financial incentives to a large number of myopic agents. Our framework is applicable in a variety of settings of interest to the operations research commu-

nity, including the design of demand-response programs for electricity consumes, the personalized design of a weight loss program, and adaptive logistics allocation for franchises. The framework we develop involves the definition of a behavioral model, the estimation of model parameters, and the optimization of incentives. We show (among other results) that under mild assumptions, the incentives computed by our approach converge to the optimal incentives that would be computed knowing full information about the agents. We evaluated our approach for personalizing the design of a weight loss program, and showed via simulation that our approach can improve outcomes with reduced treatment cost.

Appendix A: Complete MILP Formulation of MLE Problem

$$\min \frac{\sqrt{2}}{\sigma_1} \sum_{i=1}^{n_w} \zeta_{w,t_i} + \frac{\sqrt{2}}{\sigma_2} \sum_{i=1}^{n_u} \zeta_{u,t_i}$$

$$\text{s.t.} \quad -\zeta_{w,t_i} < \tilde{w}_{t_i} - w_{t_i} < \zeta_{w,t_i}$$

$$\forall 1 < t_i < n_w$$

$$(41)$$

.t.
$$-\zeta_{w,t_i} \le w_{t_i} - w_{t_i} \le \zeta_{w,t_i}$$
 $\forall 1 \le t_i \le n_w$ (42)

$$-\zeta_{u,t_i} \le \tilde{u}_{\tau_i} - u_{\tau_i} \le \zeta_{u,t_i} \qquad \qquad \forall 1 \le t_i \le n_u \tag{43}$$

$$-b(aw_1 + bu_0 + f_0 + k) - r(u_0 - u_b) = 0$$
(44)

$$-(aw_{1} + bu_{0} + f_{0} + k) - r(f_{0} - F_{b,1}) = 0$$

$$-b(aw_{4} + bu_{4} + f_{4} + k) - r(u_{4} - u_{4}) = 0$$

$$\forall 1 \le t \le m$$
(45)
$$(45)$$

$$-(aw_t + bu_t + f_t + k) - r(f_t - f_{b,t}) = 0 \qquad \forall 1 \le t \le m \qquad (10)$$
$$\forall 1 \le t \le m \qquad (47)$$

$$-2b(aw_t + bu_t + f_t + k) - 2r(u_t - u_b) + \lambda_{1,t} = 0 \qquad \forall m \le t \le n$$
(48)

$$-2(aw_{t} + bu_{t} + f_{t} + k) - 2(f_{t} - f_{b,t}) = 0 \qquad \forall m \le t \le n \qquad (49)$$

$$(g_{t} - \epsilon) - Mx_{1,t} \le u_{t} \le g_{t} - \epsilon + M(1 - x_{1,t}) \qquad \forall m \le t \le n \qquad (50)$$

$$(g_t - \epsilon) - M(1 - x_{2,t}) \le u_t \le g_t + \epsilon + M(1 - x_{2,t}) \qquad \forall m \le t \le n$$

$$(g_t + \epsilon) - M(1 - x_{3,t}) \le u_t \le g_t + \epsilon + Mx_{3,t} \qquad \forall m \le t \le n$$

$$(52)$$

$$p_t - M(1 - x_{t,1}) \le \lambda_{1,t} \le M(1 - x_{3,t}) \qquad \forall m \le t \le n$$
(53)

$$0 \le \lambda_{1,t} \le p_t \qquad \qquad \forall m \le t \le n \tag{54}$$

$$x_{t,1} + x_{t,2} + x_{t,3} = 1 \qquad \qquad \forall m \le t \le n \tag{55}$$

$$p_{t+1} \ge \gamma p_t + \delta d_{t+1} \qquad \qquad \forall m \le t \le n \tag{56}$$

$$p_{t+1} \le \gamma p_t + \delta d_{t+1} + M(1 - x_{1,t}) \qquad \qquad \forall m \le t \le n \tag{57}$$

$$p_{t+1} \ge \gamma p_t + \delta d_{t+1} + \mu - M x_{1,t} \qquad \forall m \le t \le n$$
(58)

 $p_{t+1} \le \gamma p_t + \delta d_{t+1} + \mu \qquad \qquad \forall m \le t \le n \tag{59}$

$$F_{b,t+1} = (1-\alpha)F_{b,t} + \alpha f_{b,t} \qquad \forall t \tag{60}$$

$$f_{b,t+1} = \gamma(f_{b,t} - F_{b,t}) + F_{b,t} - \beta d_t \qquad \qquad \forall t \qquad (61)$$

x_{t+}	$_{1,1} \ge x_{t,1} - d_{t+1} - \mathbb{1}(g_{t+1} - g_t < 0)$	$\forall m \leq t \leq n$	(62)
x_{t+}	$_{1,2} \le x_{t,2} - d_{t+1} + \mathbb{1}(g_{t+1} - g_t < 0)$	$\forall m \leq t \leq n$	(63)
x_{t+}	$x_{t,3} \le x_{t,3} - d_{t+1} + \mathbb{1}(g_{t+1} - g_t < 0)$	$\forall m \leq t \leq n$	(64)
w_t ,	$u_t, f_t, F_{b,t}, f_{b,t}, p_t \ge 0$	$\forall t$	(65)
$x_{t,1}$	$, x_{t,2}, x_{t,3} \in \mathbb{B}$	$\forall t$	(66)

$$w_{t,1}, w_{t,2}, w_{t,3} \subset \mathbb{I}$$

Appendix B: Complete MILP Formulation of MAP Problem

$$\min \frac{\sqrt{2}}{\sigma_1} \sum_{i=1}^{n_w} \zeta_{w,t_i} + \frac{\sqrt{2}}{\sigma_2} \sum_{i=1}^{n_u} \zeta_{u,t_i} - \sum_{x \in \theta} \sum_{i=1}^{m_x} z_{i,x} \log \phi_i^x \tag{67}$$

$$s.t. - \zeta_{u,t} \leq \tilde{w}_t - w_t \leq \zeta_{u,t_i} \qquad \forall 1 \leq t_i \leq n. \tag{68}$$

s.t.
$$-\zeta_{w,t_i} \le \tilde{w}_{t_i} - w_{t_i} \le \zeta_{w,t_i}$$
 $\forall 1 \le t_i \le n_w$ (68)

$$-\zeta_{u,t_i} \le \tilde{u}_{\tau_i} - u_{\tau_i} \le \zeta_{u,t_i} \qquad \qquad \forall 1 \le t_i \le n_u \tag{69}$$

$$-b(aw_{1}+bu_{0}+f_{0}+k)-r(u_{0}-u_{b}) = 0$$

$$-(aw_{1}+bu_{0}+f_{0}+k)-r(f_{0}-F_{b,1}) = 0$$
(70)
(71)

$$-b(aw_t + bu_t + f_t + k) - r(u_t - u_b) = 0 \qquad \forall 1 \le t \le m$$
(72)

$$-(aw_t + bu_t + f_t + k) - r(f_t - f_{b,t}) = 0 \qquad \forall 1 \le t \le m \qquad (73)$$
$$-2b(aw_t + bu_t + f_t + k) - 2r(u_t - u_b) + \lambda_{1,t} = 0 \qquad \forall m \le t \le n \qquad (74)$$

$$-2(aw_t + bu_t + f_t + k) - 2(f_t - f_{b,t}) = 0 \qquad \forall m \le t \le n$$
(75)

$$(g_t - \epsilon) - Mx_{1,t} \le u_t \le g_t - \epsilon + M(1 - x_{1,t}) \qquad \forall m \le t \le n$$
(76)

$$(g_t - \epsilon) - M(1 - x_{2,t}) \le u_t \le g_t + \epsilon + M(1 - x_{2,t}) \qquad \forall m \le t \le n$$

$$(77)$$

$$(g_t - \epsilon) - M(1 - x_{2,t}) \le u_t \le g_t + \epsilon + M\pi \qquad \forall m \le t \le n$$

$$(77)$$

$$(g_t + \epsilon) - M(1 - x_{3,t}) \le u_t \le g_t + \epsilon + M x_{3,t} \qquad \forall m \le t \le n$$

$$p_t - M(1 - x_{t,1}) \le \lambda_{1,t} \le M(1 - x_{3,t}) \qquad \forall m \le t \le n$$

$$(78)$$

$$0 \le \lambda_{1,t} \le p_t \qquad \qquad \forall m \le t \le n \qquad (80)$$

$$x_{t,1} + x_{t,2} + x_{t,3} = 1 \qquad \forall m \le t \le n \tag{81}$$

$$p_{t+1} \ge \gamma p_t + \delta d_{t+1} \qquad \qquad \forall m \le t \le n \tag{82}$$

$$p_{t+1} \le \gamma p_t + \delta d_{t+1} + M(1 - x_{1,t}) \qquad \forall m \le t \le n$$

$$m \ge \gamma m + \delta d_{t+1} + M(1 - x_{1,t}) \qquad \forall m \le t \le n$$
(83)

$$p_{t+1} \ge \gamma p_t + \delta d_{t+1} + \mu - M x_{1,t} \qquad \forall m \le t \le n$$
(84)

$$p_{t+1} \le \gamma p_t + \delta d_{t+1} + \mu \qquad \qquad \forall m \le t \le n \tag{85}$$

$$F_{b,t+1} = (1-\alpha)F_{b,t} + \alpha f_{b,t} \qquad \forall t \qquad (86)$$

$$f_{b,t+1} = \gamma(f_{b,t} - F_{b,t}) + F_{b,t} - \beta d_t \qquad \qquad \forall t \qquad (87)$$

 $x_{t+1,1} \ge x_{t,1} - d_{t+1} - \mathbb{1}(g_{t+1} - g_t < 0)$ $\forall m \leq t \leq n$ (88)d + 1(a)~ 0) \sim x_t

(70)

$x_{t+1,3} \le x_{t,3} - d_{t+1} + \mathbb{1}(g_{t+1} - g_t < 0)$	$\forall m \leq t \leq n$	(90)
$z_{i,x}h_{lb,i}^x \le x_i \le z_{i,x}h_{ub,i}^x$	$\forall x \forall i$	(91)
$\sum_{i=1}^{m_x} z_{i,x} = 1$	$\forall x \forall i$	(92)
$\sum_{i=1}^{m_x} x_i = x$	$\forall x \forall i$	(93)
$z_{i,x} \in \mathbb{B}$	$\forall x \forall i$	(94)
$w_t, u_t, f_t, F_{b,t}, f_{b,t}, p_t \geq 0$	$\forall t$	(95)
		(06)

$$x_{t,1}, x_{t,2}, x_{t,3} \in \mathbb{B} \tag{96}$$

Appendix C: Complete MILP Formulation of Personalized Treatment Plan Design Problem

min w_n		(97)
s.t. $F_{b,t+1} = (1 - \alpha)F_{b,t} + \alpha f_{b,t}$	$\forall t > T$	(98)
$f_{b,t+1} = \gamma(f_{b,t} - F_{b,t}) + F_{b,t} - \hat{\beta}d_t$	$\forall t > T$	(99)
$p_{t+1} = \gamma p_t + \hat{\delta} d_t + \hat{\mu} (1 - x_{1,t})$	$\forall t > T$	(100)
$d_t \leq \mathbb{1}\left[\mod(t,7) = 1 \right]$	$\forall t > T$	(101)
$d_t \leq 1-d_\tau \forall \tau > T, \tau+1 \leq t \leq \tau+6$	$\forall t > T$	(102)
$-2b(aw_t + bu_t + f_t + k) - 2r(u_t - u_b) + \lambda_{1,t} + \lambda_{4,t} = 0$	$\forall t > T$	(103)
$-2(aw_t + bu_t + f_t + k) - 2(f_t - f_{b,t}) + \lambda_{3,t} = 0$	$\forall t > T$	(104)
$(g_t - \epsilon) - Mx_{1,t} \le u_t \le g_t - \epsilon + M(1 - x_{1,t})$	$\forall t > T$	(105)
$(g_t-\epsilon)-M(1-x_{2,t})\leq u_t\leq g_t+\epsilon+M(1-x_{2,t})$	$\forall t > T$	(106)
$(g_t + \epsilon) - M(1 - x_{3,t}) \le u_t \le g_t + \epsilon + Mx_{3,t}$	$\forall t > T$	(107)
$p_t - M(1 - x_{t,1}) \le \lambda_{1,t} \le M(1 - x_{3,t})$	$\forall t > T$	(108)
$0 \le f_t \le M(1-x_{f,t}); 0 \le u_t \le M(1-x_{u,t})$	$\forall t > T$	(109)
$0 \le \lambda_{3,t} \le M x_{f,t}$	$\forall t > T$	(110)
$0 \le \lambda_{4,t} \le M x_{u,t}$	$\forall t > T$	(111)
$0 \le \lambda_{1,t} \le p_t$	$\forall t > T$	(112)
$x_{t,1} + x_{t,2} + x_{t,3} = 1$	$\forall t > T$	(113)
$g_{t+1} - g_t \le M(1 - g_{ind,t})$	$\forall t > T$	(114)
$g_{t+1} - g_t \geq -Mg_{ind,t}$	$\forall t > T$	(115)
$x_{t+1,1} \ge x_{t,1} - d_{t+1} - g_{ind,t}$	$\forall t > T$	(116)

 $x_{t+1,2} \le x_{t,2} - d_{t+1} + g_{ind,t} \qquad \forall t > T \qquad (117)$

$$x_{t+1,3} \le x_{t,3} - d_{t+1} + g_{ind,t} \qquad \forall t > T \qquad (118)$$

$$g_{ind,t} \in \mathbb{B} \tag{119}$$

$$x_{t,1}, x_{t,2}, x_{t,3}, x_{f,t}, x_{u,t} \in \mathbb{B}$$
(120)

$$d_t = \bar{d}_t; g_t = \bar{g}_t; w_t = \hat{w}_t; u_t = \hat{u}_t; f_t = \hat{f}_t; \theta_t = \hat{\theta}_t$$
 (121)

Appendix D: Benchmarking Performance Tables

Average Runtimes for Candidate Treatment Plan Calculation (in seconds)							
	Date of Calculation During the Program						
		15	30	60	90	120	Average
	54	21.715	9.3363	8.8984	9.5618	10.003	11.903
	63	21.715	7.7624	14.714	18.736	16.692	15.924
	72	21.715	8.8894	21.239	26.186	16.091	18.824
	81	21.715	11.408	43.147	28.112	23.284	25.533
	90	21.715	10.274	23.935	13.421	33.6	20.589
	99	21.715	9.3366	26.251	14.16	19.855	18.263
	108	21.715	9.4194	24.208	12.774	21.757	17.975
	117	21.715	9.9031	22.962	9.9226	31.865	19.273
Visit Dudant	126	21.715	10.217	21.749	10.199	28.664	18.509
Visit Budget	135	21.715	9.973	16.307	23.129	18.893	18.003
	144	21.715	11.208	14.626	14.851	8.9203	14.264
	153	21.715	12.978	13.913	13.892	8.0501	14.11
	162	21.715	12.403	14.674	14.434	10.965	14.838
	171	21.715	13.305	12.102	11.585	23.78	16.497
	180	21.715	14.879	12.116	11.964	18.473	15.829
	189	21.715	11.731	11.835	12.221	21.715	15.843
	Average	21.715	10.814	18.917	15.322	19.538	17.261

Table 1

Average Runtimes for MAP Calculation (in seconds)							
	Date of Calculation During the Program						
		15	30	60	90	120	Average
	54	16.365	11.416	9.537	11.479	7.6851	11.296
	63	16.365	13.249	12.816	17.157	11.493	14.216
	72	16.365	11.913	14.15	13.846	12.884	13.832
	81	16.365	17.448	16.11	11.936	17.237	15.819
	90	16.365	12.251	9.9218	8.9877	16.463	12.798
	99	16.365	9.9473	10.59	9.3482	13.666	11.983
	108	16.365	10.412	10.199	9.377	14.746	12.22
	117	16.365	10.696	9.5879	8.9027	23.187	13.748
Visit Dudant	126	16.365	10.36	10.737	9.6123	23.524	14.12
Visit Budget	135	16.365	9.7085	10.241	12.753	18.221	13.458
	144	16.365	9.3022	13.258	12.135	11.654	12.543
	153	16.365	8.5183	11.297	11.527	11.152	11.772
	162	16.365	9.4638	8.4174	11.279	14.812	12.068
	171	16.365	8.0878	11.092	9.8865	18.003	12.687
	180	16.365	8.8929	7.5454	6.5939	12.739	10.427
	189	16.365	7.679	7.6401	6.9321	16.365	10.996
	Average	16.365	10.584	10.821	10.735	15.239	12.749

Table 2

Average Puptimes for Knappent Calculation (in seconds)								
Average Runtimes for Rhapsack Calculation (in seconds)								
	Date of Calculation During the Program							
		15	30	60	90	120	Average	
	54	0.21791	0.19278	0.14949	0.1666	0.19679	0.18471	
	63	0.19295	0.16084	0.19613	0.23577	0.22246	0.20163	
	72	0.21302	0.2516	0.17474	0.18052	0.3125	0.22648	
	81	0.25345	0.3542	0.18128	0.17449	0.25997	0.24468	
	90	0.21698	0.17154	0.168	0.16942	0.21394	0.18798	
	99	0.17072	0.17607	0.15058	0.15699	0.20012	0.1709	
	108	0.17086	0.17178	0.17235	0.15172	0.19029	0.1714	
	117	0.17958	0.17073	0.16465	0.16171	0.20232	0.1758	
Visit Dudant	126	0.18599	0.24335	0.17326	0.16726	0.20706	0.19538	
VISIT Budget	135	0.20403	0.16143	0.16626	0.16681	0.19607	0.17892	
	144	0.17733	0.14174	0.16112	0.1618	0.18708	0.16581	
	153	0.18841	0.15411	0.17753	0.17965	0.18822	0.17758	
	162	0.17892	0.2167	0.18539	0.17419	0.21756	0.19455	
	171	0.1771	0.29787	0.21357	0.2248	0.21789	0.22625	
	180	0.20535	0.18508	0.22285	0.16565	0.19392	0.19457	
	189	0.18554	0.2104	0.19875	0.17776	0.20053	0.1946	
	Average	0.19488	0.20376	0.1785	0.17595	0.21292	0.1932	



Acknowledgments

The authors gratefully acknowledge the support of NSF Award CMMI-1450963, UCSF Diabetes Family Fund for Innovative Patient Care-Education and Scientific Discovery Award, K23 Award (NR011454), and the UCSF Clinical and Translational Science Institute (CTSI) as part of the Clinical and Translational Science Award program funded by NIH UL1 TR000004.

References

- Adlakha, S., R. Johari. 2013. Mean field equilibrium in dynamic games with strategic complementarities. Operations Research 61(4) 971–989.
- Ahuja, R., J. Orlin. 2001. Inverse optimization. Operations Research 49(5) 771-783.
- Astrom, K., B. Wittenmark. 1995. Adaptive Control. Addison-Wesley.
- Aswani, A., H. Gonzalez, S. Sastry, C. Tomlin. 2013. Provably safe and robust learning-based model predictive control. Automatica 49(5) 1216–1226.
- Aswani, A., P. Kaminsky, Y. Mintz, E. Flowers, Y. Fukuoka. 2016. Behavioral modeling in weight loss interventions. Available at SSRN: https://ssrn.com/abstract=2838443.
- Aswani, A., Z.-J. Shen, A. Siddiq. 2015. Inverse optimization with noisy data. arXiv:1507.03266 Submitted.
- Aswani, Anil, Aurélien Ouattara. 2016. Duality approach to bilevel programs with a convex lower level. arXiv preprint arXiv:1608.03260.
- Aswani, Anil, Claire Tomlin. 2009. Monotone piecewise affine systems. IEEE Transactions on Automatic Control 54(8) 1913–1918.
- Ayer, T., O. Alagoz, N. Stout. 2012. A POMDP approach to personalize mammography screening decisions. Operations Research 60(5) 1019–1034.
- Ayer, T., O. Alagoz, N. Stout, E. Burnside. 2015. Heterogeneity in women's adherence and its role on optimal breast cancer screening policies. *Management Science*.
- Azar, K., L. Lesser, B. Laing, J. Stephens, M. Aurora, L. Burke, L. Palaniappan. 2013. Mobile applications for weight management: theory-based content analysis. AJPM 45(5) 583–589.

Ban, G., C. Rudin. 2016. The big data newsvendor: Practical insights from machine learning.

- Bandura, A. 1998. Health promotion from the perspective of social cognitive theory. *Psychology and Health* **13**(4) 623–649.
- Bastani, Hamsa, Mohsen Bayati. 2015. Online decision-making with high-dimensional covariates. Available at SSRN 2661896.
- Bertsimas, D., V. Goyal. 2012. On the power and limitations of affine policies in two-stage adaptive optimization. *Mathematical programming* 134(2) 491–531.
- Bertsimas, D., V. Gupta, I. Paschalidis. 2014. Data-driven estimation in equilibrium using inverse optimization. Mathematical Programming 1–39.
- Bertsimas, D., A. O'Hair. 2013. Personalized diabetes management: A robust optimization approach. Submitted.
- Bickel, P., K. Doksum. 2006. Mathematical Statistics: Basic Ideas And Selected Topics, vol. 1. 2nd ed. Pearson Prentice Hall.
- Bitmead, R. 1984. Persistence of excitation conditions and the convergence of adaptive schemes. *IEEE Transactions on Information Theory* **30**(2) 183–191.
- Blanchet, J.H., G. Gallego, V. Goyal. 2013. A markov chain approximation to choice modeling. EC. Citeseer, 103–104.
- Boukhtouta, A., J. Berger, W.B. Powell, A. George. 2011. An adaptive-learning framework for semicooperative multi-agent coordination. 2011 IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning (ADPRL). IEEE, 324–331.
- Boyd, S., L. Vandenberghe. 2004. Convex Optimization. Cambridge University Press.
- Breiman, Leo, et al. 2001. Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical Science* **16**(3) 199–231.
- Brock, D., S. Wartman. 1990. When competent patients make irrational choices. NEJM 322(22) 1595–1599.
- Callier, F.M., C.A. Desoer. 1994. *Linear System Theory*. Springer Texts in Electrical Engineering, Springer New York.
- Cawley, John. 2004. An economic framework for understanding physical activity and eating behaviors. *AJPM* **27**(3, Supplement) 117–125.
- Conner, M., P. Norman. 1996. Predicting health behaviour: research and practice with social cognition models. Open University Press.
- Costa, L.M., G. Kariniotakis. 2007. A stochastic dynamic programming model for optimal use of local energy resources in a market environment. *Power Tech*, 2007 IEEE Lausanne. IEEE, 449–454.
- Craig, John J, Ping Hsu, S Shankar Sastry. 1987. Adaptive control of mechanical manipulators. *The International Journal of Robotics Research* **6**(2) 16–28.

- Darby, S. 2010. Smart metering: What potential for householder engagement? Building Research and Information **38**(5) 442–457.
- Dempe, S. 2002. Foundations of Bilevel Programming. Springer.
- DeNegre, S., T.K. Ralphs. 2009. A branch-and-cut algorithm for bilevel integer programming. Proceedings of the Eleventh INFORMS Computing Society Meeting. 65–78.
- Deng, R., Z. Yang, M. Y. Chow, J. Chen. 2015. A survey on demand response in smart grids: Mathematical models and approaches. *IEEE Transactions on Industrial Informatics* 11(3) 570–582.
- Deo, S., T. Jiang, S. Iravani, K. Smilowitz, S. Samuelson. 2013. Improving health outcomes through better capacity allocation in a community-based chronic care model. *Operations Research* 61(6) 1277–1294.
- Diabetes Prevention Program Research Group. 2002. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *NEJM* **346**(6) 393–403.
- Diabetes Prevention Program Research Group. 2003. Costs associated with the primary prevention of type 2 diabetes mellitus in the diabetes prevention program. *Diabetes Care* 26(1) 36–4F7.
- Diabetes Prevention Program Research Group. 2009. 10–year follow–up of diabetes incidence and weight loss in the diabetes prevention program outcomes study. *Lancet* **374**(9720) 1677–1686.
- Engineer, F., P. Keskinocak, L. Pickering. 2009. Or practice catch-up scheduling for childhood vaccination. Operations Research 57(6) 1307–1319.
- Frazier, P.I., J. Wang. 2016. Bayesian optimization for materials design. Information Science for Materials Discovery and Design. Springer, 45–75.
- Fudenberg, Drew, David Levine, Eric Maskin. 1994. The folk theorem with imperfect public information. *Econometrica* 62(5) 997–1039.
- Fukuoka, Y., C. Gay, K. Joiner, E. Vittinghoff. 2015. A novel diabetes prevention intervention using a mobile app: A randomized controlled trial with overweight adults at risk. AJPM 49(2) 223–237.
- Fukuoka, Y., J. Komatsu, L. Suarez, E. Vittinghoff, W. Haskell, T. Noorishad, K. Pham. 2011. The mPED randomized controlled clinical trial: applying mobile persuasive technologies to increase physical activity in sedentary women protocol. *BMC Public Health* 11(933).
- Gafni, Amiram. 1990. Correspondence to "competent patients and irrational choices". *NEJM* **323**(19) 1353–1355.
- Gal-Or, Esther. 1995. Maintaining quality standards in franchise chains. Management Science 41(11) 1774– 1792.
- George, A., W.B. Powell. 2007. An adaptive-learning framework for semi-cooperative multi-agent coordination .
- Grant, M., S. Boyd. 2014. CVX: Matlab software for disciplined convex programming, version 2.1. http://cvxr.com/cvx.

- Gupta, D., L. Wang. 2008. Revenue management for a primary-care clinic in the presence of patient choice. Operations Research 56 576.
- Gurobi Optimization, Inc. 2015. Gurobi optimizer reference manual. URL http://www.gurobi.com.
- Helm, Jonathan E., Mariel S. Lavieri, Mark P. Van Oyen, Joshua D. Stein, David C. Musch. 2015. Dynamic forecasting and control algorithms of glaucoma progression for clinician decision support. *Operations Research*.
- Heuberger, C. 2004. Inverse combinatorial optimization: A survey on problems, methods, and results. *Journal* of Combinatorial Optimization 8(3) 329–361.
- Iyer, K., R. Johari, C.C. Moallemi. 2014. Information aggregation and allocative efficiency in smooth markets. Management Science 60(10) 2509–2524.
- Iyer, K., R. Johari, M. Sundararajan. 2011. Mean field equilibria of dynamic auctions with learning. ACM SIGecom Exchanges 10(3) 10–14.
- James, J.T. Moore, J.F. Bard. 1990. The mixed integer linear bilevel programming problem. Oper. Res. 38(5) 911–921. doi:10.1287/opre.38.5.911.
- Jiang, L., S. Low. 2011. Real-time demand response with uncertain renewable energy in smart grid. Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on. IEEE, 1334–1341.
- Kaut, Michal, Stein W Wallace. 2003. Evaluation of scenario-generation methods for stochastic programming
- Keshavarz, A., Y. Wang, S. Boyd. 2011. Imputing a convex objective function. IEEE Multi-Conference on Systems and Control. 613–619.
- Kim, T.T., H.V. Poor. 2011. Scheduling power consumption with price uncertainty. IEEE Transactions on Smart Grid 2(3) 519–527.
- Kucukyazici, B., V. Verter, N. Mayo. 2011. An analytical framework for designing community-based care for chronic diseases. POMS 20 474.
- Laffont, J.-J., D. Martimort. 2002. The Theory of Incentives: The Principal-Agent Model. Princeton.
- Lal, R. 1990. Improving channel coordination through franchising. Marketing Science 9(4) 299–318.
- Lee, M., O. Aslam, B. Foster, D. Kathan, C. Young. 2014. Assessment of demand response and advanced metering. *Federal Energy Regulatory Commission, Staff Rep*.
- Leff, H., M. Dada, S. Graves. 1986. An lp planning model for a mental health community support system. Management Science 32 139.
- Li, N., L. Chen, S. H. Low. 2011. Optimal demand response based on utility maximization in power networks. 2011 IEEE power and energy society general meeting. IEEE, 1–8.

- Li, P., A. E. B. Lim, J. G. Shanthikumar. 2012. Decentralized control of a stochastic multi-agent queueing system. *IEEE Transactions on Automatic Control* 57(11) 2762–2777. doi:10.1109/TAC.2012.2192358.
- Li, Susan X. 1997. Transaction-efficiency analysis of franchising arrangements through chance crossconstrained game theory. Computers & operations research 24(10) 919–935.
- Liu, N., S. Ziya, V. Kulkarni. 2010. Dynamic scheduling of outpatient appointments under patient no-shows and cancellations. M&SOM 12 347.
- Lopez, M., A. Gonzalez-Barrera, E. Patten. 2013. Closing the digital divide: Latinos and technology adoption. Tech. rep., Pew Research Center.
- Lorca, Alvaro, Xu Andy Sun. 2015. Adaptive robust optimization with dynamic uncertainty sets for multiperiod economic dispatch under significant wind. *IEEE Transactions on Power Systems* **30**(4) 1702– 1713.
- Mason, J., B. Denton, S. Smith, N. Shah. 2013. Using electronic health records to monitor and improve adherence to medication. Working paper.
- Mifflin, M., S. St Jeor, L. Hill, B. Scott, S. Daugherty, Y. Koh. 1990. A new predictive equation for resting energy expenditure in healthy individuals. "The American Journal of Clinical Nutrition 51(2) 241–247.
- Mignone, Domenico, Giancarlo Ferrari-Trecate, Manfred Morari. 2000. Stability and stabilization of piecewise affine and hybrid systems: An lmi approach. Decision and Control, 2000. Proceedings of the 39th IEEE Conference on, vol. 1. IEEE, 504–509.
- Mohsenian-Rad, A., A. Leon-Garcia. 2010. Optimal residential load control with price prediction in real-time electricity pricing environments. *IEEE transactions on Smart Grid* 1(2) 120–133.
- Molderink, A., V. Bakker, M.G.C. Bosman, J.L. Hurink, G.J.JM Smith. 2010. Management and control of domestic smart grid technology. *IEEE transactions on Smart Grid* 1(2) 109–119.
- Moore, J.T., J.F. Bard. 1992. An algorithm for the discrete bilevel programming problem. *Naval Research Logistics* **39**(3) 419?435.
- Negoescu, D., K. Bimpikis, M. Brandeau, D. Iancu. 2014. Dynamic learning of patient response types: An application to treating chronic diseases.
- O'Neill, D., M. Levorato, A. Goldsmith, Urbashi U. Mitra. 2010. Residential demand response using reinforcement learning. Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on. IEEE, 409–414.
- O'Reilly, G., D. Spruijt-Metz. 2013. Current mHealth technologies for physical activity assessment and promotion. AJPM 45(4) 501–507.
- Osband, I., C. Blundell, A. Pritzel, B. Van Roy. 2016. Deep exploration via bootstrapped DQN. arXiv preprint arXiv:1602.04621.
- Osband, I., B. Van Roy. 2015. Bootstrapped Thompson sampling and deep exploration. $arXiv \ preprint$ arXiv:1507.00300.

- Pagoto, S., K. Schneider, M. Jojic, M. DeBiasse, D. Mann. 2013. Evidence–based strategies in weight–loss mobile apps. AJPM 45(5) 576–582.
- Palensky, P., D. Dietrich. 2011. Demand side management: Demand response, intelligent energy systems, and smart loads. *IEEE transactions on industrial informatics* **7**(3) 381–388.
- PG&E. 2016. Save energy and money. URL https://www.pge.com/en_US/residential/ save-energy-money.
- Radner, Roy. 1985. Repeated principal-agent games with discounting. *Econometrica* 53(5) 1173–1198.
- Ralphs, T.K., A. Hassanzadeh. 2014. On the value function of a mixed integer linear optimization problem and an algorithm for construction. Tech. rep.
- Ratliff, Lillian J, Roy Dong, Henrik Ohlsson, S Shankar Sastry. 2014. Incentive design and utility learning via energy disaggregation. *IFAC Proceedings Volumes* 47(3) 3158–3163.
- Rockafellar, R Tyrrell, Roger J-B Wets. 2009. Variational analysis, vol. 317. Springer Science & Business Media.
- Rubin, Paul H. 1978. The theory of the firm and the structure of the franchise contract. The Journal of Law & Economics 21(1) 223-233.
- Saghezchi, F.B., F.B. Saghezchi, A. Nascimento, J. Rodriguez. 2015. Game-theoretic based scheduling for demand-side management in 5g smart grids. 2015 IEEE Symposium on Computers and Communication (ISCC). IEEE, 8–12.
- Samadi, P., A. Mohsenian-Rad, R. Schober, V.W.S. Wong, J. Jatskevich. 2010. Optimal real-time pricing algorithm based on utility maximization for smart grid. Smart Grid Communications (SmartGrid-Comm), 2010 First IEEE International Conference on. IEEE, 415–420.
- Samadi, P., H. Mohsenian-Rad, R. Schober, V.W.S. Wong. 2012. Advanced demand side management for the future smart grid using mechanism design. *IEEE Transactions on Smart Grid* 3(3) 1170–1180.
- Savelsbergh, M., K. Smilowitz. 2016. Stratified patient appointment scheduling for mobile community-based chronic disease management programs. *IIE Transactions on Healthcare Systems Engineering* 6(2) 65–78. doi:10.1080/19488300.2016.1156200.
- Vahn, G. 2015. The data-driven (s, s) policy: Why you can have confidence in censored demand data. Available at SSRN.
- Vogel, S., P. Lachout. 2003a. On continuous convergence and epi-convergence of random functions. part i: Theory and relations. *Kybernetika* **39**(1) [75]–98.
- Vogel, S., P. Lachout. 2003b. On continuous convergence and epi-convergence of random functions. part ii: Sufficient conditions and applications. *Kybernetika* **39**(1) 99–118.
- Wang, W.-Y., D. Gupta. 2011. Adaptive appointment systems with patient preferences. M&SOM 13 373.

- Wijaya, T.K., T.G. Papaioannou, X. Liu, K. Aberer. 2013. Effective consumption scheduling for demand-side management in the smart grid using non-uniform participation rate. Sustainable Internet and ICT for Sustainability (SustainIT), 2013. IEEE, 1–8.
- Wolsey, L., G. Nemhauser. 1999. Integer and Combinatorial Optimization. Wiley-Interscience.
- Wu, S., K. Ell, S. Gross-Schulman, L. Sklaroff, W. Katon, A. Nezu, P.-J. Lee, I. Vidyanti, C.-P. Chou, J. Guterman. 2013. Technology-facilitated depression care management among predominantly Latino diabetes patients within a public safety net care system: Comparative effectiveness trial design. *Contemporary Clinical Trials*.
- Xie, W., Y. Zhao, Z. Jiang, P. Chow. 2016. Optimizing product service system by franchise fee contracts under information asymmetry. *Annals of Operations Research* **240**(2) 709–729.
- Zhou, Zhengyuan, Nicholas Bambos, Peter Glynn. 2016. Dynamics on Linear Influence Network Games Under Stochastic Environments. Springer International Publishing, 114–126.