Submitted to *Management Science*
manuscript MS-0001-1922.65

Appointment Scheduling Under Schedule-Dependent Patient No-Show Behavior

Qingxia Kong
School of Business, Universidad Adolfo Ibanez, q.kong@uai.cl

Shan Li
Zicklin School of Business, Baruch College, City University of New York, shan.li@baruch.cuny.edu

Nan Liu
Department of Health Policy and Management, Mailman School of Public Health, Columbia University, nan.liu@columbia.edu

Chung-Piaw Teo
Department of Decision Sciences, NUS Business School, bizteocp@nus.edu.sg

Zhenzhen Yan
Department of Decision Sciences, NUS Business School, a0109727@nus.edu.sg

This paper studies an appointment scheduling problem under schedule-dependent patient no-show behavior. The problem is motivated by our studies of independent datasets from countries in two continents which identify a significant time-of-day effect on patient show-up probabilities. We deploy a distributionally robust model, which minimizes the worst case total expected cost of patient waiting and service provider's idle and overtime, by optimizing the scheduled arrival times of patients. We show that this model under schedule-independent patient show-up behavior can be reformulated as a copositive program and then be approximated by semidefinite programs. These formulations are obtained by a new technique that uses a completely positive program to equivalently represent a linear program with uncertainties present in both the objective function and the right-hand side of the constraint sets. To tackle the case when patient no-shows are endogenous on the schedule, we construct a set of dual prices to guide the search for a good schedule and use the technique iteratively to obtain a near optimal solution. Our computational studies reveal a significant reduction in total expected cost by taking into account the time-of-day variation in patient show-up probabilities as opposed to ignoring it.

Key words: Distributionally Robust Optimization; Copositive Program; Appointment Scheduling; Patient No-shows

History:

1. Introduction

Consider the following situation: Amy has an appointment with her dentist at noon. However, she is requested to attend a business meeting announced last minute that is scheduled at 12:30pm. As a result, she cannot attend the appointment as she is supposed to. If she has had an 8am appointment, she would have been able to see her dentist first and come to work right after.

Such patient nonattendance (or commonly known as "no-show") behavior frequently arises in clinic appointment scheduling. Due to the uncertainties patient no-show brings into the picture, its prevalence in different medical specialties and geographic regions as well as its potential detrimental impact on patient health outcomes and service provider revenues (Moore et al. 2001, Ulmer and Troxler 2004), patient no-show is a crucial factor for ambulatory care providers, such as primary care doctors, dentists and physical therapists, to consider when designing appointment templates. A provider's daily appointment template specifies the expected number of patients to be seen in a day and the scheduled arrival times of these patients.

To mitigate the effects of patient no-shows, it is common that an appointment template allows over-booking appointment time slots (i.e., scheduling two or more patients into the same time slot). Over-booking will certainly reduce service provider's idle time, and therefore increase throughput by seeing more patients per day. It will, however, increase service provider's overtime and patient's waiting time, and in turn may hurt service provider's satisfaction and patient's experience. As healthcare moves towards more patient-centered, payers shift from the traditional pay-for-service scheme to pay-for-performance by linking reimbursement rates to service providers with patient satisfaction rating (Press Ganey 2008). At the same time, the booming of social media websites significantly increases information transparency in the healthcare market (McCormack 2013), and leads to a soaring competition among healthcare providers in their service quality. For outpatient care providers, it thus becomes more important than ever to adopt an appointment template that achieves the best tradeoff between capacity utilization and patient experience.

A significant amount of operations research efforts have been devoted to investigating the optimal appointment scheduling under patient no-show behavior. Some recent literature on this topic includes Kaandorp and Koole (2007), Robinson and Chen (2010), Hassin and Mendel (2008), LaGanga and Lawrence (2012), Luo et al. (2012), Jiang et al. (2015). Interested readers are referred to Cayirli and Veral (2003) and Denton and Gupta (2003) for a review of the earlier literature. Most, if not all, of this prior literature on appointment scheduling assumes that patient show-up probabilities (or the distributions) are exogenously determined.

As we illustrated in our earlier example, however, in many situations whether or not a patient will show up for an appointment can depend on the time-of-day of her appointment. Working professionals, like Amy, usually have less control of their availability as their work day progresses.

As a result, in that specific occasion, Amy has to miss her noon appointment, but she would have come for an 8am appointment.

There is a surge of interest in using data analytics to improve no-show prediction in healthcare by incorporating more complex factors such as the time-of-day effect. For instance, Gabriel Belfort and nine teammates (at a MIT Hacking Medicine event) built a prototype for the start-up "Smart Scheduling." Using hundreds of patient demographics and punctuality data, their system is able to predict cancelled or missed appointments up to 70% accuracy (c.f. The Boston Globe, July 14, 2014). This allows clinics to efficiently target reminders and double-book appointments, to provide better service availability and to improve patient experience.

A few studies have also pointed out that patient no-show rates may depend on their appointment times of day, but the patterns are not uniform. Lacy et al. (2004) laid out a few reasons for patient no-shows, some of which are related to time (for example, trouble getting off work, transportation, etc.). Moore et al. (2001) reported that morning appointments are more likely to be kept than afternoon slots. LaGanga and Lawrence (2008) showed that no-show rates may vary by appointment slots. Another prospective study of nonattendance in a physiotherapy clinic in Ireland showed that late afternoon slots produce a lower no-show rate compared to morning and early afternoon slots (French et al. 2005).

In order to gain more insights on this phenomenon, we use two independent large datasets from countries in two continents to systematically analyze the impact of appointment time-of-day on patient show-up probabilities. Controlling for patient-level and provider-level factors, we find significant empirical evidences that patient show-up probabilities indeed depend on their appointment times of day. Specifically, we find that patients in a US community healthcare facility are more likely to show up for their appointments at the beginning or the end of the day in weekdays. In a Chilean pediatric practice, however, show-up rates tend to be lower in the early morning. Such different temporal effects of appointment times may be explained by the differences in patient populations and culture. More importantly, the temporal effect size can be quite significant. For instance, our analysis shows that for a US patient scheduled on Wednesday, her show-up probability can increase from 56% to 81% when given an appointment at 8am rather than at noon. These interesting findings motivate our research questions in this paper: (1) how to design an appointment template when time-of-day affects patient no-show behavior; and (2) how much efficiency gain/cost reduction can be achieved by accounting for such time-of-day effect compared to ignoring it?

Specifically, we consider a fixed set of patients to be scheduled in a given clinic session for a single service provider. A clinic session is referred to a consecutive time window during which a service provider serves patients without taking a break. We focus on the design of appointment template that specifies the scheduled arrival times of these patients. As an appointment template is usually

determined before appointments are actually made and without the knowledge of each potential patient's individual characteristics, we do not consider the impact of individual characteristics on patient show-up probabilities in our model. Patient show-up probabilities depend on their scheduled arrival times. All patients, if show up, arrive on time; and no walk-ins are allowed. The overall goal of our model is to design an appointment template that increases throughput by seeing more patients per day (or in other words, limits service provider's idle time), but not at the expense of overwhelming patient waiting time and staff overtime. Following the convention of the literature on this topic, our objective is to minimize the sum of service provider's idle cost and overtime cost as well as patient's waiting cost.

From a practical point of view, the scheduler may not have sufficient data to confidently estimate the *exact* probability distribution of patient no-shows. In contrast, estimating only the first two moments of show-up rates is much less cumbersome. Therefore, we deploy a two-stage stochastic optimization framework from a distributionally robust perspective to solve the appointment scheduling problem mentioned above. In the second stage, we evaluate the total cost given an schedule and the realization of patient no-show status. In the first stage, instead of assuming a specific distribution of show-up rates, we use a set of distributions with given first and second moment information to find out the worst case optimal schedule that minimizes the maximum cost among the family of distributions. Such distributionally robust solutions guarantee the schedule to perform well under all possible distributions. This approach is also versatile enough to handle various salient features of the scheduling problems.

To solve this scheduling problem, we encounter several *unaddressed* technical challenges in the optimization literature. First, incorporating patient no-show behavior demands solving a completely positive program in which uncertainties occur both in the objective function and the right-hand side (RHS) of the constraint sets. In addition, because patient show-up probabilities depend on time of day, uncertainties in the system related to patient no-shows are actually *endogenous* on the schedule – our decision variables. Standard stochastic programming approach does not work here due to such schedule-dependent show-up probabilities. Specifically, we cannot generate random samples to guide the design of the schedule without knowing the schedule. To the best of our knowledge, Pflug G. (1990) was the first to address exogenous uncertainty where the underlying stochastic process depends on the optimization decisions. See Goel and Grossmann (2006) for a review of this area of research. These problems are often approached using a scenario tree representation and a mixed integer programming approach to handle the discrete number of scenarios.

To tackle this challenge, we develop a new modeling technique that enables us to reformulate such a problem with patient no-show. We first solve the appointment scheduling problem with

static show-up rate (i.e., schedule-independent), and then apply the method iteratively to tackle the case with endogenous patient no-show behavior (i.e., schedule-dependent).

To test our proposed methods, we carry out extensive numerical studies. We show that, comparing to the front-loading pattern observed in the optimal schedule with a static show-up rate (Zacharias and Pinedo 2014), when patient show-up probabilities increase over time (e.g., in the case presented by the Chile dataset), the optimal schedule still observes a front-loading pattern but it is postponed. However, when patient show-up probabilities decrease over time (e.g., in the case presented by the US dataset), it is better to spread out patients rather than front loading the system. In both situations, we find significant reductions in the total expected cost by explicitly taking into account the impact of schedule-dependent patient no-show probabilities.

In summary, this paper makes three main contributions to the literature. First, we use two large datasets from countries in two continents to study and quantify the impact of appointment time-of-day on patient show-up probabilities, controlling for patient-level and provider-level factors. We identify significant evidences on the temporal effect of appointment times on patient show-up probabilities in both datasets. Second, comparing to the "classic" front-loading schedule pattern arising from assuming a constant patient show-up rate over time (Zacharias and Pinedo 2014), our model reveals an optimal schedule with different patterns in cognizance of time-varying patient no-show behavior. More importantly, we demonstrate a significant cost reduction that can result from the schedules derived from our model. Third, from a methodological perspective, we develop a general technique that uses completely positive program to equivalently represent a distributionally robust linear program (LP) with uncertainties present in both the objective function and the RHS of the constraint sets. By doing so, we are able to reformulate such a technically challenging problem as a completely positive program that can be approximated by semidefinite programs. This paper offers a general approach to solve problems with this structure.

Two papers most relevant to ours are Kong et al. (2013) and Zacharias and Pinedo (2014). Compared to models developed in these two papers, ours is much more general. The modeling technique of this paper is inspired by Kong et al. (2013), which considers an appointment system with random service durations and assumes that all patients show up for appointments. They develop a linear copositive program to solve the appointment scheduling problem under the worst case distribution. From a technical point of view, our model is much more challenging, as in their formulation the uncertainty only appears in the objective function, but in our problem the uncertainties of patient show-up rate are present in both the objective function and the RHS of the constraint sets. The modeling technique developed in this work can be used to solve general problems of this kind. To derive the exact optimal schedule, Zacharias and Pinedo (2014) require constant, i.e., time-homogeneous, patient no-show behavior. In contrast, our model allows for schedule-dependent

show-up rates. We develop an iterative method to solve this problem, and our computational results suggest that significant efficiency gain can be achieved when accounting for the time-of-day variation in patient no-show behavior. Furthermore, we show that even though our model aims to solve the optimal schedule for the worse-case distribution, such robust schedules can also be near-optimal in terms of the total expected costs when compared to the optimal schedules generated by Zacharias and Pinedo (2014). Another paper that is also relevant is LaGanga and Lawrence (2012), which is the only work that we know considering time-varying show-up probabilities in appointment scheduling. In their model, patient service times are deterministic and are equal to the length of an appointment slot. The decision is to identify the number of patients scheduled for each slot. Given the combinatorial nature of this problem, they use complete enumeration and develop a heuristic approach to solve it. Our work significantly advances theirs by providing a unified optimization framework for a more general class of the problems and by developing algorithms to solve the model efficiently.

The rest of the paper is organized as follows. Section 2 presents a predictive analysis of the time-of-day effect on patient show-up probabilities while controlling for other factors. Section 3 develops the new approach needed to analyze a distributionally robust LP with uncertainties in both the objective and the RHS of constraints. Section 4 introduces the appointment scheduling model with schedule-dependent show-up probabilities, and applies the new approach developed in Section 3 to this model. Section 5 discusses our numerical results, and Section 6 draws our concluding remarks.

2. Time-of-day Effects on Patient Show-up Probabilities

Previous literature has shown that patient characteristics (e.g., gender, age, new or established patient) and provider-level factors (e.g., provider type and relationship with patients) are important predictors for patient no-shows; see, e.g., Ulmer and Troxler (2004). We hypothesize that, controlling for these factors above, time-of-day also has a significant impact on patient attendance behavior, for potential reasons to be discussed soon. In this section, we use two datasets on patients appointment records, one from the US and the other from Chile, to investigate our hypothesis.

2.1. US data

The US data contain patient appointment records from a large urban community health center located in New York City. This center offers comprehensive medical and dental care to the local community and has more than twenty healthcare providers including physicians, nurse practitioners, nutritionists and care managers. The annual visits to this center amount up to more than 25,000.

Our data are extracted from the EMR (Electronic Medical Record) system of this center. This large dataset spans over three years ranging from January 2011 to December 2013. When analyzing

this dataset, we focus on adult primary care visits, i.e., visits to internists, family medicine doctors or nurse practitioners. We exclude walk-in patients from the analysis because they did not schedule their appointments in advance. Among scheduled visits, we exclude those mandated by school or work, e.g., visits for PPD skin tests or vaccine shots, because these visits have much higher show-up probabilities compared to other regular visits. The final dataset contains 35,094 patient visits made by 4,142 distinct patients.

Because some patients made multiple visits, our data has a panel data structure, for which we develop a mixed-effects logistic regression model to account for potential within-subject correlation. Patient visit status (show-up vs. no-show or cancellation) is the dependent variable, and time-of-day is the independent variable of interest. We also control for a number of other potential factors available in the dataset, including patient age, gender, visit type (new vs. established), provider type (family medicine, internal medicine or nurse practitioner) and day of week. For the age variable, we dichotomize patients into younger and elder patients based on a median split at age 52 for ease of interpretation. We use a random intercept to capture individual patient effect.

A full regression model reveals that gender effect is not significant, i.e., male and female patients have statistically the same attendance behavior. The difference between visits to family medicine and internal medicine is not significant either. This is not surprising, as these are all visits to physicians who usually practice in a similar manner. The difference in patient no-show rates due to provider practice manner, if any, should present in comparison between the visits to physicians and nurses. We also find that patient show-up probabilities are (similar and) higher in Tuesday, Wednesday and Thursday compared to other days in the week, controlling for other factors. Thus, we group days into two categories: Tuesday to Thursday, and other days in the week. In our analysis, we model appointment time as a categorical variable in the regression to explore the temporal effect of appointment time. Our final regression model, after excluding non-significant predictors, takes the following form.

$$\mathbf{logit}(p_{ij}) = \beta_0 + \boldsymbol{\beta}_1 AppHour_{ij} + \beta_2 MidWeek_{ij} + \beta_3 Physician_{ij} + \beta_5 Young_{ij} + \alpha_i,$$

in which p_{ij} is the probability of patient i showing up for his/her jth appointment in the dataset ; β_0 is the fixed intercept and $\alpha_i \sim N(0, \sigma_\alpha^2)$ represents the unobserved individual random effect with σ_α^2 being its variance to be estimated; AppHour is a categorical variable for different appointment times in a day; MidWeek=1 if it is Tuesday, Wednesday or Thursday; Physician=1 if the patient sees an physician (not a nurse practitioner); Young = 1 if patient's age is lower than the median age 52 of the sample. As we will discuss later, patient show-up patterns appear quite different in weekdays compared to Saturday, so we develop two regression models, one using full data and

the other using non-Saturday visits. The estimated model coefficients are shown in Table 7 of the Appendix A.

We use the likelihood-ratio test to assess the significance of appointment times which consist of multiple levels, and this test lends strong empirical support for our hypothesis that appointment times have a significant impact on patient show-up rates, controlling for other factors ($p < 0.05$ for the full data model, and $p < 0.01$ for the model excluding Saturday visits). The impact of other factors on patient show-up probabilities is discussed in the Appendix A.

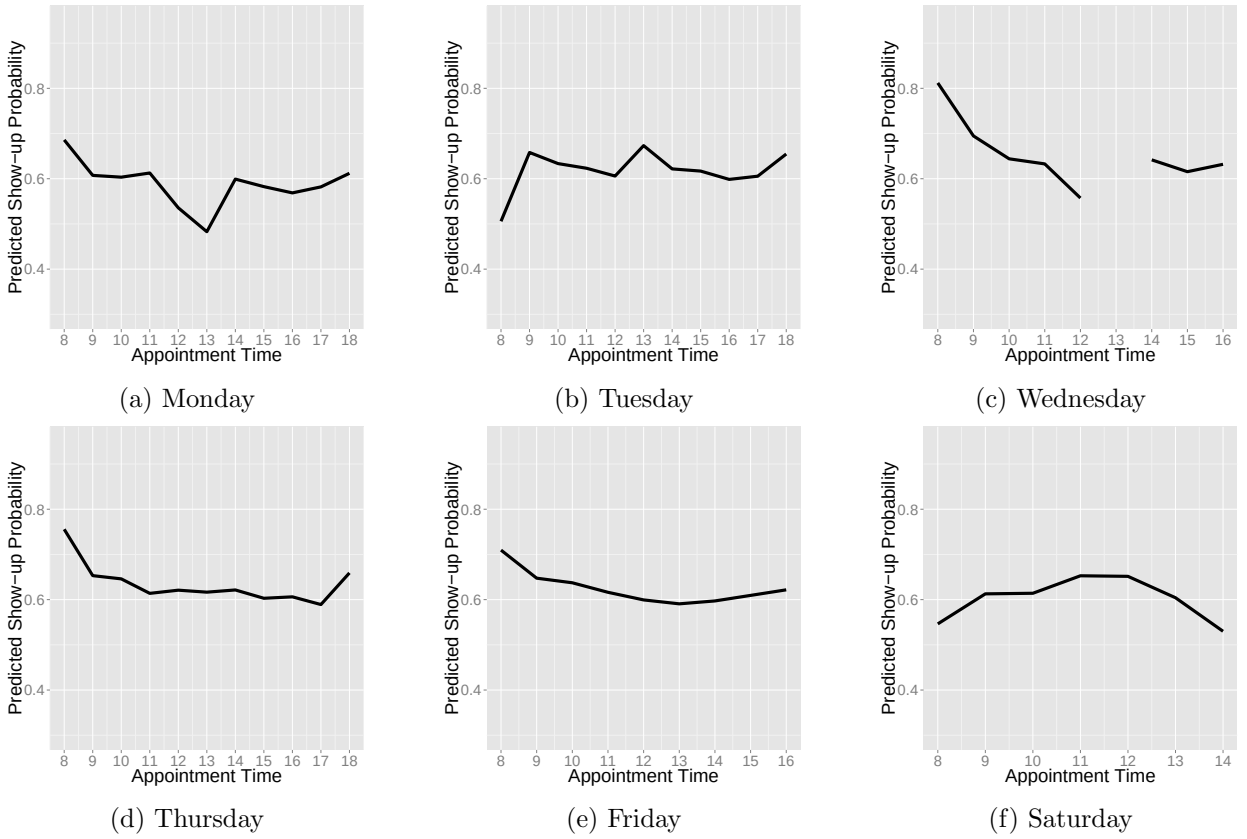
To explore the effect size of time-of-day, we plot, by day of week, the average marginal show-up probabilities over different hour of day in Figure 1. Specifically, for each day of week, we fit a separate mixed-effects logistic regression model; and then for each appointment hour, we use the fitted model to predict the show-up probability for each patient in our dataset for that day of week, holding his/her other characteristics unchanged. We then average the predicted show-up probabilities over all patients in the data for that day. This average is shown as the curved bold line in figures 1a to 1f, representing how the "expected" show-up rate of a random patient in this population changes should s/he be scheduled at different times of a day.

We observe that, in weekdays, the show-up probabilities tend to be higher either in early morning (except for Tuesday) or in late afternoon. However, Saturday exhibits a different pattern: show-up probabilities peak in the middle of the day (a dome-shape). This difference may be explained by people's different life schedules during weekdays and weekends. During weekdays, early morning slots may be the most "convenient" ones from patients' perspectives because attending these appointments slots has the least interruption to one's work/life and thus these slots are less likely to be missed. On Saturday, however, people tend to have a relaxed schedule (and a late breakfast), and thus early morning slots are associated with lower show-up probabilities.

2.2. Chile data

The Chile data consist of patient appointment records from an ENT (ear, nose and throat) department of a public teaching pediatric hospital located in Santiago, Chile. This hospital offers various services including speciality consultation, emergence, and surgical/medical hospitalization. In 2013, this hospital had around 16,500 discharges, 5,500 major surgeries and 90,000 outpatient visits. Our ENT dataset covers one year period from October 2012 to October 2013. During this period, appointments are scheduled from 8am to 4pm Monday through Friday, and there are 7,352 patient visits made by 3,302 distinct patients.

We explore this dataset using a similar approach as for the US data. We control for the following potential predictors in our regression model: age, gender, status (initial visit vs. follow-up visit), provider type (speech specialists, physicians and surgeons), distance from residence to hospital,



(a) Monday (b) Tuesday (c) Wednesday

(d) Thursday (e) Friday (f) Saturday

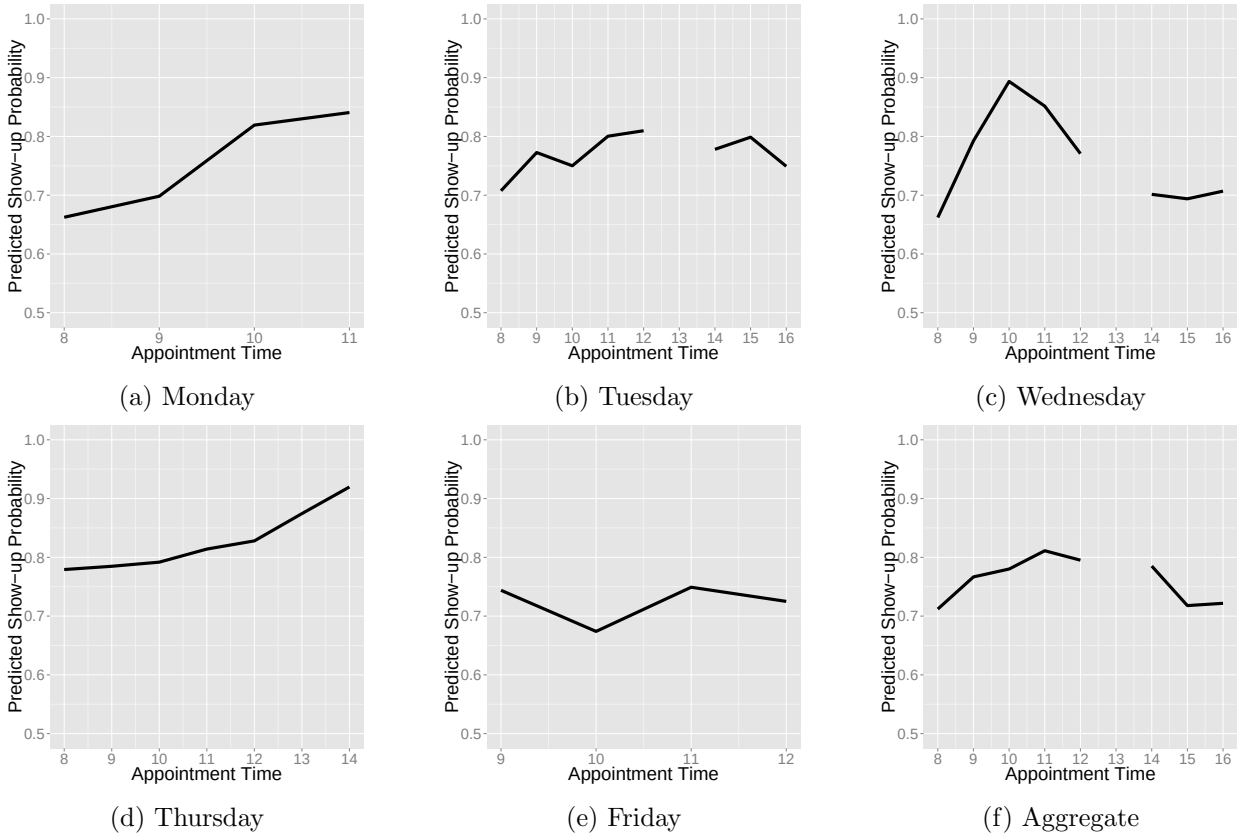
Figure 1 Sample-Average Probabilities of Show-up over Different Times of Day (the US Data).

day of week and time of day. We use a random intercept to capture individual patient effect. For the age variable, we dichotomize patients into younger and elder patients based on a median split at age 6. Our dataset has some information on patients' residence, using which we are able to group patients into three categories: close to hospital (these are patients who live in counties very close to the hospital location); inside the city (these are patients who live in the same metropolitan area where the hospital is located), and outside the city (these are patients who do not live in the metropolitan area of the hospital).

Our regression analysis reveals that gender and age effects are not significant. Distance, thought to have an impact on patient no-show rates, does not appear to be significant either. One possible explanation for these three factors not being significant is that parents were doing their best to bring their children to see the service provider regardless of their residence location or their children' gender and age. We find that patients who visit speech specialists and surgeons tend to have similar attendance behavior, and thus we group these two types of patients in a single category. We also find that patient show-up probabilities are higher on Thursdays but lower on Fridays compared to other days of the week. As a result, we group days into three categories: Monday to Wednesday, Thursday and Friday.

We refer the readers to the Appendix A for details of the final model (see Table 8), and the interpretation of the results. Here, we focus on the temporal effect of appointment times on patient show-up probabilities in different days of week. As we did for the US data, we plot the sample-average marginal show-up probabilities over different times of the day for any given day of week, as well as the aggregated marginal show-up probability over all days of week (because the daily patterns look similar in the Chile data); see Figure 2. Note that the office hour in the Chilean practice is different from that in the US, and very few patients visit at 1pm so we exclude those.

We observe an interesting pattern of show-up probabilities in this Chile dataset in contrast to the US data. In general, patient show-up rates increase over time since early morning, peak at the middle of the day, and then decrease (see Figure 2f). This pattern is different from the weekday pattern in the US data, but similar to Saturday there. There may be a few explanations. First, this is a pediatric population, and parents may need extra preparation time for the visit. Thus, mid of a day appears to be the most convenient times. Second, it is possible due to the less-work-oriented Latino culture and the fact that Latinos usually have a more relaxed attitudes towards time (Flores and Vega 1998).



(a) Monday (b) Tuesday (c) Wednesday

(d) Thursday (e) Friday (f) Aggregate

Figure 2 Sample-Average Probabilities of Show-up over Different Times of Day (Chile Data).

3. An Approach of Completely Positive Decomposition

As we demonstrated in Section 2 that, appointment time-of-day is an important factor that affects the variation in patient show-up rates. We are interested in finding out the structure of the optimal policies to schedule patients by incorporating this new empirical evidence.

We develop a copositive programming reformulation to solve the no-show problem, based on an approach first developed in Kong et al. (2013). However, our problem is more challenging due to two reasons: first, incorporating patient no-show behavior demands a completely positive model with uncertainties in both the objective function and RHS of the constraints; second, patient show-up probabilities depend on our scheduling decision variables. To address the second issue, we use the dual prices associated with the moment cones in the copositive program to guide the search for the optimal appointment schedule, and apply this method iteratively to tackle the problem with endogenous patient no-show behavior (i.e., schedule-dependent).

We first develop in this section the theories on how to use completely positive program to solve LPs with uncertainties in both objective and the RHS of constraints. Later we will apply these results to solve the problem of appointment scheduling with no-shows in Section 4.

3.1. Linear Optimization with Uncertainties in both Objective and Right-Hand Side

We consider a general LP with uncertainties occurring in both the objective $\tilde{\mathbf{c}}$ and the RHS of the constraint sets $\tilde{\mathbf{b}}$. For ease of exposition, we assume the objective function $\tilde{\boldsymbol{c}}$ is a linear function of $\tilde{\boldsymbol{b}}$. i.e.,

$$\tilde{c}_i = c_i(\tilde{\mathbf{b}}) = \mathbf{k}_i^{\mathsf{T}}\tilde{\mathbf{b}} + l_i, \forall i \in \{1, 2, ..., n\},$$

where $\mathbf{k}_i \in R^{m_1}$ is the coefficient vector[1]. We consider the following linear optimization problem:

$$\begin{array}{ll} Z_P(\tilde{\mathbf{b}}) = \max & \mathbf{c}(\tilde{\mathbf{b}})^{\mathsf{T}}\mathbf{x} \\ s.t. & A\mathbf{x} = \tilde{\mathbf{b}}, \quad H\mathbf{x} = \mathbf{d}, \quad \mathbf{x} \geq \mathbf{0} \end{array} \tag{1}$$

where $A := (\mathbf{a}_1\ \mathbf{a}_2\ ...\ \mathbf{a}_{m_1})^{\top}$, $H := (\mathbf{h}_1\ \mathbf{h}_2\ ...\ \mathbf{h}_{m_2})^{\top}$.

Without loss of generality, we assume that our linear optimization problem satisfies the following conditions:

(1) The feasible region is bounded;

(2) if $H\mathbf{x} = \mathbf{0}$ and $\mathbf{x} \geq \mathbf{0}$, then $\mathbf{x} = \mathbf{0}$.

Note that the first condition can be used to construct a redundant deterministic constraint $\sum_{i=1}^{n} x_i \leq M$ that we can add to the model to ensure that the second condition holds.

[1] Throughout this paper, we use boldface notation to denote vectors. For example, we use $\tilde{\mathbf{b}}$ to denote $(\tilde{b}_1, \tilde{b}_2, ..., \tilde{b}_{m_1})$.

We assume further that the distribution of $\tilde{\mathbf{b}}$ lies in a set of multiple distributions supported on $\mathcal{R}_+^{m_1}$ with finite first moment $\boldsymbol{\mu}$ and finite second moment Σ, denoted as $\tilde{\mathbf{b}} \sim (\boldsymbol{\mu}, \Sigma)^+$. We solve the following distributionally robust optimization problem:

$$(P): \quad Z_P = \sup_{\tilde{\mathbf{b}} \sim (\boldsymbol{\mu}, \Sigma)^+} E\left[Z_P(\tilde{\mathbf{b}})\right] \tag{2}$$

3.2. Completely Positive Decomposition

Before showing the main theorem, we first introduce some necessary notation and briefly review related concepts.

A completely positive cone is defined as

$$\begin{aligned}\mathcal{CP}_n &:= \{A \in S_n | \exists V \in \mathcal{R}_+^{n \times m}, \text{such that } A = VV^\top\} \\ &:= \{A \in S_n | \exists \mathbf{v_1}, \mathbf{v_2}, ..., \mathbf{v_k} \in \mathcal{R}_+^n, \text{such that } A = \sum_{i=1}^{k} \mathbf{v_i}\mathbf{v_i}^\top\},\end{aligned}$$

where S_n is $n \times n$ symmetric matrices.

A copositive cone is defined as

$$\mathcal{CO}_n := \{A \in S_n | \forall \mathbf{v} \in \mathcal{R}_+^n, \mathbf{v}^\mathsf{T}\mathbf{A}\mathbf{v} \geq 0\}$$

A copositive cone is the dual of a completely positive cone.

In the following sections of this paper, we use $X \geq_{cp} 0$ (resp. $X \geq_{co} 0$) to represent $X \in \mathcal{CP}_n$ (resp. $X \in \mathcal{CO}_n$). For more information on completely positive cone and copositive cone, we refer interested readers to Berman A (2003).

Let $\mathbf{x}(\tilde{\mathbf{b}})$ denote the optimal solution of problem (1) obtained under $\tilde{\mathbf{b}}$. Let

$$\begin{aligned}\mathbf{p} &:= E[\mathbf{x}(\tilde{\mathbf{b}})] \\ X &:= E[\mathbf{x}(\tilde{\mathbf{b}})\mathbf{x}(\tilde{\mathbf{b}})^\mathsf{T}] \\ Y &:= E[\mathbf{x}(\tilde{\mathbf{b}})\tilde{\mathbf{b}}^\mathsf{T}]\end{aligned}$$

We observe that

$\begin{pmatrix} 1 & \boldsymbol{\mu}^\mathsf{T} & \mathbf{p}^\mathsf{T} \\ \boldsymbol{\mu} & \Sigma & Y^\mathsf{T} \\ \mathbf{p} & Y & X \end{pmatrix} = E\left[\begin{pmatrix} 1 \\ \tilde{\mathbf{b}} \\ \mathbf{x}(\tilde{\mathbf{b}}) \end{pmatrix}\begin{pmatrix} 1 \\ \tilde{\mathbf{b}} \\ \mathbf{x}(\tilde{\mathbf{b}}) \end{pmatrix}^\mathsf{T}\right]$ is a completely positive matrix, as $\begin{pmatrix} 1 \\ \tilde{\mathbf{b}} \\ \mathbf{x}(\tilde{\mathbf{b}}) \end{pmatrix} \in \mathcal{R}_+^{m_1+n+1}$.

Furthermore, since $\tilde{c}_i = c_i(\tilde{\mathbf{b}}) = \mathbf{k}_i^\mathsf{T}\tilde{\mathbf{b}} + l_i, \forall i \in \{1, 2, ..., n\}$, we have

$$E\left[Z_P(\tilde{\mathbf{b}})\right] = E\left[\sum_i (\mathbf{k}_i^\mathsf{T}\tilde{\mathbf{b}} + l_i)x_i(\tilde{\mathbf{b}})\right] = \sum_i \left(\mathbf{k}_i^\mathsf{T} E[\tilde{\mathbf{b}}x_i(\tilde{\mathbf{b}})] + l_i E[x_i(\tilde{\mathbf{b}})]\right),$$

According to Natarajan et al. (2011), the constraint $\mathbf{h}_j^\mathsf{T} x(\tilde{\mathbf{b}}) = d_j$ is well studied and can be formulated as conic constraint by lifting, i.e.

$$E\left[(\mathbf{h}_j^\mathsf{T} x(\tilde{\mathbf{b}}))^2\right] = d_j^2$$

For the constraint $\mathbf{a}_i^\mathsf{T} x(\tilde{\mathbf{b}}) = \tilde{b}_i, \forall i$, we have

$$E\left[(\mathbf{a}_i^\mathsf{T} x(\tilde{\mathbf{b}}))^2\right] = E\left[(\mathbf{a}_i^\mathsf{T} x(\tilde{\mathbf{b}}))\tilde{\mathbf{b}}_i\right] = E\left[\tilde{\mathbf{b}}_i^2\right], \forall i$$

Then we can further infer

$$\sum_i E\left[(\mathbf{a}_i^\mathsf{T} x(\tilde{\mathbf{b}}))^2\right] = \sum_i E\left[(\mathbf{a}_i^\mathsf{T} x(\tilde{\mathbf{b}}))\tilde{\mathbf{b}}_i\right] = \sum_i E\left[\tilde{\mathbf{b}}_i^2\right],$$

and written in matrix form, we have

$$(A^\mathsf{T} A) \bullet E\left[x(\tilde{\mathbf{b}})x(\tilde{\mathbf{b}})^\mathsf{T}\right] = A^\mathsf{T} \bullet E\left[x(\tilde{\mathbf{b}})\tilde{\mathbf{b}}^\mathsf{T}\right] = I \bullet E\left[\tilde{\mathbf{b}}\tilde{\mathbf{b}}^\mathsf{T}\right]$$

Let $\boldsymbol{K}$ denote the matrix $\begin{pmatrix}\mathbf{k}_1 & \mathbf{k}_2 & ... & \mathbf{k}_n\end{pmatrix}$ and $\mathbf{l} = (l_1, \ldots, l_n)$. We now consider the following completely positive program Z_C, obtained by reformulating the problem Z_P using the variables $\mathbf{X}, \mathbf{Y}$ and $\mathbf{p}$:

$$\begin{array}{lll} (C): & Z_C = \max & K \bullet Y^\mathsf{T} + \mathbf{l}^\mathsf{T}\mathbf{p} \\ & s.t. & \begin{pmatrix} 1 & \boldsymbol{\mu}^\mathsf{T} & \mathbf{p}^\mathsf{T} \\ \boldsymbol{\mu} & \Sigma & Y^\mathsf{T} \\ \mathbf{p} & Y & X \end{pmatrix} \geq_{cp} 0 \\ & & (A^\mathsf{T} A) \bullet X = A^\mathsf{T} \bullet Y \\ & & (A^\mathsf{T} A) \bullet X = I \bullet \Sigma \\ & & H\mathbf{p} = \mathbf{d} \\ & & diag(HXH^\mathsf{T}) = diag(\mathbf{d}\mathbf{d}^\mathsf{T}) \end{array} \tag{3}$$

Note that $diag(M)$ denotes a vector of the diagonal elements of matrix M. The main result derived in this section is the following:

THEOREM 1. $Z_C = Z_P$

To show that (C) is actually equivalent to (P), we need to construct a (non-negative) distribution obtained from (C) that satisfies the moment conditions, with corresponding objective value Z_P. The construction hinges on the following observations: consider any completely positive decomposition of matrix

$$\begin{pmatrix} 1 & \mu^\top & \mathbf{p}^\mathsf{T} \\ \mu & \Sigma & \mathbf{Y}^\top \\ \mathbf{p} & \mathbf{Y} & \mathbf{X} \end{pmatrix} = \sum_{k\in\kappa} \begin{pmatrix} \alpha_k \\ \beta_k \\ \boldsymbol{\gamma}_k \end{pmatrix} \begin{pmatrix} \alpha_k \\ \beta_k \\ \boldsymbol{\gamma}_k \end{pmatrix}^\mathsf{T}$$

where $\alpha_k \in \mathcal{R}_+$, $\boldsymbol{\beta}_k, \gamma_k \in \mathcal{R}_+^n$. Let $\kappa_+ = \{k \in \kappa \mid \alpha_k > 0\}$, $\kappa_0 = \{k \in \kappa \mid \alpha_k = 0\}$. The constraints in (C) ensure that

- $A\boldsymbol{\gamma}_k = \boldsymbol{\beta}_k, \forall k \in \kappa$
- $H\frac{\gamma_k}{\alpha_k} = \mathbf{d}, \forall k \in \kappa_+$
- $\boldsymbol{\gamma}_k = \mathbf{0}, \forall k \in \kappa_0$.

Our approach is motivated by the construction in Natarajan et al. (2011), which uses completely positive decomposition to obtain such a desired distribution in the limit. The difference is that, their approach only deals with problems with uncertainties occurring in the objective function alone, whereas in our model, uncertainties are present in both the objective function and the RHS of the constraint sets. Therefore, our model further requires that each constraint with uncertainties in the RHS has to hold in each completely positive decomposition.

4. The Model

In this section, we present a stylized mathematic model to understand the effect on the optimal appointment schedule when patient no-show rates depend on the schedule. Let $N = \{1, 2, ..., m\}$ be the index set of all patients, where m denote the number of patients scheduled to arrive in a day. The number of appointment slots available per day is n, each of unit length. We assume $m \geq n$ for the purpose of focusing on the overbooking effect. The basic assumptions of our appointment scheduling model are listed as follows:

- The service sequence is fixed.
- Patients arrive punctually at the scheduled appointment times, if they show up.
- There is a single service provider in the clinic. The service provider arrives at the same time with the first patient and operates with a work conserving policy (i.e., server does not idle as long as there are patients waiting in the queue).
- Walk-in patients are not considered.

We define s_i as the length of time slot scheduled for ith patient in the sequence, indicating the arrival interval between the ith and $i+1$'th patient. We create a dummy patient, who does not consume any consultation time, arriving at the beginning of the appointment session (i.e., time 0), and thus the time allowance for this dummy patient, denoted as s_0, indicates the first patient's arrival time (and also the arrival time of the service provider). We also add a $m+1$'th patient to arrive at the end of the clinical session to capture the amount of overtime. All patients are scheduled to arrive before the $m+1$'th patient.

Let u_i denote patient i's consultation time. We use $\tilde{b}_i(\mathbf{s}) \in \{0, 1\}$ to denote the show-up status of the ith patient, with $\tilde{b}_i(\mathbf{s}) = 1$ if the ith patient shows up, and 0 otherwise. Note that the show-up state is a function of the schedule $\mathbf{s}$.

We consider three types of costs in our model: (i) the waiting cost of patients, (ii) the idle cost, and (iii) the overtime cost of service provider. The scheduler determines the length of time slot s_i for patient i.

The unit waiting cost for each patient is denoted by c_i, $\forall i = 1, ..., m$. If the last patient is completed after the nth slot, then an overtime cost c_O is charged per unit of time. If the service provider is idle sometime during the day, that incurs an idle cost of c_I per unit of time.

Let W_i denote the waiting time of the ith patient in the system. Since W_{m+1} denotes the amount of overtime work, and $\sum_{i=1}^{m} \tilde{b}_i(\mathbf{s})u_i$ denotes the amount of work brought into the system, the total cost can be represented as

$$f(\mathbf{s}, \tilde{\mathbf{b}}(\mathbf{s})) := \sum_{i=1}^{m} c_i(\tilde{b}_i(\mathbf{s})W_i) + c_O W_{m+1} + c_I(n + W_{m+1} - s_0 - \sum_{i=1}^{m} \tilde{b}_i(\mathbf{s})u_i), \tag{4}$$

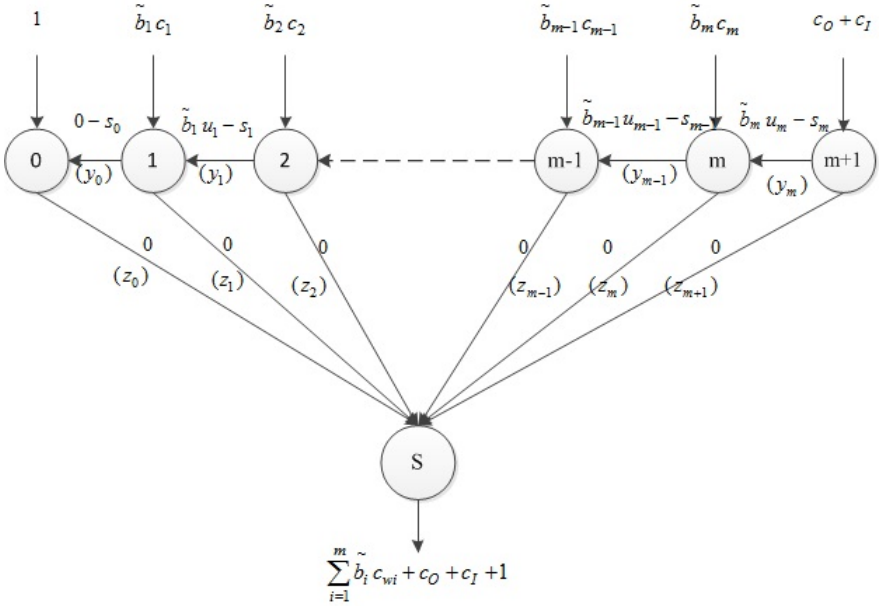
in which $f(\mathbf{s}, \tilde{\mathbf{b}}(\mathbf{s}))$ denotes the total cost incurred given a schedule $\mathbf{s}$ and the show-up state of each patient $\tilde{\mathbf{b}}(\mathbf{s})$.

Using the recursion

$$W_i = \max\left\{0, W_{i-1} + \tilde{b}_{i-1}(\mathbf{s})u_{i-1} - s_{i-1}\right\} \quad i = 2, ..., m+1, \tag{5}$$

we can use a network flow approach to model the total cost function, as shown in Figure 3[2]. In order to capture patient no-show in a network flow representation, we change the inflow to node i to $c_i\tilde{b}_i(\mathbf{s}), \forall i = 1, ..., m$. Therefore if patient i does not show up, the inflow coming into node i becomes 0. In addition, the inflow to node $(m+1)$ is $c_I + c_O$ which corresponds to the term $(c_O + c_I)W_{m+1}$ in (4).

Figure 3 Network Flow Representation



[2] We add an auxiliary patient who arrives at the end of the appointment session (i.e., node (m+1)) to represent the service provider's overtime. We also create a dummy patient arriving at the beginning of appointment session (i.e., node 0), her/his scheduled slot length determines the arrival of the first patient. In Kong et al. (2013), all the inflow to network is deterministic as patient no-show is not considered there.

In the rest of the paper, for ease of exposition we assume that each consultation duration uses exactly one appointment slot, i.e., $u_i = 1$[3] Using this network structure, our problem can be reformulated as

$$\begin{aligned} f(\mathbf{s},\tilde{\mathbf{b}}(\mathbf{s})) = \max \sum_{i=1}^{m}(\tilde{b}_i(\mathbf{s})-s_i)y_i + c_I(n-\sum_{i=1}^{m}\tilde{b}_i(\mathbf{s})) - s_0 y_0 & \\ s.t. \qquad z_0 - y_0 &= 1 \\ z_i - y_i + y_{i-1} &= \tilde{b}_i(\mathbf{s})c_i, \quad \forall i = 1,2,...,m \\ z_{m+1} + y_m &= c_O + c_I \\ \mathbf{y},\mathbf{z} &\geq 0 \end{aligned} \tag{6}$$

To ensure that the formula satisfies the two assumptions proposed in Section 3, we need to add a redundant constraint $\sum_{i=1}^{m}\tilde{b}_i(\mathbf{s}) \leq m$ into the set of constraints in the cone. To see how this constraint helps to justify the two assumptions, we refer the readers to Appendix C.

We deploy a two-stage stochastic optimization framework to solve the appointment scheduling problem in the case of schedule-independent show-up rate from the distributionally robust perspective. Specifically, we consider the following model:

$$\min_{\mathbf{s}\in\Omega_{\mathbf{s}}}\left\{\sup_{\tilde{\mathbf{b}}(\mathbf{s})\sim(\boldsymbol{\mu}(\mathbf{s}),\Sigma(\mathbf{s}))^+}\left\{E[f(\mathbf{s},\tilde{\mathbf{b}}(\mathbf{s}))]\right\}\right\} \tag{7}$$

where $E[f(\mathbf{s},\tilde{\mathbf{b}})]$ is the expected sum of service provider's idle cost and overtime cost as well as patient's waiting cost in the second stage when a schedule $\mathbf{s}$ is given, and $\Omega_{\mathbf{s}}$ is the set of constraints on the schedule in the first stage.

The first step is to calculate worst case expected cost of the second stage problem. For any schedule $\mathbf{s}$, we consider the maximization problem

$$Z_p(\mathbf{s}) := \sup_{\tilde{\mathbf{b}}(\mathbf{s})\sim(\boldsymbol{\mu}(\mathbf{s}),\Sigma(\mathbf{s}))^+}\left\{E[f(\mathbf{s},\tilde{\mathbf{b}}(\mathbf{s}))]\right\} \tag{8}$$

4.1. Reformulation

We can apply our results from Section 3 to solve problem (8). We first present model (6) in a general form as follows:

$$\begin{aligned} f(\mathbf{s},\tilde{\mathbf{b}}(\mathbf{s})) = \max \sum_i (\mathbf{k}_i^{\mathsf{T}}\tilde{\mathbf{b}}(\mathbf{s}))x_i - \mathbf{s}^{\mathsf{T}}\mathbf{x} + \mathbf{f}^{\mathsf{T}}\tilde{\mathbf{b}}(\mathbf{s}) + c_0 & \\ s.t. \qquad A\mathbf{x} = \mathbf{c}_w \circ \tilde{\mathbf{b}}(\mathbf{s}) & \\ H\mathbf{x} = \mathbf{d} & \\ \mathbf{1}_m^{\mathsf{T}}\tilde{\mathbf{b}}(\mathbf{s}) + s_l = m & \\ \mathbf{x} \geq \mathbf{0} & \end{aligned}$$

To obtain an equivalent completely positive model as we showed in Section 3, we first define the following notation:

$$\begin{array}{lll} \mathbf{p} := E[\mathbf{x}(\tilde{\mathbf{b}}(\mathbf{s}))] & Y := E[\mathbf{x}(\tilde{\mathbf{b}}(\mathbf{s}))\tilde{\mathbf{b}}(\mathbf{s})^{\mathsf{T}}] & X := E[\mathbf{x}(\tilde{\mathbf{b}}(\mathbf{s}))\mathbf{x}(\tilde{\mathbf{b}}(\mathbf{s}))^{\mathsf{T}}] \\ ss := E[s_l(\tilde{\mathbf{b}}(\mathbf{s}))^2] & \mathbf{y}_\mu := E[\tilde{\mathbf{b}}(\mathbf{s})s_l(\tilde{\mathbf{b}}(\mathbf{s}))] & \mathbf{y}_{\mathbf{x}} := E[\mathbf{x}(\tilde{\mathbf{b}}(\mathbf{s}))s_l(\tilde{\mathbf{b}}(\mathbf{s}))] \\ p_s := E[s_l(\tilde{\mathbf{b}}(\mathbf{s}))] & \mu(\mathbf{s}) := E[\tilde{\mathbf{b}}(\mathbf{s})] & \boldsymbol{\Sigma}(\mathbf{s}) := E[\tilde{\mathbf{b}}(\mathbf{s})\tilde{\mathbf{b}}(\mathbf{s})^{\top}] \end{array}$$

[3] Our method can be adopted to handle the case when service durations are random.

Let $\boldsymbol{K}$ denote the matrix $\begin{pmatrix}\mathbf{k}_1 & \mathbf{k}_2 & ... & \mathbf{k}_{2m+3}\end{pmatrix}$. Based on our results in Section 3, we consider the following completely positive program

$$\begin{array}{rrcl} Z_C(\mathbf{s}) = \max & K \bullet Y^\top - \mathbf{s}^\top\mathbf{p} + \mathbf{f}^\top\boldsymbol{\mu}(\mathbf{s}) + c_0 & & \\ s.t. & \begin{pmatrix} 1 & \boldsymbol{\mu}(\mathbf{s})^\top & \mathbf{p}^\top & p_s \\ \boldsymbol{\mu}(\mathbf{s}) & \Sigma(\mathbf{s}) & Y^\top & \mathbf{y}_\mu \\ \mathbf{p} & Y & X & \mathbf{y}_\mathbf{x} \\ p_s & \mathbf{y}_\mu{}^\top & \mathbf{y}_\mathbf{x}{}^\top & ss \end{pmatrix} & \geq_{cp} & 0 \\ & (A^\top A) \bullet X & = & A^\top \bullet ((\mathbf{1}\mathbf{c}_\mathbf{w}{}^\top) \circ Y) \\ & (A^\top A) \bullet X & = & \Lambda(\mathbf{c}_\mathbf{w}\mathbf{c}_\mathbf{w}{}^\top) \bullet \Sigma(\mathbf{s}) \\ & H\mathbf{p} & = & \mathbf{d} \\ & diag(HXH^\top) & = & diag(\mathbf{d}\mathbf{d}^\top) \\ & \begin{pmatrix}\mathbf{1}_m^\top & 1\end{pmatrix}\begin{pmatrix}\boldsymbol{\mu}(\mathbf{s}) \\ p_s\end{pmatrix} & = & m \\ & \begin{pmatrix}\mathbf{1}_m^\top & 1\end{pmatrix}\begin{pmatrix}\Sigma(\mathbf{s}) & \mathbf{y}_\mu \\ \mathbf{y}_\mu^\top & ss\end{pmatrix}\begin{pmatrix}\mathbf{1}_m \\ 1\end{pmatrix} & = & m^2 \end{array} \tag{9}$$

We observe that our second stage problem is a maximization problem, while the first stage minimizes the total cost by making schedule decisions. In the literature, one approach to tackle such a min-max problem is to take the dual of the inner maximization problem, and as a result, reformulate the min-max problem as a min-min problem, so that the two stages of the problem can be combined into one. We use $\alpha_0 \in \mathcal{R}$, $\boldsymbol{\beta_0} \in \mathcal{R}^m$, $\Gamma_0 \in \mathcal{R}^{m\times m}$ to denote the dual variables corresponding to moment constraints; and α_1, α_2,$\mathbf{w}^{(1)}$, $\mathbf{w}^{(2)}$, η_1, η_2 to denote the dual variables corresponding to each of the equality constraints in (9). Let $\Lambda(\cdot)$ be the operation of taking the diagonal matrix and $Diag(\cdot)$ be the operation that converts a vector to a diagonal matrix. Define

$$W := \begin{pmatrix} \alpha_0 & \frac{1}{2}(\boldsymbol{\beta}_0 + \eta_1\mathbf{1}_m)^\top & \frac{1}{2}\mathbf{w}^{(1)\top}H & \frac{1}{2}\eta_1 \\ \frac{1}{2}(\boldsymbol{\beta}_0 + \eta_1\mathbf{1}_m) & \Gamma_0 - \alpha_2\Lambda(\mathbf{c}_\mathbf{w}\mathbf{c}_\mathbf{w}{}^\top) + \eta_2\mathbf{1}_m\mathbf{1}_m^\top & -\frac{1}{2}\alpha_1 A \circ (\mathbf{c}_\mathbf{w}\mathbf{1}^\top) & \eta_2\mathbf{1}_m \\ \frac{1}{2}H^\top\mathbf{w}^{(1)} & -\frac{1}{2}(\alpha_1 A \circ (\mathbf{c}_\mathbf{w}\mathbf{1}^\top))^\top & A^\top A(\alpha_1+\alpha_2) + H^\top Diag(\mathbf{w}^{(2)})H & \mathbf{0} \\ \frac{1}{2}\eta_1 & \eta_2\mathbf{1}_m^\top & \mathbf{0}^\top & \eta_2 \end{pmatrix}$$

and $C(\mathbf{s}) := \begin{pmatrix} 0 & \mathbf{0}^\top & \mathbf{0}^\top & -\frac{1}{2}\mathbf{s}^\top & 0 \\ \mathbf{0} & 0 & 0 & \frac{1}{2}K & \mathbf{0} \\ \mathbf{0} & 0 & 0 & 0 & \mathbf{0} \\ -\frac{1}{2}\mathbf{s} & \frac{1}{2}K^\top & 0 & 0 & \mathbf{0} \\ 0 & \mathbf{0}^\top & \mathbf{0}^\top & \mathbf{0}^\top & 0 \end{pmatrix}$

Then the dual of the second stage problem can be written as

$$\begin{array}{rl} Z_D(\mathbf{s}) = & \min \alpha_0 + \boldsymbol{\mu}(\mathbf{s})^\top\boldsymbol{\beta}_0 + \Sigma(\mathbf{s}) \bullet \Gamma_0 + \mathbf{d}^\top\mathbf{w}^{(1)} + \mathbf{d}^\top Diag(\mathbf{w}^{(2)})\mathbf{d} + \eta_1 m + \eta_2 m^2 + \mathbf{f}^\top\boldsymbol{\mu}(\mathbf{s}) + c_0 \\ s.t. & W - C(\mathbf{s}) \geq_{co} 0 \end{array} \tag{10}$$

Note that additional constraints on $\mathbf{s}$ can be added to the cone when required. The above conic programming model therefore provides a unified approach to study many different classes of appointment scheduling systems.

4.2. Iterative Procedure

We use $\mathbf{g}$ to denote the scheduled arrival time of each patient. As we stated before,

$$g_i = s_0 + s_1 + ... + s_{i-1}, \forall i = 1, .., m,$$

where s_0 represents the scheduled time slot for the dummy patient. We define $L \in \mathcal{R}^{(m+1)\times(m+1)}$ and $L_{i+1,j} = 1, i = 1, ..., m; j = 1, ..., i$. Assuming $g_0 = 0$, we can then rewrite the relationship between $\mathbf{s}$ and $\mathbf{g}$ in matrix form as

$$\mathbf{g} = L\mathbf{s}$$

Let $p(t)$ denote a patient's show-up probability if he is scheduled to arrive at time t. For simplicity, we assume that the show-up probability depends solely on the time of arrival and does not depend on patient demographic features. In this case, for any given schedule $\mathbf{s}$,

$$\mu(\mathbf{s}) = \mathbf{p}(\mathbf{g}), \quad \Sigma(\mathbf{s}) = \mathbf{p}(\mathbf{g})\mathbf{p}(\mathbf{g})^\top + Diag(\mathbf{p}(\mathbf{g}) \circ (\mathbf{1}_m - \mathbf{p}(\mathbf{g})))$$

We consider a linear function of show-up probability $\mathbf{p}(\mathbf{g})$ to illustrate our approach[4] . Without loss of generality, we assume $p(g) = a + bg,\ 0 \le g \le n$, where $p(g)$ is linearly increasing in g if $b > 0$ and decreasing otherwise. In this way, we can model $\Sigma(\mathbf{s})$ as a quadratic function in $\mathbf{s}$ which can be easily modelled using the conic approach.

In the schedule-dependent case, the two stage problem is formulated as

$$\begin{array}{ll} \min & \alpha_0 + \boldsymbol{\mu}(\mathbf{s})^\mathsf{T}\boldsymbol{\beta}_0 + \Sigma(\mathbf{s}) \bullet \Gamma_0 + \mathbf{d}^\mathsf{T}\mathbf{w}^{(1)} + \mathbf{d}^\mathsf{T} Diag(\mathbf{w}^{(2)})\mathbf{d} + \eta_1 m + \eta_2 m^2 + \mathbf{f}^\mathsf{T}\boldsymbol{\mu}(\mathbf{s}) + c_0 \\ s.t. & W - C(\mathbf{s}) \succeq_{co} 0 \\ & \mathbf{s} \in \boldsymbol{\Omega}_\mathbf{s} \end{array} \tag{11}$$

Note that the model is non-convex due to the product term $\boldsymbol{\mu}(\mathbf{s})^\mathsf{T}\boldsymbol{\beta}_0$ and $\Sigma(\mathbf{s}) \bullet \Gamma_0$ in the objective function. The constraint set, however, is still linear conic, which enables us to apply an iterative method to solve this problem. The main idea is to separate the two sets of decision variables $(\boldsymbol{\mu}(\mathbf{s}), \Sigma(\mathbf{s}))$ and $(\boldsymbol{\beta}_0, \Gamma_0)$. We fix the value of one pair and solve the above linear conic programs to arrive at a local equilibrium solution in an iterative manner. Note that the objective value obtained this way decreases monotonically.

To do that, our first step is to use the average show-up probability over all slots. We denote the corresponding first and second moments of this average probability as $(\boldsymbol{\mu}_0, \Sigma_0)$. This reduces the problem to the schedule-independent case. We can solve (12) to obtain an optimal schedule $\mathbf{s}_0$ and the corresponding $\boldsymbol{\beta}_{01}$, Γ_{01} by setting $(\boldsymbol{\mu}(\mathbf{s}), \Sigma(\mathbf{s})) = (\boldsymbol{\mu}_0, \Sigma_0)$. We refer $\mathbf{s}_0$ as the static schedule because we use average show-up rates across time of the day. After that, we start the iteration

[4] If $\mathbf{p}(\mathbf{g})$ is nonlinear in $\mathbf{g}$, we can approximate $\mathbf{p}(\mathbf{g})$ by its gradient to apply this iterative method on more general show up function.

from $\boldsymbol{\beta}_{01}$, Γ_{01} and incorporate time-dependent show-up probabilities into the model to generate a new pair of first and second moments, and then obtain the new schedule along the way.

Specifically, after fixing $(\boldsymbol{\beta}_0, \Gamma_0) = (\boldsymbol{\beta}_{01}, \Gamma_{01})$, the next step is to obtain a new schedule $\mathbf{s}_1$ and the corresponding moments $(\boldsymbol{\mu}(\mathbf{s}_1), \Sigma(\mathbf{s}_1))$. In the case of $p(g(\mathbf{s}_1)) = a + bg(\mathbf{s}_1), 0 \le g(\mathbf{s}_1) \le n$, the function of the first moment on patient's arrival time can be presented as $\boldsymbol{\mu}(\mathbf{s}_1) = a\mathbf{1}_m + bg(\mathbf{s}_1)$, and the second moment can be simply written as $\Sigma(\mathbf{s}_1) = \boldsymbol{\mu}(\mathbf{s}_1)\boldsymbol{\mu}(\mathbf{s}_1)^\mathsf{T} + Diag(\boldsymbol{\mu}(\mathbf{s}_1) \circ (\mathbf{1}_m - \boldsymbol{\mu}(\mathbf{s}_1)))$. We next solve the following quadratic programming problem

$$\begin{array}{ll} \min & \alpha_0 + \big(a\mathbf{1}_m + bg(\mathbf{s})\big)^\mathsf{T}\boldsymbol{\beta}_0 + \big(\boldsymbol{\mu}(\mathbf{s})\boldsymbol{\mu}(\mathbf{s})^\mathsf{T} + Diag(\boldsymbol{\mu}(\mathbf{s}) \circ (\mathbf{1}_m - \boldsymbol{\mu}(\mathbf{s}))\big) \bullet \Gamma_0 \\ & + \mathbf{d}^\mathsf{T}\mathbf{w}^{(1)} + \mathbf{d}^\mathsf{T} Diag(\mathbf{w}^{(2)})\mathbf{d} + \eta_1 m + \eta_2 m^2 + \mathbf{f}^\mathsf{T}\big(a\mathbf{1}_m + bg(\mathbf{s})\big) + c_0 \\ s.t. & W - C(\mathbf{s}) \succeq_{co} 0 \\ & \mathbf{s} \in \boldsymbol{\Omega}_\mathbf{s} \end{array} \tag{12}$$

This model is still nonlinear, but the objective now has a quadratic form $g(\mathbf{s})g(\mathbf{s})^\mathsf{T}$ and $g(\mathbf{s}) = \mathbf{L}\mathbf{s}$. We can reformulate this quadratic problem as a conic problem by replacing

$$\mathbf{s}\mathbf{s}^\top = Z_{ss},\ \mathbf{g}(\mathbf{s})\mathbf{g}(\mathbf{s})^\top = Z_{gg},$$

and reformulate the first stage constraints in conic form. Specifically, in the case of continuous slot length, the first stage constraints $\mathbf{s} \in \Omega_s = \left\{\mathbf{s} \in \mathcal{R}_+^{m+1} \mid \sum_{i=0}^{m} s_i = n, s_i \ge 0\right\}$ can be reformulated in conic form as

$$\begin{array}{c} \begin{pmatrix} 1 & \mathbf{s}^\mathsf{T} & \mathbf{g}^\mathsf{T} \\ \mathbf{s} & Z_{ss} & Z_{gs}^\mathsf{T} \\ \mathbf{g} & Z_{gs} & Z_{gg} \end{pmatrix} \succeq_{cp} 0 \\ \mathbf{1}_{m+1}^\mathsf{T}\mathbf{s} = n \\ \mathbf{1}_{m+1}^\mathsf{T} Z_{ss}\mathbf{1}_{m+1} = n^2 \\ \big(\ L\ -I_{m+1}\ \big)\begin{pmatrix} \mathbf{s} \\ \mathbf{g} \end{pmatrix} = \mathbf{0} \\ diag\left(\big(\ L\ -I_{m+1}\ \big)\begin{pmatrix} Z_{ss} & Z_{gs}^\mathsf{T} \\ Z_{gs} & Z_{gg} \end{pmatrix}\big(\ L\ -I_{m+1}\ \big)^\mathsf{T}\right) = \mathbf{0} \end{array} \tag{13}$$

By solving (13) we can obtain a new schedule $\mathbf{s}_1$ and a corresponding show-up rate and moment conditions. This leads to the next iteration of our numerical procedure to obtain a new (β_0, Γ_0).

4.3. Fixed Slot Length

In practice, healthcare appointment systems often adopt a scheduling template with fixed-length appointment slots (for example, each appointment slot is 20-minute long in many clinics). In that case, the length of appointment slot is no longer continuous. Instead, the constraint set $\Omega_\mathbf{s}$ can be presented as $\Omega_\mathbf{s} = \left\{\mathbf{s} \in \mathcal{R}_+^{m+1} \mid \sum_{i=0}^{m} s_i = n, s_i \in \{0,1\}\right\}$, where s_i is a binary variable that indicates whether or not the ith patient is scheduled the same time as patient $i+1$. If the ith patient is scheduled the same time as the $i+1$st patient, $s_i = 0$; otherwise $s_i = 1$. By doing so, we obtain the number of patients scheduled at each appointment slot.

The binary constraint set adds more difficulties to the problem. Based on Burer (2009), which presents an equivalent completely positive representation for quadratic program with binary variables, we can capture the binary constraints using a completely positive program.

Denote

$$\Omega_{Cs} := \left(\begin{pmatrix} 1 & \mathbf{s}^\top & \mathbf{s_l}^\top \\ \mathbf{s} & Z_{ss} & Y^\top \\ \mathbf{s_l} & Y & Z_{ll} \end{pmatrix} \geq_{cp} 0 \;\middle|\; \begin{array}{cc} \mathbf{1}_{m+1}^\top \mathbf{s} & = n \\ \mathbf{1}_{m+1}^\top Z_{ss} \mathbf{1}_{m+1} & = n^2 \\ diag(Z_{ss}) & = \mathbf{s} \\ \begin{pmatrix} \mathbf{e}_i \\ \mathbf{e}_i \end{pmatrix}^\top \begin{pmatrix} \mathbf{s} \\ \mathbf{s_l} \end{pmatrix} & = 1, \forall i \\ \begin{pmatrix} \mathbf{e}_i \\ \mathbf{e}_i \end{pmatrix}^\top \begin{pmatrix} Z_{ss} & Y^\top \\ Y & Z_{ll} \end{pmatrix} \begin{pmatrix} \mathbf{e}_i \\ \mathbf{e}_i \end{pmatrix} & = 1, \forall i \end{array} \right).$$

We can add Ω_{Cs} to constrain the set of feasible schedule $\mathbf{s}$ of model (10), and round the solution to the conic program to 0-1 solution for the schedule $\mathbf{s}$.

5. Computational Studies

We briefly introduce the choice of cost parameters used in our experiments. Three types of costs are involved in determining the optimal schedule: patient's waiting cost rate c_w, service provider's idle cost rate c_I, and overtime cost rate c_O. What matters in the optimization is the ratios among these cost rates, but not their magnitudes. Thus we set $c_I = 1.0$ without loss of generality. One classic way to measure the value of time is its opportunity cost, which is typically assumed to be the wage rate (Becker 1965). As a typical primary care physician's income is about \$220,942 and the median personal income in the US is \$24,062, we consider $c_w = 0.1$ as a base case in our study. However, we vary $c_w \in [0.05, 0.5]$ to study the impact of different waiting costs. To set the overtime cost rate, we adopt the US federal government mandate that overtime salary rate should be at least 1.5 times of the regular time salary rate, and thus we set $c_O = 1.5$. Note that the ranges of these cost parameters are also similar to those considered in the previous literature, e.g., Zacharias and Pinedo (2014).

5.1. Demonstration of the Iteration Method

This section uses a simple example of a 12-slot clinic session to illustrate the iteration method developed above. We assume the consultation time is deterministic and each patient spends exactly one time slot with the service provider. We set $c_w = 0.1, c_I = 1$, $c_O = 1.5$; and consider two patterns of show-up probabilities: the decreasing case

$$p(t) = 0.9 - \frac{0.8}{12}t, \; t = 0, 1, \cdots, 12, \tag{14}$$

in which patient show-up probabilities linearly decrease from 0.9 to 0.1 over time and the increasing case

$$p(t) = 0.1 + \frac{0.8}{12}t, \; t = 0, 1, \cdots, 12, \tag{15}$$

Table 1 Performance of the Iterative Method under Decreasing Show-up Probabilities

m	13	14	15	16	17	18	19	20
Static schedule	13.1657	14.4350	13.2988	12.3526	11.7105	11.0076	10.6449	10.7975
Iterative method	5.9862	5.5848	5.3458	5.2119	5.0513	5.0388	5.1173	5.2168
Improvement (%)	54.53	61.31	59.80	57.81	56.87	54.22	51.93	51.68

Table 2 Performance of the Iterative Method under Increasing Show-up Probabilities

m	13	14	15	16	17	18	19	20
Static schedule	13.2002	13.2924	12.7287	12.1040	11.6465	11.8348	10.9509	10.7546
Iterative method	7.8485	7.7065	7.5828	7.4635	7.3813	7.8811	7.1959	7.1141
Improvement (%)	40.54	42.02	40.43	38.34	36.62	33.41	34.29	33.85

where patient show-up rates linearly increase from 0.1 to 0.9 over time. We vary the number of patients to be scheduled $m \in \{13, 14, \ldots, 20\}$ to test the impact of different overbooking levels.

To implement the iterative method described in Section 4.2, we first use the average show-up probability over all slots (which is 0.5 in both the increasing the decreasing cases) to get the corresponding optimal schedule. We call this schedule the *static* schedule because show-up probabilities are static over time. To iterate, we solve for the next schedule using patient show-up probabilities calculated based on the last obtained schedule.

For each m, we perform 500 iterations. We report the average worse case expected cost in the last 100 iterations; see row 2 in Tables 1 and 2. We also show the worse case expected cost of the static schedule (that ignores schedule-dependent no-shows) while schedule-dependent no-shows indeed present and the percentage improvement made by the iterative method that considers schedule-dependent no-shows; see rows 1 and 3 in both tables. We observe that the iterative method converges quickly; the average coefficients of variation in the last 100 iterations is less than 0.1% for all cases we tested. In addition, the optimal schedule obtained under the iterative method makes a significant improvement $(30\% - 60\%)$ over the static schedule in terms of the worse case expected cost.

5.2. Analysis of Schedule Patterns

In this section, we study how show-up probabilities and patient's waiting cost affect the optimal schedule.

5.2.1. Impact of Show-up Probabilities We assume that there are $n = 12$ slots in the clinic session and consider both increasing and decreasing show-up rates (14) and (15) described in Section 5.1. We use $m = 18$ to illustrate the impact of show-up probabilities on the optimal schedule. Figure 4 plots the optimal schedules under both sets of show-up probabilities as well as the optimal static schedule (assuming show-up probability is 0.5 throughout the day). The horizontal axis indicates the index of patients, and the vertical axis shows the scheduled arrival

time for each patient. As expected, the static schedule demonstrates an obvious pattern of "front loading." In particular, the first four patients are scheduled to come at the beginning of the clinic session. This observation is consistent with the pattern of an optimal schedule under the total expected cost criteria when patient show-up probabilities are constant over time; see, e.g., Hassin and Mendel (2008).

When show-up probabilities increase over time, we still observe a front-loading pattern in the optimal schedule but it is postponed. That is, a group of patients are asked to arrive together not at the beginning of the session but later in the day. This is likely due to the low show-up probabilities at the beginning of the session, and thus it makes more sense to have patients come later to avoid idle time of the service provider[5]. This finding suggests that service providers may want to proactively delay the start of their service time if observing a high level of no-show probabilities at the beginning of the day.

When show-up probabilities decrease over time, however, patients are in general scheduled to arrive later compared to the static schedule and the schedule under increasing show-up probabilities. This is because higher show-up probabilities at the beginning of the session are likely to build up wait lines of patients, if they were scheduled densely at the beginning of the session. As a result, the optimal schedule in this case tends to "smooth" the workload later into the day.

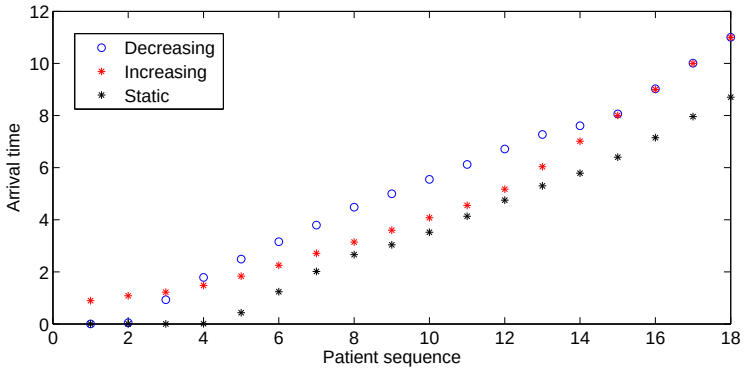
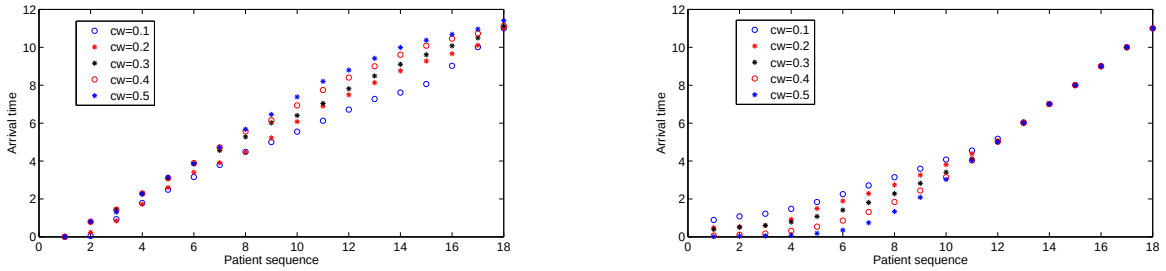


Figure 4 Optimal Schedules under Different Show-up Probabilities

5.2.2. Schedule under Different Waiting Costs In this section, we vary the waiting costs from 0.1 to 0.5 to examine its effect on the optimal schedule. This analysis sheds light on the design of appointment templates for patient populations with different valuations of waiting. Similar to the previous section, we fix $n = 12, m = 18$ and consider both increasing and decreasing show-up probabilities (14) and (15). Figures 5a and 5b show the optimal schedules under these two sets

[5] Note that the service provider comes together with the first patient.

of show-up probabilities, respectively. A close comparison of these two figures reveals that under decreasing show-up probabilities, a larger waiting cost rate leads to less front-loading. In contrast, with increasing show-up probabilities, a larger waiting cost rate results in more significant front-loading. One intuitive explanation is that as the waiting cost gets higher, we want to avoid patient waiting by possibly reducing the "expected" number of patients who show up. Thus, the optimal schedule tends to assign more patients to appointment slots with lower show-up probabilities. As a result, when waiting cost increases, fewer patients are assigned early in the session when show-up probabilities peak at the beginning the session; but more patients get assigned early in the session when show-up probabilities bottom at the beginning of the session.



(a) Decreasing Show-up Rates (b) Increasing show-up rate

Figure 5 Schedule under Different Waiting Costs

5.3. Value of Incorporating Schedule-Dependent No-Shows

Our optimal schedule is determined based on minimizing the *worst case* total cost. It is therefore of natural interest to examine its performance in terms of the total expected cost. To do so, we run simulations (10,000 replications) to calculate the expected total costs under our robust optimal schedule (which takes into account schedule-dependent no-shows) and the static schedule (which ignores that). Similar to above, we consider two patterns of show-up probabilities (14) and (15).

Table 3 shows the results associated with the decreasing show-up rates (14). The first two rows show the total expected costs under the static schedule and the robust optimal schedule over a range of overbooking levels. The third row represents the average cost reduction percentage. We observe a significant improvement (3.24% to 54.92% reduction) in total expected cost due to explicit consideration of schedule-dependent no-shows. Furthermore, when the overbooking level increases, the improvement becomes more significant. In Section 5.2.2, we note that the static schedule assigns quite a few patients early in the session when the overbooking level is high ($m = 18$). As this does not recognize the fact that show-up rates are higher (than average) at the beginning of the session, it leads to more waiting costs and thus larger total expected costs. Therefore, the cost difference becomes more significant when more patients need to be scheduled.

Table 4 shows the results under increasing show-up rates (15). We see a significant improvement (12.37% to 27.04%) in total expected cost by explicitly taking into account the impact of patient increasing show-up rates. In contrast to the case of decreasing show-up rates, the cost reduction is higher when the overbooking level is lower. As discussed in Section 5.2.1, patients will be postponed to come under increasing show-up probabilities. With fewer patients to be scheduled, such a postponement is likely to be more effective. If, however, there are a larger number of patients to be scheduled, postponement will lead to longer patient wait later in the day and thus has a smaller room to make improvement.

Table 3 Comparison of Total Expected Costs under Decreasing Show-up Rates

m	13	14	15	16	17	18	19	20
Static schedule	5.2443	5.3194	5.5416	5.8990	6.5564	7.4330	8.6695	10.0161
Iterative method	5.0742	4.8016	4.7510	4.7976	4.5542	4.1783	4.2586	4.5149
Improvement (%)	3.24	9.73	14.27	18.67	30.54	43.79	50.88	54.92

Table 4 Comparison of Total Expected Costs under Increasing Show-up Rates

m	13	14	15	16	17	18	19	20
Static schedule	8.0677	7.9187	7.5349	7.1644	6.8241	6.4858	6.1691	5.9093
Iterative method	5.9578	5.7775	5.6014	5.4852	5.3661	5.2704	5.1990	5.1784
Improvement (%)	26.15	27.04	25.66	23.44	21.37	18.74	15.73	12.37

5.4. Case Studies

In this section, we apply our methodology to two case studies inspired by real data. In particular, we consider designing appointment templates under patient show-up probabilities found in our US and Chile datasets. We are interested in fixed-interval-length appointment templates. Since the schedule proposed by the copositive model is fractional, we develop a rounding heuristic for generating binary schedules during the iterations, i.e., to get a fix-interval-length appointment template in each step of the iterative method (see the Appendix D).

To evaluate the performance of our rounding heuristics, we benchmark our binary schedules with those presented in Zacharias and Pinedo (2014) that provide the lowest total expect cost. Similar to their setup, we let $c_I = 1$, $c_O = 1.5$,$c_w = 0.1$, $n = 12$, $m = 18$; and vary patient no-show probability from 0 to 0.8.[6] See Appendix E for detailed schedules.

For ease of cost comparison, Figure 6 illustrates the total expected total cost, waiting cost, overtime cost and idle cost under our schedules and those in Zacharias and Pinedo (2014). We see that even though our model focuses on the worst case perspective, the resulting schedules perform

[6] Note that the patient show-up probability here is constant over time of day.

fairly close to those in Zacharias and Pinedo (2014) in terms of total expected cost. Furthermore, our model replicates the key insights obtained by Zacharias and Pinedo (2014) that there exists an optimal level of no-show probability to achieve the lowest total expected cost for a given number of patients to be scheduled. These results indicate that our rounding heuristics can generate fix-interval-length appointment templates that are near optimal in the sense of total expected cost.

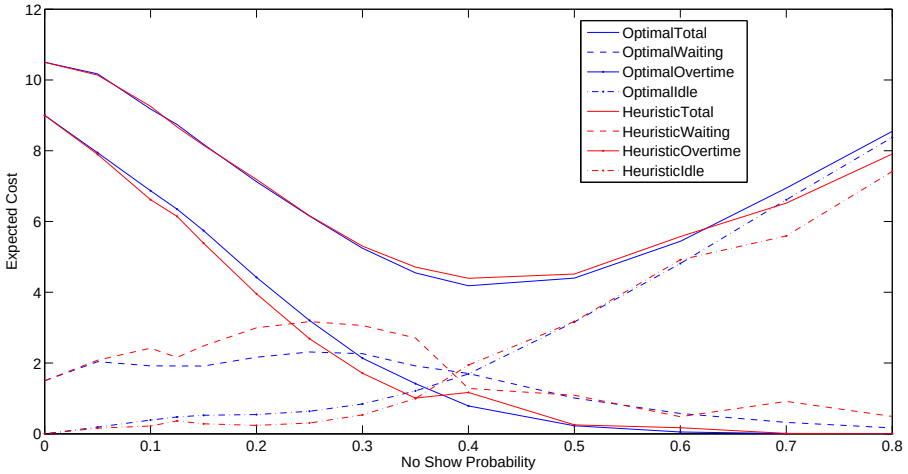


Figure 6 Cost Decomposition under Different Show-up Probabilities

5.4.1. The US Case Instead of modeling each single day, we consider the overall range of show-up probabilities in the US dataset. From the predictive analysis in Section 2, the show-up probability ranges approximately between 0.8 to 0.4. As service providers often take a lunch break at noon, it makes sense to consider a half day. Specifically, we consider the morning sessions in which patient show-up rates decrease in most days in the US dataset. We assume that patient show-up probabilities linearly decrease from 0.8 to 0.4. In our calculations, we use the optimal schedule obtained from Zacharias and Pinedo (2014) with a static show-up probability 0.6 as the starting point of the iteration method. Assuming $c_w = 0.1, c_I = 1, c_O = 1.5$, we apply the iterative method and rounding heuristics described previously to get the robust optimal schedule. We then use simulations to estimate the corresponding total expected costs, and compare these costs with those under the static schedule derived based on Zacharias and Pinedo (2014) and assuming a constant no-show probability over time. Table 5 presents the detailed results.

We observe a significant reduction (7%-13%) in total expected costs when schedules are obtained by incorporating the schedule-dependent no-show behaviors. More importantly, this improvement is relatively insensitive to the waiting cost rate and overbooking level, suggesting that service

Table 5 Case Study Results for the US data

$c_w = 0.05$	Total Cost	**Improvement**	1	2	3	4	5	6	7	8	9	10	11	12	m
Static	3.3715		3	2	1	2	1	2	1	2	1	1	1	1	18
Iterative	3.1381	**6.92**%	1	2	1	2	1	2	2	1	2	1	2	1	18
$c_w = 0.1$	Total Cost	**Improvement**	1	2	3	4	5	6	7	8	9	10	11	12	m
Static	4.0613		3	1	2	1	2	1	1	2	1	1	1	1	17
Iterative	3.6167	**10.95**%	1	1	2	1	2	1	2	1	2	1	2	1	17
$c_w = 0.15$	Total Cost	**Improvement**	1	2	3	4	5	6	7	8	9	10	11	12	m
Static	4.2574		2	2	1	1	2	1	2	1	1	1	1	1	16
Iterative	3.8740	**9.01**%	1	1	2	1	2	1	1	2	1	2	1	1	16
$c_w = 0.2$	Total Cost	**Improvement**	1	2	3	4	5	6	7	8	9	10	11	12	m
Static	4.7362		2	1	2	1	2	1	1	2	1	1	1	1	16
Iterative	4.2864	**9.5**%	1	1	2	1	2	1	1	2	1	2	1	1	16
$c_w = 0.25$	Total Cost	**Improvement**	1	2	3	4	5	6	7	8	9	10	11	12	m
Static	5.2880		2	1	2	1	2	1	1	2	1	1	1	1	16
Iterative	4.6898	**9.5**%	1	1	2	1	2	1	1	2	1	2	1	1	16
$c_w = 0.3$	Total Cost	**Improvement**	1	2	3	4	5	6	7	8	9	10	11	12	m
Static	5.0494		2	1	1	2	1	1	2	1	1	1	1	1	15
Iterative	4.5597	**9.7**%	1	1	1	2	1	1	2	1	1	2	1	1	15
$c_w = 0.4$	Total Cost	**Improvement**	1	2	3	4	5	6	7	8	9	10	11	12	m
Static	5.7294		2	1	1	2	1	1	1	2	1	1	1	1	15
Iterative	5.0577	**11.72**%	1	1	1	2	1	1	2	1	1	2	1	1	15
$c_w = 0.5$	Total Cost	**Improvement**	1	2	3	4	5	6	7	8	9	10	11	12	m
Static	5.6311		2	1	1	1	1	2	1	1	1	1	1	1	14
Iterative	4.9934	**11.33**%	1	1	1	1	2	1	1	1	2	1	1	1	14
$c_w = 0.6$	Total Cost	**Improvement**	1	2	3	4	5	6	7	8	9	10	11	12	m
Static	5.5526		2	1	1	1	1	1	1	1	1	1	1	1	13
Iterative	4.8491	**12.67**%	1	1	1	1	1	1	2	1	1	1	1	1	13

providers may always benefit significantly from taking into account the schedule-dependent no-show behavior regardless of patient valuation of waiting and patient demand level. In addition, the optimal robust schedule exhibits a quite different pattern compared to the static schedule. It does not exhibit a front-loading pattern; instead, it spreads out the overbooked slots throughout the session.

5.4.2. The Chile Case We follow a similar rationale above to conduct a case study based on Chile data. We observe the show-up probability in Chile data ranges roughly from 0.5 to 0.9, and increases over time of day. In contrast to the US data case, we see more front-loading in the robust optimal schedule compared to the static schedule; see Table 6. This seems to contradict with our earlier finding in Section 5.2.1 that front-loading would be postponed when show-up probabilities increase over time. But note that in our case study we do not allow flexible arrival times, and only consider fixed-interval-length schedules. Enforcing this integer constraint turns out to assign more patients to the slots with lower show-up probability, in particular, the first slot. As a result,

Table 6 Case Study Results for Chile data

$c_w = 0.05$	Total Cost	**Improvement**	1	2	3	4	5	6	7	8	9	10	11	12	m
Static	2.6908		2	2	1	1	2	1	1	1	1	1	1	1	15
Iterative	2.6678	**0.85**%	2	2	2	1	1	1	1	1	1	1	1	1	15
$c_w = 0.1$	Total Cost	**Improvement**	1	2	3	4	5	6	7	8	9	10	11	12	m
Static	3.0306		2	1	2	1	1	2	1	1	1	1	1	1	15
Iterative	2.9821	**1.60**%	2	2	2	1	1	1	1	1	1	1	1	1	15
$c_w = 0.15$	Total Cost	**Improvement**	1	2	3	4	5	6	7	8	9	10	11	12	m
Static	3.4120		2	1	1	1	2	1	1	1	1	1	1	1	14
Iterative	3.2540	**1.60**%	2	1	2	1	1	1	1	1	1	1	1	1	14
$c_w = 0.2$	Total Cost	**Improvement**	1	2	3	4	5	6	7	8	9	10	11	12	m
Static	3.4779		2	1	1	1	2	1	1	1	1	1	1	1	14
Iterative	3.3947	**1.60**%	2	1	2	1	1	1	1	1	1	1	1	1	14
$c_w = 0.25$	Total Cost	**Improvement**	1	2	3	4	5	6	7	8	9	10	11	12	m
Static	3.5388		2	1	1	1	1	2	1	1	1	1	1	1	14
Iterative	3.7190	**1.60**%	2	1	2	1	1	1	1	1	1	1	1	1	14
$c_w = 0.3$	Total Cost	**Improvement**	1	2	3	4	5	6	7	8	9	10	11	12	m
Static	3.6287		2	1	1	1	1	1	1	1	1	1	1	1	13
Iterative	3.6287	**0**%	2	1	1	1	1	1	1	1	1	1	1	1	13
$c_w = 0.4$	Total Cost	**Improvement**	1	2	3	4	5	6	7	8	9	10	11	12	m
Static	3.7449		2	1	1	1	1	1	1	1	1	1	1	1	13
Iterative	3.7449	**0**%	2	1	1	1	1	1	1	1	1	1	1	1	13

both the robust optimal schedule and the static schedule present (relatively similar) front-loading patterns, and thus their performances do not differ too much.

6. Conclusion

In this paper, we study an appointment scheduling problem in the presence of schedule-dependent patient no-show behavior. The problem is well motivated by the studies of two independent datasets from countries in two continents. Specifically, we find that patients in a US adult practice are more likely to attend their appointments at the beginning or the end of the day in weekdays. In contrast, in a Chilean pediatric clinic facility, patients are less likely to show up for early morning appointments. We incorporate these interesting findings into the problem of appointment template design, analyze the pattern of the optimal schedule and the efficiency gain achieved by accounting for such time-of-day effects.

One critical difficulty here is that we can not generate random samples without knowing the schedule, if system uncertainties (i.e., patient no-show rates) are endogenous on our decision variables (i.e., the schedule). To tackle this challenge, we develop a distribution-free two stage stochastic programming problem formulation and show that the problem can be reduced into a single stage conic programming problem. The dual prices obtained from the conic program can be used iteratively to guide the algorithm to search for a good scheduling solution in the case when no-show rate is schedule-dependent.

We present the distributionally robust solutions for the situations when the appointment slots are continuous-length or fixed-length, respectively. Through an extensive set of numerical experiments we show that the optimal schedule generated under the worst case optimization also gives near-optimal solutions in terms of the total expected cost. Furthermore, the iterative method converges quickly, and results in a significant reduction (30% - 60%) in total expected cost as opposed to ignoring the time-of-day effect on patient no-show rate.

Based on realistic examples motivated by our two large datasets, our model reveals that an optimal appointment schedule in cognizance of schedule-dependent no-show rates may exist different patterns compared to the classic "front-loading" or "dome-shaped" patterns reported in the literature that assumes constant show-up rate. Specifically, we show that health care providers should spread out the overbooked slots when patient show-up rate decreases over time (e.g., the case presented by our US dataset), but the optimal schedule should present a postponed front-loading pattern when patient show-up rate increases over time (e.g., the case presented by our Chile dataset). By taking into account the time-of-day effect, the overall cost reduction can be up to 13% in our case studies.

In summary, our study empirically demonstrates how appointment time-of-day impact patient no-show behavior. Relying on only the first and second moment information, our modeling approaches are suitable for a clinic provider who has a limited amount of data on patient attendance behavior. Our results offer insights on the structure of the optimal appointment template that corresponds to different time-varying patterns of patient no-show rates, and our model can be used by any ambulatory care provider to design his or her appointment template when facing a specific patient no-show pattern.

Our work also points to several avenues for future research. First, patients may have different schedule-dependent attendance behaviors depending on their work and life styles, e.g., employed vs. retired. Taking into account such heterogeneities, how can one design an appointment template? The other interesting direction is to consider the impact of additional uncertainties that may incur in the system, e.g., patient unpunctual arrivals. In addition, our model focuses on the design of a daily appointment template. One may consider evaluating the performance of these appointment templates in a rolling-horizon setting in which patients arrive randomly each day. Technical analysis of such systems is likely to be difficult, but a simulation study may lead to meaningful results.

Acknowledgments

References

Becker, G. S. 1965. A theory of the allocation of time. *The economic journal* **75**(299) 493–517.

Berman A, Shaked-Monderer N. 2003. *Completely Positive Matrices*. World Scientific, Singapore, Republic of Singapore.

Burer, S. 2009. On the copositive representation of binary and continuous nonconvex quadratic programs. *Mathematical Programming* **120** 479–495.

Cayirli, T., E. Veral. 2003. Outpatient scheduling in health care: A review of literature. *Production and Operations Management* **12**(4) 519–549.

Denton, B., D. Gupta. 2003. A sequential bounding approach for optimal appointment scheduling. *IIE Transactions* **35**(11) 1003–1016.

Flores, G., L. R¿ Vega. 1998. Barriers to health care access for latino children: a review. *FAMILY MEDICINE-KANSAS CITY-* **30** 196–205.

French, H., E. McGrane, G. Cooke. 2005. A prospective study of non-attendance to a physiotherapy outpatient department. *Physiother Ireland* **26** 16–22.

G., Pflug. 1990. On-line optimization of simulated markovian processes. *Math. Oper. Res* **15**(3) 381–395. *Math. Oper. Res*. 1990.

Goel, Vikas, Ignacio E. Grossmann. 2006. A class of stochastic programs with decision dependent uncertainty. *Mathematical Programming* **108**(2) 355–394. *Mathematical Programming*. September, 2006.

Hassin, R., S. Mendel. 2008. Scheduling arrivals to queues: A single-server model with no-shows. *Management Science* **54**(3) 565–572.

Jiang, R., S. Shen., Y. Zhang. 2015. Distributionally robust appointment scheduling with random no-shows and service durations. Working paper. University of Michigan, Ann Arbor.

Kaandorp, G. C., G. Koole. 2007. Optimal outpatient appointment scheduling. *Health Care Management Science* **10**(3) 217–229.

Kong, Q., C.Y. Lee, C.P. Teo, Z. Zheng. 2013. Scheduling arrivals to a stochastic service delivery system using copositive cones. *Operations research* **61**(3) 711–726.

Lacy, N. L., A. Paulman, M. D. Reuter, B. Lovejoy. 2004. Why we dont come: patient perceptions on no-shows. *The Annals of Family Medicine* **2**(6) 541–545.

LaGanga, L., S. R. Lawrence. 2012. Appointment overbooking in health care clinics to improve patient service and clinic performance. *Production and Operations Management* Forthcoming.

LaGanga, L. R., S. R. Lawrence. 2008. Clinic no-shows and overbooking: Reflections and new directions in appointment yield management. *Proceedings of Decision Sciences Institute Annual Conference, Baltimore, Maryland*.

Laurant, M., D. Reeves, R. Hermens, J. Braspenning, R. Grol, B. Sibbald, et al. 2005. Substitution of doctors by nurses in primary care. *Cochrane Database Syst Rev* **2**(2).

Lerner, B. 2007. When patients do not follow up? *The New York Times*. November 13, 2014.

Luo, J., V. G. Kulkarni, S. Ziya. 2012. Appointment scheduling under patient no-shows and service interruptions. *Manufacturing and Service Operations Management* Forthcoming.

McCormack, M. 2013. How to treat patient wait time woes. *Software Advice: IndustryView* Retrieved on June 24, 2015, `http://www.softwareadvice.com/medical/industryview/how-to-treat-patient-wait-time-woes/`.

Moore, C. G., P. Wilson-Witherspoon, J. C. Probst. 2001. Time and money: Effects of no-showsat a family practice residency clinic. *Family Medicine* **33**(7) 522–527.

Natarajan, K., C.P. Teo, Z. Zheng. 2011. Mixed zero-one linear programs under objective uncertainty: a completely positive representation. *Operations Research* **59** 713–728.

Press Ganey. 2008. The impact of patient satisfaction on pay-for-performance in medical practices. *White Papers.* Retrieved on June 24, 2015, `http://pressganey.com`.

Robinson, L. W., R. R. Chen. 2010. A comparison of traditional and open-access policies for appointment scheduling. *Manufacturing & Service Operations Management* **12**(2) 330–346.

Ulmer, T., C. Troxler. 2004. The economic cost of missed appointments and the open access system. *Community Health Scholars.* University of Florida, Gainsville, FL.

Zacharias, C., M. Pinedo. 2014. Appointment scheduling with no-shows and overbooking. *Production and Operations Management* **23**(5) 788–801.

Appendix A: Predictive Analysis

Table 7 Estimated Regression Coefficients for Patient Show-up Probabilities (US data).

	All days			Weekdays only		
	Est. Coef.	95% CI	p-value	Est. Coef.	95% CI	p-value
Intercept	0.871	[0.624, 1.119]	N.A.	0.896	[0.635, 1.159]	N.A.
9am	-0.134	[-0.382, 0.114]		-0.165	[-0.427, 0.098]	
10am	-0.163	[-0.411, 0.085]		-0.207	[-0.469, 0.052]	
11am	-0.177	[-0.425, 0.072]		-0.246	[-0.510, 0.017]	
12pm	-0.281	[-0.536, -0.027]		-0.378	[-0.649, -0.108]	
1pm	-0.217	[-0.500, 0.067]		-0.255	[-0.579, 0.070]	
2pm	-0.216	[-0.465, 0.033]	<0.05	-0.266	[-0.529, -0.002]	<0.01
3pm	-0.247	[-0.497, 0.004]		-0.307	[-0.572, -0.042]	
4am	-0.273	[-0.533, -0.014]		-0.332	[-0.606, -0.059]	
5pm	-0.289	[-0.561, -0.017]		-0.346	[-0.632, -0.060]	
6pm	-0.080	[-0.424, 0.264]		-0.138	[-0.495, 0.219]	
MidWeek	0.116	[0.068, 0.165]	< 0.01	0.118	[0.067, 0.169]	< 0.01
Physician	-0.167	[-0.224, -0.110]	< 0.01	-0.136	[-0.198, -0.074]	< 0.01
Young	-0.450	[-0.520, -0.380]	< 0.01	-0.436	[-0.510, -0.363]	< 0.01
Obs.	35094			32157		
Patient num.	4142			4038		
AIC	45175.93			41427.56		
LogLik	-22572.96			-20698.78		

Interpretation of Table 7: Patients who schedule visits during Tuesday, Wednesday and Thursday are more likely to show up with 12%($= e^{0.116} - 1$) higher in odds. This may be explained by the fact that people are busy catching up with work on Monday and planning for the weekend on Friday, and thus have more free time in the middle of the week. We also find that an appointment with physicians is less likely to be retained compared to a nurse practitioner's appointment (17%$= 1 - e^{-0.167}$ lower in odds). This is consistent with earlier findings that patients are in general more satisfied with care provided by nurse practitioners as they usually spend much more time with patients than physicians (Laurant et al. 2005). In addition, young patients are less likely to show up compared to their older counterparts (36%$= 1 - e^{-0.45}$ lower in odds), likely due to that younger people is more risk-taking and thus is more likely not to follow physicians' advice.

Interpretation of Table 8: In contrast to the US data, patients are more likely to show up ($34.9\% = e^{0.299} - 1$ higher in odds) in an appointment with physicians. This can be probably explained by the fact that Chile does not have nurse practitioners so patients spend most of their consultation time with physicians. Besides, show-up rate for initial visits is 158.6% ($= e^{0.95} - 1$) higher in odds compared to follow-up visits. This is consistent with the major concern that patients tend to think of follow-up no longer necessary once they feel better and thus not come for follow-up appointments (Lerner 2007).

Table 8 Estimated Regression Coefficients for Patient Show-up Probabilities (Chile Data).

	Est. Coef.	**95% CI**	**p-value**
Intercept	0.582	[0.387, 0.777]	N.A.
9am	0.298	[0.093, 0.504]	
10am	0.378	[0.155, 0.601]	
11am	0.578	[0.345, 0.810]	
12pm	0.472	[0.202, 0.742]	<0.01
14pm	0.409	[0.104, 0.714]	
15pm	0.031	[-0.272, 0.332]	
16pm	0.051	[-0.321, 0.422]	
Thursday	0.235	[0.092, 0.378]	<0.01
Friday	-0.371	[-0.554, -0.188]	
Physician	0.299	[0.023, 0.575]	< 0.05
Initial visit	0.950	[0.785, 1.114]	< 0.01
Observations	7281		
Patient #	3166		
AIC	7686.80		
LogLik	-3830.400		

Appendix B: Proofs of the Results

LEMMA 1. $Z_C \geq Z_P$

Proof of Lemma 1

Proof: For each feasible solution $\mathbf{x}(\tilde{\mathbf{b}})$ of Z_P, it is trivial to see that from $\mathbf{h}_i^\mathsf{T}\mathbf{x}(\tilde{\mathbf{b}}) = d_i$, we have $\mathbf{h}_i^\mathsf{T}\mathbf{x}(\tilde{\mathbf{b}})\mathbf{x}^\mathsf{T}(\tilde{\mathbf{b}})\mathbf{h}_i = d_i^2$, by taking expectation, we can derive

$$\mathbf{h}_i^\mathsf{T} p = d_i \quad \forall i \in I_2 = \{1, ..., m_2\}$$

$$\mathbf{h}_i^\mathsf{T} X \mathbf{h}_i = d_i^2 \quad \forall i \in I_2 = \{1, ..., m_2\}$$

As for constraint $A\mathbf{x}(\tilde{\mathbf{b}}) = \tilde{\mathbf{b}}$, note that it is equivalent to

$$\begin{aligned}(A\mathbf{x}(\tilde{\mathbf{b}}))^\mathsf{T} A\mathbf{x}(\tilde{\mathbf{b}}) - 2\tilde{\mathbf{b}}^\mathsf{T}(A\mathbf{x}(\tilde{\mathbf{b}})) + \tilde{\mathbf{b}}^\mathsf{T}\tilde{\mathbf{b}} &= (A\mathbf{x}(\tilde{\mathbf{b}}) - \tilde{\mathbf{b}})^\mathsf{T}(A\mathbf{x}(\tilde{\mathbf{b}}) - \tilde{\mathbf{b}}) \\ &= 0\end{aligned} \tag{16}$$

Rewriting (16) as

$$\begin{aligned}&trace((A^\mathsf{T} A\mathbf{x}(\tilde{\mathbf{b}}))\mathbf{x}(\tilde{\mathbf{b}})^\mathsf{T}) - 2trace((A\mathbf{x}(\tilde{\mathbf{b}}))\tilde{\mathbf{b}}^\mathsf{T}) + trace(\tilde{\mathbf{b}}\tilde{\mathbf{b}}^\mathsf{T}) \\ &= trace(A^\mathsf{T} A(\mathbf{x}(\tilde{\mathbf{b}})\mathbf{x}(\tilde{\mathbf{b}}))^\mathsf{T}) - 2trace(A(\mathbf{x}(\tilde{\mathbf{b}})\tilde{\mathbf{b}}^\mathsf{T})) + trace(\tilde{\mathbf{b}}\tilde{\mathbf{b}}^\mathsf{T}) \\ &= 0\end{aligned} \tag{17}$$

Hence $A\mathbf{x}(\tilde{\mathbf{b}}) = \tilde{\mathbf{b}}$ can be equivalently written as $(A^\mathsf{T} A) \bullet (\mathbf{x}(\tilde{\mathbf{b}})\mathbf{x}(\tilde{\mathbf{b}})^\mathsf{T}) - 2A^\mathsf{T} \bullet (\mathbf{x}(\tilde{\mathbf{b}})\tilde{\mathbf{b}}^\mathsf{T}) + \tilde{b} \bullet \tilde{b}^\top = 0$. By taking expectation we can get $(A^\top A) \bullet X - 2A^\top \bullet Y + I \bullet \Sigma = 0$. Note that the objective can be written as $K \bullet Y^\mathsf{T} + \mathbf{l}^\mathsf{T}\mathbf{p} = trace(K^\mathsf{T}(\tilde{\mathbf{b}}\mathbf{x}^\mathsf{T})) + \mathbf{l}^\mathsf{T}\mathbf{p} = trace((K^\mathsf{T}\tilde{\mathbf{b}})\mathbf{x}^\mathsf{T}) + \mathbf{l}^\mathsf{T}\mathbf{p} = \mathbf{x}^\mathsf{T}(K^\mathsf{T}\tilde{\mathbf{b}}) + \mathbf{l}^\mathsf{T}\mathbf{p} = E[\sum_{i=1}^{n} \mathbf{k}_i^\mathsf{T}\tilde{\mathbf{b}}x_i + \sum_{i=1}^{n} l_i x_i] = E[\sum_{i=1}^{n} c_i(\tilde{\mathbf{b}})x_i + \sum_{i=1}^{n} l_i x_i]$. Therefore it is clear that Z_C is a relaxation of Z_P. Hence $Z_C \geq Z_P$. Q.E.D.

LEMMA 2. *Let (Y,X) be a feasible solution to Z_C, and consider any completely positive decomposition of matrix*

$$\begin{pmatrix} \Sigma & Y^\mathsf{T} \\ Y & X \end{pmatrix} = \sum_{k\in\kappa} \begin{pmatrix} \boldsymbol{\beta}_k \\ \boldsymbol{\gamma}_k \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_k \\ \boldsymbol{\gamma}_k \end{pmatrix}^\mathsf{T} \qquad \boldsymbol{\beta}_k \in \mathcal{R}_+^{m_1}, \boldsymbol{\gamma}_k \in \mathcal{R}_+^n, \forall k \in \kappa,$$

then

$$A\boldsymbol{\gamma}_k = \boldsymbol{\beta}_k, \forall k \in \kappa.$$

Proof of Lemma 2

Proof: From the decomposition, we have

$$Y = \sum_{k\in\kappa} \boldsymbol{\gamma}_k \boldsymbol{\beta}_k^\mathsf{T}, \quad X = \sum_{k\in\kappa} \boldsymbol{\gamma}_k \boldsymbol{\gamma}_k^\mathsf{T}, \quad \Sigma = \sum_{k\in\kappa} \boldsymbol{\beta}_k \boldsymbol{\beta}_k^\mathsf{T}$$

Then from $(A^\mathsf{T}A) \bullet X = A^\mathsf{T} \bullet Y$, we have

$$(A^\mathsf{T}A) \bullet \sum_{k\in\kappa} \boldsymbol{\gamma}_k \boldsymbol{\gamma}_k^\mathsf{T} = A^\mathsf{T} \bullet \sum_{k\in\kappa} \boldsymbol{\gamma}_k \boldsymbol{\beta}_k^\mathsf{T}$$

which can be equivalently rewritten as:

$$\sum_{k\in\kappa} (A^\mathsf{T}A) \bullet \boldsymbol{\gamma}_k \boldsymbol{\gamma}_k^\mathsf{T} = \sum_{k\in\kappa} A^\mathsf{T} \bullet \boldsymbol{\gamma}_k \boldsymbol{\beta}_k^\mathsf{T}$$

Hence,

$$\sum_{k\in\kappa} (A\boldsymbol{\gamma}_k)^\mathsf{T} A\boldsymbol{\gamma}_k = \sum_{k\in\kappa} \boldsymbol{\beta}_k^\mathsf{T} (A\boldsymbol{\gamma}_k) \tag{18}$$

And similarly from $(A^\mathsf{T}A) \bullet X = I \bullet \Sigma$, we can derive

$$\sum_{k\in\kappa} (A^\mathsf{T}A) \bullet \boldsymbol{\gamma}_k \boldsymbol{\gamma}_k^\mathsf{T} = \sum_{k\in\kappa} \boldsymbol{\beta}_k \bullet \boldsymbol{\beta}_k^\mathsf{T}$$

Hence,

$$\sum_{k\in\kappa} (A\boldsymbol{\gamma}_k)^\mathsf{T} A\boldsymbol{\gamma}_k = \sum_{k\in\kappa} \boldsymbol{\beta}_k^\mathsf{T} \boldsymbol{\beta}_k \tag{19}$$

Combining (18) and (19), we get

$$\begin{aligned} &\sum_{k\in\kappa} (A\boldsymbol{\gamma}_k - \boldsymbol{\beta}_k)^\mathsf{T} (A\boldsymbol{\gamma}_k - \boldsymbol{\beta}_k) \\ &= \sum_{k\in\kappa} (A\boldsymbol{\gamma}_k)^\mathsf{T} A\boldsymbol{\gamma}_k - 2\sum_{k\in\kappa} \boldsymbol{\beta}_k^\mathsf{T} A\boldsymbol{\gamma}_k + \sum_{k\in\kappa} \boldsymbol{\beta}_k^\mathsf{T} \boldsymbol{\beta}_k = 0 \end{aligned}$$

As $(A\boldsymbol{\gamma}_k - \boldsymbol{\beta}_k)^\mathsf{T} (A\boldsymbol{\gamma}_k - \boldsymbol{\beta}_k) \geq 0$, we have for every $k \in \kappa$,

$$(A\boldsymbol{\gamma}_k - \boldsymbol{\beta}_k)^\mathsf{T} (A\boldsymbol{\gamma}_k - \boldsymbol{\beta}_k) = 0,$$

which implies $A\boldsymbol{\gamma}_k = \boldsymbol{\beta}_k, \forall k \in \kappa$. Q.E.D.

LEMMA 3. *(Natarajan et al. (2011)) Let $(\mathbf{p},X)$ be a feasible solution to Z_C, and consider any completely positive decomposition of matrix*

$$\begin{pmatrix} 1 & \mathbf{p}^\mathsf{T} \\ \mathbf{p} & X \end{pmatrix} = \sum_{k\in\kappa} \begin{pmatrix} \alpha_k \\ \boldsymbol{\gamma}_k \end{pmatrix} \begin{pmatrix} \alpha_k \\ \boldsymbol{\gamma}_k \end{pmatrix}^\mathsf{T}$$

where $\alpha_k \in \mathcal{R}_+$, $\boldsymbol{\gamma}_k \in \mathcal{R}_+^n, \forall k \in \kappa$, denote $\kappa_+ = \{k \in \kappa \mid \alpha_k > 0\}$, $\kappa_0 = \{k \in \kappa \mid \alpha_k = 0\}$, then (1) $H\frac{\boldsymbol{\gamma}_k}{\alpha_k} = \mathbf{d}, \forall k \in \kappa_+$; (2)$\boldsymbol{\gamma}_k = \mathbf{0}, \forall k \in \kappa_0$.

Proof of lemma 3

Proof: From the decomposition, we can rewrite the constraints

$$\begin{aligned} \mathbf{h}_i^\mathsf{T}\mathbf{p} &= d_i \quad \forall i=1,...,m_2 \\ \mathbf{h}_i^\mathsf{T} X \mathbf{h}_i &= d_i^2 \quad \forall i=1,...,m_2 \end{aligned}$$

as follows:

$$\begin{aligned} \mathbf{h}_i^\mathsf{T} \sum_{k\in\kappa} \alpha_k \boldsymbol{\gamma}_k &= d_i \\ \mathbf{h}_i^\mathsf{T} \sum_{k\in\kappa} \boldsymbol{\gamma}_k \boldsymbol{\gamma}_k^\mathsf{T} \mathbf{h}_i &= d_i^2 \end{aligned} \tag{20}$$

note that $\sum_{k\in\kappa} \alpha_k^2 = 1$, then based on (20), we have

$$(\sum_{k\in\kappa} \alpha_k^2)\sum_{k\in\kappa}(\mathbf{h}_i^\mathsf{T}\boldsymbol{\gamma}_k)^2 = (\sum_{k\in\kappa}\alpha_k \mathbf{h}_i^\mathsf{T}\boldsymbol{\gamma}_k)^2$$

By the equality condition of Cauchy-Schwartz inequality, $\exists\zeta$ such that $\zeta\alpha_k = \mathbf{h}_i^\mathsf{T}\boldsymbol{\gamma}_k$. Note $d_i = \sum_{k\in\kappa}\alpha_k \mathbf{h}_i^\mathsf{T}\boldsymbol{\gamma}_k = \sum_{k\in\kappa}\alpha_k\zeta\alpha_k = \zeta\sum_{k\in\kappa}\alpha_k^2 = \zeta$, so for $k\in\kappa_+$, we get $\mathbf{h}_i^\mathsf{T}\frac{\boldsymbol{\gamma}_k}{\alpha_k} = d_i, \forall i=1,...,m_2$, for $k\in\kappa_0$, we have $\mathbf{h}_i^\mathsf{T}\boldsymbol{\gamma}_k = 0, \forall i=1,...,m_2$. Based on Assumption 2, we can derive $\boldsymbol{\gamma}_k = \mathbf{0}$ for $k\in\kappa_0$. Q.E.D.

Proof of Theorem 1

Proof: We have shown Z_C is a relaxation of Z_P in Lemma 1. Let $\begin{pmatrix} 1 & \boldsymbol{\mu}^\mathsf{T} & \mathbf{p}^\mathsf{T} \\ \boldsymbol{\mu} & \Sigma & Y^\mathsf{T} \\ \mathbf{p} & Y & X \end{pmatrix}$ denote an optimal solution of Z_C, do the completely positive decomposition:

$$\begin{pmatrix} 1 & \boldsymbol{\mu}^\mathsf{T} & \mathbf{p}^\mathsf{T} \\ \boldsymbol{\mu} & \Sigma & Y^\mathsf{T} \\ \mathbf{p} & Y & X \end{pmatrix} = \sum_{k\in\kappa} \begin{pmatrix} \alpha_k \\ \boldsymbol{\beta}_k \\ \boldsymbol{\gamma}_k \end{pmatrix} \begin{pmatrix} \alpha_k \\ \boldsymbol{\beta}_k \\ \boldsymbol{\gamma}_k \end{pmatrix}^\mathsf{T}$$

and define $\kappa_+ := \{k\in\kappa \mid \alpha_k > 0\}$, $\kappa_0 := \{k\in\kappa \mid \alpha_k = 0\}$

From Lemma 2 and Lemma 3, we have

$$A\boldsymbol{\gamma}_k = \boldsymbol{\beta}_k, \quad \forall k\in\kappa \quad \text{and} \quad H\frac{\boldsymbol{\gamma}_k}{\alpha_k} = \mathbf{d}, \qquad \forall k\in\kappa_+, \qquad \boldsymbol{\gamma}_k = \mathbf{0}, \quad \forall k\in\kappa_0$$

It follows that $\boldsymbol{\beta}_k = \mathbf{0}, \forall k\in\kappa_0$ and $\frac{\boldsymbol{\gamma}_k}{\alpha_k}$ is a feasible solution to the original LP for all $k\in\kappa_+$.

The optimal solution of Z_C can be decomposed as

$$\begin{pmatrix} 1 & \boldsymbol{\mu}^\mathsf{T} & \mathbf{p}^\mathsf{T} \\ \boldsymbol{\mu} & \Sigma & Y^\mathsf{T} \\ \mathbf{p} & Y & X \end{pmatrix} = \sum_{k\in\kappa_+} \alpha_k^2 \begin{pmatrix} 1 \\ \frac{\boldsymbol{\beta}_k}{\alpha_k} \\ \frac{\boldsymbol{\gamma}_k}{\alpha_k} \end{pmatrix} \begin{pmatrix} 1 \\ \frac{\boldsymbol{\beta}_k}{\alpha_k} \\ \frac{\boldsymbol{\gamma}_k}{\alpha_k} \end{pmatrix}^\mathsf{T} \tag{21}$$

And for $k\in\kappa_+$, $\frac{\boldsymbol{\gamma}_k}{\alpha_k}$ satisfies all the constraints in Z_P.

Finally, we construct a distribution based on the decomposition (21). Define $P((\tilde{\mathbf{b}}^*, \mathbf{x}^*(\tilde{\mathbf{b}}^*)) = (\frac{\boldsymbol{\beta}_k}{\alpha_k}, \frac{\boldsymbol{\gamma}_k}{\alpha_k})) = \alpha_k^2, \forall k\in\kappa_+$, then $\mathbf{x}^*(\tilde{\mathbf{b}}^*)$ is feasible solution of Z_P.

First note that it is a valid distribution, because

$$\sum_{k\in\kappa_+} \alpha_k^2 = 1$$

Also,

$$\begin{aligned}
E[\tilde{\mathbf{b}}^*] &= \sum_{k\in\kappa_+} \tfrac{\boldsymbol{\beta}_k}{\alpha_k}\alpha_k^2 = \sum_{k\in\kappa_+} \alpha_k\boldsymbol{\beta}_k = \boldsymbol{\mu}\\
E[\tilde{\mathbf{b}}^*\tilde{\mathbf{b}}^{*\mathsf{T}}] &= \sum_{k\in\kappa_+} \tfrac{\boldsymbol{\beta}_k}{\alpha_k}\tfrac{\boldsymbol{\beta}_k^{\mathsf{T}}}{\alpha_k}\alpha_k^2 = \sum_{k\in\kappa_+} \boldsymbol{\beta}_k\boldsymbol{\beta}_k^{\mathsf{T}} = \Sigma\\
E[\mathbf{x}^*(\tilde{\mathbf{b}}^*)] &= \sum_{k\in\kappa_+} \tfrac{\boldsymbol{\gamma}_k}{\alpha_k}\alpha_k^2 = \sum_{k\in\kappa_+} \alpha_k\boldsymbol{\gamma}_k = \mathbf{p}\\
E[\mathbf{x}^*(\tilde{\mathbf{b}}^*)\mathbf{x}^*(\tilde{\mathbf{b}}^*)^{*\mathsf{T}}] &= \sum_{k\in\kappa_+} \tfrac{\boldsymbol{\gamma}_k}{\alpha_k}\tfrac{\boldsymbol{\gamma}_k{}^{\mathsf{T}}}{\alpha_k}\alpha_k^2 = \sum_{k\in\kappa_+} \boldsymbol{\gamma}_k\boldsymbol{\gamma}_k^{\mathsf{T}} = X\\
E[\mathbf{x}^*(\tilde{\mathbf{b}}^*)\tilde{\mathbf{b}}^{*\mathsf{T}}] &= \sum_{k\in\kappa_+} \tfrac{\boldsymbol{\gamma}_k}{\alpha_k}\tfrac{\boldsymbol{\beta}_k{}^{\mathsf{T}}}{\alpha_k}\alpha_k^2 = \sum_{k\in\kappa_+} \boldsymbol{\gamma}_k\boldsymbol{\beta}_k^{\mathsf{T}} = Y
\end{aligned}$$

Thus,

$$\begin{aligned}
Z_P &= \sup_{\tilde{\mathbf{b}}\sim(\boldsymbol{\mu},\Sigma)} E[Z_P(\tilde{\mathbf{b}})]\\
&\geq E[Z_P(\tilde{\mathbf{b}}^*)]\\
&\geq E[\mathbf{c}(\tilde{\mathbf{b}}^*)^{\mathsf{T}}\mathbf{x}^*(\tilde{\mathbf{b}}^*)]\\
&= K\bullet Y^{\mathsf{T}} + \mathbf{l}^{\mathsf{T}}\mathbf{p}\\
&= Z_C
\end{aligned}$$

i.e., $Z_C = Z_P$. Q.E.D.

Appendix C: Assumption Justification

In this part we illustrate how the redundant constraint $\sum_{i=1}^{m}\tilde{b}_i + s_l = m$ helps to ensure (6) satisfies Assumption 2.

Consider any completely positive decomposition of an optimal solution of $Z_C(\mathbf{s})$:

$$\begin{pmatrix} 1 & \boldsymbol{\mu}^{\mathsf{T}} & \mathbf{p}^{\mathsf{T}} & p_s\\ \boldsymbol{\mu} & \Sigma & Y^{\mathsf{T}} & \mathbf{y}_\mu\\ \mathbf{p} & Y & X & \mathbf{y}_\mathbf{x}\\ p_s & \mathbf{y}_\mu{}^{\mathsf{T}} & \mathbf{y}_\mathbf{x}{}^{\mathsf{T}} & ss \end{pmatrix} := \begin{pmatrix} 1 & \boldsymbol{\mu}^{\mathsf{T}} & \mathbf{p}_\mathbf{z}{}^{\mathsf{T}} & \mathbf{p}_\mathbf{y}{}^{\mathsf{T}} & p_s\\ \boldsymbol{\mu} & \Sigma & Y_z^{\mathsf{T}} & Y_y^{\mathsf{T}} & \mathbf{y}_\mu\\ \mathbf{p}_\mathbf{z} & Y_z & Z_{zz} & Z_{yz}^{\mathsf{T}} & \mathbf{y}_\mathbf{z}\\ \mathbf{p}_\mathbf{y} & Y_y & Z_{yz} & Z_{yy} & \mathbf{y}_\mathbf{y}\\ p_s & \mathbf{y}_\mu{}^{\mathsf{T}} & \mathbf{y}_\mathbf{z}{}^{\mathsf{T}} & \mathbf{y}_\mathbf{y}{}^{\mathsf{T}} & ss \end{pmatrix} = \sum_{k\in\kappa} \begin{pmatrix} \alpha_k\\ \boldsymbol{\beta}_k\\ \boldsymbol{\gamma}_k\\ \boldsymbol{\theta}_k\\ s_k \end{pmatrix}\begin{pmatrix} \alpha_k\\ \boldsymbol{\beta}_k\\ \boldsymbol{\gamma}_k\\ \boldsymbol{\theta}_k\\ s_k \end{pmatrix}^{\mathsf{T}}$$

where $\alpha_k \in \mathcal{R}_+, \boldsymbol{\beta}_k \in \mathcal{R}_+^m, \boldsymbol{\gamma}_k \in \mathcal{R}_+^{m+1}, \boldsymbol{\theta}_k \in \mathcal{R}_+^m, s_k \in \mathcal{R}_+$. Based on Lemma 2, the first two sets of constraints of $Z_C(\mathbf{s})$ guarantee that

$$A\begin{pmatrix} \boldsymbol{\gamma}_k\\ \boldsymbol{\theta}_k \end{pmatrix} = \mathbf{c}_\mathbf{w} \circ \boldsymbol{\beta}_k$$

Applying the same approach in Lemma 3, we can prove for $k \in \kappa_+$, the third and forth set of constraints imply $\gamma_{k0} - \theta_{k0} = 0$ and $\gamma_{k,m+1} + \theta_{km} = 0$, then $\gamma_{k,m+1} = 0, \theta_{km} = 0$ since $\boldsymbol{\gamma}_k \in \mathcal{R}_+^{m+1}, \boldsymbol{\theta}_k \in \mathcal{R}_+^m$. While it does not imply $\boldsymbol{\gamma}_k = \mathbf{0}, \boldsymbol{\theta}_k = \mathbf{0}, \forall k \in \kappa_0$. To ensure that, we rewrite the last two sets of constraints, which is the conic formulation of $\sum_{i=1}^{m}\tilde{b}_i + s_l = m$, based on the decomposition:

$$\begin{aligned}
\begin{pmatrix} \mathbf{1}_m^{\mathsf{T}} & 1 \end{pmatrix} \sum_{k\in\kappa} \alpha_k \begin{pmatrix} \boldsymbol{\beta}_k\\ s_k \end{pmatrix} &= m\\
\begin{pmatrix} \mathbf{1}_m^{\mathsf{T}} & 1 \end{pmatrix} \sum_{k\in\kappa} \begin{pmatrix} \boldsymbol{\beta}_k\\ s_k \end{pmatrix}\begin{pmatrix} \boldsymbol{\beta}_k\\ s_k \end{pmatrix}^{\mathsf{T}}\begin{pmatrix} \mathbf{1}_m\\ 1 \end{pmatrix} &= m^2
\end{aligned}$$

Similar to Lemma 3, by applying equality condition of Cathy-Schwartz inequality, we have

$$\begin{pmatrix} \mathbf{1}_m^{\mathsf{T}} & 1 \end{pmatrix} \begin{pmatrix} \frac{\boldsymbol{\beta}_k}{\alpha_k}\\ \frac{s_k}{\alpha_k} \end{pmatrix} = m, \forall k \in \kappa_+, \qquad \boldsymbol{\beta}_k = \mathbf{0}, s_k = 0, \forall k \in \kappa_0.$$

combined with $\theta_{km}=0, \forall k \in \kappa_0$, $\gamma_{k0}-\theta_{k0}=0$ and $A_1\begin{pmatrix}\boldsymbol{\gamma}_k\\ \boldsymbol{\theta}_k\end{pmatrix}=(\mathbf{c_w}\circ\boldsymbol{\beta}_k)\forall k\in\kappa$, we can prove

$$\boldsymbol{\gamma}_k=\mathbf{0}, \boldsymbol{\theta}_k=\mathbf{0}, \forall k\in\kappa_0$$

Q.E.D.

Appendix D: Iterative Method for Fixed Slot Length

Following the same procedure as the iterative method in the case of continuous slot length, we first use the average show-up probabilities over time, i.e. letting $(\boldsymbol{\mu}(\mathbf{s}),\Sigma(\mathbf{s}))=(\boldsymbol{\mu}_0,\Sigma_0)$, to solve (22) and obtain an optimal schedule $\mathbf{s}_0$ and the corresponding $\boldsymbol{\beta}_{01}$, Γ_{01}. And again refer the static schedule as $\mathbf{s}_0$. Note that the obtained $\mathbf{s}$ in the following completely positive cone represents the expected schedule obtained under different show-up scenario, which is fractional. To obtain a binary solution, we need to apply some rounding heuristic. We defer the discussion of rounding method at the end of this part.

$$\begin{aligned}
Z_D = \min\ & \alpha_0+\boldsymbol{\mu}(\mathbf{s})^\mathsf{T}\boldsymbol{\beta}_0+\Sigma(\mathbf{s})\bullet\Gamma_0+\mathbf{d}^\mathsf{T}\mathbf{w}^{(1)}+\mathbf{d}^\mathsf{T}Diag(\mathbf{w}^{(2)})\mathbf{d}+\eta_1 m+\eta_2 m^2+\mathbf{f}^\mathsf{T}\boldsymbol{\mu}(\mathbf{s})+c_0 \\
s.t.\quad & W-C(\mathbf{s}) && \geq_{co} 0\\
& \mathbf{1}_{m+1}^\mathsf{T}\mathbf{s} && = n\\
& \mathbf{1}_{m+1}^\mathsf{T}Z_{ss}\mathbf{1}_{m+1} && = n^2\\
& \begin{pmatrix}\mathbf{e}_i\\ \mathbf{e}_i\end{pmatrix}^\mathsf{T}\begin{pmatrix}\mathbf{s}\\ \mathbf{s_l}\end{pmatrix} && = 1,\quad \forall i\\
& \begin{pmatrix}\mathbf{e}_i\\ \mathbf{e}_i\end{pmatrix}^\mathsf{T}\begin{pmatrix}Z_{ss} & Y^\mathsf{T}\\ Y & Z_{ll}\end{pmatrix}\begin{pmatrix}\mathbf{e}_i\\ \mathbf{e}_i\end{pmatrix} && = 1,\quad \forall i\\
& diag(Z_{ss}) && = \mathbf{s}
\end{aligned} \tag{22}$$

Similarly, after fixing $(\boldsymbol{\beta}_0,\Gamma_0)=(\boldsymbol{\beta}_{01},\Gamma_{01})$, the next step is to obtain a new schedule $\mathbf{s}_1$ and the corresponding moments $(\boldsymbol{\mu}(\mathbf{s}_1),\Sigma(\mathbf{s}_1))$. In the case of $p(g(\mathbf{s}_1))=a+bg(\mathbf{s}_1), 0\leq g(\mathbf{s}_1)\leq n$, the function of the first moment on patient's arrival time can be written as $\boldsymbol{\mu}(\mathbf{s}_1)=a\mathbf{1}_m+bg(\mathbf{s}_1)$. And the second moment can be simply written as $\Sigma(\mathbf{s}_1)=\boldsymbol{\mu}(\mathbf{s}_1)\boldsymbol{\mu}(\mathbf{s}_1)^\mathsf{T}+Diag(\boldsymbol{\mu}(\mathbf{s}_1)\circ(\mathbf{1}_m-\boldsymbol{\mu}(\mathbf{s}_1)))$, where $Diag(\cdot)$ denote the operation that converts a vector to a diagonal matrix. Then we can obtain a new $\mathbf{s}$ and $\mathbf{g}$ by solving (23).

$$\begin{aligned}
Z_1(\boldsymbol{\beta}_{01},\Gamma_{01}) = \min\ & \alpha_0+a\mathbf{1}_m^\mathsf{T}\boldsymbol{\beta}_{01}+a^2(\mathbf{1}_m\mathbf{1}_m^\mathsf{T})\bullet(\Gamma_{01}-\Lambda(\Gamma_{01}))+b(\boldsymbol{\beta}_{01}^\mathsf{T}+a\mathbf{1}_m^\mathsf{T}\Gamma_{01}+a\mathbf{1}_m^\mathsf{T}\Gamma_{01}^\mathsf{T}\\
& -2adiag(\Gamma_{01})^\mathsf{T}+diag(\Gamma_{01})^\mathsf{T}+b\mathbf{f}^\mathsf{T})\hat{\mathbf{g}}+b^2(\Gamma_{01}-\Lambda(\Gamma_{01}))\bullet\hat{Z}_{gg}+adiag(\Gamma_{01})^\mathsf{T}\mathbf{1}_m\\
& +\mathbf{d}^\mathsf{T}\mathbf{w}^{(1)}+\mathbf{d}^\mathsf{T}Diag(\mathbf{w}^{(2)})\mathbf{d}+\eta_1 m+\eta_2 m^2+a\mathbf{f}^\mathsf{T}\mathbf{1}_m+c_0\\
s.t.\quad & W-C(\mathbf{s}) && \geq_{co} 0\\
& \begin{pmatrix}1 & \mathbf{s}^\mathsf{T} & \mathbf{g}^\mathsf{T} & \mathbf{s_l}^\mathsf{T}\\ \mathbf{s} & Z_{ss} & Z_{gs}^\mathsf{T} & Z_{ls}^\mathsf{T}\\ \mathbf{g} & Z_{gs} & Z_{gg} & Z_{lg}^\mathsf{T}\\ \mathbf{s_l} & Z_{ls} & Z_{lg} & Z_{ll}\end{pmatrix} && \geq_{cp} 0\\
& \begin{pmatrix} L & -I_{m+1}\end{pmatrix}\begin{pmatrix}\mathbf{s}\\ \mathbf{g}\end{pmatrix} && = \mathbf{0}\\
& diag(\begin{pmatrix} L & -I_{m+1}\end{pmatrix}\begin{pmatrix}Z_{ss} & Z_{gs}^\mathsf{T}\\ Z_{gs} & Z_{gg}\end{pmatrix}\begin{pmatrix} L & -I_{m+1}\end{pmatrix}^\mathsf{T}) && = \mathbf{0}\\
& \mathbf{1}_{m+1}^\mathsf{T}\mathbf{s} && = n\\
& \mathbf{1}_{m+1}^\mathsf{T}Z_{ss}\mathbf{1}_{m+1} && = n^2\\
& \begin{pmatrix}\mathbf{e}_i\\ \mathbf{e}_i\end{pmatrix}^\mathsf{T}\begin{pmatrix}\mathbf{s}\\ \mathbf{s_l}\end{pmatrix} && = 1,\quad \forall i\\
& \begin{pmatrix}\mathbf{e}_i\\ \mathbf{e}_i\end{pmatrix}^\mathsf{T}\begin{pmatrix}Z_{ss} & Z_{sl}^\mathsf{T}\\ Z_{sl} & Z_{ll}\end{pmatrix}\begin{pmatrix}\mathbf{e}_i\\ \mathbf{e}_i\end{pmatrix} && = 1,\quad \forall i\\
& diag(Z_{ss}) && = \mathbf{s}
\end{aligned} \tag{23}$$

Where $\hat{\mathbf{g}}$ denotes the arrival time for all the patients excluding the dummy one, $\hat{Z}_{gg}$ denotes the right lower $m \times m$ submatrix of Z_{gg}. Then we use the new obtained $\mathbf{s}$ to update the moment information.

D.1. Rounding Heuristic

Note that each completely positive decomposition of $\begin{pmatrix} 1 & \mathbf{s}^\intercal \\ \mathbf{s} & Z_s \end{pmatrix}$ represents the schedule under one patients show-up scenario. We then use spectral decomposition to approximate the completely positive decomposition and take each eigenvector to generate a schedule. This procedure results in $m+1$ fractional schedules. We then apply two rounding methods described below to get a set of binary schedules.

(1) Consecutive Rounding method

For each $i \in \{1, 2.., n\}$, if $\sum_{j=1}^{t-1} s_{0j} < i$ and $\sum_{j=1}^{t} s_{0j} \geq i$, then set $s_{0t} = 1$.

(2) Ranking based rounding method

1. Sort the fractional schedule $\mathbf{s}$ in an increasing order, denoted by $\left(s_{(1)}, ..., s_{(m)}\right)$.
2. Set the s_i corresponding to the $m-n$ smallest ones $\left(s_{(1)}, ..., s_{(m-n)}\right)$ to be 0, others to be 1.

Appendix E: Schedule Comparison with Zacharias and Pinedo (2014)

This appendix presents the schedules generated by our rounding heuristics (Table 9 and 10) and the optimal schedules given by Zacharias and Pinedo (2014)(Table 11 and 12). q represents patient no-show probability.

c_w				q	=	0.2							m
0.01	3	1	1	1	1	1	1	1	1	1	1	1	14
0.05	3	1	1	1	1	1	1	1	1	1	1	1	14
0.1	2	1	1	1	1	1	1	1	1	1	1	1	13
0.15	2	1	1	1	1	1	1	1	1	1	1	1	13
0.2	1	1	1	1	1	1	1	1	1	1	1	1	12
0.25	1	1	1	1	1	1	1	1	1	1	1	1	12
0.3	1	1	1	1	1	1	1	1	1	1	1	1	12
0.4	1	1	1	1	1	1	1	1	1	1	1	1	12
0.5	1	1	1	1	1	1	1	1	1	1	1	1	12
0.6	1	1	1	1	1	1	1	1	1	1	1	1	12
0.7	1	1	1	1	1	1	1	1	1	1	1	1	12

Table 9 Schedules Given by Rounding Heuristics under $q = 0.2$

w	q = 0.3												m	q = 0.4												m
0.01	5	1	1	1	1	1	1	1	1	1	1	1	16	7	1	1	1	1	1	1	1	1	1	1	1	18
0.05	3	1	1	2	1	2	1	1	1	1	1	1	16	4	1	1	1	1	1	3	1	1	1	1	1	17
0.1	3	1	1	2	1	1	1	1	1	1	1	1	15	4	1	1	1	2	1	1	1	1	1	1	1	16
0.15	2	1	1	1	1	2	1	1	1	1	1	1	14	2	1	1	1	2	2	1	1	1	1	1	1	15
0.2	2	1	1	1	1	2	1	1	1	1	1	1	14	2	2	1	1	1	1	2	1	1	1	1	1	15
0.25	2	1	1	1	1	1	1	1	1	1	1	1	13	2	2	1	1	1	1	2	1	1	1	1	1	15
0.3	2	1	1	1	1	1	1	1	1	1	1	1	13	2	1	1	1	1	2	1	1	1	1	1	1	14
0.4	1	1	1	1	1	1	1	1	1	1	1	1	12	2	1	1	1	1	1	1	1	1	1	1	1	13
0.5	1	1	1	1	1	1	1	1	1	1	1	1	12	2	1	1	1	1	1	1	1	1	1	1	1	13
0.6	1	1	1	1	1	1	1	1	1	1	1	1	12	2	1	1	1	1	1	1	1	1	1	1	1	13
0.7	1	1	1	1	1	1	1	1	1	1	1	1	12	1	1	1	1	1	1	1	1	1	1	1	1	12

Table 10 **Schedules Given by Rounding Heuristics under $q = 0.3$ and $q = 0.4$**

w	q = 0.2												m
0.01	3	1	1	1	1	1	1	1	1	1	1	1	14
0.05	2	1	2	1	1	1	1	1	1	1	1	1	14
0.1	2	1	1	1	1	1	1	1	1	1	1	1	13
0.15	2	1	1	1	1	1	1	1	1	1	1	1	13
0.2	2	1	1	1	1	1	1	1	1	1	1	1	13
0.25	1	1	1	1	1	1	1	1	1	1	1	1	12
0.3	1	1	1	1	1	1	1	1	1	1	1	1	12
0.4	1	1	1	1	1	1	1	1	1	1	1	1	12
0.5	1	1	1	1	1	1	1	1	1	1	1	1	12
0.6	1	1	1	1	1	1	1	1	1	1	1	1	12
0.7	1	1	1	1	1	1	1	1	1	1	1	1	12

Table 11 **Optimal Schedules under $q = 0.2$((Zacharias and Pinedo 2014))**

w	q = 0.3												m	q = 0.4												m
0.01	3	2	1	2	1	1	1	1	1	1	1	1	16	4	2	1	2	1	2	1	1	1	1	1	1	18
0.05	2	2	1	1	2	1	1	1	1	1	1	1	15	3	2	1	2	1	2	1	2	1	1	1	1	18
0.1	2	1	2	1	1	2	1	1	1	1	1	1	15	3	1	2	1	2	1	1	2	1	1	1	1	17
0.15	2	1	1	1	2	1	1	1	1	1	1	1	14	2	2	1	1	2	1	2	1	1	1	1	1	16
0.2	2	1	1	1	2	1	1	1	1	1	1	1	14	2	1	2	1	2	1	1	2	1	1	1	1	16
0.25	2	1	1	1	1	2	1	1	1	1	1	1	14	2	1	2	1	2	1	1	2	1	1	1	1	16
0.3	2	1	1	1	1	1	1	1	1	1	1	1	13	2	1	1	2	1	1	2	1	1	1	1	1	15
0.4	2	1	1	1	1	1	1	1	1	1	1	1	13	2	1	1	2	1	1	1	2	1	1	1	1	15
0.5	1	1	1	1	1	1	1	1	1	1	1	1	12	2	1	1	1	1	2	1	1	1	1	1	1	14
0.6	1	1	1	1	1	1	1	1	1	1	1	1	12	2	1	1	1	1	1	1	1	1	1	1	1	13
0.7	1	1	1	1	1	1	1	1	1	1	1	1	12	1	1	1	1	1	1	1	1	1	1	1	1	12

Table 12 **Optimal Schedules under $q = 0.3$ and $q = 0.4$((Zacharias and Pinedo 2014))**