2-2019

Effects of rescheduling on patient no-show behavior in outpatient clinics

Jiayi LIU
University of Science and Technology of China

Jingui XIE
University of Science and Technology of China

Kum Khiong YANG
Singapore Management University, kkyang@smu.edu.sg

Zhichao ZHENG
Singapore Management University, DANIELZHENG@smu.edu.sg
DOI: https://doi.org/10.1287/msom.2018.0724

Effects of Rescheduling on Patient No-show Behavior in Outpatient Clinics

Jiayi Liu
School of Management, University of Science and Technology of China, Hefei, An Hui, China 230026, ljy723@mail.ustc.edu.cn

Jingui Xie
School of Management, University of Science and Technology of China, Hefei, An Hui, China 230026, xiej@ustc.edu.cn

Kum Khiong Yang
Lee Kong Chian School of Business, Singapore Management University, Singapore 178899, kkyang@smu.edu.sg

Zhichao Zheng
Lee Kong Chian School of Business, Singapore Management University, Singapore 178899, danielzheng@smu.edu.sg

We study the effects of rescheduling on no-show behavior in an outpatient appointment system for both new and follow-up patients. Previous literature has primarily focused on new patients and investigated the role of waiting time on no-show probability. We offer a more nuanced understanding of this costly phenomenon. Using comprehensive clinical data, we demonstrate that for follow-up patients, their no-show probability decreases by 10.9 percentage points if their appointments were rescheduled at their own request, but increases by 6.2 percentage points if they were rescheduled by the clinic. New patients, in contrast, are less sensitive to who initiates rescheduling. Their no-show probability decreases by 2.3 percentage points if their appointments were rescheduled at their own request, and increases by 3.2 percentage points—but is statistically insignificant at the 10% level—if they were rescheduled by the clinic. New patients are more concerned about waiting time compared to follow-up patients. For patients whose appointments were not rescheduled, new patients' no show probability decreases by 1.3 percentage points if their waiting time is reduced by one week, but the waiting time has a small and statistically insignificant effect on follow-up patients' no show probability. Using data-driven simulation, we conduct counterfactual investigation of the impact of allowing active rescheduling on the performance of appointment systems. In particular, allowing the flexibility of patient rescheduling can reduce the overall no-show rate and increase system utilization, but at a cost of increased wait time for new patients. If patients are able to reschedule at least one week in advance, new patients' wait time is largely reduced, while the no-show rate remains the same; this is equivalent to the effect of a 5% increase in clinic's capacity.

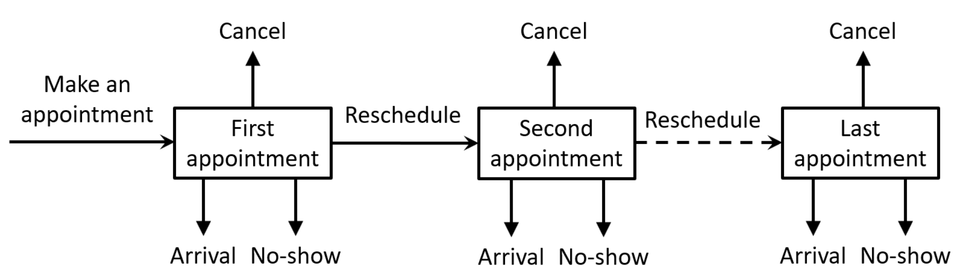
Key words: Appointment scheduling; rescheduling; no-show; econometrics; data-driven simulation

1. Introduction

Enhancing the patient experience is a priority for efforts to improve health care. To achieve this goal, one important component is to give patients more flexibility to schedule or reschedule their appointments in appointment systems (Feldman et al. 2014, Liu et al. 2017). On the negative side, offering patients the flexibility of rescheduling may lead to idle physician time if rescheduled appointments cannot be filled by other patients; it also increases the complexity of the operation. To the best of our knowledge, few studies have examined the effects of rescheduling on the performance of appointment systems. We aim to fill this gap by empirically quantifying these effects.

In this paper, we term appointments that are rescheduled by patients *active rescheduling*. An appointment can also be rescheduled by the clinic, which we term *passive rescheduling*; these are usually due to physicians' being on leave (for study or conferences) or to an unexpected increase in emergency cases. As a result, a patient's ultimate attendance may follow a series of (re)scheduling decisions, rather than a static process. As shown in Figure 1, a patient may arrive, cancel, reschedule, or simply not show up for an appointment (known as "no-show").

Figure 1 Appointment process



No-show is a key indicator of the effectiveness of planning and scheduling for most healthcare appointment systems. No-show has far-reaching effects on clinic efficiency and healthcare cost: Efficiency suffers because physicians are underutilized, and patients have to wait an unnecessarily long time for appointments. For example, Pesata et al. (1999) estimated a loss of more than one million dollars resulting from 14,000 missed appointments in a pediatric practice, and Moore et al. (2001) found that no-shows wasted 25.4% of scheduled time in a family medicine practice and cost clinics 3% to 14% of their anticipated daily revenue.

An important distinction between this paper and previous empirical literature on patient no-show behavior is that we consider both new and follow-up patients. New patients, who visit the clinics for the first time with new conditions, typically desire timely access to service. Many clinics also include the average waiting time of new patients as a key performance indicator for their service quality. On the other hand, follow-up patients, also known as repeat or return patients, usually visit the clinic for follow-up treatment or consultation after previous visits. Their conditions are relatively more stable and the schedule for their visits is more flexible. Therefore, follow-up patients may care more about convenience or preferred appointment times. Such non-negligible differences between these two patient groups may lead to distinctive behaviors and require separate analysis. Specifically, follow-up patients should be granted more attention, in light of a more affluent and aging population that demands better and more frequent follow-up care (WHO 2011, 2014).

In a nutshell, we aim to *quantify the effect of active and passive rescheduling on patients' no-show behavior, for both new and follow-up patients.*

We collect a comprehensive data set consisting of 749,880 scheduled appointments over a 3-year period (Jan 2011–Dec 2013) for a group of outpatient specialty clinics in Singapore, which includes diabetes, dermatology, colorectal surgery, and oncology. In these clinics, patients are required to schedule their appointments ahead of time via phone, text messages, or the online system.

At the patient level, we show that the two types of patients respond to rescheduling and waiting time in notably different ways. For follow-up patients, no-show probability decreases by 10.9 percentage points if the appointment has been rescheduled actively, and increases by 6.2 percentage points if it has been rescheduled passively. In contrast, average no-show probability across follow-up patients whose appointments were not rescheduled is 27.8%, which suggests that the impact of rescheduling is significant. For new patients, no-show probability decreases by 2.3 percentage points if the appointment has been rescheduled actively, and increases by 3.2 percentage points—but is statistically insignificant at the 10% level—if it has been rescheduled passively. Compared to follow-up patients, new patients are much less sensitive to rescheduling, but sensitive to waiting time. For patients whose appointments were not rescheduled, new patients' no show probability decreases by 1.3 percentage points if their waiting time is reduced by one week. In contrast, waiting time has a small and statistically insignificant effect on follow-up patients' no show probability.

At the system level, we conduct counterfactual analysis using data-driven simulation to assess the impact of offering patients the flexibility of active rescheduling on the performance of appointment systems. We simulate the operation of a median-size clinic in our data set, and find that allowing active rescheduling would reduce the no-show rate by 11.02% (from 21.69% to 19.30%), and increase system utilization from 85.66% to 87.32%. On the negative side, allowing active rescheduling increases the average wait time for new patients by 6.67 business days. However, if patients reschedule at least one week before the appointment day, much of the downside is averted, with the no-show rate decreasing to 19.14%, system utilization increasing to 87.84%, and the average wait time for new patients increasing by only 1.56 business days. This is similar to the effects of a 5% increase in the clinic's capacity, with the no-show rate decreasing to 19.46%, and the average wait time for new patients increasing by 1.11 business days. Solely increasing the system's capacity, however, results in lower system utilization of 83.81%. Furthermore, when the percentage of follow-up patients increases, the positive impact of early active rescheduling becomes even more significant.

The rest of the paper is organized as follows. We review the relevant literature in Section 2. Next, we develop our hypotheses in Section 3, and describe our data in Section 4. We then introduce our econometric model and present the empirical results in Section 5. Finally, we discuss the implications of our findings through data-driven simulation and counterfactual analysis in Section 6 and conclude in Section 7.

2. Literature Review

Our work draws on three streams of research: (i) empirical studies on no-show behavior, (ii) appointment scheduling in operations management, and (iii) econometric methodology.

An extensive empirical literature explores factors that correlate with patients' no-show behavior, these studies typically conduct statistical regressions on historical data to identify important "predictors" (e.g., Deyo and Inui 1980, Daggy et al. 2010, Gupta and Wang 2012, Norris et al. 2014). Such studies have found that younger patients are less likely to keep appointments (Bean and Talaga 1992), Medicaid patients may miss appointments due to transportation difficulties (Lowes 2005), and bad weather is correlated with lower attendance (Morse et al. 1984). In particular, waiting time has commonly been cited as an important and controllable factor that affects no-shows: An extensive stream of literature demonstrated the negative relationship between waiting time and patient attendance (e.g., Bean and Talaga 1995, Bodenheimer and Pham 2010, Compton et al. 2006, Daggy et al. 2010, Gallucci et al. 2005, Grunebaum et al. 1996, Liu et al. 2010). A notable example of this stream of literature is Osadchiy and KC (2017), who argue that patients who choose to reschedule an appointment may have an inherently different sensitivity to waiting than those who do not, and therefore the effect of waiting time may be underestimated. The authors introduce an unobservable variable—willingness to wait, and use a nonparametric method to tackle this sample-selection problem. Their findings demonstrate an even more important role of waiting time: On average, a one-day reduction in waiting time increases throughput by 5.7%.

Some studies have used interviews to investigate self-reported reasons for no-show (e.g., Campbell et al. 2000, Gany et al. 2011, Martin et al. 2005). For example, Lacy et al. (2004) explore the emotional aspects of patients' no-show through semistructured interviews, and identified three major reasons for not showing up: discomfort experienced during the appointment, the perception that the health care system disrespects their time and beliefs, and misconceptions about the consequences of missed appointments. Other common self-reported reasons include forgetting the appointment (Campbell et al. 2000, Neal et al. 2005); competing priorities or conflicts (Campbell et al. 2000, Martin et al. 2005, Gany et al. 2011); and feeling better and no longer needing the appointment (Corfield et al. 2008).

A stream of literature in operations research and operations management has been devoted to an approach to system design in appointment scheduling that explicitly considers patient no-show (e.g., Hassin and Mendel 2008, Liu et al. 2010, Liu and Ziya 2014). Green and Savin (2008) propose a queuing model to quantify the effect of waiting time on no-show and examine the implications of patient patience on the panel size. LaGanga and Lawrence (2012) and Muthuraman and Lawley (2008) investigate the use of overbooking to compensate for patient no-shows in an appointment system and, in turn, maximize capacity utilization and patient service. Other studies of appointment scheduling (Gupta and Denton 2008, Wang and Gupta 2011, Feldman et al. 2014) have considered patients' preferences for when they would like to be seen. Another approach gaining popularity in many primary clinics is the "open-access" scheduling policy, under which each patient is offered an appointment on the same day he or she calls; research in this vein (Murray and

Tantau 2000, Williams et al. 2008, Cameron et al. 2010) suggests that clinics' operational efficiency will be enhanced, in the sense that an open-access system allows patients to be seen sooner, and can thus virtually eliminate no-shows. However, because of the stochastic nature of daily demand, it is likely that providers will have to either work overtime, squeeze in additional appointments during the day, or delegate some tasks to nurses when faced with high demand. This, in turn, will cause the clinic to incur additional costs and/or reduce the quality of services provided (Robinson and Chen 2010). Therefore, the effectiveness of an open-access policy in different environments is debatable. In summary, these scheduling policies can be seen as mechanisms to mitigate the negative impact caused by no-shows.

Clinics also engage in other practices aimed at reducing no-shows; these include sending appointment reminders by mail or text message, financial incentives such as transport vouchers, and exit-interview education. While some applaud the effectiveness of these interventions (Macharia et al. 1992, Guy et al. 2012), critics argue that they only modestly reduce no-show (Bean and Talaga 1992, George and Rubin 2003, Johnson et al. 2007). Moreover, the effects of these strategies appear to vary significantly by patient population and baseline no-show rates (Hashim et al. 2001, Geraghty et al. 2008), and some can be costly and require intensive management, such as operating a call center.

Our study contributes to the above literature by (i) identifying and quantifying an important operational factor—rescheduling—on patient no-show behavior, which not only enhances our understanding of this costly phenomenon but also allows researchers to enrich operations research models in the design of appointment scheduling systems; and (ii) establishing a connection between no-show behavior and the dynamic appointment-rescheduling process, which offers important insights for policies that aim to reduce no-shows.

To evaluate the impact of rescheduling on patient no-show behavior, one challenge is that some unobserved factors may simultaneously affect the decision to reschedule and patients' no-show probability. This endogeneity complicates our attempt to identify the causal effect of rescheduling on no-show probability. A standard approach to deal with endogeneity is to find valid and relevant instrumental variable(s) (KC and Terwiesch 2011, 2012, Freeman et al. 2016, Freeman and Scholtes 2017, Hu et al. 2017, Chan et al. 2017). We adopt a *recursive bivariate probit model* that allows us to estimate of the treatment effect a binary endogenous variable has on binary outcomes in the presence of unobservables (Maddala

1983). Kim et al. (2014) recently use this model to measure the impact of ICU admission on patient outcomes, and Freeman et al. (2016) to examine the effect of workload during a service episode on gatekeepers' behavior choices. In addition, we use *propensity score matching* and *near-far matching* to conduct robustness checks (Sudhir and Talukdar 2015, Baiocchi et al. 2010, Lu et al. 2011, Lorch et al. 2012, Yang et al. 2014, Hu et al. 2017).

3. Hypothesis Development

In this section, we first provide formal definitions of different waiting times and then develop our hypotheses.

3.1. Definition of waiting times

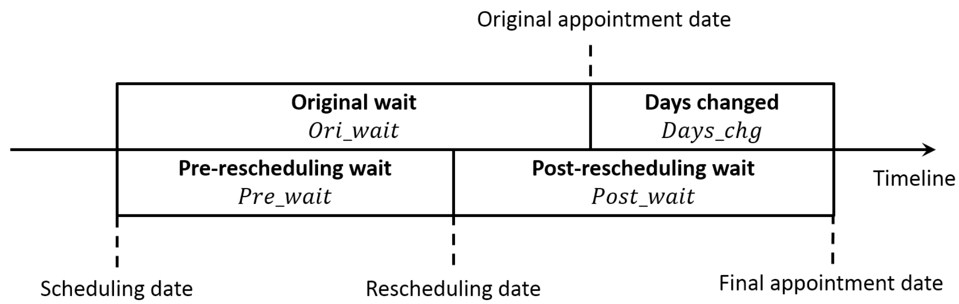
We define four specifications for waiting time, as follows. First, the time interval between the *scheduling date* when the appointment is first requested and the *appointment date* when the appointment is first due is called **original wait** (Ori_wait). As a follow-up appointment in our sample is scheduled at the end of the previous visit, the original wait of a follow-up patient is equivalent to our measure of **treatment cycle**. Note that the original wait is immediately known when an appointment is made, but it will differ from the realized waiting time if the appointment is rescheduled. Since our primary goal is to explore the impact of rescheduling on patients' no-show behavior, we focus on appointments that are rescheduled exactly once. We define the time interval between the *final appointment date* and the *original appointment date* as **days changed** ($Days\_chg$), which is positive if the original appointment is moved back and negative if it is moved forward.

In addition, rescheduling divides the total waiting time into two parts—before and after the *rescheduling date*. We define **pre-rescheduling wait** (Pre_wait) as the time the patient has already waited before rescheduling, and **post-rescheduling wait** ($Post\_wait$) as the time the patient must still wait after rescheduling until the final appointment date is due. Figure 2 illustrates these four different waits for a patient whose appointment is rescheduled once. The original appointment date is created on scheduling date, and the final appointment date is created on rescheduling date.

3.2. Hypotheses

Next, we develop our hypotheses regarding the effects of waiting time and rescheduling on patients' no-show behavior.

Figure 2 **Definitions of different waits**



3.2.1. Waiting time The effect of waiting time on no-shows has been studied extensively in the literature. Psychologically, "wait" is shown to be associated with monetary equivalent disutility (Leclerc et al. 1995); stress (Suck and Holling 1997); and increasing "pain" (Janakiraman et al. 2011). Several studies have examined how customers behave in the presence of queues in various settings, such as fast-food outlets (Allon et al. 2011); call centers (Akşin et al. 2013); and hospital emergency departments (Batt and Terwiesch 2015). In all situations, individuals display an aversion to waiting.

In an outpatient appointment setting, Gallucci et al. (2005) identified waiting time as a key driver of no-shows. Longer waits may increase the probability that a patient will recover before the appointment, make an appointment with another provider, or simply forget the appointment (Bodenheimer and Pham 2010). Daggy et al. (2010) found that appointments scheduled more than two weeks in advance were more than twice as likely to become no-show. Compton et al. (2006) estimated that no-show probability for the first appointment after psychiatric hospitalization increased by 4% for each day of wait. Similar results have been reported in Bean and Talaga (1995), Grunebaum et al. (1996), Gallucci et al. (2005), Liu et al. (2010), etc.

Based on these studies, we hypothesize that patients with longer waiting times are less likely to show up for their appointments. Note that in previous studies, "waiting time" refers to the *original wait* for non-rescheduled appointments, while in our context it refers to *post-rescheduling wait* for rescheduled appointments[1]. Since we consider both types of waiting time, we have Hypotheses 1A and 1B. Meanwhile, depending on whether a rescheduled appointment has been moved forward or back, the actual waiting time can be shortened or prolonged, which translates into lower or higher no-show probability, respectively. This leads to Hypothesis 1C.

[1] As discussed earlier, previous studies have treated rescheduled appointments as new ones.

Hypothesis 1A. *No-show probability increases with original wait.*

Hypothesis 1B. *No-show probability increases with post-rescheduling wait.*

Hypothesis 1C. *No-show probability decreases if the appointment is moved forward and increases if it is moved back.*

3.2.2. Rescheduling Patients have preferences regarding their appointment times (Ryan and Farrar 2000, Rubin et al. 2006, Cheraghi-Sohi et al. 2008, Hole 2008). A stream of literature has conducted discrete choice experiments to examine patient preference and choice behavior, and found that although patients prefer appointments that are sooner rather than later, the preference for shorter waiting time can be outweighed, in some cases, by a more convenient day or time (Cheraghi-Sohi et al. 2008, Feldman et al. 2014, Liu et al. 2017). For example, some people prefer a particular day of the week, while others might book appointments for a specific time of day, even at the cost of extra waiting.

Originally booked appointment times do not always reflect patients' true preferences, because the desired slots may not be available (Wang and Gupta 2011). As a result, some patients may choose to reschedule their appointments when more desirable slots become available. Meanwhile, even if some patients are able to initially book appointments for their preferred times, this may be outweighed by work or family commitments that arise and conflict with their appointments. According to patients' self-report, competing priorities or schedule conflicts are among the most frequently cited factors that discourage them from attending their appointments (Campbell et al. 2000, Martin et al. 2005, Gany et al. 2011). Medical staff also report that patients' non-attendance is mainly due to "patient" factors, such as their daily schedules (Martin et al. 2005). Therefore, if patients can reschedule their appointments to more suitable times, we expect that no-show risk would largely be curbed. This suggests that active rescheduling allows patients to update their appointments to more preferred times, which could significantly reduce the likelihood of no-show.

In contrast, appointments rescheduled by clinics are mainly the results of physicians' unplanned leave or an increase in urgent cases. In general, if a physician is unavailable for a specific period of time, all of the appointments scheduled for that time period must be rescheduled by the clinic. Since it is hard to accommodate all patients' preferences in the rescheduling process, some patients may have their appointments moved to times or days they are inherently less able to attend. Further, they may feel that their needs are not being adequately served, which may also reduce their likelihood to attend.

Hypothesis 2A. *No-show probability decreases when the appointment is actively rescheduled.*

Hypothesis 2B. *No-show probability increases when the appointment is passively rescheduled.*

4. Data

In this section we describe our data set and present a descriptive analysis, followed by discussion of our data-selection process.

4.1. Descriptive Statistics

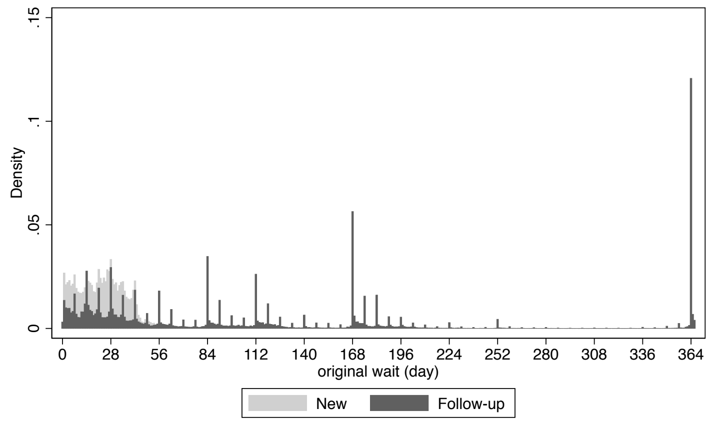
We employ a large data set consisting of 778,441 scheduled appointments for a group of specialty clinics in Singapore over a course of 3 years. The data set includes the date that each appointment was created, as well as the date and time of the actual appointment. We can also observe various sources of patient-level heterogeneity that may influence patient behavior, including age, race, nationality, religion, government-subsidy status and visit type. Clinical information includes clinic code, specialty, and source of referral. Exogenous factors are also considered; for instance, travel distance is calculated by the great-circle distance between the patient's residence and the clinic, with the longitude and latitude of each derived from the postal code. In the data set, arrival (A), no-show (NS), cancellation (C), and rescheduling (R) account for 45%, 17%, 9%, and 29% of all appointments, respectively. Summary statistics are presented in Table 1.

Table 1 Basic description of the data set

Visit Type	%	Nationality	%	Subsidy	%	Race	%
New	27.1	Singaporean	79.8	Sub	48.1	Chinese	67.5
Follow-up	72.9	Foreigner	20.2	Non-Sub	51.9	Malay	12.3
						Indian	10.9
						Others	9.3
	Mean	St.Dev	Median				
Age (years)	33.0	18.8	35.0				
Distance (miles)	9.8	4.3	10.4				
Original wait (days)	97.5	118.8	43.0				

Specifically, the distribution of original wait is much more widely dispersed for follow-up patients than for new patients (see Figure 3). For follow-up patients, a strong periodic pattern is evident, with peaks at multiples of 7 days: 28, 84, 168, and 364 days (i.e., roughly

Figure 3 **Original wait for new and follow-up patients**



1 month, 3 months, 6 months, and 1 year); these are the most common cycles, and account for 3%, 4%, 6%, and 12% of all follow-up appointments, respectively.

We now describe the basic characteristics of rescheduling. As shown in Figure 4(a), patients tend to reschedule when the original appointments are imminent: Among all the actively rescheduled appointments, 65% were rescheduled within 2 weeks of their original appointment date. In contrast, if the appointments were rescheduled by the clinic, no obvious patterns are found. In addition, Figure 5 shows the time interval between the new appointment date and the original appointment date. For the actively rescheduled appointments, 20% were moved forward, 73% moved back, and 7% rescheduled to a different time slot within the original appointment date. An interesting feature of Figure 5(a) is the periodic pattern, with peaks at multiples of 7 days, which can be explained by the fact that many physicians are available for consultation only on a specific day of the week; many patients may also prefer a specific day of the week due to their schedules. As for rescheduled appointments initiated by the clinic, half were rescheduled to a different time slot within the original appointment date, 12% were moved forward, and 38% were moved back.

Regarding the relationship between rescheduling and waiting time, it seems plausible that the longer the waiting time, the greater the chance that an appointment will be rescheduled by either the patient or the clinic. However, we observe in Figure 6 that the likelihood of active rescheduling appears to be much less relevant to waiting time, compared with the increasing trend shown in passive rescheduling. This is probably because many patients choose to reschedule only when conflicts between their personal schedules and

Figure 4 When does rescheduling occur? Distribution of time intervals between rescheduling date and original appointment date

(a) Active rescheduling **(b) Passive rescheduling**

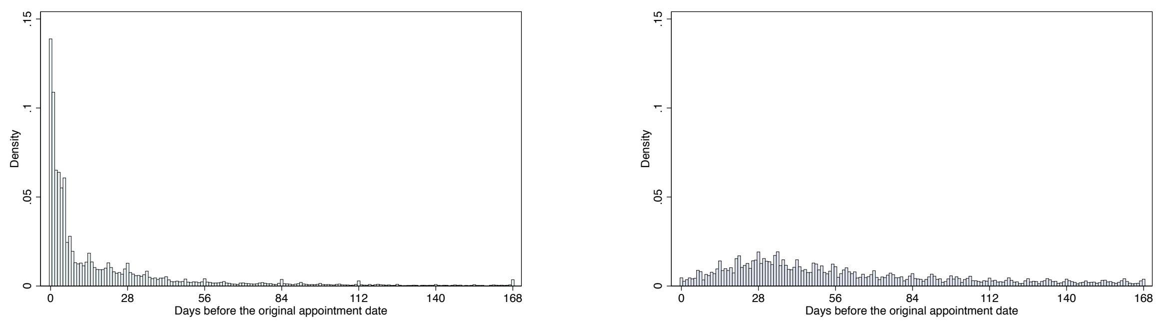
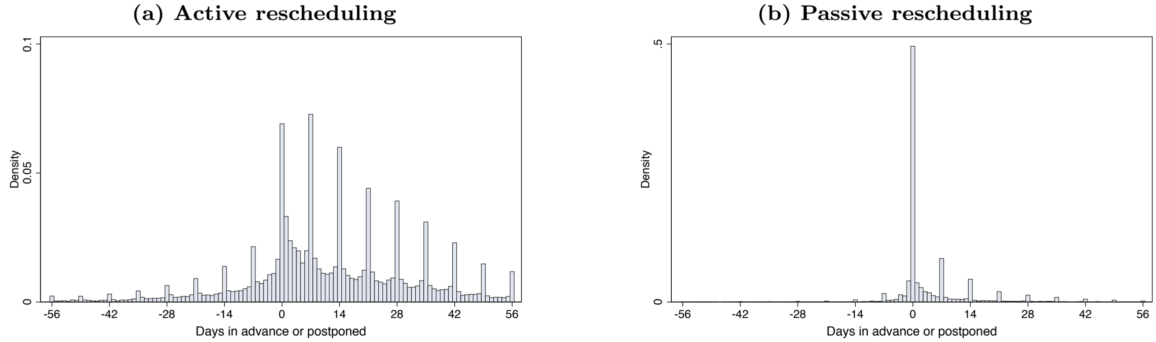


Figure 5 Are appointments moved forward or back? Distribution of time intervals between new appointment date and original appointment date

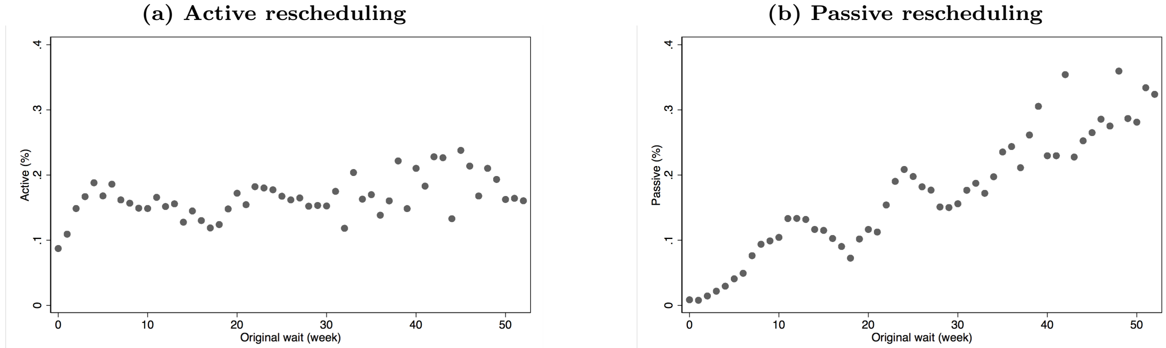
(a) Active rescheduling **(b) Passive rescheduling**



original appointment dates become real and imminent, regardless of the waiting time (see Figure 4a).

Figure 6 Rescheduling and waiting time

(a) Active rescheduling **(b) Passive rescheduling**



4.2. Data Selection

Figure 7 illustrates our data-selection process. We first eliminate appointments for anonymous patients, which reduces our sample size by 3.7%. In addition, some appointments are cancelled directly without booking a new appointment. As our main concern is whether the patient will show up, we only include appointment records with arrival or no-show as the final status. Therefore, cancellations, as well as appointments that lack rescheduling information, are excluded from our analysis. In the remaining sample, 86% of follow-up appointments are made at the end of the current appointment, based on the treatment cycle suggested by the physician. Follow-up appointments scheduled at a later time are usually open appointments following the last visit[2], or simply the result of uncertainty about the physician's or clinic's schedule. As the data does not record the reasons for later rescheduling, we remove these later-scheduled appointments from our analysis[3].

For each of the remaining patients, we can observe their appointments in time sequence. For instance, the records for patient ID 6 in Table 2 display a typical pattern for a follow-up patient's attendance at appointments—i.e., R–R–A, which represents a series of two rescheduled appointments before final attendance, and R–A, which represents final attendance with only one rescheduling. Sequential appointments (represented by R–A, R–R–A, etc.) constitute sets of appointments in which the last appointment of each set (A or NS) is termed the **final appointment**, and the first appointment rescheduled in a set is the **original appointment** (recall Figure 2). Some appointments have not been rescheduled (such as appointments for patient ID 89757), which are termed **direct appointments** (or **non-rescheduled appointments**). We only keep direct appointments and final appointments, along with information about the rescheduling process. Finally, to isolate the impact of rescheduling on patient no-show behavior, we exclude appointments that have been rescheduled more than once and limit our observations to appointments that are not rescheduled or rescheduled only once.

[2] Open appointments refer to appointments without a specific date or time that clinics give to the patient in case he or she needs additional health care services, usually due to unforeseen issues related to the health condition being treated. A typical example of an open appointment is when the patient is asked to contact the clinic within 3 months after the last treatment or consultation if necessary, and the patient will be given higher priority when scheduling such an appointment.

[3] These include appointments made by the same patients for the same specialty clinics within one week from last no-shows. We remove them from our analysis since we cannot tell whether they are new appointments or rescheduled ones for the no-show appointments. Nevertheless, the no-show rates for these appointments are higher than newly scheduled appointments for both new and follow-up patients. Therefore, by not treating these appointments as direct appointments (to be defined later), our estimated impact of active rescheduling on no-show is more conservative.

Figure 7 Data selection process

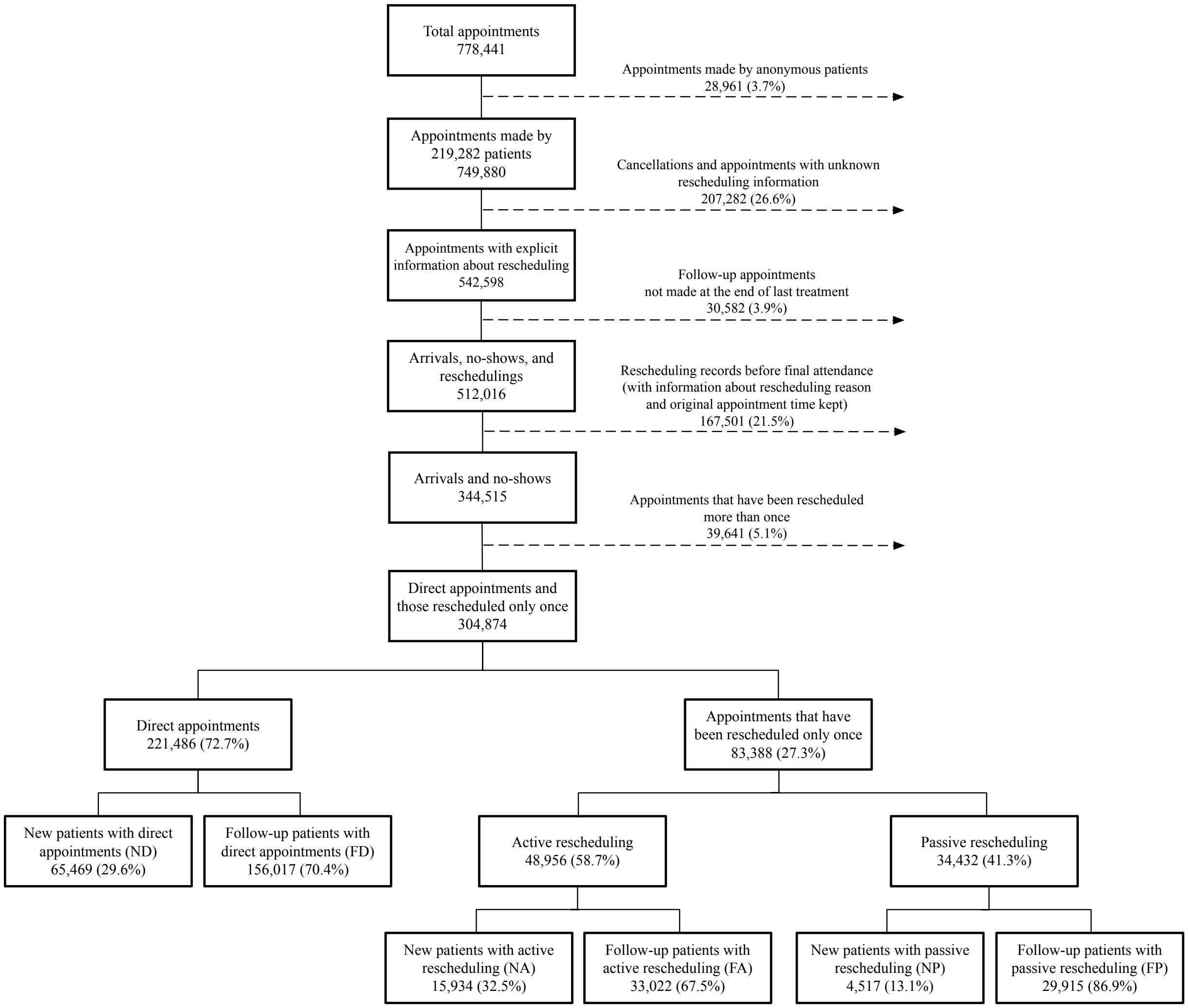


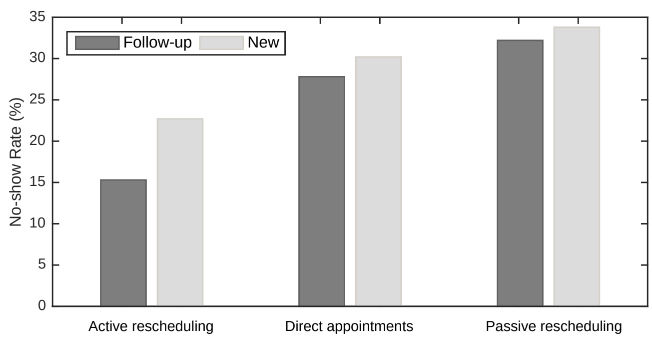
Table 2 Records of two typical patients

Patient ID	Appointment created date	Appointment date	Attendance status	Rescheduled reason	Attribute
6	14-Dec-2011	15-May-2012	R	Patients' request	Original app
6	6-May-2012	22-May-2012	R	Patients' request	.
6	16-May-2012	21-May-2012	A	n.a.	Final app
6	21-May-2012	20-Nov-2012	R	Doc's leave	Original app
6	13-Sep-2012	20-Nov-2012	A	n.a.	Final app
89757	21-Sep-2011	22-Feb-2012	A	n.a.	Direct app
89757	22-Feb-2012	23-May-2012	A	n.a.	Direct app
89757	23-May-2012	23-Jul-2012	NS	n.a.	Direct app

Note: "R" refers to rescheduling, "A" and "NS" refers to arrival and no-show respectively.

Our final data set consists of 304,874 appointments, with 72.7% being direct appointments and 27.3% rescheduled appointments. Of the latter, 58.3% are actively rescheduled and 41.3% passively rescheduled. The numbers of observations in each subsample are given in the bottom of Figure 7, and Figure 8 depicts the no-show rate for each subsample. However, it is important to note that all appointments excluded from the analysis sample are still included in our estimation of capacity, which we will discuss in the following section.

Figure 8 **No-show rate of each subsample**



5. Econometric Analysis

In this section, we construct our econometric models to test the hypotheses proposed in Section 3. We address the hypotheses on waiting time first, followed by the hypotheses on rescheduling.

5.1. Waiting Time

5.1.1. Econometric Model

As the observed outcome variable is whether a scheduled patient shows up, this requires a model of binary choice such as the logit, probit, skewed logit, or complimentary log log (Nagler 1994, Greene 2012). Selecting the best binary-choice model a *priori* is difficult, as each has its own theoretical or practical advantages and disadvantages. In testing and comparing the different models, we find that all models produce similar results for the coefficients of interest. We present our results below using a probit model. The two basic models for patient i are as follows:

$$NS_i = \mathbb{I}\{NS_i^* > 0\}, \text{ and}$$
$$NS_i^* = \boldsymbol{X}_i\beta_1 + \gamma_1 Ori\_wait_i + \gamma_2 Days\_chg_i + \nu_{h(i)} + \varepsilon_i, \tag{1}$$

Table 3 Correlation matrix of different waits

	Ori_wait	Days_chg	Pre_wait	Post_wait
Ori_wait	1.000			
Days_chg	-0.083^{*}	1.000		
Pre_wait	0.776^{*}	0.061^{*}	1.000	
Post_wait	0.472^{*}	0.132^{*}	0.057^{*}	1.000

Note: * Significant at the 1% level.

$$NS_i^* = \boldsymbol{X}_i\beta_2 + \gamma_3 Pre\_wait_i + \gamma_4 Post\_wait_i + \nu_{h(i)} + \epsilon_i, \tag{2}$$

where NS_i denotes the outcome of no-show with $NS_i = 1$ indicating no-show and $NS_i = 0$ indicating arrival; $\mathbb{I}$ denotes the indicator function; NS_i^* is a latent variable that represents the propensity for no-show; $\boldsymbol{X}_i$ is a vector of control variables including patient-level characteristics (age, class, race, nation, and distance from residence to the clinic), and appointment-related control variables (clinic specialty, season, day of week, period of day); $\nu_{h(i)}$ is the clinic fixed effect; and ε_i and ϵ_i are assumed to have a standard normal distribution.

The correlation matrix of waiting time is shown in Table 3. We observe that Ori_wait is negatively associated with $Days\_chg$ (with correlation$= -0.083$), and Pre_wait is positively associated with $Post\_wait$ (with correlation$= 0.057$). The weak correlations suggest that the estimation of Model (1) and (2) should not have significant multicollinearity issues.

Another estimation issue worth noting is that the t-statistics for demographic information are potentially overstated, as there is a lack of independence across appointments for a given patient. Therefore, we estimate the model using the Huber/White/sandwich robust method with adjustment for within-cluster correlations for each patient (Wooldridge 2010). This method adjusts the covariance matrix for potential correlation in errors between multiple visits by a single individual, and also adjusts for potential misspecification of the model's functional form.

5.1.2. Estimation Results

Table 4 summarizes the results of Models (1) and (2), for both new and follow-up patients. We only report the coefficients of interest for the various waiting times. Full estimation results are presented in Appendix A. The column "samples" indicate the subsamples used in the models to estimate the effects of waiting time. The average marginal effect in Table 4 shows the average expected absolute change in no-show probability (for each subsample) when an appointment's waiting time is increased by one week. For new patients, the effects of waiting time are all statistically significant and positive. For non-rescheduled appointments (ND), the predicted probability of no-show increases by 1.3 percentage points with a one-week increase in the waiting time. If the appointment was rescheduled (i.e., NA, NP), the effect of Ori_wait is outweighed by $Days\_chg$, which suggests that the patient's no-show behavior is more affected by how many days the appointment is moved forward or back rather than by the original wait. For example, for patients who have actively rescheduled their appointments, no-show probability increases by 1.0 percentage point with an additional week of original wait, while it increases by 2.7 percentage points with their appointments rescheduled to one week later. In addition, no-show increases significantly with Pre_wait and $Post\_wait$ for both actively and passively rescheduled appointments. Therefore, for new patients, Hypotheses 1A, 1B, and 1C are all well supported.

For follow-up patients, in contrast, the effects of waiting time are mostly insignificant. Specifically, no-show probability is not significantly affected by the length of treatment cycle (Ori_wait) for all subsamples. For active rescheduling (FA), rescheduling one-week into the future leads to a 0.8-percentage-point increase in the no-show probability, but is only significant at the 10% level; for passive rescheduling (FP), how many days the appointment is moved forward or back does not have any significant impact on no-show probability. In terms of the effects for Pre_wait and $Post\_wait$, the magnitude of the former is much less notable than the latter, indicating that the effect of pre-rescheduling wait is outweighed by post-rescheduling wait. This result appears to resonate with the idea that rescheduling may psychologically separate the waiting experience, known as the "landmark effect" (Dai et al. 2014). In particular, the marginal effect of Pre_wait is negative if the appointment is actively rescheduled, which indicates that patients may fall into the sunk-cost fallacy—that is, the more time one has already waited before rescheduling, the less likely one will miss the appointment. In summary, results for follow-up patients do not support Hypotheses 1A, 1B, or 1C. This suggests that waiting time may not be a primary concern for follow-up patients[4].

[4] An alternative explanation for the lack of an effect of waiting time on the no-show rate for follow-up patients could be that some follow-up patients switch to care-givers who offer a more frequent or convenient schedule. This means that our sample may be limited to patients who are insensitive to the wait offered in the first place, similar to the effect studied by Osadchiy and KC (2017). We expect, however, that the impact of such events may be low in our sample, since treatment plans for most specialties usually follow standard protocols, and treatment cycles

Table 4 **Average marginal effect of waiting time on no-show**

	Subsample	Model 1 *Ori_wait*		Model 1 *Days_chg*		Model 2 *Pre_wait*		Model 2 *Post_wait*	
New	NA	0.010^{***}	(0.001)	0.027^{***}	(0.001)	0.024^{***}	(0.002)	0.026^{***}	(0.002)
	NP	0.019^{***}	(0.003)	0.025^{***}	(0.007)	0.020^{***}	(0.006)	0.023^{***}	(0.008)
	ND	0.013^{***}	(0.000)						
Follow-up	FA	-0.001	(0.001)	0.008^{*}	(0.005)	-0.001^{***}	(0.000)	0.006^{***}	(0.000)
	FP	0.001	(0.001)	0.006	(0.005)	0.000	(0.000)	0.005^{***}	(0.000)
	FD	-0.002	(0.002)						

Note: Clustered robust standard errors are shown in parentheses. Controls not shown include age, class, race, nation, distance, clinic, specialty, season, day of week, and period of day. All ***wait*** is counted in **weeks**.
Subsamples: NA (New Active), NP (New Passive), ND (New Direct), FA (Follow-up Active), FP (Follow-up Passive), FD (Follow-up Direct).
* p<0.1, ** p<0.05, *** p<0.01

5.2. Rescheduling

5.2.1. Econometric Model

In evaluating the effect of active rescheduling on no-show probability, potential endogeneity issues arise. Some unobserved factors may simultaneously affect patients' tendency to reschedule and their no-show probability. For example, patients who choose to actively reschedule may inherently be more likely to show up for their appointments, since this signals that the appointment is important enough for them to reach out and reschedule rather than just not show up. Therefore, it could potentially generate a negative bias in the estimate of the effect of active rescheduling.

To address these concerns, we use a recursive bivariate probit model to account for non-random assignment. The model simultaneously estimates (i) an appointment's propensity to be rescheduled and (ii) the patient's final attendance outcome, taking into account that the error terms might be correlated. To obtain consistent estimates, it is desirable to identify at least one instrumental variable (IV) that affects the decision to reschedule but has no direct effect on no-show behavior—a condition known as "exclusion restriction" (Wooldridge 2010).

Although we do not have an IV that can be used immediately for active rescheduling, we can construct an IV by observing the *appointment availability* in the data set. The design of this IV is based on the fact that most patients choose to reschedule their appointment

are typically long enough (as suggested by our data) that slot availability may not be the bottleneck. Furthermore, through our communication with the clinics, we understand that they typically give higher priority to follow-up visits or treatments to provide consistent and quality care, and overbooking is possible if necessary. Nevertheless, such events, unfortunately, are not observed in our data set, and thus our results are subject to this potential bias due to sample selection.

within two weeks of the original appointment date (see Figure 4a), and they tend to reschedule the appointment to the same weekday as the original appointment date (see Figure 5a), due to their preference and physician's availability on those days. Therefore, the number of available slots in the clinic on the same weekday within the weeks closest to the original appointment date could be correlated to the rescheduling observed in the data set. For example, if an appointment was originally scheduled for September 15, the number of available slots in the clinic for September 1, 8, 15, 22, and 29, observed on September 1, could be correlated to whether this appointment would be rescheduled. Following this idea, we construct an indicator variable, $APPAVLB$, which equals to one if the average appointment availability of these five days is higher than 20%. Otherwise, $APPAVLB = 0$ represents the case in which the clinic is relatively congested on one's preferred days with low appointment availability. Formal definition of this IV is presented in Appendix B.1. We also follow standard procedures to justify the validity of our IV, and conduct various tests of the IV for under-, and weak identification. All the results support the validity of our IV with details provided in Appendix B.2.

We model the decision to actively reschedule, $Active_i$, via a latent variable model:

$$\begin{aligned} Active_i &= \mathbb{I}\{Active_i^* > 0\}, \text{ and} \\ Active_i^* &= \boldsymbol{X}_i\theta_1 + \alpha_1 APPAVLB_i + \omega_{h(i)} + \xi_i, \end{aligned} \tag{3}$$

where $Active_i = 1$ indicates appointment i was actively rescheduled and $Active_i = 0$ otherwise; $Active_i^*$ is a latent variable that represents an appointment's propensity to be actively rescheduled; $\boldsymbol{X}_i$ is a vector of control variables for patient characteristics and appointment information; $APPAVLB_i$ is the instrumental variable (IV) for active rescheduling; $\omega_{h(i)}$ is the clinic fixed effect; and ξ_i represents unobservable factors that affect the choice to actively reschedule. The outcome of no-show NS_i is modeled as follows:

$$\begin{aligned} NS_i &= \mathbb{I}\{NS_i^* > 0\}, \text{ and} \\ NS_i^* &= \boldsymbol{X}_i\beta_1 + \gamma_1 Active_i + \nu_{h(i)} + \varepsilon_i, \end{aligned} \tag{4}$$

where NS_i^* is a latent variable that represents the propensity for no-show; $\nu_{h(i)}$ is the clinic fixed effect; and ε_i captures unobservable factors that affect no-show behavior. To account for the endogeneity in the decision to actively reschedule, represented by $Active_i$, we allow the error term ε_i to be correlated with the unobservable factors that affect rescheduling—ξ_i in Equation (3)—by assuming that the random vector (ξ_i, ε_i) follows a standard bivariate normal distribution with correlation coefficient ρ, which will be estimated along with other model parameters. Note that this will require joint estimation of the rescheduling decision model (3) and the attendance outcome model (4). Such a bivariate probit (BiProbit) model can be estimated via the full maximum likelihood estimation (Cameron and Trivedi 1998, Greene 2012). Endogeneity can be tested through a likelihood ratio test of the correlation coefficient ρ being zero.

Passive rescheduled appointments, in contrast, are mainly due to physicians' being on leave (for study or conferences) or to an unexpected surge in urgent cases. Clinic management teams confirmed that all passive rescheduling was performed by central clerks who had limited or no understanding of patients' medical conditions and would not be selective in rescheduling their appointments. In other words, the clinics' decisions to reschedule appointments should be exogenous, at least to a large extent. Therefore, we use a probit model to estimate the effect of passive rescheduling, $Passive_i$ on patient's no-show NS_i as follows:

$$\begin{aligned} NS_i &= \mathbb{I}\{NS_i^* > 0\}, \text{ and} \\ NS_i^* &= \boldsymbol{X}_i\beta_2 + \gamma_2 Passive_i + \nu_{h(i)} + \upsilon_i. \end{aligned} \tag{5}$$

To further confirm that passive rescheduling is exogenous, we use a recursive bivariate probit model for passive rescheduling and no-show. While it is desirable to identify at least one instrumental variable (IV) that affects the clinic's decision to reschedule but has no direct effect on patients' no-show behavior, this is not necessary—i.e., it is possible to estimate the BiProbit model without IVs (Greene 2012, Maddala 1983, Winship et al. 1988, Wilde 2000), and the likelihood ratio test of zero correlation ($\rho = 0$) can then be used as a Hausman endogeneity test (Knapp and Seaks 1998).

The correlation coefficient between passive rescheduling and the no-show models' error terms is insignificant at the 10% level ($\rho = -0.066$, p-value $= 0.1093$), which shows no evidence of endogeneity in the form of omitted variable bias[5]. We acknowledge, however,

[5] However, as Monfardini and Radice (2008) point out, the availability of IVs is of paramount importance, especially when error terms are misspecified (i.e., by deviating from the assumption of bivariate normality). Hence, we cannot rule out the possibility that given an appropriate IV being identified, the correlation between two error terms could be non-zero.

Table 5 Average marginal effect of rescheduling on no-show

		Subsample	AME	ρ	Test $\rho=0$
New	Active (No IV)	NA+ND	-0.095^{***} (0.003)	—	—
	Active (IV)	NA+ND	-0.023^{***} (0.005)	-0.180^{***} (0.048)	0.00
	Passive	NP+ND	0.032 (0.023)	—	—
Follow-up	Active (No IV)	FA+FD	-0.121^{***} (0.003)	—	—
	Active (IV)	FA+FD	-0.109^{***} (0.008)	-0.081^{**} (0.036)	0.03
	Passive	FP+FD	0.062^{***} (0.017)	—	—

Note: "AME" is the average marginal effect. Clustered robust standard errors are shown in parentheses. Controls not shown include age, class, race, nation, distance, clinic, specialty, season, day of week, and period of day.
Subsamples: NA (New Active), NP (New Passive), ND (New Direct), FA (Follow-up Active), FP (Follow-up Passive), FD (Follow-up Direct).
* p<0.1, ** p<0.05, *** p<0.01

the limitation in our analysis in that other potential endogeneity issues could be present that we cannot account for—due to lack of a satisfactory IV—despite our best efforts.

5.2.2. Estimation Results

Table 5 summarizes the results of the no-show models, for both new and follow-up patients. We only report the coefficients of interest, i.e., active rescheduling and passive rescheduling. Full estimation results are presented in Appendix A.

The column "samples" indicate the subsamples used in the models to estimate the effect of rescheduling. For example, two subsamples, NA and ND are used in model (3) to estimate the effect of active rescheduling on patient no-show (with and without IV). The average marginal effect (AME) shown in the table is the average expected absolute change in no-show probability (among all appointments in one sample) when an appointment has been rescheduled (actively or passively). The column "Test $\rho = 0$" shows the p-values of the null hypothesis that the decision to reschedule is exogenous; this test is equivalent to a likelihood ratio test that the correlation coefficient between rescheduling and the no-show models' errors (ρ) is zero. The null hypothesis is rejected in all models—that is, accounting for the endogeneity with IV is important for obtaining consistent estimates of the impact of rescheduling on patients' no-show behavior.

The naive estimates without IV ignores the endogeneity in the decision to reschedule. The effects of active rescheduling are all significantly negative, for both new and follow-up patients, which suggests that active rescheduling is associated with a significant reduction in no-show probability. With the introduction of IV, the effects of active rescheduling are still negative and statistically significant. However, the magnitude of the effect is reduced, especially for new patients. This is consistent with our expectation that endogeneity could generate a negative bias in the effect of active rescheduling. After accounting for endogeneity, the marginal effect of active rescheduling is reduced to -2.3% for new patients and -10.9% for follow-up patients.

Regarding passive rescheduling, the estimate shows that on average, it increases no-show for new patients by 3.2 percentage points—which is not significant at the 10% level[6]—and by 6.2 percentage points—significant at the 1% level—for follow-ups. Note that this is considered an unbiased estimate, because, as discussed in Section 5.2, it is an exogenous decision whether a clinic chooses to reschedule an appointment.

In summary, for follow-up patients, no-show decreases by 10.9 percentage points if the appointment has been actively rescheduled and increases by 6.2 percentage points if it has been passively rescheduled. For comparison, the average no-show probability is 27.8% across all follow-up patients whose appointments were not rescheduled. For new patients, in contrast, who initiates rescheduling has an insignificant or very weak impact on no-show probability. Hypothesis 2A claims that no-show probability decreases when the appointment is actively rescheduled, and Hypothesis 2B claims that no-show probability increases when the appointment is passively rescheduled. Therefore, Hypotheses 2A and 2B are both supported for follow-up patients. Hypothesis 2A is weakly supported and Hypothesis 2B is not supported for new patients.

5.3. Robustness Checks

To confirm the robustness of the results discussed in previous sections, we conducted a similar analysis using two matching methods to account for observed and unobserved heterogeneity, and obtained similar results. The first is the *propensity score matching* (PSM) method, which aims to remove any bias in the estimated treatment effect due to differences or imbalances in the observed covariates across treatment groups. For example, in measuring the effect of IT adoption on productivity, Sudhir and Talukdar (2015) note that there are obvious selection concerns because IT adoption is not random; some firms may avoid IT adoption to avoid transparency. To address these concerns, the authors use PSM to ensure that the inferences about productivity differences between adopters and

[6] While the problem of *rare events* in maximum likelihood probit regression could be present (since passive reschedulings account for only 6.45% of the data used in the regression), this concern could be greatly alleviated given the large number of "rare events" (i.e., 4,517 data points in NP). To be more conservative, we use an alternative penalized likelihood approach—the-Firth method—to reduce potential bias (Firth 1993, Heinze and Schemper 2002). Consistently, the effect of passive rescheduling on new patients remains insignificant.

nonadopters are limited to firms that are comparable in their propensity to adopt IT. Moreover, we also implemented the *near-far matching* method, which is a nonparametric, matching-based instrumental variable methodology (Baiocchi et al. 2010). The main idea is to match observations that are near in the covariates and far on the instrument in order to reduce model dependence and, potentially, strengthen the instrument (Lu et al. 2011, Lorch et al. 2012, Yang et al. 2014, Hu et al. 2017).

We also ran the same model estimation with alternative definitions of the IV (including the continuous version) and alternative definitions of no-show (e.g., cancellation within a specific number of days before the original appointment), and all of the qualitative insights are consistent. In addition, we analyze different subsamples (e.g., high-volume specialties, appointments with the most common treatment cycles, different age groups, different government-subsidy conditions, etc.), and our main results still hold. Furthermore, we test the possibility of an alternative explanation for our results as a consequence of a patient's "sickness" level, which was not supported by our data.

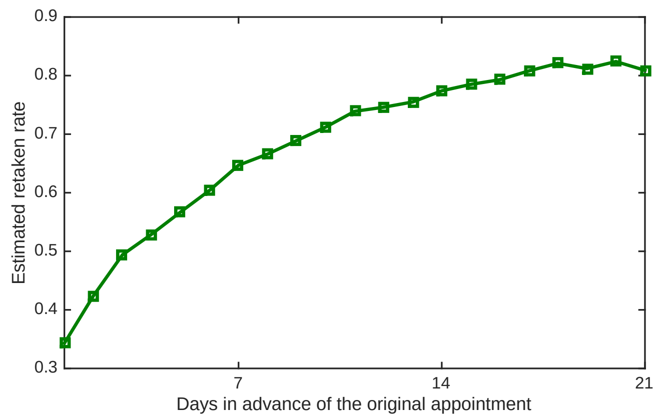
Details of these robustness checks are presented in Appendix C.

6. Discussion and Implications

We have demonstrated that active rescheduling could reduce no-show probability significantly for follow-up patient, which translates into increased system utilization. At the same time, however, a rescheduled appointment also frees up a slot that may not be filled, which results in physician idleness and increases wait time. Therefore, it is not clear whether rescheduling has a net positive or negative impact. In this section, we use data-driven simulation and counterfactual analysis to provide some insights into the question whether clinics should provide opportunities for active rescheduling.

Intuitively, the earlier an appointment is rescheduled, the higher the likelihood that the freed-up slot will be filled. Since we have no access to the physician's information for each appointment, we assume that a newly open slot is filled if an appointment with the same starting time is subsequently booked. Based on this assumption, for slots freed up by rescheduling one week (or two weeks) in advance, the likelihood that they would be retaken increased significantly, to around 65% (or 77%) (see Figure 9). We also conduct counterfactual experiments[7] to investigate the benefits if all rescheduling occurs one week or two weeks before the original appointment date.

Figure 9 **How many slots freed up by active rescheduling can be retaken?**



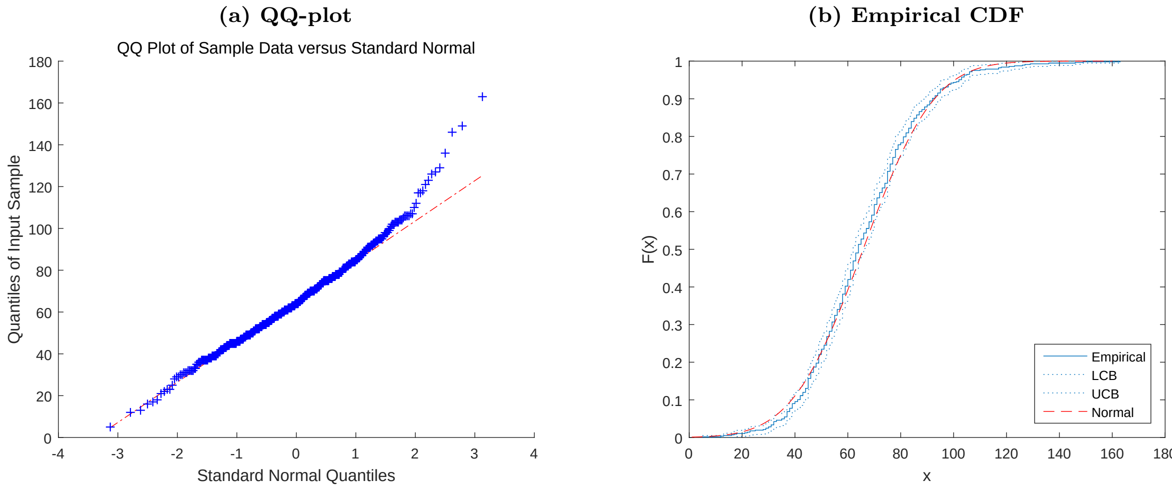
To address the above questions, we simulate the appointment system of a median-size clinic in our data set. Results from simulating different clinics are consistent with the findings presented below. We first collect all of the appointment records for this clinic, then remove the records for weekends and holidays, as these are unusual, ad hoc appointments. We also remove 10% of the appointments at the beginning of the study period and 10% at the end of the period due to censoring[8]. Finally, we keep 37,294 appointments requests for 567 workdays from January 2012 to April 2014. For the chosen clinic, the average number of daily appointment requests is 65.77, and the standard deviation is 21.07. Of all appointment requests, 25.34% are made by new patients and 74.66% by follow-up patients. Following the common practice in the literature for simulating healthcare systems, we assume fixed capacity level and use a stationary Gaussian process to model the daily appointment requests (Kim et al. 2014, Freeman et al. 2016, Kim et al. 2017, Mandelbaum et al. 2017). We estimate daily capacity by counting the arrivals and no-shows for each day and use the average daily capacity, 47, as the fixed capacity level. To validate the Gaussian arrival assumption in our clinic, we depict the QQ-plot in Figure 10(a). We also plot the

[7] Note that the assumption that a newly open slot is filled if an appointment with the same starting time is subsequently booked is used only to construct Figure 9. This assumption is not required or used in our simulation model and counterfactual analysis.

[8] We do not have earlier appointment records for the appointments made in the beginning period of our data set. Hence, we cannot separate direct appointments from rescheduled ones. Similarly, we do not have data on future rescheduling for appointments made in the ending period of our data set.

empirical cumulative distribution function of daily appointment requests against the fitted normal distribution $\mathcal{N}(65.86, 21.07^2)$ with 95% confidence interval in Figure 10(b). We can observe that in general, the normal assumption provides a good fit to the empirical data.

Figure 10 **Distribution of daily appointment requests**



Note: In figure (b), "LCB" stands for lower confidence bound, and "UCB" stands for upper confidence bound.

We use a bootstrapping procedure to generate daily appointment request samples and the subsequent rescheduling process (if any) by random sampling with replacement from the clinic data. The proportions of new and follow-up patients are generated according to the empirical proportions for the benchmark case, or different percentages for counterfactual analysis. New patients are assigned to the first available slot—which is consistent with the observation in Osadchiy and KC (2017) that new patients tend to choose time slots that are not far from the shortest wait—and their outcomes are randomly sampled from the pool of new patients with the same or nearest original waits. Follow-up patients are assigned to days according to waiting times from the bootstrapping samples, and similarly for all of their outcomes. For each round of simulation, we simulate the system for a period of 12 years, with the first and last year removed for final analysis. We report average performance measures and their 95% confidence intervals (CI) for 30 rounds of such simulation. Performance measures include (1) percentages of different appointment outcomes (arrival, no-show, active and passive rescheduling); (2) average wait time for all patients, new patients and follow-up patients; and (3) average system utilization, defined as number of daily arrivals over capacity. Results from the data and simulation are summarized in the second and third columns of Table 6 under "Data" and "Simulation: Rescheduling," respectively. We can see that our high-fidelity data-driven simulation mimics the clinic operation well in all of these performance measures. The simulation model is also verified by the clinic management team.

Table 6 Simulation results

	Data	Simulation				
		Rescheduling	No active rescheduling	Active rescheduling ahead 1 week	Active rescheduling ahead 2 weeks	Increse capacity (5%)
Arrival	53.33%	52.03% (51.97%, 52.09%)	55.2% (55.13%, 55.27%)	52.29% (52.24%, 52.35%)	52.24% (52.17%, 52.32%)	52.11% (52.07%, 52.16%)
No-show	18.56%	19.30% (19.26%, 19.35%)	21.69% (21.62%, 21.77%)	19.14% (19.08%, 19.20%)	18.99% (18.93%, 19.05%)	19.46% (19.43%, 19.48%)
Cancellation	11.71%	11.96% (11.92%, 12.00%)	13.23% (13.2%, 13.26%)	11.94% (11.90%, 11.98%)	11.95% (11.91%, 11.98%)	11.82% (11.79%, 11.85%)
Rescheduling	16.41%	16.71% (16.68%, 16.74%)	9.85% (9.82%, 9.87%)	16.62% (16.57%, 16.68%)	16.82% (16.77%, 16.87%)	16.61% (16.58%, 16.64%)
- Active	7.28%	7.45% (7.41%, 7.49%)	0% (0%, 0%)	7.49% (7.42%, 7.55%)	7.72% (7.67%, 7.78%)	7.50% (7.48%, 7.53%)
- Passive	9.13%	9.26% (9.22%, 9.29%)	9.85% (9.82%, 9.87%)	9.14% (9.11%, 9.17%)	9.10% (9.07%, 9.12%)	9.11% (9.09%, 9.13%)
Average wait time						
All	95.66	95.21 (95.02, 95.39)	91.91 (91.75, 92.08)	92.67 (92.50, 92.84)	92.55 (92.30, 92.79)	93.44 (93.31, 93.56)
New	37.53	37.38 (37.01, 37.74)	30.71 (30.42, 31.00)	32.27 (31.86, 32.68)	30.39 (30.01, 30.77)	31.82 (31.47, 32.18)
Follow-up	118.77	117.92 (117.76, 118.07)	116.27 (116.10, 116.44)	116.35 (116.15, 116.56)	116.54 (116.3, 116.78)	117.73 (117.61, 117.86)
Utilization	87.56%	87.32% (87.06%, 87.58%)	85.66% (85.42%, 85.90%)	87.84% (87.60%, 88.09%)	87.74% (87.52%, 87.97%)	83.81% (83.60%, 84.03%)

Note: Numbers inside the parentheses indicate 95% confidence intervals.

As shown in Table 6, not allowing active rescheduling would increase the patient no-show rate from 19.30% (95%CI: 19.26%, 19.35%) to 21.69% (95%CI: 21.62%, 21.77%)—a 12.4% increase, which results in an annual loss of S$55,017 on consultation fees alone (not including other fees and charges) for the clinic studied if specialist consultation costs S$140 per visit[9]. On the other hand, not allowing active rescheduling would reduce new patients' wait time by 6.67 days[10] on average, which would lead to a lower no-show rate for new patients according to our empirical results. Although overall system utilization increases from 85.66% (95%CI: 85.42%, 85.90%) to 87.32% (95%CI: 87.06%, 87.58%) by allowing active rescheduling, the clinic should be aware of the negative impact on the wait time for new patients.

[9] Consultation fees for specialists in Singapore public hospitals depend on the seniority of the doctor, and usually vary from S$100 to S$160; major public hospitals charge S$140 for the first visit with a consultant.

[10] In our simulation, one day means one business day since weekends and holidays are not considered in our model.

As discussed earlier, allowing active rescheduling may result in unfilled freed-up slots and disrupt the availability of appointment slots for new patients, especially if rescheduling occurs close to the original appointment day. What if all active rescheduling happens at least one or two weeks before the original appointment day? We conduct this counterfactual analysis by bootstrapping the pool of patients who rescheduled one or two weeks in advance for actively rescheduled appointments. Our results in Table 6 suggests that if all active rescheduling could occur at least one week before the original appointment day, the clinic could significantly reduce the negative impact of active rescheduling on new patients' wait time—from 37.38 days (95%CI: 37.01, 37.74) to 32.27 days (95%CI: 31.86, 32.68), compared to 30.71 days (95%CI: 30.42, 31.00) with no active rescheduling—while still enjoying the benefits of lower no-show rates and higher system utilization. This is equivalent to increasing capacity by 5% in the original system; however, purely increasing capacity will lead to lower system utilization—only 83.81% (95%CI: 83.60%, 84.03%). If all active rescheduling could occur even earlier—at least two weeks before the original appointment day—the clinic could achieve similar wait time performance for new patients compared to the system with no active rescheduling. Therefore, our counterfactual analysis reveals the importance of early rescheduling on system performance.

We conduct additional simulations and counterfactual analysis (see Appendix D) and confirm that the findings discussed above hold when the mix of new and follow-up patients changes or there is more active rescheduling. In addition, we observe that when the number of follow-up patients increases, the impact of early active rescheduling is more significant. For example, if the percentage of follow-up patients is 90%, allowing active rescheduling reduces the overall patient no-show rate from 19.21% to 17.64% and increases system utilization from 86.44% to 87.40%. The benefits of increasing capacity by 5% can thus be obtained by ensuring that all active rescheduling occurs only one week ahead.

7. Concluding Remarks

This study contributes to our understanding of patients' no-show behavior by examining the waiting time and rescheduling process in an outpatient appointment system. In previous studies, waiting time has been widely identified as an important factor that influences no-show behavior. Specifically, we show that the effects of waiting time differ significantly for different patient groups. For follow-up patients—whose waiting time is largely determined by the treatment-cycle length—no-show rates are insensitive to waiting time in our sample. More importantly, we find that rescheduling has a significant impact on patient no-show behavior, and that patients of different types respond to rescheduling in notably different ways. For follow-up patients, no-show probability decreases by 10.9 percentage points if their appointments were actively rescheduled, and increases by 6.2 percentage points if their appointments were passively rescheduled. In marked contrast, the average no-show probability is 27.8% across all follow-up patients whose appointments were not rescheduled. Interestingly, whether a follow-up appointment is rescheduled to an earlier or later slot does not affect the no-show rate. In contrast, new patients always desire faster access and prefer shorter waiting times. If their appointments are rescheduled, new patients care much more whether their appointments are moved forward or back, rather than who initiates the rescheduling.

A limitation of our study is the lack of physicians' information. As a result, we can only estimate appointment capacity and availability based on our observed data. For the same reason, we are unable to control the heterogeneity across physicians. For instance, prior studies have shown that physicians with greater expertise (e.g., faculty versus resident) are less likely to experience patient no-shows (Bennett and Baxley 2009, Tseng 2010); also, patients are willing to endure longer waiting times for highly rated physicians (Osadchiy and KC 2017). Likewise, patient rescheduling behavior may be influenced by physicians' characteristics, which is worth exploring in future research. Another limitation comes from our assumptions in the simulation and counterfactual analysis. For example, it is challenging to estimate behavioral changes and outcomes for patients who used to reschedule within one or two weeks if they are not allowed to do so. We acknowledge that the bootstrapping method used in our simulation may be unrealistic. A better approach is to conduct field experiments and surveys to understand behavioral changes due to different rescheduling policies.

Our simulation results show that encouraging patients to actively reschedule their appointments is likely to improve system performance. Identifying how clinics could able to achieve this, would be an intriguing avenue for future research. Another future research direction would be to construct a well-justified structural model that balances the costs and benefits of appointment rescheduling from the patient's perspective. Such a model would provide deeper insights into the effect of rescheduling on patient behavior and system performance, and allow more detailed and analytical counterfactual study. We believe that our results in this paper, together with future research on the mechanisms responsible for the effect of rescheduling on patient behavior (such as patient preference for certain appointment times and other psychological factors) will yield useful insights on how to build such models for further analysis. Our results provide a first glimpse into the issue of rescheduling in the appointment systems and our preliminary counterfactual analysis identifies a significant impact of patient rescheduling on system performance—which, we hope, will be sufficient to justify future research in this direction.

References

Akşin, Zeynep, Barış Ata, Seyed Morteza Emadi, Che-Lin Su. 2013. Structural estimation of callers' delay sensitivity in call centers. *Management Science* **59**(12) 2727–2746.

Allon, Gad, Awi Federgruen, Margaret Pierson. 2011. How much is a reduction of your customers' wait worth? An empirical study of the fast-food drive-thru industry based on structural estimation methods. *Manufacturing & Service Operations Management* **13**(4) 489–507.

Baiocchi, Mike, Dylan S Small, Scott Lorch, Paul R Rosenbaum. 2010. Building a stronger instrument in an observational study of perinatal care for premature infants. *Journal of the American Statistical Association* **105**(492) 1285–1296.

Batt, Robert J, Christian Terwiesch. 2015. Waiting patiently: An empirical study of queue abandonment in an emergency department. *Management Science* **61**(1) 39–59.

Bean, Andrew B, James Talaga. 1995. Predicting appointment breaking. *Marketing Health Services* **15**(1) 29.

Bean, Andrew G, James Talaga. 1992. Appointment breaking: Causes and solutions. *Marketing Health Services* **12**(4) 14.

Bennett, Kevin J, Elizabeth G Baxley. 2009. The effect of a carve-out advanced access scheduling system on no-show rates. *Family Medicine* **41**(1) 51.

Bodenheimer, Thomas, Hoangmai H Pham. 2010. Primary care: Current problems and proposed solutions. *Health Affairs* **29**(5) 799–805.

Cameron, C, P Trivedi. 1998. *Models for count data*. Cambridge University Press.

Cameron, Stewart, Laura Sadler, Beverley Lawson. 2010. Adoption of open-access scheduling in an academic family practice. *Canadian Family Physician* **56**(9) 906–911.

Campbell, Jennifer D, Ronald A Chez, Tina Queen, Annette Barcelo, Ellen Patron. 2000. The no-show rate in a high-risk obstetric clinic. *Journal of Women's Health & Gender-Based Medicine* **9**(8) 891–895.

Chan, Carri W, Linda V Green, Lijian Lu, Suparerk Lekwijit, Gabriel J Escobar. 2017. Assessing the impact of service intensity on customers: An empirical investigation of hospital step-down units. *Working Paper* .

Cheraghi-Sohi, Sudeh, Arne Risa Hole, Nicola Mead, Ruth McDonald, Diane Whalley, Peter Bower, Martin Roland. 2008. What patients want from primary care consultations: a discrete choice experiment to identify patients priorities. *The Annals of Family Medicine* **6**(2) 107–115.

Compton, Michael T, Bruce E Rudisch, Jason Craw, Tina Thompson, Dwight Antonio Owens. 2006. Predictors of missed first appointments at community mental health centers after psychiatric hospitalization. *Psychiatric Services* **57**(4) 531–537.

Corfield, Lorraine, Alexis Schizas, Andrew Williams, A Noorani. 2008. Non-attendance at the colorectal clinic: a prospective audit. *The Annals of The Royal College of Surgeons of England* **90**(5) 377–380.

Daggy, Joanne, Mark Lawley, Deanna Willis, Debra Thayer, Christopher Suelzer, Po-Ching DeLaurentis, Ayten Turkcan, Santanu Chakraborty, Laura Sands. 2010. Using no-show modeling to improve clinic performance. *Health Informatics Journal* **16**(4) 246–259.

Dai, Hengchen, Katherine L Milkman, Jason Riis. 2014. The fresh start effect: Temporal landmarks motivate aspirational behavior. *Management Science* **60**(10) 2563–2582.

Deyo, Richard A, Thomas S Inui. 1980. Dropouts and broken appointments: A literature review and agenda for future research. *Medical Care* **18**(11) 1146–1157.

Feldman, Jacob, Nan Liu, Huseyin Topaloglu, Serhan Ziya. 2014. Appointment scheduling under patient preference and no-show behavior. *Operations Research* **62**(4) 794–811.

Firth, David. 1993. Bias reduction of maximum likelihood estimates. *Biometrika* **80**(1) 27–38.

Freeman, Michael, Nicos Savva, Stefan Scholtes. 2016. Gatekeepers at work: An empirical analysis of a maternity unit. *Management Science, forthcoming* .

Freeman, Michael, Stefan Scholtes. 2017. Gatekeeping under uncertainty: An empirical study of referral errors in the emergency department. *Working Paper* .

Gallucci, Gerard, Wayne Swartz, Florence Hackerman. 2005. Impact of the wait for an initial appointment on the rate of kept appointments at a mental health center. *Psychiatric Services* **56**(3) 344–346.

Gany, Francesca, Julia Ramirez, Serena Chen, Jennifer CF Leng. 2011. Targeting social and economic correlates of cancer treatment appointment keeping among immigrant chinese patients. *Journal of Urban Health* **88**(1) 98–103.

George, Ajay, Greg Rubin. 2003. Non-attendance in general practice: A systematic review and its implications for access to primary health care. *Family Practice* **20**(2) 178–184.

Geraghty, M, F Glynn, M Amin, J Kinsella. 2008. Patient mobile telephone 'text' reminder: a novel way to reduce non-attendance at the ent out-patient clinic. *The Journal of Laryngology & Otology* **122**(3) 296–298.

Green, Linda V, Sergei Savin. 2008. Reducing delays for medical appointments: A queueing approach. *Operations Research* **56**(6) 1526–1538.

Greene, William H. 2012. *Econometric Analysis. 7th edition*. Prentice Hall.

Grunebaum, Michael, Philip Luber, Mark Callahan, Andrew C Leon, Mark Olfson, Laura Portera. 1996. Predictors of missed appointments for psychiatric consultations in a primary care clinic. *Psychiatric Services* .

Gupta, Diwakar, Brian Denton. 2008. Appointment scheduling in health care: Challenges and opportunities. *IIE Transactions* **40**(9) 800–819.

Gupta, Diwakar, Wen-Ya Wang. 2012. Patient appointments in ambulatory care. *Handbook of Healthcare System Scheduling*. Springer, 65–104.

Guy, Rebecca, Jane Hocking, Handan Wand, Sam Stott, Hammad Ali, John Kaldor. 2012. How effective are short message service reminders at increasing clinic attendance? a meta-analysis and systematic review. *Health Services Research* **47**(2) 614–632.

Hashim, Muhammad Jawad, Peter Franks, Kevin Fiscella. 2001. Effectiveness of telephone reminders in improving rate of appointments kept at an outpatient clinic: A randomized controlled trial. *The Journal of the American Board of Family Practice* **14**(3) 193–196.

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

Heinze, Georg, Michael Schemper. 2002. A solution to the problem of separation in logistic regression. *Statistics in medicine* **21**(16) 2409–2419.

Hole, Arne Risa. 2008. Modelling heterogeneity in patients' preferences for the attributes of a general practitioner appointment. *Journal of Health Economics* **27**(4) 1078–1094.

Hu, Wenqi, Carri W Chan, José R Zubizarreta, Gabriel J Escobar. 2017. An examination of early transfers to the ICU based on a physiologic risk score. *Manufacturing & Service Operations Management, forthcoming* .

Janakiraman, Narayan, Robert J Meyer, Stephen J Hoch. 2011. The psychology of decisions to abandon waits for service. *Journal of Marketing Research* **48**(6) 970–984.

Johnson, Bradley J, James W Mold, J Michael Pontious. 2007. Reduction and management of no-shows by family medicine residency practice exemplars. *The Annals of Family Medicine* **5**(6) 534–539.

KC, Diwas Singh, Christian Terwiesch. 2011. The effects of focus on performance: Evidence from california hospitals. *Management Science* **57**(11) 1897–1912.

KC, Diwas Singh, Christian Terwiesch. 2012. An econometric analysis of patient flows in the cardiac intensive care unit. *Manufacturing & Service Operations Management* **14**(1) 50–65.

Kim, Song-Hee, Carri W Chan, Marcelo Olivares, Gabriel Escobar. 2014. ICU admission control: An empirical study of capacity allocation and its implication for patient outcomes. *Management Science* **61**(1) 19–38.

Kim, Song-Hee, Ward Whitt, Won Chul Cha. 2017. A data-driven model of an appointment-generated arrival process at an outpatient clinic. *Working Paper* .

Knapp, Laura Greene, Terry G Seaks. 1998. A Hausman test for a dummy variable in probit. *Applied Economics Letters* **5**(5) 321–323.

Lacy, Naomi L, Audrey Paulman, Matthew D Reuter, Bruce Lovejoy. 2004. Why we don't come: patient perceptions on no-shows. *The Annals of Family Medicine* **2**(6) 541–545.

LaGanga, Linda R, Stephen R Lawrence. 2012. Appointment overbooking in health care clinics to improve patient service and clinic performance. *Production and Operations Management* **21**(5) 874–888.

Leclerc, France, Bernd H Schmitt, Laurette Dube. 1995. Waiting time and decision making: Is time like money? *Journal of Consumer Research* **22**(1) 110–119.

Liu, Nan, Stacey R Finkelstein, Margaret E Kruk, David Rosenthal. 2017. When waiting to see a doctor is less irritating: Understanding patient preferences and choice behavior in appointment scheduling. *Management Science, forthcoming* .

Liu, Nan, Serhan Ziya. 2014. Panel size and overbooking decisions for appointment-based services under patient no-shows. *Production and Operations Management* **23**(12) 2209–2223.

Liu, Nan, Serhan Ziya, Vidyadhar G Kulkarni. 2010. Dynamic scheduling of outpatient appointments under patient no-shows and cancellations. *Manufacturing & Service Operations Management* **12**(2) 347–364.

Lorch, Scott A, Michael Baiocchi, Corinne E Ahlberg, Dylan S Small. 2012. The differential impact of delivery hospital on the outcomes of premature infants. *Pediatrics* **130**(2) 270–278.

Lowes, R. 2005. A no-show showdown letter. *Medical Economics* **82**(8) 66–67.

Lu, Bo, Robert Greevy, Xinyi Xu, Cole Beck. 2011. Optimal nonbipartite matching and its statistical applications. *The American Statistician* **65**(1) 21–30.

Macharia, William M, Gladys Leon, Brian H Rowe, Barbara J Stephenson, R Brian Haynes. 1992. An overview of interventions to improve compliance with appointment keeping for medical services. *JAMA* **267**(13) 1813–1817.

Maddala, GS. 1983. *Limited-dependent and qualitative variables in econometrics*. Cambridge University Press.

Mandelbaum, Avishai, Petar Momcilovic, Nikolaos Trichakis, Sarah Kadish, Ryan Leib, Craig A Bunnell. 2017. Data-driven appointment-scheduling under uncertainty: The case of an infusion unit in a cancer center. *Working Paper* .

Martin, Chris, Tracey Perfect, Greg Mantle. 2005. Non-attendance in primary care: The views of patients and practices on its causes, impact and solutions. *Family Practice* **22**(6) 638–643.

Monfardini, Chiara, Rosalba Radice. 2008. Testing exogeneity in the bivariate probit model: A monte carlo study. *oxford Bulletin of Economics and Statistics* **70**(2) 271–282.

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

Morse, Dale L, Molly P Coulter, Rudolph J Napodano, Ho-Ling Hwang, Charles Lawrence. 1984. Broken appointments at a neighborhood health center: emphasis on weather. *Medical Care* **22**(9) 813–817.

Murray, Mark, Catherine Tantau. 2000. Same-day appointments: Exploding the access paradigm. *Family Practice Management* **7**(8) 45–50.

Muthuraman, Kumar, Mark Lawley. 2008. A stochastic overbooking model for outpatient clinical scheduling with no-shows. *IIE Transactions* **40**(9) 820–838.

Nagler, Jonathan. 1994. Scobit: an alternative estimator to logit and probit. *American Journal of Political Science* **38**(1) 230–255.

Neal, Richard D, Mahvash Hussain-Gambles, Victoria L Allgar, Debbie A Lawlor, Owen Dempsey. 2005. Reasons for and consequences of missed appointments in general practice in the uk: questionnaire survey and prospective review of medical records. *BMC Family Practice* **6**(1) 47.

Norris, John B, Chetan Kumar, Suresh Chand, Herbert Moskowitz, Steve A Shade, Deanna R Willis. 2014. An empirical investigation into factors affecting patient cancellations and no-shows at outpatient clinics. *Decision Support Systems* **57** 428–443.

Osadchiy, Nikolay, Diwas KC. 2017. Are patients patient? The role of time to appointment in patient flow. *Production and Operations Management* **26**(3) 469–490.

Pesata, Virginia, Geri Pallija, Adele A Webb. 1999. A descriptive study of missed appointments: Families' perceptions of barriers to care. *Journal of Pediatric Health Care* **13**(4) 178–182.

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

Rubin, Greg, Angela Bate, Ajay George, Phil Shackley, Nicola Hall. 2006. Preferences for access to the gp: a discrete choice experiment. *The British Journal of General Practice* **56**(531) 743–748.

Ryan, Mandy, Shelley Farrar. 2000. Using conjoint analysis to elicit preferences for health care. *British Medical Journal* **320**(7248) 1530–1533.

Suck, Reinhard, Heinz Holling. 1997. Stress caused by waiting: A theoretical evaluation of a mathematical model. *Journal of Mathematical Psychology* **41**(3) 280–286.

Sudhir, K, Debabrata Talukdar. 2015. The peter pan syndrome in emerging markets: The productivity-transparency trade-off in it adoption. *Marketing Science* **34**(4) 500–521.

Tseng, Fen-Yu. 2010. Non-attendance in endocrinology and metabolism patients. *Journal of the Formosan Medical Association* **109**(12) 895–900.

Wang, Wen-Ya, Diwakar Gupta. 2011. Adaptive appointment systems with patient preferences. *Manufacturing & Service Operations Management* **13**(3) 373–389.

WHO. 2011. *Global health and aging*. World Health Organization.

WHO. 2014. *Chronic diseases and health promotion*. World Health Organization.

Wilde, Joachim. 2000. Identification of multiple equation probit models with endogenous dummy regressors. *Economics letters* **69**(3) 309–312.

Williams, Marian E, James Latta, Persila Conversano. 2008. Eliminating the wait for mental health services. *The Journal of Behavioral Health Services & Research* **35**(1) 107–114.

Winship, Christopher, Robert D Mare, J Scott Long. 1988. *Endogenous switching regression models for the causes and effects of discrete variables*. Sage Press.

Wooldridge, Jeffrey M. 2010. *Econometric analysis of cross section and panel data*. MIT press.

Yang, Fan, José R Zubizarreta, Dylan S Small, Scott Lorch, Paul R Rosenbaum. 2014. Dissonant conclusions when testing the validity of an instrumental variable. *The American Statistician* **68**(4) 253–263.

Appendix to
Effects of Rescheduling on Patient No-show Behavior in Outpatient Clinics

A Full Estimation Results

A.1 Waiting Time

In §5.1 of the main paper, we explore the role of waiting time on no-show for new and follow-up patients. The two probit models for patient i are as follows:

$$\begin{aligned} NS_i &= \mathbb{I}\{NS_i^* > 0\}, \text{ and} \\ NS_i^* &= \boldsymbol{X}_i\beta_1 + \gamma_1 Ori\_wait_i + \gamma_2 Days\_chg_i + \nu_{h(i)} + \varepsilon_i, \end{aligned} \tag{1}$$

$$NS_i^* = \boldsymbol{X}_i\beta_2 + \gamma_3 Pre\_wait_i + \gamma_4 Post\_wait_i + \nu_{h(i)} + \epsilon_i, \tag{2}$$

where NS_i denotes the outcome of no-show with $NS_i = 1$ indicating no-show and $NS_i = 0$ indicating arrival; $\mathbb{I}$ denotes the indicator function; NS_i^* is a latent variable that represents the propensity for no-show; $\boldsymbol{X}_i$ is a vector of control variables including patient-level characteristics (age, subsidized condition, race, and nationality), and appointment-related control variables (clinic specialty, season, day of week, period of day); $\nu_{h(i)}$ is the clinic fixed effect; and ε_i and ϵ_i are assumed to have a standard normal distribution.

The full results for Table 4 in the main paper are shown in Table 1 (for new patients) and Table 2 (for follow-up patients) in this section. Model (1) and Model (2) represent the two probit models introduced above in Equations (1) and (2), respectively. Note that within each model, different columns correspond to results estimated from different subsamples. For example, columns (a) to (c) of Model (1) present the results estimated from subsamples NA (New Active), NP (New Passive), and ND (New Direct), respectively.

Table 1: The effect of waiting time (new patients)

	Model (1)			Model (2)	
	(a)	(b)	(c)	(a)	(b)
Age					
(Base: [0,15])					
(15,30]	-0.013	-0.057	0.102**	-0.138	0.024
	(0.080)	(0.258)	(0.042)	(0.102)	(0.328)
(30,45]	-0.114	-0.234	-0.137***	-0.303***	-0.142
	(0.083)	(0.266)	(0.043)	(0.105)	(0.343)
(45,60]	-0.282***	-0.803***	-0.403***	-0.403***	-0.959***
	(0.086)	(0.283)	(0.045)	(0.109)	(0.365)
(60,]	-0.395***	-0.657*	-0.359***	-0.641***	-0.883*
	(0.118)	(0.354)	(0.054)	(0.149)	(0.461)
Subsidized Condition					
(Base: Sub)					
Non-sub	0.217***	0.194***	0.297***	0.093	0.219***
	(0.049)	(0.028)	(0.099)	(0.064)	(0.020)
Race					
(Base: R1)					
R2	0.404***	0.528***	0.501***	0.506***	0.457***
	(0.036)	(0.131)	(0.017)	(0.044)	(0.164)
R3	0.420***	0.371***	0.377***	0.410***	0.297*
	(0.039)	(0.125)	(0.017)	(0.048)	(0.156)
R4	0.312***	0.312**	0.199***	0.271***	0.378**
	(0.048)	(0.141)	(0.021)	(0.062)	(0.184)
Nationality					
(Base: Citizen)					
Non-citizen	-0.078	-0.664***	-0.386***	-0.134	-0.604
	(0.157)	(0.107)	(0.056)	(0.149)	(0.463)
Specialty					
(Base: S1)					
S2	-0.145	0.122	0.043	-0.033	0.035
	(0.144)	(0.106)	(0.060)	(0.175)	(0.126)
S3	-0.142**	-0.489**	-0.084***	-0.018**	-0.613**
	(0.060)	(0.207)	(0.025)	(0.076)	(0.267)
S4	-0.117**	-0.277	0.129***	0.014	-0.017
	(0.056)	(0.221)	(0.027)	(0.075)	(0.189)
S5	-0.033	-0.104	0.414***	-0.032	-0.149
	(0.152)	(0.114)	(0.068)	(0.197)	(0.152)
S6	-0.156**	-0.197	-0.061**	-0.097	0.112
	(0.065)	(0.171)	(0.027)	(0.082)	(0.222)
S7	-0.048	-0.166	0.204***	-0.112	0.133
	(0.150)	(0.132)	(0.071)	(0.189)	(0.188)

S8	-0.052	-0.003	0.465***	-0.237*	0.212
	(0.094)	(0.156)	(0.048)	(0.122)	(0.178)
S9	-0.028	-0.118	-0.193***	-0.373**	0.032
	(0.121)	(0.089)	(0.056)	(0.159)	(0.091)
S10	-0.049	-0.152	-0.148***	-0.067	-0.110
	(0.084)	(0.137)	(0.034)	(0.105)	(0.132)
S11	0.290	0	-0.107	-0.094	0
	(0.680)	(.)	(0.198)	(0.168)	(.)
S12	-0.130	0.049	0.133***	-0.074	0.586**
	(0.085)	(0.052)	(0.032)	(0.106)	(0.278)
S13	-0.061	-0.219	-0.185***	0.016	-0.246
	(0.095)	(0.179)	(0.044)	(0.119)	(0.278)
Season					
(Base: Spring)					
Summer	0.001	0.007	0.007	0.048	0.020
	(0.035)	(0.117)	(0.016)	(0.045)	(0.152)
Autumn	-0.025	0.113	0.017	0.007	0.032
	(0.035)	(0.114)	(0.016)	(0.045)	(0.149)
Winter	-0.039	-0.031	0.029*	-0.003	-0.247
	(0.038)	(0.135)	(0.018)	(0.048)	(0.179)
Week of Day					
(Base: Wed)					
Mon	0.040	0.054	0.028	0.040	-0.075
	(0.039)	(0.127)	(0.018)	(0.049)	(0.165)
Tue	-0.084**	0.138	-0.017	-0.098*	0.061
	(0.042)	(0.136)	(0.019)	(0.052)	(0.174)
Thu	-0.046	0.021	-0.018	-0.053	-0.031
	(0.041)	(0.145)	(0.019)	(0.052)	(0.189)
Fri	-0.086**	0.016	-0.017	-0.090*	-0.021
	(0.042)	(0.136)	(0.019)	(0.052)	(0.174)
Period of Day					
(Base: Morning)					
Noon	-0.026	-0.142	-0.036	0.056	-0.170
	(0.062)	(0.196)	(0.029)	(0.078)	(0.247)
Afternoon	-0.002	-0.307***	-0.069***	-0.022	-0.422***
	(0.127)	(0.095)	(0.013)	(0.034)	(0.124)
Evening	0.092	-0.160	-0.245***	-0.001	-0.142
	(0.129)	(0.375)	(0.061)	(0.166)	(0.558)
Ori_wait	0.082***	0.053***	0.135***		
	(0.010)	(0.011)	(0.006)		
Days_chg	0.144***	0.297***			
	(0.038)	(0.014)			
Pre_wait				0.147***	0.102***
				(0.013)	(0.039)
Post_wait				0.162***	0.128***
				(0.012)	(0.025)
Sample	NA	NP	ND	NA	NP
Observations	15,934	4,517	65,469	15,934	4,517
Log-Likelihood	-7179.59	-2218.06	-39539.50	-7980.57	-2758.27
Pseudo-R^2	0.108	0.134	0.099	0.100	0.131

AUC	0.677	0.689	0.688	0.676	0.682

Note: Clustered robust standard errors are shown in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2: The effect of waiting time (follow-up patients)

	Model (1)			Model (2)	
	(a)	(b)	(c)	(a)	(b)
Age					
(Base: [0,15])					
(15,30]	0.083	0.029	0.092***	0.097	0.063
	(0.065)	(0.099)	(0.026)	(0.064)	(0.098)
(30,45]	-0.126*	-0.227**	-0.152***	-0.110*	-0.192*
	(0.067)	(0.101)	(0.026)	(0.065)	(0.101)
(45,60]	-0.301***	-0.679***	-0.469***	-0.304***	-0.641***
	(0.068)	(0.102)	(0.027)	(0.067)	(0.101)
(60,]	-0.315***	-0.794***	-0.530***	-0.356***	-0.739***
	(0.080)	(0.112)	(0.031)	(0.078)	(0.110)
Subsidized Condition					
(Base: Sub)					
Non-sub	0.028	0.039	0.023	0.110***	0.120***
	(0.030)	(0.028)	(0.029)	(0.031)	(0.029)
Race					
(Base: R1)					
R2	0.331***	0.342***	0.359***	0.340***	0.342***
	(0.026)	(0.036)	(0.010)	(0.026)	(0.036)
R3	0.351***	0.418***	0.431***	0.368***	0.417***
	(0.026)	(0.037)	(0.010)	(0.026)	(0.037)
R4	0.198***	0.181***	0.199***	0.183***	0.187***
	(0.034)	(0.043)	(0.014)	(0.034)	(0.043)
Nationality					
(Base: Citizen)					
Non-citizen	0.015	-0.035	-0.197***	-0.127	-0.083
	(0.030)	(0.121)	(0.043)	(0.101)	(0.091)
Specialty					
(Base: S1)					
S2	-0.092	0.088	0.153***	-0.045	0.108
	(0.149)	(0.107)	(0.039)	(0.087)	(0.107)
S3	-0.085**	-0.024	-0.104***	-0.086**	-0.036
	(0.039)	(0.047)	(0.015)	(0.037)	(0.048)
S4	0.073**	0.009	-0.042***	0.049	0.018
	(0.036)	(0.021)	(0.014)	(0.037)	(0.051)
S5	0.371***	0.124	0.237***	0.358***	0.140
	(0.086)	(0.118)	(0.034)	(0.086)	(0.129)
S6	0.013	0.019	-0.048***	-0.027	0.024
	(0.042)	(0.021)	(0.015)	(0.040)	(0.039)
S7	0.232**	0.094	0.354***	0.232**	0.083

	(0.095)	(0.068)	(0.035)	(0.094)	(0.087)
S8	0.301***	0.010	0.505***	0.239***	0.034
	(0.073)	(0.025)	(0.029)	(0.073)	(0.054)
S9	0.092	-0.255**	0.182***	0.113	-0.192
	(0.083)	(0.124)	(0.033)	(0.083)	(0.124)
S10	-0.083	-0.021	-0.067***	-0.117**	-0.003
	(0.081)	(0.049)	(0.021)	(0.048)	(0.060)
S11	-0.654	-0.449	-0.318***	-0.798*	-0.454
	(0.435)	(0.281)	(0.086)	(0.431)	(0.276)
S12	0.120**	-0.060	0.017	0.064	-0.053
	(0.050)	(0.052)	(0.017)	(0.048)	(0.051)
S13	-0.023	0.188**	0.067**	-0.015	0.217***
	(0.065)	(0.086)	(0.027)	(0.064)	(0.047)
Season					
(Base: Spring)					
Summer	0.023	-0.016	-0.023**	0.012	-0.007
	(0.024)	(0.018)	(0.010)	(0.023)	(0.030)
Autumn	0.014	-0.007	0.011	0.002	0.007
	(0.024)	(0.021)	(0.010)	(0.024)	(0.030)
Winter	0.030	0.010	0.011	0.013	0.017
	(0.026)	(0.017)	(0.010)	(0.025)	(0.032)
Week of Day					
(Base: Wed)					
Mon	0.053**	-0.058*	0.018*	0.037	-0.061*
	(0.027)	(0.033)	(0.010)	(0.026)	(0.033)
Tue	-0.029	-0.018	-0.026**	-0.027	-0.009
	(0.034)	(0.016)	(0.011)	(0.026)	(0.036)
Thu	-0.003	-0.052	-0.035***	-0.018	-0.043
	(0.038)	(0.034)	(0.011)	(0.027)	(0.034)
Fri	-0.036	-0.004	-0.047***	-0.044	-0.014
	(0.029)	(0.026)	(0.012)	(0.029)	(0.036)
Period of Day					
(Base: Morning)					
Noon	-0.012	0.012	0.050***	-0.004	0.034
	(0.043)	(0.034)	(0.017)	(0.040)	(0.033)
Afternoon	0.014	-0.034	0.002	0.009	-0.018
	(0.019)	(0.024)	(0.007)	(0.039)	(0.024)
Evening	0.076	-0.171	-0.042	0.086	-0.171
	(0.082)	(0.113)	(0.038)	(0.079)	(0.112)
Ori_wait	-0.002	0.009	-0.004		
	(0.002)	(0.007)	(0.004)		
Days_chg	0.143*	0.037			
	(0.078)	(0.026)			
Pre_wait				-0.009***	0.003
				(0.001)	(0.003)
Post_wait				0.038***	0.018***
				(0.008)	(0.005)
Sample	FA	FP	FD	FA	FP
Observations	33,022	29,915	156,017	33,022	29,915
Log-Likelihood	-13400.79	-14287.01	-84703.17	-13766.46	-14473.14

Pseudo-R^2	0.087	0.083	0.081	0.085	0.086
AUC	0.662	0.655	0.652	0.663	0.670

Note: Clustered robust standard errors are shown in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

A.2 Rescheduling

In §5.2 of the main paper, we evaluate the effect of active rescheduling on no-show probability. Specifically, the decision to actively reschedule, $Active_i$, is modeled as follows:

$$\begin{aligned} Active_i &= \mathbb{I}\{Active_i^* > 0\}, \text{ and} \\ Active_i^* &= \boldsymbol{X}_i\theta_1 + \alpha_1 APPAVLB_i + \omega_{h(i)} + \xi_i, \end{aligned} \tag{3}$$

where $Active_i = 1$ indicates appointment i was actively rescheduled and $Active_i = 0$ otherwise; $Active_i^*$ is a latent variable that represents an appointment's propensity to be actively rescheduled; $\boldsymbol{X}_i$ is a vector of control variables for patient characteristics and appointment information; $APPAVLB_i$ is the instrumental variable (IV) for active rescheduling (see §5.2.1 in the main paper and Section B for more details); $\omega_{h(i)}$ is the clinic fixed effect; and ξ_i represents unobservable factors that affect the choice to actively reschedule. The outcome of no-show, NS_i, is modeled as follows:

$$\begin{aligned} NS_i &= \mathbb{I}\{NS_i^* > 0\}, \text{ and} \\ NS_i^* &= \boldsymbol{X}_i\beta_1 + \gamma_1 Active_i + \nu_{h(i)} + \varepsilon_i, \end{aligned} \tag{4}$$

where NS_i^* is a latent variable that represents the propensity for no-show; $\nu_{h(i)}$ is the clinic fixed effect; and ε_i captures unobservable factors that affect no-show behavior.

To account for the potential endogeneity of active rescheduling, we allow the error terms in the above two equations to be correlated, assuming that the random vector (ξ_i, ε_i) follows a standard bivariate normal distribution with correlation coefficient ρ, which will be estimated along with other model parameters. The presence of endogeneity can be tested through a likelihood ratio test of the correlation coefficient ρ being zero.

In this section we first show the full results of estimating the univariate probit model, i.e., Model (4). Labels (d) (e) (f) and (g) in Table 3 denote the results estimated using different subsamples. Next, we present the full results of the recursive bivariate probit model, i.e., estimating Model (3) and Model (4) simultaneously.

A.2.1 Probit Models (without IV)

Table 3: Univariate probit models

	New patients		Follow-ups	
	(d)	(e)	(f)	(g)
Age				
(Base: [0,15])				
(15,30]	0.123**	0.125**	0.145***	0.143***
	(0.060)	(0.060)	(0.037)	(0.037)
(30,45]	-0.321***	-0.322***	-0.269***	-0.276***
	(0.062)	(0.062)	(0.039)	(0.039)
(45,60]	-0.709***	-0.706***	-0.829***	-0.844***
	(0.065)	(0.066)	(0.039)	(0.039)
(60,]	-0.631***	-0.626***	-0.958***	-0.976***
	(0.081)	(0.082)	(0.047)	(0.048)
Subsidized condition				
(Base: Sub)				
Non-sub	0.397***	0.377***	0.068***	0.069***
	(0.030)	(0.031)	(0.017)	(0.018)
Race				
(Base: R1)				
R2	0.849***	0.838***	0.586***	0.592***
	(0.024)	(0.024)	(0.015)	(0.015)
R3	0.692***	0.685***	0.703***	0.712***
	(0.025)	(0.025)	(0.015)	(0.015)
R4	0.470***	0.467***	0.337***	0.342***
	(0.030)	(0.030)	(0.019)	(0.019)
Nationality				
(Base: Citizen)				
Non-citizen	0.054	0.060**	0.210***	0.211***
	(0.029)	(0.029)	(0.017)	(0.017)
Specialty				
(Base: S1)				
S2	-0.045	-0.015	0.186***	0.187***
	(0.092)	(0.093)	(0.056)	(0.056)
S3	-0.130***	-0.126***	-0.151***	-0.152***
	(0.038)	(0.037)	(0.023)	(0.023)
S4	0.035	0.013	-0.082***	-0.083***
	(0.038)	(0.038)	(0.020)	(0.021)
S5	0.502***	0.442***	0.379***	0.352***
	(0.098)	(0.098)	(0.049)	(0.049)
S6	-0.149***	-0.166***	-0.077***	-0.093***
	(0.041)	(0.041)	(0.022)	(0.023)
S7	0.197*	0.207**	0.445***	0.424***

	(0.103)	(0.104)	(0.049)	(0.049)
S8	0.548***	0.541***	0.716***	0.686***
	(0.067)	(0.068)	(0.042)	(0.042)
S9	-0.153*	-0.133	0.231***	0.232***
	(0.081)	(0.081)	(0.048)	(0.048)
S10	-0.194***	-0.195***	-0.084***	-0.077**
	(0.052)	(0.053)	(0.030)	(0.030)
S11	0.076	0.142	-0.692***	-0.676***
	(0.324)	(0.325)	(0.151)	(0.151)
S12	0.189***	0.177***	-0.003	-0.016
	(0.048)	(0.048)	(0.026)	(0.025)
S13	-0.245***	-0.226***	0.102***	0.117***
	(0.069)	(0.067)	(0.039)	(0.039)
Season				
(Base: Spring)				
Summer	0.025	0.008	-0.027*	-0.030**
	(0.024)	(0.024)	(0.014)	(0.014)
Autumn	0.030	0.017	0.019	0.014
	(0.024)	(0.024)	(0.014)	(0.014)
Winter	0.028	0.020	0.024	0.021
	(0.026)	(0.026)	(0.015)	(0.015)
Week of Day				
(Base: Wed)				
Mon	0.057**	0.053**	0.029*	0.027*
	(0.027)	(0.027)	(0.016)	(0.015)
Tue	-0.042	-0.037	-0.041**	-0.039**
	(0.028)	(0.028)	(0.016)	(0.016)
Thu	-0.038	-0.034	-0.049***	-0.054***
	(0.028)	(0.028)	(0.016)	(0.016)
Fri	-0.027	-0.024	-0.074***	-0.074***
	(0.028)	(0.029)	(0.017)	(0.017)
Period of Day				
(Base: Morning)				
Noon	-0.064	-0.059	0.072***	0.072***
	(0.042)	(0.043)	(0.026)	(0.026)
Afternoon	-0.099***	-0.103***	-0.004	-0.006
	(0.018)	(0.019)	(0.011)	(0.011)
Evening	-0.242***	-0.218**	-0.099*	-0.072
	(0.091)	(0.092)	(0.053)	(0.054)
Ori_wait	0.081***	0.149***		
	(0.007)	(0.007)		
Active	-0.695***		-0.745***	
	(0.080)		(0.024)	
Passive		0.330		0.171***
		(0.273)		(0.028)
Sample	NA+ND	NP+ND	FA+FD	FP+FD
Observations	81,403	69,986	189,039	185,932
Log-Likelihood	-37821.64	-31338.67	-101208.13	-115080.16
Pseudo-R^2	0.082	0.081	0.085	0.083
AUC	0.707	0.703	0.723	0.720

Note: Clustered robust standard errors are shown in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

A.2.2 Bivariate Probit Models (with IV for Active Rescheduling)

Table 4: Bivariate probit models

	New patients		Follow-ups	
	NS	Active	NS	Active
Instrument				
APPAVLB		0.458***		0.183***
		(0.043)		(0.013)
Age				
(Base: [0,15])				
(15,30]	-0.005	0.126*	0.071**	-0.002
	(0.065)	(0.071)	(0.035)	(0.037)
(30,45]	-0.259***	0.144**	-0.169***	0.028
	(0.067)	(0.073)	(0.036)	(0.038)
(45,60]	-0.476***	0.094	-0.483***	-0.061
	(0.070)	(0.075)	(0.037)	(0.038)
(60,]	-0.499***	-0.085	-0.530***	-0.204***
	(0.086)	(0.090)	(0.043)	(0.043)
Race				
(Base: R1)				
R2	0.513***	-0.140***	0.323***	-0.108***
	(0.026)	(0.029)	(0.015)	(0.015)
R3	0.391***	-0.147***	0.426***	-0.033**
	(0.028)	(0.031)	(0.016)	(0.016)
R4	0.236***	-0.062	0.185***	0.009
	(0.035)	(0.039)	(0.020)	(0.020)
Subsidized Condition				
(Base: Sub)				
Non-sub	0.244***	-0.222***	-0.027**	0.165***
	(0.033)	(0.036)	(0.013)	(0.012)
Nationality				
(Base: Citizen)				
Non-citizen	-0.456***	-0.114	0.188***	-0.185***
	(0.090)	(0.098)	(0.018)	(0.017)
Specialty				
(Base: S1)				
S2	-0.091	-0.834***	0.159**	0.222***
	(0.106)	(0.130)	(0.071)	(0.050)
S3	-0.082**	-0.126***	-0.088***	0.055**
	(0.039)	(0.044)	(0.022)	(0.023)
S4	0.069*	0.285***	-0.087***	0.046**
	(0.040)	(0.043)	(0.019)	(0.019)

S5	0.466***	0.154	0.158***	-0.170***
	(0.110)	(0.123)	(0.053)	(0.054)
S6	-0.080*	0.068	-0.036*	-0.131***
	(0.043)	(0.047)	(0.020)	(0.022)
S7	0.003	-0.003	0.150***	-0.272***
	(0.112)	(0.120)	(0.050)	(0.055)
S8	0.322***	0.224***	0.423***	-0.148***
	(0.072)	(0.080)	(0.040)	(0.042)
S9	-0.156*	-0.236**	-0.005	-0.131***
	(0.089)	(0.108)	(0.048)	(0.050)
S10	-0.075	-0.032	-0.022	0.105***
	(0.057)	(0.060)	(0.030)	(0.028)
S11	0.440	-0.511	-0.218	-0.170
	(0.401)	(0.477)	(0.195)	(0.181)
S12	0.313***	-0.250***	0.049	-0.105***
	(0.054)	(0.062)	(0.026)	(0.026)
S13	-0.111	-0.521***	0.044	0.024
	(0.071)	(0.090)	(0.041)	(0.042)
Season				
(Base: Spring)				
Summer	0.011		-0.034***	
	(0.025)		(0.011)	
Autumn	0.041		-0.017	
	(0.026)		(0.013)	
Winter	0.061**		-0.017	
	(0.028)		(0.013)	
Week of Day				
(Base: Wed)				
Mon	0.035		0.012	
	(0.028)		(0.014)	
Tue	-0.004		-0.041***	
	(0.029)		(0.015)	
Thu	0.010		-0.049***	
	(0.028)		(0.015)	
Fri	0.016		-0.069***	
	(0.030)		(0.015)	
Period of Day				
(Base: Morning)				
Noon	-0.031		0.030	
	(0.044)		(0.023)	
Afternoon	-0.091***		0.008	
	(0.020)		(0.010)	
Evening	0.040		0.036	
	(0.095)		(0.062)	
Ori_wait	0.029***			
	(0.003)			
Active	-0.108***		-0.485***	
	(0.035)		(0.068)	
Sample	NA+ND		FA+FD	
Observations	81,403		189,039	

Log-likelihood	-61832.37	-91543.61
ρ	-0.180***(0.048)	-0.081**(0.036)

Note: Clustered robust standard errors are shown in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

B IV Specification and Validation

B.1 IV for Active Rescheduling

To formally define our IV for active rescheduling, we estimate the daily capacity, C_t, by counting the arrivals and no-shows for each clinic[1] on day t, since we do not have exact data on the physicians' shift. The number of available slots for day t observed on day s ($s < t$), denoted as $A_{t,s}$, is then computed as the difference between the capacity of day t and the number of patients already scheduled by day s for day t. Next, the *appointment availability* for day t observed on day s, denoted as $AVLB_{t,s}$, is defined as the ratio between the number of available slots and the capacity, i.e., $AVLB_{t,s} = A_{t,s}/C_t$. For each appointment, we use the same weekday 2 weeks before and after the appointment date (recall the example above), and compute the appointment availability for the 5 days (i.e., t = September 1, 8, 15, 22, and 29) 2 weeks before the appointment date (i.e., s = September 1). The *average appointment availability*, $\overline{APPAVLB}$, is then computed across these five dates. Finally, we define $APPAVLB$ as an indicator variable that equals 1 when average appointment availability is higher than 20% and 0 otherwise. Intuitively, $APPAVLB = 0$ represents the case in which the clinic is relatively congested on one's preferred days with low appointment availability. Table 5 summarizes the proportions of patients from different subsamples that are exposed to the binary IV.

[1]The computations hereafter are all done within the same clinic, so we omit the superscript or subscript for clinic.

Table 5: Patients from different subsamples exposed to IV

	New		Follow-up	
	NA	ND	FA	FD
$APPAVLB = 1$	2,938 (14.8%)	16,891 (85.2%)	11,056 (24.3%)	34,445 (75.7%)
$APPAVLB = 0$	5,206 (8.5%)	56,368 (91.5%)	21,667 (15.1%)	121,871 (84.9%)

B.2 Validity Tests

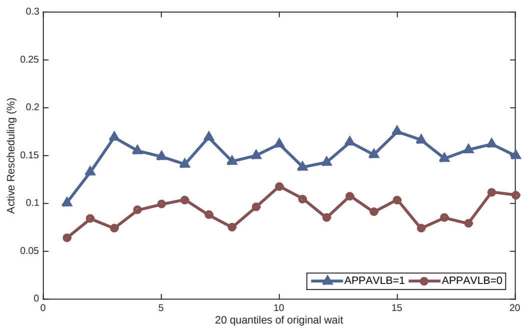
The appropriateness of $APPAVLB$ as an IV for active scheduling must be tested to satisfy two conditions: (i) it is correlated with patients' decision to reschedule, and (ii) it is independent of unobservable factors that may influence no-show behavior.

We first test the condition that $APPAVLB$ is correlated with the decision to actively reschedule. In Figure 1, we divide patients into 20 groups based on the lengths of their original waits and compare the percentage of active reschedulings by patients who observed a high level of appointment availability versus those who observed a low level. We find that based on this coarse comparison, greater appointment availability is associated with an increased fraction of rescheduled appointments. Using a series of Probit regression models, we also find, at the 0.1% significance level, that patients are more likely to reschedule their appointments when appointment availability is high.

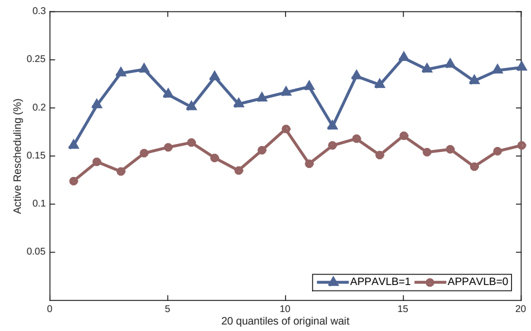
Furthermore, the second condition requires that a valid IV is uncorrelated with *unobserved* factors that may influence no-show behavior. While this condition cannot be test statistically, we provide some evidence for the validity of IV by demonstrating that *observed* factors such as patients' age and treatment cycle (which could be considered as proxies for their severity) are equally distributed between the two levels of $APPAVLB$ (0 and 1). We also use a patient's previous rescheduling records as a proxy for the patient's attitude/personality and compute its correlation with the instrument. The resulting correlation coefficient is not significantly different from zero.

Figure 1: Appointment availability and active rescheduling

(a) New patients



(b) Follow-up patients

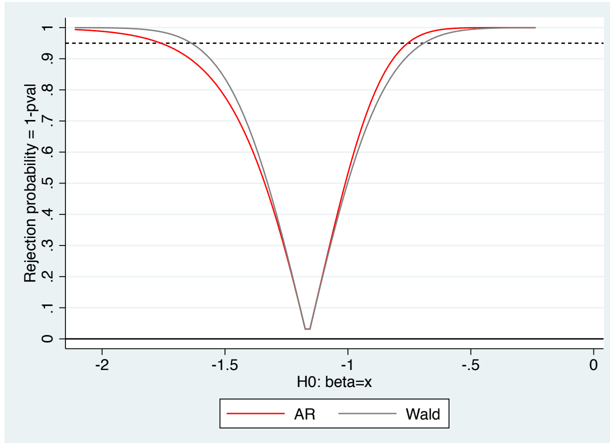


To perform formal tests for underidentification and weak identification, we note that the majority of "weak IV" tests are based on a linear IV regression model where the dependent variable in the outcome equation and the endogenous variable are continuous. Following Freeman et al. (2016), we first treat both no-show and active rescheduling as continuous and estimate the model via *ivreg2* command in Stata 14.0 (Baum et al. 2002). Note that the coefficients estimated using continuous model specification are qualitatively consistent with our main results.

For the above model, the first-stage Wald statistics is 8.30 (p-value $= 0.0040$). This shows strong evidence to reject the null hypothesis of under-identification at the 1% significance level and therefore the excluded instrument can be considered as "relevant". Turning next to the issue of weak identification, Stock and Yogo (2005) tabulated critical values for the first-stage F-statistic to test whether instruments are weak. For a single endogenous regressor, assuming the model to be estimated under limited information maximum likelihood, the critical F-values are 16.38, 8.96, and 6.66 for maximum biases of 10%, 15%, and 20%, respectively. In our case, the estimated F-statistic is 46.92, indicating a maximum bias of lower than 10%.

We then create a confidence interval and p-values based on inverting the Anderson-Rubin test statistic (Mikusheva et al. 2006, Finlay et al. 2016). This confidence interval is robust to weak instruments and optimal in the case of a single instrument (Moreira 2009). Figure 2 shows the confidence intervals under Wald and Anderson-Rubin test statistics: A 95% Anderson-Rubin confidence interval excludes zero, indicating statistical significance at the 5% level. This greatly alleviates our concern of weak instrument.

Figure 2: Anderson-Rubin Confidence Interval



C Robustness Checks

We present the detailed analysis of robustness check in this section.

C.1 Propensity Score Matching

One could argue that patients who choose to reschedule their appointments may self-select to be different from those who do not reschedule. For instance, patients' active rescheduling may signal their strong willingness to attend the appointments, and this self-selection effect could confound our results. A standard way to control for self-selection effects is the propensity score matching (PSM) method pioneered by Rosenbaum and Rubin (1983). The idea is to obtain a sample of the control group (i.e., patients who do not reschedule) that matches the treatment group (i.e., patients who actively reschedule their appointment) on observable dimensions. Such a sample matching approach drastically reduces differences between the control group and the treatment group, thus controlling, to a certain extent, the self-selection effect.

We create a dummy variable to indicate whether a patient belongs to the active group and use a logit model with this dummy variable as the dependent variable and a set of observable variables as independent variables. We then perform a nearest-neighbor matching algorithm based on the propensity score calculated in the previous step to create a matched control group. In the results of Table 6, only 4 of the covariates is well balanced before matching; and the control group is significantly different from the treatment group on dimensions such as clinic specialty, subsidy status, and race. However, after matching, the absolute value of bias is less than 5% and t-test is insignificant for all covariates, which suggests that the balance assumption is satisfied. This can also be demonstrated by the large p-values in Table 7 after matching.

Table 6: The result of balance checking (a)

	[U=unmatched]	New				Follow-up			
Variable	[M=matched]	Treatment	Control	% bias	p-value	Treatment	Control	% bias	p-value
Age (years)	U	38.094	38.144	-0.4	0.750	41.405	42.658	-9.9	0.000
	M	38.094	37.996	0.7	0.622	41.405	41.272	1.0	0.228
Race2	U	0.129	0.137	-2.1	0.077	0.103	0.116	-2.9	0.000
	M	0.129	0.129	0.0	0.982	0.102	0.102	0.3	0.761
Race3	U	0.117	0.139	-6.5	0.000	0.108	0.104	1.1	0.103
	M	0.117	0.115	0.7	0.642	0.108	0.108	-0.2	0.810
Race4	U	0.093	0.121	-9.3	0.000	0.077	0.076	0.1	0.839
	M	0.093	0.096	-1.1	0.416	0.077	0.078	-0.6	0.522
Nation2	U	0.212	0.288	-17.5	0.000	0.173	0.178	-1.2	0.092
	M	0.212	0.217	-1.0	0.491	0.173	0.177	-0.8	0.344
Subsidy2	U	0.412	0.549	-27.6	0.000	0.503	0.428	15.1	0.000
	M	0.412	0.414	-0.4	0.809	0.503	0.498	1.0	0.281
Distance (miles)	U	9.812	9.619	4.5	0.000	9.986	9.757	2.8	0.000
	M	9.812	9.881	-1.6	0.272	9.876	9.876	0.0	0.986
Specialty1	U	0.019	0.023	-2.6	0.032	0.030	0.019	7.4	0.000
	M	0.019	0.021	-1.4	0.338	0.030	0.029	0.6	0.533
Specialty3	U	0.166	0.187	-5.7	0.000	0.151	0.129	6.4	0.000
	M	0.166	0.164	0.3	0.841	0.151	0.146	1.6	0.078
Specialty4	U	0.424	0.306	24.7	0.000	0.407	0.441	-6.9	0.000
	M	0.424	0.427	-0.7	0.673	0.407	0.417	-2.0	0.025
Specialty6	U	0.125	0.136	-3.3	0.006	0.103	0.096	2.2	0.002
	M	0.125	0.125	-0.1	0.946	0.103	0.105	-0.6	0.509
Specialty8	U	0.061	0.073	-4.9	0.000	0.081	0.053	10.8	0.000
	M	0.061	0.061	0.0	0.975	0.081	0.078	0.9	0.339
Specialty10	U	0.000	0.001	-2.7	0.046	0.001	0.004	-4.8	0.000
	M	0.000	0.000	0.0	1.000	0.001	0.001	0.8	0.189
Specialty12	U	0.051	0.084	-13.1	0.000	0.061	0.113	-18.8	0.000
	M	0.051	0.054	-1.1	0.422	0.061	0.058	1.1	0.147
Specialty13	U	0.038	0.036	0.9	0.435	0.030	0.025	3.0	0.000
	M	0.038	0.038	-0.3	0.846	0.030	0.032	-1.5	0.101
Ori_wait (weeks)	U	3.71	3.67	0.9	0.488	17.876	14.733	18.4	0.000
	M	3.71	3.73	-0.8	0.575	17.876	17.816	0.3	0.705

Table 7: Result of balance checking (b)

	New			Follow-up		
	MeanBias	LR χ^2	$p > \chi^2$	MeanBias	LR χ^2	$p > \chi^2$
Unmatched	4.7	864.90	0.000	7.0	2367.06	0.000
Matched	0.7	7.24	0.968	0.8	16.83	0.397

Figures 3(a) and 3(b) show the kernel density functions of the treatment group and the control group, based on pre- and post-matching of the two groups, respectively. The sufficient overlap between the groups' propensity scores validates the use of PSM. After matching, the groups' kernel density functions are closer, which demonstrates that the characteristics of the variables in the two groups are similar after matching. We also employed radius matching and kernel matching algorithms, and the results are similar.

Figure 3: Kernel density of treatment and control groups

(a) Before matching **(b) After matching**

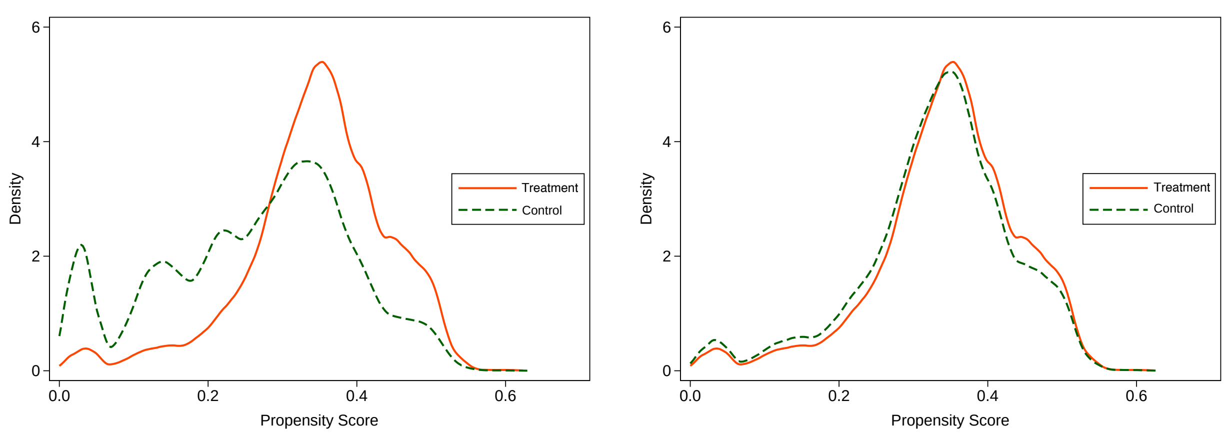


Table 8 shows the estimated effect of active rescheduling before and after matching. For new patients, the impact of active rescheduling is reduced significantly after matching. For follow-up patients, the estimated average treatment effect on the treated (ATT) after matching is significant at the 1% level across all three matching methods, varying from -0.097 to -0.107. This indicates that in general, no-show probability decreases by 9.7 to 10.7 percentage points for patients who actively rescheduled their appointments.

Table 8: Impact of active rescheduling on no-show

	New		Follow-up	
	ATT	S.E.	ATT	S.E.
Pre-matching	-0.075^{***}	0.007	-0.124^{***}	0.004
Nearest-neighbor matching	-0.025^{**}	0.011	-0.107^{***}	0.005
Radius matching	-0.022^{***}	0.008	-0.097^{***}	0.006
Kernel matching	-0.021^{**}	0.009	-0.104^{***}	0.005

Note: "Pre-matching" refers to the sample without matching the active group and control group.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The results are consistent with those reported in the paper, which demonstrates that after accounting for observable heterogeneity, patients who actively reschedule are more likely to attend their appointments. In other words, our main results are robust with respect to the self-selection effect discussed previously. In addition, we add the propensity scores as a control variable to the bivariate probit model used in the paper, and find that the main results remain qualitatively the same even when this new control variable is included.

Passive rescheduling was performed by central clerks who had limited or no understanding on patients' medical conditions, thus we expect that there would be no self-selection effect. This is confirmed by the result of Hausman endogeneity test (see Section 5.2 in the paper). Nevertheless, there might be other potential endogeneity issues that we cannot account for. Hence we repeat the analysis on active rescheduling and conduct the propensity score matching for robustness check. As shown in Table 9, passive rescheduling—after matching—increases no-show probability by 5.2 to 5.8 percentage for follow-up patients, while the effects are insignificant for new patients. These results are qualitatively consistent with our main findings.

Table 9: Impact of passive rescheduling on no-show

	New		Follow-up	
	ATT	S.E.	ATT	S.E.
Pre-matching	0.035	0.023	0.069***	0.014
Nearest-neighbor matching	0.027	0.033	0.058***	0.020
Radius matching	0.022	0.029	0.053***	0.021
Kernel matching	0.021	0.031	0.052**	0.024

Note: "Pre-matching" refers to the sample without matching the active group and control group.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

C.2 Near-Far Matching

In Section C.1, we estimated treatment effects using PSM. However, PSM is solely based on observable characteristics, and assumes that all variables that influence the decision to actively reschedule and no-show behavior are observed in the data set. Clearly, this is a strong assumption. One might also be concerned about the unobserved heterogeneity. In this section we draw upon the literature on design of observational studies (Rosenbaum 2010) and use recent advancements in the methodology of near-far matching (Baiocchi et al. 2010, Hu et al. 2017).

In instrumental variable settings, the goal of matching is to find a matched sample that is balanced on the observed covariates and imbalanced (or separated) on the instrument. The first goal attempts to reduce biases due to imbalances in observed covariates and model misspecification, whereas the second goal aims at strengthening the instrument. This is achieved by near-matching on the covariates and far-matching on the instrument (Baiocchi et al. 2010). This design parallels a matched-pair randomized controlled trial of patients who are encouraged versus patients who are not encouraged to actively reschedule. By combining IV with this matched-pairs design, we improve the equality of matched groups, and thus reduce model dependence and potentially strengthen the instrument (see Lu et al. 2011, Lorch et al. 2012, Yang et al. 2014).

We also use $\overline{APPAVLB}$ as the instrument. The only difference is that we use the continuous variable instead of the indicator variable. We first examine the quality of the matched sample by looking at balance tables for all covariates used in near matching. Results show that the covariate balance has been achieved after matching: Standardized difference are all less than 0.1, which indicates that our matched sample is well-balanced, thereby reducing model dependence and allowing for more robust estimates of the effect of active rescheduling.

Table 10: Balance Table for Measured Covariates

	New Patients			Follow-up patients		
	Encouraged	Unencouraged	Std. Diff	Encouraged	Unencouraged	Std. Diff
Age (years)	28.40	28.40	0.00	34.02	34.03	0.00
Citizen (1/0)	0.79	0.79	0.00	0.84	0.84	0.00
Race						
R1	67.36%	67.36%	0.00	72.87%	72.86%	0.00
R2	14.11%	14.11%	0.00	11.04%	11.05%	0.00
R3	10.25%	10.25%	0.00	9.43%	9.45%	0.00
R4	8.27%	8.27%	0.00	6.66%	6.64%	0.00
Subsidized (1/0)	0.66	0.66	0.00	0.53	0.53	0.00
Distance (km)						
Within 5	16.32%	16.32%	0.00	15.26%	15.26%	0.00
5 to 10	29.91%	29.91%	0.00	31.46%	31.46%	0.00
10 to 15	43.14%	43.14%	0.00	42.24%	42.42%	0.00
More than 15	10.63%	10.63%	0.00	11.04%	11.04%	0.00
Ori_wait (weeks)	3.50	3.51	0.01	16.78	16.65	0.01
Specialty						
S1	12.99%	12.99%	0.00	13.93%	13.93%	0.00
S2				1.71%	1.71%	0.00
S3	4.34%	4.34%	0.00	14.53%	14.53%	0.00
S4	33.97%	33.97%	0.00	21.53%	21.53%	0.00
S5				2.68%	2.68%	0.00
S6	9.90%	9.90%	0.00	11.50%	11.50%	0.00
S7				3.34%	3.34%	0.00
S8	28.70%	28.70%	0.00	14.39%	14.39%	0.00
S9	1.13%	1.13%	0.00	2.68%	2.68%	0.00
S10	4.44%	4.44%	0.00	6.93%	6.93%	0.00
S12	4.26%	4.26%	0.00	5.29%	5.29%	0.00
S13	0.28%	0.28%	0.00	1.48%	1.48%	0.00

Note: The "Encouraged" group corresponds to the patients who observed a high *APPAVLB* and are matched to patients with similar covariates but observing a low *APPAVLB* ("Unencouraged" group).

Table 11 summarizes the estimation results after near-far matching. For both types of patients, the instrument is highly significant at the 1% level. Being encouraged to actively reschedule increases the probability of active rescheduling by 35 and 39 percentage points for new patients and follow-up patients, respectively.

Note that ρ becomes insignificant after matching, which suggests that model dependence is decreased. For follow-up patients, no-show probability decreases by 10.5 percentage points if the appointment is actively rescheduled, while for new patients, the reduction in no-show probability is only 1.7 percentage points. These results are consistent with our main analysis.

Table 11: Estimation results using the strengthened IV after matching

NS	IV	AME on $Pr(Active)$	ρ	Active	AME on $Pr(NS)$
New	$0.67^{***}(0.022)$	$0.35^{***}(0.007)$	$-0.025(0.036)$	$-0.078^{***}(0.026)$	$-0.017^{***}(0.006)$
Follow-up	$0.82^{***}(0.019)$	$0.39^{***}(0.005)$	$0.021(0.022)$	$-0.478^{***}(0.075)$	$-0.105^{***}(0.008)$

Note: AME is the average marginal effect. Standard errors are shown in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

C.3 Alternative IV and No-Show Definitions

Similar to Kim et al. (2014), Hu et al. (2017), Chan et al. (2017), we use the binary instrument variable to handle the endogeniety problem in our main paper. We now consider different IV specifications for active rescheduling, $APPAVLB$.

Specifically, we consider: (1) estimation of appointment availability at 3 days and 1 week before the original appointment date and (2) the use of different appointment availability thresholds (15% and 25%). The results are qualitatively robust over these alternative cut-off values.

We also conduct the same analysis using a continuous version of the IV, i.e., $\overline{APPAVLB}$. The results are presented in Table 12 together with the original indicator version. We can observe that the estimated AMEs are similar from these two versions.

Table 12: Continuous and indicator IVs

	IV	Model	Estimate	AME	ρ	Test $\rho = 0$
New	No IV	Probit	-0.695^{***}(0.080)	-0.095		
	Continuous IV	BiProbit	-0.123^{*} (0.068)	-0.062	-0.027^{**} (0.011)	0.00
	Indicator IV	BiProbit	-0.108^{***}(0.035)	-0.056	-0.053^{***}(0.014)	0.01
Follow-up	No IV	Probit	-0.745^{***}(0.024)	-0.121		
	Continuous IV	BiProbit	-0.103^{***}(0.031)	-0.083	-0.025^{*} (0.014)	0.07
	Indicator IV	BiProbit	-0.485^{***}(0.068)	-0.109	-0.081^{**} (0.036)	0.03

Note: AME is the average marginal effect. Standard errors are shown in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

We explore different definitions of no-show by treating cancellations and reschedulings that occur within 1 day of the original appointment date as no-shows, because a 1-day period is generally inadequate to rebook the freed up slot. We repeat the analytical procedures, and key results remain consistent. We also relabel cancellations and reschedulings that occur within other periods before the original appointment (2-5 days) as no-shows, and qualitative results are unchanged.

In addition, we analyze different subsamples—high-volume specialties, appointments with the most common treatment cycles, different age groups, and different government-subsidy conditions—and our main results still hold.

C.4 Subsample Analysis

We repeat our analysis on different subsamples of the data. The coefficients of interest—as shown in Table 13—are consistent with those obtained from analysis of the full sample (see Table 5 in the main paper).

Table 13: The effect of rescheduling on subsamples

		Specialty							
		S1	S3	S4	S6	S8	S10	S12	S13
New	*Active*	−0.199***	−0.301***	−0.284***	−0.266***	−0.564***	−0.141***	−0.384***	−0.383***
		(0.047)	(0.052)	(0.026)	(0.047)	(0.025)	(0.063)	(0.074)	(0.090)
	Passive	0.212**	0.092	0.182	-0.125	-0.062	0.009	0.150	0.005
		(0.100)	(0.114)	(0.147)	(0.126)	(0.065)	(0.140)	(0.139)	(0.291)
Follow-up	*Active*	−0.329***	−0.258***	−0.561***	−0.252***	−0.551***	−0.305***	−0.679***	−0.794***
		(0.028)	(0.028)	(0.017)	(0.031)	(0.024)	(0.037)	(0.041)	(0.062)
	Passive	0.294***	0.398***	0.129***	0.337***	−0.260***	0.330***	0.138***	0.465***
		(0.030)	(0.036)	(0.023)	(0.278)	(0.051)	(0.050)	(0.039)	(0.084)

		Age					Subsidized condition	
		[0,15)	[15,30)	[30,45)	[45,60)	[60,)	Sub	Non-sub
New	*Active*	−0.522***	−0.364***	−0.191**	−0.087**	−0.288***	−0.363***	−0.283***
		(0.022)	(0.029)	(0.026)	(0.036)	(0.089)	(0.017)	(0.023)
	Passive	-0.026	0.064	0.097	0.208*	0.240	0.041	0.063
		(0.058)	(0.077)	(0.078)	(0.118)	(0.241)	(0.052)	(0.054)
Follow-up	*Active*	−0.691***	−0.404***	−0.468***	−0.529***	−0.265***	−0.473***	−0.623***
		(0.019)	(0.021)	(0.016)	(0.020)	(0.048)	(0.013)	(0.014)
	Passive	0.009***	0.317***	0.344***	0.141***	0.105***	0.047*	0.312***
		(0.031)	(0.027)	(0.019)	(0.023)	(0.049)	(0.018)	(0.015)

		Day of week				
		Mon	Tue	Wed	Thu	Fri
New	*Active*	−0.265***	−0.327***	−0.316***	−0.370***	−0.352***
		(0.027)	(0.032)	(0.030)	(0.031)	(0.032)
	Passive	0.072	-0.011	0.047	0.109	0.094
		(0.079)	(0.102)	(0.083)	(0.084)	(0.079)
Follow-up	*Active*	−0.584***	−0.680***	−0.304***	−0.461***	−0.557***
		(0.019)	(0.021)	(0.021)	(0.021)	(0.023)
	Passive	0.152***	0.216***	0.230***	0.183***	0.258***
		(0.024)	(0.028)	(0.024)	(0.026)	(0.029)

		Time period				
		[8:00, 9:00)	[9:00,11:00)	[11:00, 14:00)	[14:00,16:00)	[16,00, 21:00]
New	*Active*	−0.342***	−0.395***	−0.278***	−0.394***	-0.127
		(0.024)	(0.030)	(0.024)	(0.034)	(0.116)
	Passive	0.049	-0.011	0.102	0.081	-0.274
		(0.070)	(0.086)	(0.065)	(0.090)	(0.273)
Follow-up	*Active*	−0.400***	−0.598***	−0.407***	−0.674***	−0.427***
		(0.022)	(0.017)	(0.020)	(0.017)	(0.060)
	Passive	0.218***	0.215***	0.162***	0.215***	0.276***
		(0.026)	(0.020)	(0.025)	(0.023)	(0.078)

		Treatment cycle					
		Half month	One month	Two months	Three months	Half year	One year
Follow-up	*Active*	0.116*	−0.172***	−0.463***	−0.607***	−0.678***	−0.676***
		(0.063)	(0.055)	(0.066)	(0.046)	(0.038)	(0.034)
	Passive	0.550***	0.136	0.275***	0.086	0.259***	0.330***
		(0.179)	(0.122)	(0.086)	(0.052)	(0.039)	(0.029)

C.5 Alternative Explanation

One possible concern with respect to our results is that patients with more serious condition may be more likely to show up and more inclined to reschedule (if the existing appointment is inconvenient). As a result, the reduced no-show probability is not because of patients' active rescheduling, but rather an indicator of relatively more severe condition. If this were the case, one would expect that the active patients have shorter treatment cycles, because for the same medical diagnosis, shorter interval to the next treatment indicates less stable condition and higher severity in general. For example, as part of continuity of care for Type II diabetes, patients should check fasting lipid profile once a year if normal, and more frequently (e.g., every 3 months) if needed to achieve certain goals.

We test this possibility by comparing the treatment cycles of patients with active rescheduling to those without active rescheduling. Overall, to the contrary, the average treatment cycle of patients with active rescheduling (18.7 weeks) is slightly longer than those without active rescheduling (17.7 weeks) ($t = -9.5842$, $p = 0.000$). We further conduct the same comparison for each specialty. For 5 out of the 13 specialties, there are no significant difference of treatment cycle between the two groups, while for the rest 8 specialties, patients with active rescheduling have significantly longer treatment cycle than the rest at the 1% level. In addition, we include treatment cycle as well as patient- and clinic-level control variables in the logistic regression on active rescheduling for each specialty, and there is only one specialty with negative coefficient for treatment cycle at the 10% significance level. To summarize, if treatment cycles could represent patients' sickness level, our data does not suggest a significant impact of sickness level on likelihood of active rescheduling.

An alternative mechanism is that patients who actively rescheduled their appointments could be those getting worse and thus called for a more urgent treatment. If this was the case, then we would expect that the days changed for most actively rescheduling were negative. This is, however, inconsistent with our observation that among all the actively rescheduled appointments, only 20% were moved forward, whereas 80% were either rescheduled within the same day or moved back.

D Additional Simulations

In this section, we present additional counterfactual analysis using our simulation model. First, we investigate what occurs if the mix of new and follow-up patients changes. Recall that in our data, 75% of clinic patients are follow-ups. Tables 14 and 15 show the results when the percentage of follow-up patients decreases to 60% and increases to 90%, respectively. In general, the key trade-off between allowing and not allowing active rescheduling is consistent with the case with 75% follow-up patients. When the number of follow-up patients increases, the impact of early active rescheduling is more significant. When the percentage of follow-up patients is 60%, early rescheduling must occur two weeks in advance to achieve similar benefits to increasing capacity by 5%. In contrast, if the percentage of follow-up patients is 90%, the benefits of increasing capacity by 5% can be obtained by ensuring that all active rescheduling occurs only one week ahead.

Table 14: Percentage of follow-up patients is 60%

	Rescheduling	No rescheduling	Rescheduling ahead 1 week	2 weeks	Increse capacity (5%)
Arrival	52.08% (51.70%, 52.46%)	55.66% (55.35%, 55.98%)	53.28% (52.97%, 53.6%)	53.51% (53.27%, 53.75%)	53.14% (52.89%, 53.40%)
No-show	20.13% (19.75%, 20.51%)	21.65% (21.32%, 21.98%)	19.47% (19.23%, 19.71%)	19.26% (19.07%, 19.45%)	19.65% (19.44%, 19.85%)
Cancellation	10.88% (10.75%, 11.01%)	12.62% (12.38%, 12.85%)	11.29% (11.09%, 11.5%)	11.17% (11.02%, 11.32%)	11.21% (11.09%, 11.33%)
Rescheduling	16.91% (16.70%, 17.12%)	10.03% (9.88%, 10.19%)	15.96% (15.73%, 16.18%)	16.06% (15.93%, 16.19%)	16.00% (15.83%, 16.17%)
- Active	7.39% (7.21%, 7.57%)	0% (0%, 0%)	6.35% (6.18%, 6.53%)	6.78% (6.64%, 6.91%)	7.08% (6.96%, 7.19%)
- Passive	9.52% (9.38%, 9.66%)	10.03% (9.88%, 10.19%)	9.60% (9.46%, 9.74%)	9.29% (9.16%, 9.41%)	8.93% (8.77%, 9.08%)
Average wait time					
All	89.48 (88.98, 89.98)	84.78 (84.38, 85.18)	85.99 (85.65, 86.33)	85.35 (84.93, 85.77)	85.58 (85.19, 85.97)
New	52.21 (51.26, 53.15)	43.42 (42.73, 44.11)	45.24 (44.43, 46.06)	42.99 (42.40, 43.59)	42.36 (41.80, 42.93)
Follow-up	117.73 (117.36, 118.09)	116.01 (115.61, 116.41)	116.42 (116.1, 116.75)	116.61 (116.06, 117.16)	118.09 (117.71, 118.46)
Utilization	87.58% (87.03%, 88.13%)	86.59% (86.08%, 87.11%)	88.84% (88.36%, 89.33%)	89.20% (88.80%, 89.60%)	84.85% (84.48%, 85.22%)

Note: Numbers inside the parentheses indicate 95% confidence intervals.

Next, we investigate what occurs if there is more active rescheduling. From Table 16, when the number of actively rescheduled appointments increases by 50%, the no-show rate decreases from 19.3% (95%CI: 19.26%, 19.35%) to 18.17% (95%CI: 18.02%, 18.33%), while the average wait time for new patients increases from 37.38 days (95%CI: 37.07, 37.74) to 40.54 days (95%CI: 40.05, 41.04). When the number of actively rescheduled appointments doubles, the no-show rate further decreases to 16.93% (95%CI: 16.73%, 17.13%), and the average wait time for new patients increases to 44.62 days (95%CI: 43.86, 45.38). Recall our findings that new patients care more about waiting time and are insensitive to passive rescheduling: As long as their appointments are moved forward, they are more likely to show up, even if they are rescheduled by the clinic. Therefore, we simulate the case in which

Table 15: Percentage of follow-up patients is 90%

	Rescheduling	No rescheduling	Rescheduling ahead 1 week	Rescheduling ahead 2 weeks	Increse capacity (5%)
Arrival	51.63% (51.52%, 51.74%)	55.04% (54.92%, 55.17%)	52.13% (52.04%, 52.22%)	52.26% (52.16%, 52.36%)	52.03% (51.96%, 52.11%)
No-show	17.64% (17.56%, 17.73%)	19.21% (19.11%, 19.31%)	17.12% (17.05%, 17.19%)	16.97% (16.88%, 17.05%)	17.33% (17.26%, 17.40%)
Cancellation	13.37% (13.32%, 13.43%)	14.88% (14.8%, 14.96%)	13.35% (13.29%, 13.42%)	13.37% (13.29%, 13.45%)	13.26% (13.19%, 13.32%)
Rescheduling	17.36% (17.29%, 17.42%)	10.83% (10.77%, 10.89%)	17.39% (17.32%, 17.46%)	17.40% (17.34%, 17.47%)	17.38% (17.33%, 17.43%)
- Active	7.25% (7.20%, 7.30%)	0% (0%, 0%)	7.36% (7.30%, 7.42%)	7.44% (7.39%, 7.49%)	7.33% (7.28%, 7.38%)
- Passive	10.11% (10.04%, 10.17%)	10.83% (10.77%, 10.89%)	10.03% (9.97%, 10.09%)	9.97% (9.92%, 10.02%)	10.05% (10.00%, 10.10%)
Average wait time					
All	107.55 (107.25, 107.84)	105.35 (105.03, 105.67)	105.56 (105.28, 105.83)	105.79 (105.46, 106.11)	106.94 (106.62, 107.26)
New	27.03 (26.70, 27.36)	21.96 (21.59, 22.34)	22.90 (22.53, 23.27)	21.71 (21.38, 22.04)	22.85 (22.51, 23.19)
Follow-up	117.98 (117.66, 118.29)	116.35 (116.02, 116.67)	116.32 (116.02, 116.61)	116.54 (116.18, 116.9)	117.90 (117.56, 118.23)
Utilization	87.40% (87.13%, 87.67%)	86.44% (86.14%, 86.73%)	88.19% (87.98%, 88.39%)	88.49% (88.29%, 88.7%)	84.64% (84.39%, 84.9%)

Note: Numbers inside the parentheses indicate 95% confidence intervals.

the clinic proactively offers freed-up slots to new patients with the longest wait times. This provides an optimistic estimate of the benefits of such an allocation strategy (see Table 16). We observe that actively allocating freed-up slots to new patients can indeed reduce their average wait time, and the benefit is more significant when the system has more actively rescheduled appointments.

Table 16: More active rescheduling

	Current System	More rescheduling			
		1.5 times		2 times	
		Without allocation of freed-up slots	With allocation of freed-up slots	Without allocation of freed-up slots	With allocation of freed-up slots
Arrival	52.03% (51.97%, 52.09%)	51.04% (50.88%, 51.2%)	51.06% (50.83%, 51.28%)	50.04% (49.82%, 50.25%)	49.98% (49.8%, 50.17%)
No-show	19.3% (19.26%, 19.35%)	18.17% (18.02%, 18.33%)	18.29% (18.13%, 18.45%)	16.93% (16.74%, 17.13%)	17.20% (17.04%, 17.37%)
Cancellation	11.96% (11.92%, 12.00%)	11.59% (11.5%, 11.68%)	11.41% (11.31%, 11.51%)	11.32% (11.2%, 11.45%)	11.21% (11.09%, 11.32%)
Rescheduling	16.71% (16.68%, 16.74%)	19.20% (19.1%, 19.3%)	19.24% (19.14%, 19.35%)	21.71% (21.58%, 21.84%)	21.61% (21.51%, 21.71%)
- Active	7.45% (7.41%, 7.49%)	10.47% (10.37%, 10.56%)	10.6% (10.5%, 10.7%)	13.41% (13.29%, 13.52%)	13.46% (13.37%, 13.55%)
- Passive	9.26% (9.22%, 9.29%)	8.73% (8.64%, 8.82%)	8.65% (8.56%, 8.73%)	8.30% (8.21%, 8.39%)	8.15% (8.06%, 8.24%)
Average wait time					
All	95.21 (95.02, 95.39)	97.35 (97.06, 97.64)	96.80 (96.46, 97.14)	100.07 (99.70, 100.45)	98.82 (98.46, 99.18)
New	37.38 (37.01, 37.74)	40.54 (40.05, 41.04)	39.62 (39.24, 40)	44.62 (43.86, 45.38)	41.77 (41.19, 42.35)
Follow-up	117.92 (117.76, 118.07)	119.23 (118.90, 119.56)	118.97 (118.58, 119.37)	120.98 (120.67, 121.30)	120.52 (120.07, 120.97)
Utilization	87.32% (87.06%, 87.58%)	88.53% (88.15%, 88.92%)	88.32% (87.97%, 88.67%)	89.65% (89.23%, 90.07%)	89.07% (88.69%, 89.44%)

Note: Numbers inside the parentheses indicate 95% confidence intervals.

References

Baiocchi, Mike, Dylan S Small, Scott Lorch, Paul R Rosenbaum. 2010. Building a stronger instrument in an observational study of perinatal care for premature infants. *Journal of the American Statistical Association* **105**(492) 1285–1296.

Baum, CF, ME Schaffer, Steven Stillman. 2002. IVREG2: Stata module for extended instrumental variables/2SLS and GMM estimation. *Statistical Software Components* .

Chan, Carri W, Linda V Green, Lijian Lu, Suparerk Lekwijit, Gabriel J Escobar. 2017. Assessing the impact of service intensity on customers: An empirical investigation of hospital step-down units. *Working Paper* .

Finlay, Keith, Leandro Magnusson, Mark E Schaffer, et al. 2016. WEAKIV: Stata module to perform weak-instrument-robust tests and confidence intervals for instrumental-variable (IV) estimation of linear, probit and tobit models. *Statistical Software Components* .

Freeman, Michael, Nicos Savva, Stefan Scholtes. 2016. Gatekeepers at work: An empirical analysis of a maternity unit. *Management Science, forthcoming* .

Hu, Wenqi, Carri W Chan, José R Zubizarreta, Gabriel J Escobar. 2017. An examination of early transfers to the ICU based on a physiologic risk score. *Manufacturing & Service Operations Management, forthcoming* .

Kim, Song-Hee, Carri W Chan, Marcelo Olivares, Gabriel Escobar. 2014. ICU admission control: An empirical study of capacity allocation and its implication for patient outcomes. *Management Science* **61**(1) 19–38.

Lorch, Scott A, Michael Baiocchi, Corinne E Ahlberg, Dylan S Small. 2012. The differential impact of delivery hospital on the outcomes of premature infants. *Pediatrics* **130**(2) 270–278.

Lu, Bo, Robert Greevy, Xinyi Xu, Cole Beck. 2011. Optimal nonbipartite matching and its statistical applications. *The American Statistician* **65**(1) 21–30.

Mikusheva, Anna, Brian P Poi, et al. 2006. Tests and confidence sets with correct size when instruments are potentially weak. *Stata Journal* **6**(3) 335–347.

Moreira, Marcelo J. 2009. Tests with correct size when instruments can be arbitrarily weak. *Journal of Econometrics* **152**(2) 131–140.

Rosenbaum, Paul R. 2010. *Design of Observational Studies*. Springer.

Rosenbaum, Paul R, Donald B Rubin. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* **70**(1) 41–55.

Stock, James H, Motohiro Yogo. 2005. *Testing for weak instruments in linear IV regression*. Cambridge University Press.

Yang, Fan, José R Zubizarreta, Dylan S Small, Scott Lorch, Paul R Rosenbaum. 2014. Dissonant conclusions when testing the validity of an instrumental variable. *The American Statistician* **68**(4) 253–263.