



This is a repository copy of *A probabilistic patient scheduling model for reducing the number of no-shows*.

White Rose Research Online URL for this paper:  
<https://eprints.whiterose.ac.uk/151553/>

Version: Accepted Version

---

**Article:**

Ruiz-Hernández, D., García-Heredia, D., Delgado-Gómez, D. et al. (1 more author) (2019) A probabilistic patient scheduling model for reducing the number of no-shows. *Journal of the Operational Research Society*, 71 (7). pp. 1102-1112. ISSN 0160-5682

<https://doi.org/10.1080/01605682.2019.1658552>

---

This is an Accepted Manuscript of an article published by Taylor & Francis in *Journal of the Operational Research Society* on 10th September 2019, available online:  
<http://www.tandfonline.com/10.1080/01605682.2019.1658552>

**Reuse**

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<https://eprints.whiterose.ac.uk/>

1 A probabilistic patient scheduling model for reducing the  
2 number of no-shows

3 Diego Ruiz-Hernández<sup>a</sup>     David García-Heredia<sup>b</sup>     David Delgado-Gómez<sup>b</sup>  
4 Enrique Baca-García<sup>c</sup>

<sup>a</sup>Sheffield University Management School, Sheffield, UK

<sup>b</sup>Department of Statistics, Universidad Carlos III de Madrid, Leganés, Spain

<sup>c</sup>Department of Psychiatry, Fundación Jiménez Díaz, Madrid, Spain

5 **Abstract**

6 Patients who do not attend their appointments, or “no-shows”, cause the under-  
7 utilisation of the health centres’ resources and increase the average waiting time for  
8 accessing specialty health care services. Although this problem has been addressed in  
9 different appointment scheduling models, behavioural issues associated to the patient’s  
10 socio-demographic and economic characteristics and/or his or her diagnosis, have not  
11 been widely included in scheduling optimisation models. In this article, we propose an  
12 integer linear programming model that takes into account such behavioural issues in or-  
13 der to reduce impact of no-shows in medical services. To achieve this goal, the objective  
14 function maximises the health centre’s expected revenue by using show-up probabilities  
15 estimated for each combination of patient and appointment slot. These behaviour-based  
16 probabilities are obtained using both the individual’s personal and clinical characteristics  
17 and his or her attendance history. In addition, the model takes into account the require-  
18 ments imposed by both the health centre’s management and the health authorities (e.g.  
19 distinguishing between first visits and follow-ups, among others). An extension of the  
20 model allows overbooking in some appointment slots. Experimental results show that  
21 the proposed model is capable of reducing the waiting list length by 13% and to attain  
22 an increase of about 5% in revenue when comparing to a basic model that assigns each

23 patient to the first available slot. It was also observed that when overbooking was allowed  
24 in one to three slots per day, the waiting list was reduced between 30% and 62%; and  
25 the revenue increased by 7% to 13%.

26 **Keywords:** appointment scheduling; no-shows; overbooking; healthcare; behavioural  
27 OR

## 28 1 Introduction

29 Over the past decades, there has been a considerable increase on health care expenditures  
30 worldwide. For instance, in the United States, the percentage of the GDP spent on health has  
31 increased from the 12.51% in 2000 to the 16.84% in 2015 (World Bank, 2018). A not negligible  
32 part of this expense is caused by the patients, commonly referred to as no-shows, who do not  
33 show-up for their appointments. For example, Moore et al. (2001) concluded that “over the  
34 course of a year, total revenue shortfalls [due to no-shows] could range from 3% to 14% of  
35 total clinic income”; likewise, Berg et al. (2013) estimated daily losses of about 16.5% of the  
36 revenue for a no-show rate of about 18%. In overall, McKee (2014) estimates that no-shows  
37 cost the American healthcare industry around 150 billion dollars per year. No-shows have  
38 also an important negative effect on the efficiency of health systems, causing under-utilisation  
39 of resources, long waiting lists and decreased revenue. The volume of no-shows depends on  
40 elements as disparate as the region, the patient’s socio-demographic characteristics, clinical  
41 diagnosis and prior no-show history, as well as the specialty and the type of service provided,  
42 among others (Dantas et al., 2018). In their literature review, Kheirkhah et al. (2015) refer  
43 reported no-show rates ranging from 3 to 80 percent. Along the same line, Moore et al.  
44 (2001) observed that no-shows and cancellations represent about 32.2% of scheduled time at  
45 a family planning residence clinic.

46 In order to reduce these figures, health centres utilise two alternative approaches. On  
47 one hand, the so-called active approaches include reminders and sanctions. The success of  
48 these methods is uncertain, with some research reports showing a drastic reduction in the  
49 percentage of no-shows after these measures are implemented (Molfenter, 2013), while others  
50 find no differences or, at most, a modest reduction (Hixon et al., 1999; Satiani et al., 2009).

51 This difference can be explained by the fact that the effectiveness of these methods may  
52 depend on the characteristics of the target population (Hashim et al., 2001). On the other  
53 hand, the so-called passive approaches aim at improving the current appointment system  
54 of the health centre by means of more sophisticated (and efficient) appointment assignment  
55 policies, instead of the most frequent practice of assigning the patient to the first available  
56 slot.

57 Optimising patient appointment systems has been an active subject of research over the  
58 last few decades (Cayirli and Veral, 2003; Gupta and Denton, 2008; Ahmadi-Javid et al.,  
59 2017). The patient allocation systems that have been proposed in the scientific literature  
60 present several differences, which are mainly consequence of the specific characteristics of  
61 the health centre and the type of service provided. For example, some centres establish that  
62 patients must receive their appointment at the time when this is requested, while in other  
63 cases appointments are scheduled at the end of certain period (the patient is notified later  
64 on by physical or electronic means). These two approaches are usually named as online and  
65 offline, respectively (Zacharias and Pinedo, 2014). Although online systems are the most  
66 frequently used, the rapid development of electronic appointment systems has caused an  
67 increase in the relevance of offline systems (Ahmadi-Javid et al., 2017). Another difference is  
68 whether the scheduling system admits overbookings or not, although most of the proposed  
69 systems include overbooking in their models (LaGanga and Lawrence, 2007; Chakraborty  
70 et al., 2010; Kim and Giachetti, 2006). A more detailed description of the different types  
71 of appointment systems can be found in the recent review conducted (Ahmadi-Javid et al.,  
72 2017).

73 Notwithstanding there is evidence that the probability that a patient will show-up to an  
74 appointment is closely related to his or her socio-demographic characteristics and condition  
75 (Dantas et al., 2018), traditional appointment scheduling models for medical services are  
76 usually based on the availability of slots, practitioner's timetables, and visit times, among  
77 other characteristics of the service provided. Only seldom, the proposed models take into  
78 consideration the probability that a patient will attend an appointment in a given time  
79 window. Moreover, those models tend to allocate probabilities based in generic data without

80 taking into account characteristics and behavioural traits specific to each patient.

81 This article constitutes an effort for bringing the field of behavioural operational re-  
82 search to the area of patient schedulling, by proposing an appointment planning method  
83 that takes into account each individual's probabilities of no-show (estimated from their socio-  
84 demographic characteristics, diagnose and attendance history) for each specific combination  
85 of time-slot and patient. In doing so, our work seeks to fill a gap existing in the application of  
86 OR in healthcare (comprehensive reviews include those by Brailsford et al. (2009) and Hulshof  
87 et al. (2012)) through the development of behaviourally informed approaches (Hämäläinen  
88 et al., 2013), that aim at improving the provision of medical services by including associated  
89 patient's behaviour in the modelling process.

90 In this article, an integer linear programming (ILP) model is developed for optimising the  
91 offline assignment of medical appointments in a speciality service of a public health centre.  
92 The system aims at minimising the number of no-shows, and indirectly the waiting list length.  
93 This is attained by means of an objective function that maximising the expected revenue of  
94 the health centre. The model is designed as a single server system accounting for the fact  
95 that, in general, each practitioner has his own list of patients. Finally, as the health centre  
96 may be required by law to serve a fixed proportion of new patients every week, the model  
97 includes the possibility of reserving a percentage of slots for first-visits.

98 Under certain conditions, in order to reduce the large number of practitioners' idle periods  
99 caused by no-shows, a health centre may consider the possibility of introducing overbooking  
100 in some slots. This may also have a positive impact on the length of the waiting list (mainly in  
101 centres with large incidence of no-shows). For those cases, we propose a mixed integer linear  
102 programming (MILP) formulation that extends the initial model by allowing overbooking in  
103 a limited (pre-defined) number of slots.

104 Before introducing the mathematical formulation of the system, in Section 2, we provide a  
105 brief description of some related approaches available in the literature. In section 3 we present  
106 the proposed mathematical model. In Section 4 we conduct a simulation experiment in order  
107 to test our model's performance. We conclude this article in Section 5 with a discussion of  
108 the results and pointing out future lines of research.

## 109 2 Related literature

110 As mentioned above, several models have been proposed for improving patients' access to  
111 health care. The differences in these models are mainly consequence of the heterogeneity  
112 of the requirements imposed by the health centres (e.g. online or offline scheduling, single  
113 or multiple servers or if no-shows should be taken into account) and the goals pursued (e.g.  
114 maximise the revenue or reduce the length of the waiting list). In this section we focus  
115 our discussion on the analysis of those models most closely related to our work: first, we  
116 discuss the offline mathematical programming models (either ILP or MILP) proposed for  
117 single server systems; later, we present a review of some of the most relevant works that take  
118 into account the presence of no-shows from a probabilistic perspective.

119 Conforti, Guerreiro and Guido developed various ILP models that maximise the number  
120 of patients –weighted by the severity of their illnesses- scheduled for starting a radiotherapy  
121 treatment (Conforti et al., 2008, 2011). Their models assign each patient to several time slots  
122 during a given number of weeks so that the treatment can be conducted without interruptions.  
123 This assignment is conducted taking into account the constraints generated by patients that  
124 have already started the treatment. Zhu et al. developed a similar model for scheduling the  
125 access to a Magnetic Resonance Imaging scanner (Zhu et al., 2012). Their model assigns  
126 the patients to the required time slots in a two-week schedule so the number of allocated  
127 patients, weighted by their priorities, is maximised. Their model takes into account patients'  
128 time availability. Wang and Fung developed a model aiming at maximising profit, measured  
129 as the revenue earned from the attended patients minus the cost incurred from patients'  
130 rejection (Wang and Fung, 2014). The revenue was dependent on the patients' preferences  
131 for appointment time and practitioner. Additionally, a constraint was included for limiting  
132 the degree of discrepancy between the time allocated and the patient's preferences. More  
133 recently, Wiesche, Schacht and Werner proposed a MILP model that seeks to minimise the  
134 number of assigned appointments, penalising the number of patient shifted from morning to  
135 afternoon sessions (Wiesche et al., 2017). This allowed the authors, in one hand, to increase  
136 the time availability for attending walk-ins, and to balance the physicians' workload, on the  
137 other.

138        However, the models discussed above do not consider the existence of no-shows. In this  
139 regard, Savelbersbergh and Smilowicz developed an ILP model whose objective function  
140 aimed at maximising the health condition of the population in a mobile asthma management  
141 program (Savelbersbergh and Smilowicz, 2016). The health condition was measured by the  
142 likelihood that a patient’s disease was controlled, which was strongly related to the probability  
143 that the patient showed-up to his appointment. The authors defined no-show probabilities  
144 for six different categories of patients depending on their preferences (strong or weak) for  
145 three different time windows (AM, noon, or PM) and 8 time slots in each time window. To  
146 our knowledge, this is the only offline ILP model that, although implicitly, takes into account  
147 the existence of no-shows.

148        Regarding the works that include no-show information from a probabilistic point of view,  
149 we find that most of them are developed from an on-line perspective and formulated as  
150 Stochastic Programming or Markov Decision Problems (Ahmadi-Javid et al., 2017). For ex-  
151 ample, Muthuraman and Lawley developed a stochastic overbooking model that considered  
152 each patient’s no-show probability (Muthuraman and Lawley, 2008). The objective function  
153 aimed at maximising the revenue penalised by an overbooking cost, represented by the pa-  
154 tient’s waiting time and staff’s overtime. This model was later tested by Daggy et al. on  
155 real data where the no-show probabilities were estimated applying a logistic regression to a  
156 dataset obtained from a Veterans Affairs medical centre (Daggy et al., 2010). In a different  
157 work, Glowacka, Henry and May estimated the probabilities that a patient will show-up to  
158 his or her appointment by means of an association rule mining technique (Glowacka et al.,  
159 2009). They used these probabilities to derive three manageable sets of rules for patient  
160 scheduling. Recently, Samorani and Laganga have proposed an online scheduling model that  
161 admits overbooking, and whose objective function aims at maximising the revenue penalised  
162 by the patients’ waiting time and overtime cost. Instead of a probabilistic classifier, they use  
163 a binary one to maintain their problem computational tractable (Samorani and LaGanga,  
164 2015).

165        The model proposed in this article extends the available literature in appointment schedul-  
166 ing for health centres in the following directions. First, unlike most of the mathematical

167 programming-based research, our model takes into consideration the likelihood that a pa-  
168 tient will not show-up to his or her appointment. Secondly, our formulation adopts an off-line  
169 approach that uses differentiated show-up profiles for each patient. These show-up profiles,  
170 that provide a specific show-up probability for each available slot, are obtained using socio-  
171 demographic and clinical characteristics of the patient. This is an important difference with  
172 respect to other available probabilistic work, which uses predominantly on-line approaches  
173 and/or where the no-show probabilities are either categorised (Savelsbergh and Smilowitz,  
174 2016) or binarised (Samorani and LaGanga, 2015). A third characteristic is that, unlike other  
175 works that consider first visits and follow-ups as homogeneous groups or, plainly, ignore the  
176 first visit group (Daggy et al., 2010), our formulation distinguishes among them, allowing  
177 the model, apart from satisfying a legal requirement, to exploit the different characteristics  
178 of these groups. Finally, our model allocates priorities to the patients depending on the time  
179 they have been in the waiting list.

### 180 **3 The Probabilistic Patient Scheduling Problem**

181 In this section, we introduce a probabilistic scheduling model for reducing no-shows in spe-  
182 cialty health centres that takes into consideration patient-specific probabilities of showing  
183 at each given day/time slot. The objective is maximising the centre’s expected revenue by  
184 means of a reduction in the number of no-shows. The model distinguishes between two types  
185 of patients (first visits and follow-ups) and, by using a priority value associated to each pa-  
186 tient, takes into account the time that the patient has remained in the waiting list. It also  
187 takes into account a Spanish legal constraint regarding the proportion of first visits that must  
188 be scheduled every week.

189 The following notation will be used in the mathematical formulation of the model.

190 Sets

191  $\mathcal{I}$ , days of the week;

192  $\mathcal{J}$ , time slots available in any given day;

193  $\mathcal{K}$ , set of patients to be scheduled for appointment during the reference week.



194 Parameters

- 195  $q$ , proportion of the number of available slots that must be allocated to first visits;
- 196  $d_k$ , binary parameter indicating if patient  $k \in \mathcal{K}$  has high ( $d_k = 0$ ) or low ( $d_k = 1$ ) priority  
 197 during the reference week;
- 198  $Z_k$ , binary parameter indicating if patient  $k \in \mathcal{K}$  is a first visit ( $Z_k = 1$ ) or a follow-up  
 199 ( $Z_k = 0$ );
- 200  $P_{i,j,k}$ , probability that patient  $k \in \mathcal{K}$  will show-up to an appointment in slot  $\{i, j\}$ , for all  
 201  $i \in \mathcal{I}$  and  $j \in \mathcal{J}$ ;
- 202  $w_z$ , revenue obtained either from a first visit ( $z = 1$ ), or a follow-up ( $z = 0$ ).

203 Variables

- 204  $X_{i,j,k}$ , binary variable taking value 1 if patient  $k \in \mathcal{K}$  is assigned to slot  $\{i, j\}$ , for all  $i \in \mathcal{I}$   
 205 and  $j \in \mathcal{J}$ .
- 206  $X_k^T$ , binary variable taking value 0 if patient  $k \in \mathcal{K}$  is assigned a slot in the current week  
 207 and 1 if the patient is referred back to the waiting list.

208 With this notation, and taking into account that the operator  $\lceil \cdot \rceil$  rounds a real number  
 209 to its upper integer value, the model is formulated as follows:

$$\max \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \sum_{k \in \mathcal{K}} X_{i,j,k} P_{i,j,k} (Z_k w_1 + (1 - Z_k) w_0) \quad (1)$$

$$\text{s.t.} \quad \sum_{k \in \mathcal{K}} X_{i,j,k} \leq 1, \quad \forall i \in \mathcal{I}, j \in \mathcal{J} \quad (2)$$

$$\sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} X_{i,j,k} + X_k^T = 1, \quad \forall k \in \mathcal{K} \quad (3)$$

$$\sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \sum_{k \in \mathcal{K}} X_{i,j,k} Z_k \geq \min \left\{ \sum_{k \in \mathcal{K}} Z_k, \lceil q|I||J| \rceil \right\}, \quad (4)$$

$$\sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \sum_{k \in \mathcal{K}} X_{i,j,k} (1 - d_k) \geq \min \left\{ \sum_{k \in \mathcal{K}} (1 - d_k), \right. \quad (5)$$

$$\left. |I||J| - \min \left\{ \sum_{k \in \mathcal{K}} Z_k, \lceil q|I||J| \rceil \right\} \right\},$$

$$X_{i,j,k}, X_k^T \in \{0, 1\}, \quad \forall i \in \mathcal{I}, j \in \mathcal{J}, k \in \mathcal{K}. \quad (6)$$

210 The objective function maximises the clinic's expected revenue. Notice that when  $w_0 =$   
 211  $w_1 = w$  the objective function maximises the expected showing-up rate; otherwise, the  
 212 objective maximises the expected weighted showing-up rate. Constraints (2) guarantee that  
 213 only one patient is assigned to each slot. Constraints (3) make sure that if a patient is not  
 214 allocated in the current week, he or she is returned to the waiting list. As we are working  
 215 with binary variables, constraints (3) also ensure that each patient is not allocated in more  
 216 than one slot. Constraint (4) forces to reserve a number of slots for the first time visits.  
 217 Constraint (5) guarantees that low priority patients will not be allocated to a slot as long as  
 218 there are high priority patients unallocated.

## 219 **Model with overbooking**

220 As mentioned in the Introduction, there may be cases in which performing overbooking is  
 221 considered convenient. For these situations, the baseline model is extended for allowing the  
 222 possibility of assigning two patients to the same slot, provided that the sum of their showing-  
 223 up probabilities is less than certain predetermined value. This is attained by introducing an  
 224 overbooking penalty in the objective function and a number of associated constraints. The  
 225 following additional notation is used in the extended model:

### 226 Parameters

227  $C^{ov}$ , positive parameter representing the overbooking penalty;

228  $M$ , constant satisfying  $M > \max \{w_0, w_1\}$ ;

229  $G_{i,j}$ , binary parameter taking value 1 if overbooking is allowed in slot  $\{i, j\}$ , for all  $i \in \mathcal{I}$   
 230 and  $j \in \mathcal{J}$ ;

231  $\pi_{i,j}$ , parameter imposing a bound on the sum of the showing-up probabilities for any pair of  
 232 patients simultaneously booked in slot  $\{i, j\}$ , for all  $i \in \mathcal{I}$  and  $j \in \mathcal{J}$ <sup>1</sup>.

233 Variables

234  $Y_{i,j}$ , binary variable taking value 1 if overbooking has been used in slot  $\{i, j\}$ , for all  $i \in \mathcal{I}$   
 235 and  $j \in \mathcal{J}$ ;

236  $O_{i,j}$ , binary variable taking value 1 if at least one patient has been booked in slot  $\{i, j\}$ , for  
 237 all  $i \in \mathcal{I}$  and  $j \in \mathcal{J}$ .

The model with overbooking is then given by:

$$\max \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \sum_{k \in \mathcal{K}} X_{i,j,k} P_{i,j,k} (Z_k w_1 + (1 - Z_k) w_0) - C^{ov} Y_{i,j} + M O_{i,j} \quad (7)$$

s.t. Constraints (3)-(5), and

$$Y_{i,j} \leq G_{i,j}, \quad \forall i \in \mathcal{I}, j \in \mathcal{J} \quad (8)$$

$$\sum_{k \in \mathcal{K}} X_{i,j,k} \leq 1 + Y_{i,j}, \quad \forall i \in \mathcal{I}, j \in \mathcal{J} \quad (9)$$

$$\sum_{k \in \mathcal{K}} X_{i,j,k} P_{i,j,k} \leq \pi_{i,j}, \quad \forall i \in \mathcal{I}, j \in \mathcal{J} \quad (10)$$

$$O_{i,j} \leq \sum_{k \in \mathcal{K}} X_{i,j,k}, \quad \forall i \in \mathcal{I}, j \in \mathcal{J} \quad (11)$$

$$X_{i,j,k}, X_k^T, Y_{i,j}, O_{i,j} \in \{0, 1\}, \quad \forall i \in \mathcal{I}, j \in \mathcal{J}, k \in \mathcal{K} \quad (12)$$

238 The  $C^{ov} Y_{i,j}$  term in the extended objective function, equation (7), represents a penalty  
 239 incurred when overbooking is used in slot  $\{i, j\}$ . Please notice that this term attains the  
 240 largest possible reduction in the practitioners' idle times by allocating the same slot (if  
 241 overbooking is admissible) to the pair of patients with highest sum of show-up probabilities.  
 242 This is guaranteed by the fact that the larger the weighted sum of show-up probabilities,  
 243 the larger the profit after discounting the overbooking cost for any given slot. Notice also  
 244 that if  $C^{ov} < \{w_0, w_1\} \times \min_{i,j,k} \{P_{i,j,k}\}$ , the model will always use overbooking when  $|\mathcal{K}| >$

<sup>1</sup>Notice that the probability of both patients showing-up is given by  $P_{i,j,k} \cdot P_{i,j,k'}$ , which attains a maximum at  $P_{i,j,k} = P_{i,j,k'} = \frac{\pi_{ij}}{2}$  for any given value of  $\pi_{ij}$ .

245  $|\mathcal{I}||\mathcal{J}|$ , i.e. whenever the number of patients in the waiting list is larger than the number of  
 246 available slots. Likewise, if  $C^{ov} > \max\{w_0, w_1\} \times \max_{i,j,k} \{P_{i,j,k}\}$ , the model will never use  
 247 overbooking.

248 Regarding the associated constraints, equations (8) define the slots where overbooking  
 249 is allowed. Equations (9) limit the number of overbooked patients in a given slot to two,  
 250 provided that overbooking is allowed. Finally, in order to control for the maximal probability  
 251 of overcrowding (the case where two overbooked patients show-up for the same appointment),  
 252 the sum of showing-up probabilities in an overbooked slot is bounded by parameter  $\pi_{ij}$  in  
 253 constraints (10).

254 Term  $MO_{ij}$  in equation (7), together with constraints (11) and the fact that by definition  
 255  $M > \max\{w_0, w_1\}$ , ensures that the model does not consider overbooking unless all slots are  
 256 used.

257 Additionally, our model presents the following two properties, which will be used in the  
 258 computational implementation of the model for speeding up the execution:

259 **Proposition 3.1.** *In the model with overbooking, the  $O_{i,j}$  variables always take integer values  
 260 when they are relaxed to  $0 \leq O_{i,j} \leq 1$  for all  $i \in \mathcal{I}$ ,  $j \in \mathcal{J}$ .*

261 *Proof.* Let  $O_{ij}$  be a continuous variable defined in the interval  $[0, 1]$  for all  $i \in \mathcal{I}$ ,  $j \in \mathcal{J}$ . If  
 262  $\sum_{k \in \mathcal{K}} X_{i,j,k} = 0$ , from constraint (11) it immediately follows that  $O_{ij} = 0$ . Alternatively, if  
 263  $\sum_{k \in \mathcal{K}} X_{i,j,k} > 0$  and given that the  $X_{ijk}$  are binary variables, then  $O_{ij}$  can take any value  
 264 in the interval  $[0, 1]$ . However, given that  $MO_{ij}$  appears with positive sign in the objective  
 265 function of the maximisation problem, it follows that  $O_{ij} = 1$ .  $\square$

266 **Proposition 3.2.** *In the model with overbooking, the  $Y_{i,j}$  variables always take integer values  
 267 when they are relaxed to  $0 \leq Y_{i,j} \leq 1$  for all  $i \in \mathcal{I}$ ,  $j \in \mathcal{J}$ .*

268 *Proof.* Let  $Y_{i,j}$  be a continuous variable defined in the interval  $[0, 1]$  for all  $i \in \mathcal{I}$ ,  $j \in \mathcal{J}$ . We  
 269 consider two possible scenarios:

- 270 1.  $|K| \leq |I||J|$ : From the objective function it follows directly that  $Y_{ij} = 0$  for all  $i \in$   
 271  $\mathcal{I}$ ,  $j \in \mathcal{J}$ . Double booking any slot, when a number of slots remains unallocated, will  
 272 imply unnecessarily incurring a penalty of  $C^{ov}$ .

273 2.  $|K| > |I||J|$ : Consider any given slot  $\{i, j\}$ . If overbooking is not allowed,  $G_{i,j} = 0$ ,  
274 constraints (8) guarantee that  $Y_{ij} = 0$ .

275 Assume now that overbooking is allowed and conducted at some slot  $\{i, j\}$ , i.e.  $G_{i,j} = 1$   
276 and  $\sum_{k \in \mathcal{K}} X_{i,j,k} = 2$ . Let  $Y_{i,j} = \delta$  with  $0 < \delta < 1$ , satisfying constraints (8). From  
277 constraints (9) it follows that  $\sum_{k \in \mathcal{K}} X_{i,j,k} \leq 1 + \delta$ , and given that  $X_{i,j,k} \in \{0, 1\}$  we  
278 conclude that  $\sum_{k \in \mathcal{K}} X_{i,j,k} \leq 1$ , which is a contradiction. Therefore, if slot  $\{i, j\}$  is  
279 overbooked, then necessarily  $Y_{ij} = 1$ .

280 Finally, assume that overbooking is allowed but not conducted at some slot  $\{i, j\}$ , i.e.  
281  $G_{i,j} = 1$  and  $\sum_{k \in \mathcal{K}} X_{i,j,k} = 1$ . Let  $Y_{i,j} = \delta$  with  $0 < \delta < 1$ , satisfying constraints (8). As  
282 before, constraints (9) imply that  $\sum_{k \in \mathcal{K}} X_{i,j,k} \leq 1 + \delta$ , and given that  $X_{i,j,k} \in \{0, 1\}$  it  
283 still holds that  $\sum_{k \in \mathcal{K}} X_{i,j,k} \leq 1$ . Now, given that  $C^{ov}Y_{ij}$  appears with negative sign in  
284 the objective function of the maximisation problem, it follows that  $Y_{ij} = 0$ . Therefore,  
285 if slot  $\{i, j\}$  is not overbooked, then immediately  $Y_{ij} = 0$ .

286 □

### 287 **Comment**

288 If instead of a penalty for overbooking, an expected cost for overcrowding was the driver  
289 behind the overbooking decision, the corresponding term in the objective function –and the  
290 associated constraints– will need to incorporate the overcrowding probability (the product  
291 of the attendance probabilities of the two overbooked patients). In this case, the objective  
292 function will seek to allocating the overbooked slots to the pair of patients with lowest product  
293 of show-up probabilities. Consequently, with the aim of minimising the overcrowding penalty,  
294 the system will still face large idle times (as the probability of none of the patients showing-up  
295 will still be large). Moreover, the problem will become non-linear.

### 296 **3.1 Scheduling Procedure**

297 The scheduling procedure works as follows:

- 298 1. A waiting list is available with the records of the patients waiting for appointment,  
299 including information about the number of weeks they have been in the list (sojourn)

300 and whether it is a first-time visit or not. New patients are added to the list at the  
301 time the appointment request is received and their sojourn length counter is initialised  
302 to zero.

303 2. The list of patients (henceforth referred to as the *buffer*) to be passed each week to the  
304 scheduler is built as follows:

305 (a) The system first selects the patients with largest sojourn value and assigns them  
306 high priority ( $d_k = 0$ ). This group contains both first-visits ( $Z_k = 1$ ) and follow-  
307 ups ( $Z_k = 0$ ).

308 (b) Once the high priority patients have been selected, if the number of first-visits  
309 in the buffer is still below the legal requirement, the system sequentially adds  
310 first-visits in decreasing order of sojourn length until the requirement is satisfied  
311 or no more first visits are left in the waiting list. At each iteration, all first-visits  
312 in the corresponding sojourn level are included. These patients have low priority  
313 ( $d_k = 1$ ) and  $Z_k = 1$ .

314 (c) Finally, if after including high priority patients and first-visits, the number of  
315 patients in the buffer is smaller than the number of available slots (and there  
316 are still patients in the waiting list), the system sequentially adds patients in  
317 decreasing order of sojourn length until the size of the buffer is larger or equal to  
318 the number of available slots (or the waiting list is empty). At each iteration, all  
319 patients in the corresponding sojourn level are included. These patients have low  
320 priority ( $d_k = 1$ ).

321 3. After this selection has been conducted, the system passes the list of candidates to  
322 the scheduler for solving the Probabilistic Patient Scheduling Problem with or without  
323 overbooking. Once the schedule has been obtained, the patients who did not receive  
324 an appointment are sent back to the waiting list with their original sojourn value.

325 Regarding the overbooking policy, whenever two patients show up for the same appoint-  
326 ment, subsequent appointments are delayed until either a no-show happens and the last

327 delayed patient takes that slot, in which case the original schedule is reestablished, or the  
328 day finishes and the practitioner does over-time until the list is cleared. Please notice that the  
329 over-time impact of this policy will be limited as long as the number of slots where overtime  
330 is admissible does not exceed a reasonable limit (e.g. no more than 2 or 3 slots).

## 331 4 Numerical Experiments

332 In order to evaluate the performance of our model, an experiment that reproduces the routine  
333 of a psychiatry department in a Spanish health centre was designed. In order to estimate the  
334 probabilities that the patients would show-up for their appointments, a database containing  
335 information from 47,118 visits to this department was used. In addition to the variable  
336 indicating whether the patient attended the appointment or not, this database contains  
337 several variables that have been frequently used to characterise non-shows. These variables  
338 were age (Alaeddini et al., 2011; Kopach et al., 2007), sex (Alaeddini et al., 2011), week day  
339 and time of the appointment (Glowacka et al., 2009; Daggy et al., 2010), lead time (time in  
340 queue) in weeks (Daggy et al., 2010), practitioner ID, appointment type (first visit or follow-  
341 up) (Kopach et al., 2007), number of previous appointments (Kopach et al., 2007; Daggy  
342 et al., 2010), and percentage of no-shows in previous appointments (Kopach et al., 2007;  
343 Daggy et al., 2010). The probabilities of show-up were obtained using a decision tree (Norris  
344 et al., 2014) classifier. The use of the database allowed us to obtain specific and differentiated  
345 attendance probabilities for each available appointment slot, provided the patient’s profile.  
346 The simulation is conducted as follows:

- 347 1. At the beginning of each week, we generate the set of patients who call for a new  
348 appointment. To do this, a random number is generated according to a discrete uni-  
349 form variable whose parameters are provided below. This number is used for randomly  
350 selecting a number of patients from our database. By doing this, we respect the pro-  
351 portion of first visits/follow-ups as well as the distribution of the variables representing  
352 the patients’ characteristics. The selected patients are added-up at the end of the wait-  
353 ing list. Each patient in the waiting list has assigned a sojourn value representing the

354 number of weeks that he or she has remained in the list. New arrivals are all assigned  
355 a sojourn value equal to zero.

356 2. The list of patients to be passed each week to the scheduling model is built as described  
357 in item 2 of Section 3.1.

358 3. After this selection has been performed, the Probabilistic Patient Scheduling Prob-  
359 lem is solved using the generated data<sup>2</sup>. Once the model makes the assignment, the  
360 parameters of the system are updated for the following week as follows:

361 For each appointment we randomly determine whether the patient will show-up or not  
362 depending on the patient’s estimated attendance probability given the allocated slot. If  
363 the patient shows-up to the appointment, the health centre obtains the corresponding  
364 income and the patient is removed from the system. Otherwise, the patient is either  
365 returned to the waiting list according to a predetermined probability, or eliminated  
366 from the system. If returned, the patient is put at the end of the list with sojourn  
367 value 0. This way, the experiment mimics the situation in which the patient that did  
368 not attend an appointment asks for a new one.

369 Patients who did not receive an appointment are sent back to the waiting list with their  
370 initial sojourn time.

371 4. At the end of each scheduling stage, the sojourn values of all patients in the waiting  
372 list are increased by one.

#### 373 **4.1 Simulation framework**

374 As we mentioned, our experiment reproduces the functioning of a psychiatry department  
375 week by week during one year (52 weeks). In this centre, the doctor does consultation from  
376 8:30 to 15:30 from Monday to Friday and each consultation lasts 30 minutes. Therefore,  
377 if overbooking is not considered, the doctor would attend a maximum of 70 patients. Of  
378 those, at least 30% are first visits to comply with the regulatory requirements. For each first

---

<sup>2</sup>We solved the optimisation problems using Cplex 12.7.



379 visit the centre receives 70 euros and for each revision 50 euros. At the beginning of the  
380 simulation, it is assumed that there is a 7-week waiting list to access to the medical services.  
381 For each of the simulated weeks, the following operations are performed:

382 The weekly number of new requested appointments in the simulation follows a uniform  
383 [51, 69] distribution. This choice, together with an estimated no-show rate of 24% and the  
384 60% of no-shows who are referred back to the waiting list, returns an expected number of  
385 appointment requests of 68.64 per week. These figures guarantee that the weekly number  
386 of patients asking for a new appointment is always close to the 70 slots available for each  
387 practitioner.

388 Using this scenario, we test the following scheduling approaches:

- 389 1. The probabilistic scheduling model without overbooking.
- 390 2. The extended model with overbooking in three different situations: allowing overbook-  
391 ing just at 12:00 each day, allowing overbooking at 9:00 and at 12:00; and allowing  
392 overbooking at 9:00, at 10:00 and at 12:00. The reason why we chose these hours is  
393 because, in our database, they are the time-slots with the greatest number of no-shows.  
394 From hereafter, they will be referred as one, two and three rows of overbooking, re-  
395 spectively. In all of them, the parameter that limits the maximum expected number of  
396 patients  $\pi_{i,j}$  is set to 1.5. Later, we will perform a sensitivity analysis to analyse the  
397 influence of this parameter.
- 398 3. The traditional model in which each patient is assigned to the first available slot (Daggy  
399 et al., 2010). We will refer this model as a FIFO system.

## 400 4.2 Results

401 Figure 1 shows the obtained results. In these plots, the legends “NoOver”, “Over1”, “Over2”  
402 and “Over3” stand for model without overbooking, and model with one, two and three rows  
403 of overbooking respectively.

404 Figure 1 (a) displays the number of people in the waiting list along the different weeks.  
405 It can be seen that the models which use overbooking obtain a fast reduction of the length

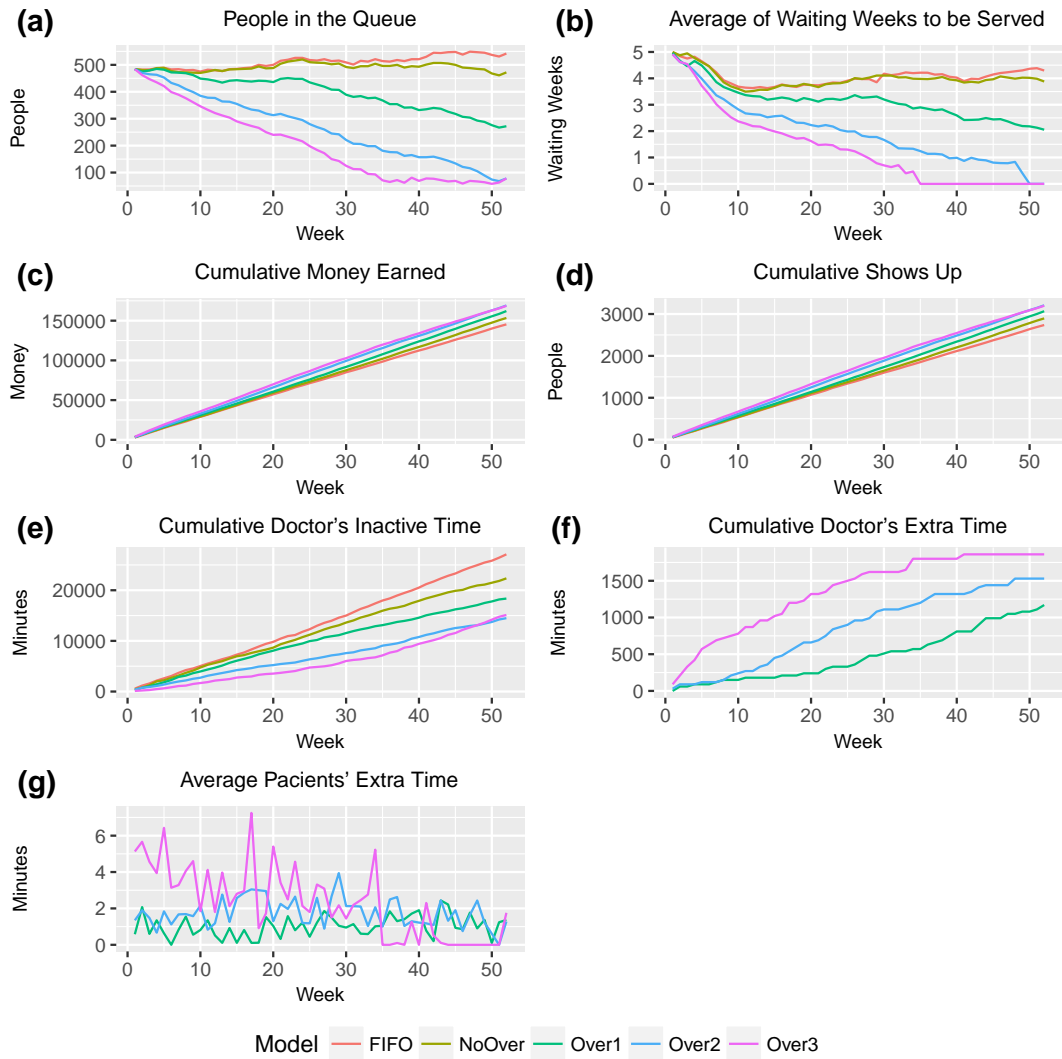


Figure 1: Simulation results.

406 of this list. It can be notice that the model which uses three rows of overbooking handles  
407 to eliminate the waiting list by week 35. After this week, the queue length is stable. One  
408 important result is that the proposed model that does not use overbooking (NoOver) is able  
409 to maintain the queue stable while the FIFO model cannot. At the end of the experiment,  
410 in week 52, the difference on the number of patients in the waiting list of these two models  
411 is 70 patients, which represents the complete schedule for a week. This result indicates that  
412 by only improving the patient assignment, without considering overbooking, is possible to  
413 avoid that the length of the waiting list increases.

414 Figure 1 (b) exhibits the mean time that the patients remain in the waiting list. These  
415 results are similar to the previous ones: the overbooking models reduce the mean time faster  
416 than the other two models, and the Over3 model stabilised around week 35. As before, the  
417 NoOver model attains to stabilised and the FIFO model do not. The drastic drop-out during  
418 the first weeks is consequence of the initial waiting list structure.

419 Figure 1 (c) shows the cumulative revenue. As expected, the models that have the greatest  
420 incomes are the overbooking models, followed by the NoOver and the FIFO models.

421 Figure 1 (d) illustrates the cumulative number of people who show up to the appointment.  
422 It can be noticed that this value is greater for the overbooking models. It is sensible to think  
423 that it is consequence of the fact that these models assign more patients, but it is also  
424 because these patients are optimally scheduled. This fact can also be appreciated in the  
425 NoOver and FIFO models. Despite they have the same number of assigned patients, the  
426 number of patients who show up is higher for the NoOver model.

427 Figure 1 (e) shows the cumulative doctor's idle time. The first interesting fact is that  
428 this value is higher in the FIFO model than in the NoOver model even though doctors in  
429 these models have assigned the same number of patients. It was observed that FIFO had  
430 an average of 3.5 empty slots per day, while NoOver just 2.8. Regarding the overbooking  
431 models, it can be observed that, at the end of the simulation, models Over2 and Over3 have  
432 a similar cumulative doctor's idle time. This is, again, consequence that after week 35, the  
433 length of the waiting list for the Over3 model is minimal.

434 Figure 1 (f) displays doctor's overtime in which the effect of adding an extra row to the

435 overbooking model can be appreciated. It is important to differentiate the curves before and  
436 after week 35 because, for the Over3 model, the waiting list is practically zero after this week  
437 as it was commented before. Therefore, after week 35, doctor barely suffer from over time in  
438 this model.

439 Figure 1 (g) represents the average time that each patient waited in the health centre to  
440 be attended. It can be observed that for the Over3 model, the patients have to wait between  
441 2 and 6 minutes, which is a 6% and 20% of the time of each slot.

442

443 Next, a sensibility analysis is conducted in order to assess the effect of parameter  $\pi$   
444 in the model's performance. To this end, the previous experiment is repeated for values  
445  $\{1.1, 1.3, 1.5, 1.7\}$  of this parameter. Table 1 shows the obtained results.

	Number in queue	Waiting weeks	Revenue (€)	Average show-ups (%)	Average empty slots	Average doctors' weekly overtime (min)	Average patients' extra waiting time (min)
FIFO	542	4.29	145,330	75.19	3.47	0	0
NoOver	472	3.88	153,430	79.53	2.86	0	0
Over1, $\pi_{i,j} = 1.1$	382	2.97	156,340	79.49	2.66	3.4	0.75
Over2, $\pi_{i,j} = 1.1$	262	1.90	161,200	77.69	2.29	5.1	2.35
Over3, $\pi_{i,j} = 1.1$	147	0.90	164,970	74.29	2.06	6.85	4.05
Over1, $\pi_{i,j} = 1.3$	280	2.14	161,990	78.48	2.33	18.85	3.95
Over2, $\pi_{i,j} = 1.3$	107	0.75	169,060	77.63	1.85	29.15	8
Over3, $\pi_{i,j} = 1.3$	84	0.15	168,190	76.40	1.91	33.45	9
Over1, $\pi_{i,j} = 1.5$	272	2.05	162,090	78.64	2.35	16.25	4.4
Over2, $\pi_{i,j} = 1.5$	78	0	169,140	77.50	1.85	36	9.6
Over3, $\pi_{i,j} = 1.5$	78	0	168,640	77.72	1.93	51.40	15.85
Over1, $\pi_{i,j} = 1.7$	263	1.87	163,270	79.20	2.26	20.55	5.05
Over2, $\pi_{i,j} = 1.7$	83	0.14	168,780	77.60	1.91	41.15	12.25
Over3, $\pi_{i,j} = 1.7$	83	0	168,530	77.40	2	63.5	20.05

Table 1: Results of the sensitivity analysis.

446 Some remarks on Table 1: i) values in the first three columns correspond to week 52; ii)

447 values displayed in columns four and five are averages over 52 weeks; and iii) average values  
448 in the last two columns are calculated over the first 35 weeks to avoid the noise caused by  
449 the exhaustion of the waiting list (please see the comments around model Over3 earlier in  
450 this section). Moreover, for the sake of clarity, we report the average number of empty slots  
451 per day instead of doctor's idle time.

452 We notice that the model without overcrowding (NoOver) increases the centre's revenue  
453 in 5.5% with respect to the current policy (FIFO), reducing the waiting list by 13% in  
454 a year. The results show that a scheduling regime that assigns appointments taking into  
455 consideration the patient's characteristics may contribute –in the health centre under study–  
456 to a reduction of about 17.5% in the number of empty slots.

457 Regarding the overbooking model, the results depend on the value assigned to parameter  
458  $\pi$  and the number of slots in which the overbooking is allowed ( $G_{ij} = 1$ ). If, for instance,  
459  $\pi = 1.1$ , it can be noticed that the impact on practitioners and patients is minimal. This is  
460 due to two main reasons: i) the small probability of overcrowding (two overbooked patients  
461 showing-up to the same appointment), 0.3 maximum; and ii) in the case of overcrowding, it  
462 occurs early enough for a no-show in later hours to compensate for the extra time devoted  
463 to attending the additional patient. For this value of  $\pi$  the revenue would increase in a  
464 range between 7% and 13% and the waiting list would be reduced from 30% to 72%. These  
465 values are consistent with the ones reported by Moore et al. (2001). We also notice that  
466 allowing overbooking always improves the health centre's revenue (with respect to the NoOver  
467 case), with the maximal revenue attained when overbooking is allowed in up to two slots  
468 (Over2). Moreover, allowing overbooking in two slots always reduces the number of empty  
469 slots, irrespectively of the value of  $\pi$ .

470 Finally, regarding the value of parameter  $\pi$ , the best results are obtained when this  
471 parameter takes values between 1.3 and 1.5. In those cases, the value of the objective  
472 function increases noticeably without imposing serious penalties on the patients, with average  
473 waiting times below 10 minutes for models Over1 and Over 2. These values return maximum  
474 overcrowding probabilities (two overbooked patients showing-up to the same appointment)  
475 of 0.42 and 0.56, respectively. This suggests that the optimal value of  $\pi$  should be such that

476 the overcrowding probability is close to 0.5.

## 477 **5 Conclusions**

478 In this article we address the problem of no-shows in specialty clinics. This problem imposes  
479 large economic costs to the health centres –mainly due to practitioners’ idle times-, and to  
480 the patients, who suffer the personal and economic impact of long waiting lists.

481 The no-shows problem is tackled in this article by proposing a scheduling strategy based  
482 on a mixed-integer programming model together with a dynamic priority allocation scheme.  
483 The proposed model aims at maximising the expected revenue of the health centre taking  
484 into account the revenue obtained from both first visit and follow-up patients. When the  
485 revenue of these two groups is the same, the objective function is equivalent to maximising the  
486 expected number of show-ups. The model takes into account several constraints imposed by  
487 both the law and the health centre’s policies; among them, allocating a minimum percentage  
488 of the available slots to first visits, or assigning priorities based on the time the patient has  
489 been in the waiting list. Our formulation can be easily adapted for considering other types of  
490 priority, as jumping the queue when the severity of the patient’s condition demands it, among  
491 others. The base model is extended for allowing the possibility of overbooking.

492 The maximisation of the expected number of show-ups is attained by using individualised  
493 show-up probabilities which depend on the patients’ socio-demographic and personal charac-  
494 teristics as well as on his or her diagnosed pathology. These probabilities are computed for  
495 each day/slot combination using a decision tree classifier on a sample of nearly 50 thousand  
496 visits.

497 Simulation experiments show that whereas the waiting lists size increases on time when  
498 a FIFO scheduling regime is used, our base model is capable of reducing the waiting list  
499 and attaining a 5% increase in revenue with respect to the FIFO regime. Experimental  
500 results also suggest that a more significant reduction in the waiting list would be attained if  
501 overbooking was applied. The magnitude of this reduction would naturally depend on the  
502 amount of doctors’ overtime that the health centre is willing to accept. It was observed that,  
503 by allowing overbooking in one time slot per day, a reduction of the waiting list of about

504 30% can be achieved at a minimum overtime cost. These results suggest a combined strategy  
505 where limited overbooking can be initially used for obtaining a significant reduction in the  
506 waiting list and, later on, switching to a regime without overbooking.

507 Our results point at two interesting research lines. The first one will aim at endogenising  
508 the number and selection of the appointment slots where overbooking is allowed. Given that  
509 not all the patients require the same consultation time, the second research line should extend  
510 the model for taking into account the expected consultation times of the different types of  
511 patient.

## 512 **References**

513 Ahmadi-Javid, A., Jalali, Z., and Klassen, K. J. (2017). Outpatient appointment systems in  
514 healthcare: A review of optimization studies. *European Journal of Operational Research*,  
515 258(1):3–34.

516 Alaeddini, A., Yang, K., Reddy, C., and Yu, S. (2011). A probabilistic model for predicting  
517 the probability of no-show in hospital appointments. *Health care management science*,  
518 14(2):146–157.

519 Berg, B. P., Murr, M., Chermak, D., Woodall, J., Pignone, M., Sandler, R. S., and Denton,  
520 B. T. (2013). Estimating the cost of no-shows and evaluating the effects of mitigation  
521 strategies. *Medical Decision Making*, 33(8):976–985.

522 Brailsford, S. C., Harper, P. R., Patel, B., and Pitt, M. (2009). An analysis of the academic  
523 literature on simulation and modelling in health care. *Journal of simulation*, 3(3):130–140.

524 Cayirli, T. and Veral, E. (2003). Outpatient scheduling in health care: a review of literature.  
525 *Production and operations management*, 12(4):519–549.

526 Chakraborty, S., Muthuraman, K., and Lawley, M. (2010). Sequential clinical scheduling with  
527 patient no-shows and general service time distributions. *IIE Transactions*, 42(5):354–366.

- 528 Conforti, D., Guerriero, F., and Guido, R. (2008). Optimization models for radiotherapy  
529 patient scheduling. *4OR: A Quarterly Journal of Operations Research*, 6(3):263–278.
- 530 Conforti, D., Guerriero, F., Guido, R., and Veltri, M. (2011). An optimal decision-making  
531 approach for the management of radiotherapy patients. *OR Spectrum*, 33(1):123–148.
- 532 Daggy, J., Lawley, M., Willis, D., Thayer, D., Suelzer, C., DeLaurentis, P.-C., Turkcan,  
533 A., Chakraborty, S., and Sands, L. (2010). Using no-show modeling to improve clinic  
534 performance. *Health Informatics Journal*, 16(4):246–259.
- 535 Dantas, L. F., Fleck, J. L., Oliveira, F. L. C., and Hamacher, S. (2018). No-shows in  
536 appointment scheduling—a systematic literature review. *Health Policy*.
- 537 Glowacka, K. J., Henry, R. M., and May, J. H. (2009). A hybrid data mining/simulation  
538 approach for modelling outpatient no-shows in clinic scheduling. *Journal of the Operational  
539 Research Society*, 60(8):1056–1068.
- 540 Gupta, D. and Denton, B. (2008). Appointment scheduling in health care: Challenges and  
541 opportunities. *IIE transactions*, 40(9):800–819.
- 542 Hämäläinen, R. P., Luoma, J., and Saarinen, E. (2013). On the importance of behavioral op-  
543 erational research: The case of understanding and communicating about dynamic systems.  
544 *European Journal of Operational Research*, 228(3):623–634.
- 545 Hashim, M. J., Franks, P., and Fiscella, K. (2001). Effectiveness of telephone reminders in  
546 improving rate of appointments kept at an outpatient clinic: a randomized controlled trial.  
547 *The Journal of the American Board of Family Practice*, 14(3):193–196.
- 548 Hixon, A. L., Chapman, R. W., and Nuovo, J. (1999). Failure to keep clinic appointments:  
549 implications for residency education and productivity. *FAMILY MEDICINE-KANSAS  
550 CITY-*, 31:627–630.
- 551 Hulshof, P. J., Kortbeek, N., Boucherie, R. J., Hans, E. W., and Bakker, P. J. (2012).  
552 Taxonomic classification of planning decisions in health care: a structured review of the  
553 state of the art in or/ms. *Health systems*, 1(2):129–175.



- 554 Kheirkhah, P., Feng, Q., Travis, L. M., Tavakoli-Tabasi, S., and Sharafkhaneh, A. (2015).  
555 Prevalence, predictors and economic consequences of no-shows. *BMC health services re-*  
556 *search*, 16(1):13.
- 557 Kim, S. and Giachetti, R. E. (2006). A stochastic mathematical appointment overbooking  
558 model for healthcare providers to improve profits. *IEEE Transactions on systems, man,*  
559 *and cybernetics-Part A: Systems and humans*, 36(6):1211–1219.
- 560 Kopach, R., DeLaurentis, P.-C., Lawley, M., Muthuraman, K., Ozsen, L., Rardin, R., Wan,  
561 H., Intrevado, P., Qu, X., and Willis, D. (2007). Effects of clinical characteristics on  
562 successful open access scheduling. *Health care management science*, 10(2):111–124.
- 563 LaGanga, L. R. and Lawrence, S. R. (2007). Clinic overbooking to improve patient access  
564 and increase provider productivity. *Decision Sciences*, 38(2):251–276.
- 565 McKee, S. (2014). Measuring the cost of patient no-shows. *Power Your Practice*.
- 566 Molfenter, T. (2013). Reducing appointment no-shows: going from theory to practice. *Sub-*  
567 *stance use & misuse*, 48(9):743–749.
- 568 Moore, C. G., Wilson-Witherspoon, P., and Probst, J. C. (2001). Time and money: effects of  
569 no-shows at a family practice residency clinic. *Family Medicine-Kansas City-*, 33(7):522–  
570 527.
- 571 Muthuraman, K. and Lawley, M. (2008). A stochastic overbooking model for outpatient  
572 clinical scheduling with no-shows. *IIE Transactions*, 40(9):820–837.
- 573 Norris, J. B., Kumar, C., Chand, S., Moskowitz, H., Shade, S. A., and Willis, D. R. (2014).  
574 An empirical investigation into factors affecting patient cancellations and no-shows at  
575 outpatient clinics. *Decision Support Systems*, 57:428–443.
- 576 Samorani, M. and LaGanga, L. R. (2015). Outpatient appointment scheduling given in-  
577 dividual day-dependent no-show predictions. *European Journal of Operational Research*,  
578 240(1):245–257.

- 579 Satiani, B., Miller, S., and Patel, D. (2009). No-show rates in the vascular laboratory: analysis  
580 and possible solutions. *Journal of Vascular and Interventional Radiology*, 20(1):87–91.
- 581 Savelsbergh, M. and Smilowitz, K. (2016). Stratified patient appointment scheduling for mo-  
582 bile community-based chronic disease management programs. *IIE Transactions on Health-  
583 care Systems Engineering*, 6(2):65–78.
- 584 Wang, J. and Fung, Y. (2014). An integer programming formulation for outpatient scheduling  
585 with patient preference. *Industrial Engineering & Management Systems*, 13(2):193–202.
- 586 Wiesche, L., Schacht, M., and Werners, B. (2017). Strategies for interday appointment  
587 scheduling in primary care. *Health care management science*, 20(3):403–418.
- 588 World Bank (2018). Current health expenditure. <https://data.worldbank.org>.
- 589 Zacharias, C. and Pinedo, M. (2014). Appointment scheduling with no-shows and overbook-  
590 ing. *Production and Operations Management*, 23(5):788–801.
- 591 Zhu, H., Hou, M., Wang, C., and Zhou, M. (2012). An efficient outpatient scheduling  
592 approach. *IEEE Transactions on Automation science and engineering*, 9(4):701–709.