Appendix

This file is the electronic companion of the paper “Managing Appointment-based Services in the Presence of Walk-in Customers” by Shan Wang, Nan Liu and Guohua Wan.

A. Additional Results of the Exploratory Study

Table A1 **Correlation Analysis of Walk-in Counts in Different Time Slots for Provider ALD**

correlation (p-value)	8am	9am	10am	11am	12pm	1pm	2pm	3pm	4pm	5pm	6pm
8am	1										
9am	-.07(.38)	1									
10am	-.06(.47)	-.02(.77)	1								
11am	-.05(.50)	-.05(.53)	-.07(.35)	1							
12pm	-.04(.61)	-.05(.49)	.10(.21)	-.07(.37)	1						
1pm	-.03(.70)	-.10(.21)	-.06(.41)	.05(.52)	-.08(.31)	1					
2pm	.01(.92)	.01(.91)	-.09(.28)	.12(.12)	.00(.95)	-.11(.17)	1				
3pm	.02(.84)	-.06(.42)	-.09(.23)	-.08(.30)	.15(.06)	-.08(.28)	.08(.32)	1			
4pm	.04(.62)	-.03(.68)	.06(.48)	.09(.26)	-.02(.83)	-.04(.58)	-.02(.82)	-.08(.28)	1		
5pm	-.03(.71)	-.09(.24)	-.11(.15)	.02(.84)	-.02(.78)	-.02(.77)	.15(.06)	.09(.27)	-.04(.60)	1	
6pm	-.02(.78)	.02(.79)	-.08(.28)	.00(.96)	.14(.07)	-.02(.82)	.10(.22)	.15(.06)	-.03(.70)	-.02(.83)	1

Notes. For each entry in the table, the first number is the correlation coefficient, and the second number in the parentheses is the associated p-value.

Table A2 **Correlation Analysis of Walk-in Counts in Different Time Slots for Provider GAR**

correlation (p-value)	8am	9am	10am	11am	12pm	1pm	2pm	3pm	4pm
8am	1								
9am	.17(.00)	1							
10am	.11(.01)	.24(.00)	1						
11am	.07(.09)	.14(.00)	.29(.00)	1					
12pm	.12(.00)	.08(.07)	.14(.00)	.18(.00)	1				
1pm	.14(.00)	.11(.01)	.08(.06)	.03(.42)	.13(.00)	1			
2pm	.04(.38)	-.01(.79)	.06(.13)	.01(.75)	.01(.77)	.01(.90)	1		
3pm	-.05(.25)	.02(.61)	-.08(.07)	-.01(.73)	-.03(.55)	.01(.88)	.05(.21)	1	
4pm	.03(.55)	.08(.05)	.06(.13)	.00(.98)	-.06(.13)	.04(.36)	.08(.07)	.06(.17)	1

Notes. For each entry in the table, the first number is the correlation coefficient, and the second number in the parentheses is the associated p-value.

Table A3 **Correlation Analysis of Walk-in Counts in Different Time Slots for Provider GED**

correlation (p-value)	9am	10am	11am	12pm	1pm	2pm	3pm	4pm	5pm	6pm
9am	1									
10am	.23(.00)	1								
11am	-.18(.01)	-.10(.15)	1							
12pm	-.15(.03)	-.21(.00)	.13(.06)	1						
1pm	-.04(.57)	-.04(.56)	.01(.88)	-.03(.69)	1					
2pm	.15(.03)	-.01(.92)	-.10(.13)	-.13(.06)	-.13(.07)	1				
3pm	-.01(.87)	-.11(.10)	.05(.42)	-.02(.80)	.00(.96)	.01(.94)	1			
4pm	-.06(.36)	-.10(.14)	.04(.58)	.03(.67)	.02(.81)	-.03(.67)	.23(.00)	1		
5pm	-.12(.09)	-.08(.26)	.01(.89)	.14(.05)	.12(.07)	-.04(.59)	-.09(.18)	-.02(.72)	1	
6pm	-.08(.21)	-.08(.26)	.05(.50)	.12(.09)	-.02(.73)	-.13(.07)	.00(.96)	.09(.17)	.12(.07)	1

Notes. For each entry in the table, the first number is the correlation coefficient, and the second number in the parentheses is the associated p-value.

Table A4 **Correlation Analysis of Walk-in Counts in Different Time Slots for Provider KNI**

correlation (p-value)	8am	9am	10am	11am	12pm	1pm	2pm	3pm	4pm	5pm
8am	1									
9am	.03(.68)	1								
10am	-.10(.11)	.05(.43)	1							
11am	.00(.95)	.04(.55)	.05(.46)	1						
12pm	.11(.09)	.00(.95)	.19(.00)	.13(.04)	1					
1pm	.10(.11)	-.05(.40)	.00(.94)	-.02(.76)	.10(.10)	1				
2pm	.03(.61)	-.12(.07)	-.12(.06)	-.06(.32)	-.10(.11)	.00(.98)	1			
3pm	.03(.65)	-.08(.20)	-.13(.04)	-.01(.90)	-.10(.12)	-.04(.54)	.19(.00)	1		
4pm	-.02(.76)	.01(.87)	-.11(.09)	-.10(.12)	.02(.73)	-.06(.31)	-.07(.26)	.07(.29)	1	
5pm	-.01(.90)	-.05(.47)	.00(.99)	-.06(.37)	-.04(.53)	.11(.10)	-.04(.53)	-.03(.64)	-.01(.87)	1

Notes. For each entry in the table, the first number is the correlation coefficient, and the second number in the parentheses is the associated p-value.

Table A5 **Correlation Analysis of Walk-in Counts in Different Time Slots for Provider LOK**

correlation (p-value)	8am	9am	10am	11am	12pm	1pm	2pm	3pm	4pm	5pm	6pm
8am	1										
9am	.06(.47)	1									
10am	-.01(.89)	-.14(.07)	1								
11am	-.09(.26)	.07(.37)	-.05(.56)	1							
12pm	.08(.32)	.04(.58)	-.22(.00)	.11(.15)	1						
1pm	-.03(.72)	-.12(.13)	-.03(.70)	-.07(.39)	.17(.03)	1					
2pm	.03(.66)	.14(.08)	.05(.49)	.06(.45)	.04(.65)	-.10(.18)	1				
3pm	-.01(.92)	.02(.80)	-.04(.57)	-.04(.64)	-.02(.76)	.01(.92)	.06(.43)	1			
4pm	.19(.01)	-.01(.89)	.06(.46)	-.11(.16)	-.06(.42)	-.06(.44)	.00(.99)	.23(.00)	1		
5pm	.27(.00)	-.01(.87)	.05(.56)	-.09(.26)	.01(.88)	-.05(.52)	.19(.01)	.11(.15)	-.03(.70)	1	
6pm	-.03(.69)	-.13(.09)	-.06(.47)	-.09(.26)	.14(.08)	-.03(.72)	-.02(.84)	.32(.00)	.28(.00)	-.06(.47)	1

Notes. For each entry in the table, the first number is the correlation coefficient, and the second number in the parentheses is the associated p-value.

Table A6 **Correlation Analysis of Walk-in Counts in Different Time Slots for Provider WAT**

correlation (p-value)	8am	9am	10am	11am	12pm	1pm	2pm	3pm	4pm	5pm	6pm
8am	1										
9am	.00(.99)	1									
10am	-.03(.59)	.01(.90)	1								
11am	-.01(.79)	-.07(.24)	.03(.61)	1							
12pm	-.08(.17)	-.06(.31)	-.13(.02)	.05(.39)	1						
1pm	-.06(.27)	-.06(.26)	-.07(.20)	-.10(.06)	-.08(.17)	1					
2pm	.04(.50)	.00(.98)	-.03(.65)	-.03(.57)	-.12(.04)	-.12(.03)	1				
3pm	-.06(.30)	-.02(.67)	-.12(.03)	-.10(.08)	.00(.95)	-.13(.02)	.14(.01)	1			
4pm	.01(.80)	.02(.67)	-.04(.48)	-.10(.09)	-.07(.22)	-.08(.16)	-.05(.41)	.05(.35)	1		
5pm	.02(.72)	-.10(.09)	-.04(.50)	.02(.78)	.01(.86)	-.07(.21)	-.03(.61)	-.02(.77)	.26(.00)	1	
6pm	.04(.45)	-.06(.29)	-.10(.09)	.06(.31)	.01(.82)	.01(.89)	-.03(.56)	-.04(.49)	.02(.76)	.14(.01)	1

Notes. For each entry in the table, the first number is the correlation coefficient, and the second number in the parentheses is the associated p-value.

B. Summary of Notations

Table B7 Summary of Notations.

Notation	**Defnition**
Constants	
T	The total number of regular time slots (i.e., the regular length of a clinic session)
$\overline{T}$	The maximum number of time slots ($\overline{T} > T$)
N_S	The maximum number of patients to be scheduled
N_W	The expected total number of walk-in patients
C_S	Unit time waiting cost per scheduled patient, normalized to be 1
C_W	Unit time waiting cost per walk-in
C_I	Unit time idling cost for the provider
C_O	Unit time overtime cost for the provider
C_D	Unit time duration cost, i.e., $C_I + C_O$
$\overline{N_t}$	A sufficiently large number
(Auxiliary) Decision variables	
x_t	The number of patients scheduled in slot t
n	The total number of scheduled patients $n = \sum_{t=1}^{T} x_t$
$\boldsymbol{x}$	The vector of x_t
y_t	The total number of patients waiting at the end of slot t
$\boldsymbol{y}$	The vector of y_t (in the matrix form, it represents the vector of y_t and y_t^s)
y_t^s	The total number of scheduled patients waiting at the end of slot t
$\boldsymbol{y}^s$	The vector of y_t^s
$z_{t,i}$	If patient i is scheduled at t then $z_{t,i} = 1$, otherwise $z_{t,i} = 0$
$\boldsymbol{z}$	The vector of $z_{t,i}$
Random variables	
β_t	The number of walk-ins in slot t, and its probability mass function (PMF) is $p_t(b) = Pr(\beta_t = b)$
$\boldsymbol{\beta}$	The vector of β_t
$\alpha_t(x_t)$	The number of show-up patients in slot t given a schedule $\boldsymbol{x}$, and its PMF is $q_t(k, x_t) = Pr(\alpha_t(x_t) = k)$
$\boldsymbol{\alpha}(\boldsymbol{x})$	The vector of $\alpha_t(x_t)$
$\Omega_0(\boldsymbol{x})$	The set of all possible scenarios given a schedule $\boldsymbol{x}$
ω_0	$\omega_0 \in \Omega_0(\boldsymbol{x})$, an arbitrary scenario in the set $\Omega_0(\boldsymbol{x})$
$\Omega(\boldsymbol{z})$	The set of all possible scenarios given a schedule $\boldsymbol{z}$
ω	$\omega \in \Omega(\boldsymbol{z})$, an arbitrary scenario in the set $\Omega(\boldsymbol{z})$
$\gamma_{t,i}(\omega)$	The indicator for patient i's show-up status at t under scenario ω
$\boldsymbol{\gamma}(\omega)$	The vector of $\gamma_{t,i}(\omega)$
Values to be calculated	
$\varGamma_S$	The expected total wait time of scheduled patients
$\varGamma_W$	The expected total wait time of walk-ins
$\varGamma_I$	The expected idle time of the provider
$\varGamma_O$	The expected over time of the provider
$\varGamma_D$	The expected duration from the beginning of the session to the departure time of the last patient or T, whichever is later
$\Pi_t(k)$	The probability of k patients waiting for services at the end of t
$\Psi_t(k)$	The probability of k scheduled patients waiting for services at the end of t
$\varUpsilon(\boldsymbol{x}, \omega_0)$	The total cost under schedule $\boldsymbol{x}$ when scenario ω_0 occurs
$\varUpsilon(\boldsymbol{z}, \omega)$	The total cost under schedule $\boldsymbol{z}$ when scenario ω occurs

C. Proofs of Analytical Results

C.1. Proof of Proposition 1

Before we start our proof, we first introduce a definition and a few ancillary results.

Definition 3. For two random variables, if $\mathbb{P}\{a \leq b\} = 1$, we say $a \leq b$, $a - b \leq 0$ or $b - a \geq 0$.

Lemma 2. *If a and b are non-negative integer random variables such that $a - b \geq 0$, then $0 \leq (a-1)^+ - (b-1)^+ \leq (a+\beta-1)^+ - (b+\beta-1)^+ \leq a - b$ for any non-negative integer random variable β.*

Proof of Lemma 2. First, let us prove $0 \leq (a-1)^+ - (b-1)^+$. (In the remaining of the appendix, if we say "when a random variable is greater than, less than or equal to some number", it means "when the realization of the random variable is greater than, less than or equal to some number".) When $a - 1 \leq 0$ and $b - 1 \leq 0$, then $(a-1)^+ = (b-1)^+ = 0$. When $a - 1 > 0$ and $b - 1 \leq 0$, $\mathbb{P}\{0 \leq (a-1)^+ - (b-1)^+\} = \mathbb{P}\{a \geq 1\} = 1$. When $a - 1 > 0$ and $b - 1 > 0$, $\mathbb{P}\{0 \leq (a-1)^+ - (b-1)^+\} = \mathbb{P}\{b \leq a\} = 1$. So, $0 \leq (a-1)^+ - (b-1)^+$.

Second, let us prove $(a-1)^+ - (b-1)^+ \leq (a+\beta-1)^+ - (b+\beta-1)^+$. When $\beta = 0$, $(a-1)^+ - (b-1)^+ = (a+\beta-1)^+ - (b+\beta-1)^+$. When $\beta \geq 1$, $(a+\beta-1)^+ - (b+\beta-1)^+ = a - b$: when $a - 1 \leq 0$ and $b - 1 \leq 0$, then $(a-1)^+ - (b-1)^+ = 0 \leq a - b$; when $a - 1 > 0$ and $b - 1 \leq 0$, $(a-1)^+ - (b-1)^+ = a - 1 \leq a - b$; when $a - 1 > 0$ and $b - 1 > 0$, $(a-1)^+ - (b-1)^+ = a - b$, thus $(a-1)^+ - (b-1)^+ \leq a - b$. So, $(a-1)^+ - (b-1)^+ \leq (a+\beta-1)^+ - (b+\beta-1)^+$.

Third, let us prove $(a+\beta-1)^+ - (b+\beta-1)^+ \leq a - b$. When $\beta = 0$, $(a+\beta-1)^+ - (b+\beta-1)^+$ becomes $(a-1)^+ - (b-1)^+$ which has been proved to be no greater than $a - b$. When $\beta \geq 1$, $(a+\beta-1)^+ - (b+\beta-1)^+ = a - b$. So, $(a+\beta-1)^+ - (b+\beta-1)^+ \leq a - b$. Q.E.D.

Lemma 3. *If a and b are non-negative integer random variables such that $0 \leq a - b \leq 1$, then $(a+\beta-1)^+ - (b+\beta-1)^+ \leq (a-1)^+ + \beta - (b+\beta-1)^+ \leq 1$ for any non-negative integer random variable β.*

Proof of Lemma 3. First, let us prove $(a+\beta-1)^+ - (b+\beta-1)^+ \leq (a-1)^+ + \beta - (b+\beta-1)^+$, i.e., $(a+\beta-1)^+ \leq (a-1)^+ + \beta$. When $\beta = 0$, $(a+\beta-1)^+ = (a-1)^+ + \beta$. When $\beta \geq 1$, $(a+\beta-1)^+ = a - 1 + \beta$, $(a-1)^+ + \beta = (a-1)^+ + \beta$, since $\mathbb{P}\{a - 1 \leq (a-1)^+\} = 1$ then $(a+\beta-1)^+ \leq (a-1)^+ + \beta$. So $(a+\beta-1)^+ - (b+\beta-1)^+ \leq (a-1)^+ + \beta - (b+\beta-1)^+$.

Second, let us prove $(a-1)^+ + \beta - (b+\beta-1)^+ \leq 1$. When $\beta = 0$, $(a-1)^+ + \beta - (b+\beta-1)^+ = (a-1)^+ - (b-1)^+ \leq a - b \leq 1$ by Lemma 2. When $\beta \geq 1$, $(a-1)^+ + \beta - (b+\beta-1)^+ = (a-1)^+ - b + 1$: when $a = 0$ then $(a-1)^+ + \beta - (b+\beta-1)^+ = 1 - b \leq 1$ since $b \geq 0$, when $a \geq 1$ then $(a-1)^+ + \beta - (b+\beta-1)^+ = a - b \leq 1$. So $(a-1)^+ + \beta - (b+\beta-1)^+ \leq 1$. Q.E.D.

Lemma 4. *If a, b, c and d are non-negative integer random variables such that $0 \leq a - b \leq c - d$, $a \leq c$ and $b \leq d$, then $0 \leq (a+\beta-1)^+ - (b+\beta-1)^+ \leq (c+\beta-1)^+ - (d+\beta-1)^+$, $(a+\beta-1)^+ \leq (c+\beta-1)^+$ and $(b+\beta-1)^+ \leq (d+\beta-1)^+$ for any non-negative integer random variable β.*

Proof of Lemma 4. By Lemma 2, we firstly have $0 \leq (a+\beta-1)^+ - (b+\beta-1)^+$, $(a+\beta-1)^+ \leq (c+\beta-1)^+$ and $(b+\beta-1)^+ \leq (d+\beta-1)^+$.

Then let us prove $(a+\beta-1)^+ - (b+\beta-1)^+ \leq (c+\beta-1)^+ - (d+\beta-1)^+$. When $\beta \geq 1$, $(a+\beta-1)^+ - (b+\beta-1)^+ = a - b \leq c - d = (c+\beta-1)^+ - (d+\beta-1)^+$. When $\beta = 0$, $(a+\beta-1)^+ - (b+\beta-1)^+ = (a-1)^+ - (b-1)^+$

and $(c+\beta-1)^+-(d+\beta-1)^+=(c-1)^+-(d-1)^+$. Then we need to prove $(a-1)^+-(b-1)^+\le(c-1)^+-(d-1)^+$. Recall that $b\le a\le c$ and $b\le d\le c$: when $b\ge 1$, then $(a-1)^+-(b-1)^+=a-b\le c-d=(c-1)^+-(d-1)^+$; when $b=0$, $a\ge 1$ and $d=0$, then $(a-1)^+-(b-1)^+=a-1\le c-1=(c-1)^+-(d-1)^+$; when $b=0$, $a\ge 1$ and $d\ge 1$, then $(a-1)^+-(b-1)^+=a-1\le c-d=(c-1)^+-(d-1)^+$; when $b=0$ and $a=0$, then $(a-1)^+-(b-1)^+=0\le(c-1)^+-(d-1)^+$ by Lemma 2; when $a=b=c=d=0$ then $(a-1)^+-(b-1)^+=(c-1)^+-(d-1)^+=0$. So $(a+\beta-1)^+-(b+\beta-1)^+\le(c+\beta-1)^+-(d+\beta-1)^+$. Q.E.D.

Lemma 5. *If a, b, c and d are non-negative integer random variables such that $0\le a-b\le c-d$, $c\le a$, $d\le b$ and $a-d\le 1$, then $0\le(a+\beta-1)^+-(b+\beta-1)^+\le(c+\beta-1)^+-(d+\beta-1)^+$, $(c+\beta-1)^+\le(a+\beta-1)^+$, $(d+\beta-1)^+\le(b+\beta-1)^+$ and $(a+\beta-1)^+-(d+\beta-1)^+\le 1$ for any non-negative integer random variable β.*

Proof of Lemma 5. By Lemma 2 and 3, we firstly have $0\le(a+\beta-1)^+-(b+\beta-1)^+$, $(c+\beta-1)^+\le(a+\beta-1)^+$, $(d+\beta-1)^+\le(b+\beta-1)^+$ and $(a+\beta-1)^+-(d+\beta-1)^+\le 1$.

Then let us prove $(a+\beta-1)^+-(b+\beta-1)^+\le(c+\beta-1)^+-(d+\beta-1)^+$. When $\beta\ge 1$, $(a+\beta-1)^+-(b+\beta-1)^+=a-b\le c-d=(c+\beta-1)^+-(d+\beta-1)^+$. When $\beta=0$, $(a+\beta-1)^+-(b+\beta-1)^+=(a-1)^+-(b-1)^+$ and $(c+\beta-1)^+-(d+\beta-1)^+=(c-1)^+-(d-1)^+$. Then we need to prove $(a-1)^+-(b-1)^+\le(c-1)^+-(d-1)^+$. Recall that $c\le a$, $d\le b$ and $a-d\le 1$: when $d\ge 1$, then $(a-1)^+-(b-1)^+=a-b\le c-d=(c-1)^+-(d-1)^+$; when $d=0$, then $a\le 1$, $b\le 1$ and $c\le 1$ since $a\le d+1$, $c\le a$ and $b\le a$, then $(a-1)^+-(b-1)^+=(c-1)^+-(d-1)^+=0$. So $(a+\beta-1)^+-(b+\beta-1)^+\le(c+\beta-1)^+-(d+\beta-1)^+$. Q.E.D.

Lemma 6. *If two random variables $a\le b$, then $\mathbb{E}[h(a)]\le\mathbb{E}[h(b)]$ for all increasing functions $h(\cdot)$ for which the expectations exist.*

Proof of Lemma 6. It follows Shaked and Shanthikumar (2007). Q.E.D.

Define the function $f(\boldsymbol{x})$: $\mathbb{Z}^T\to\mathbb{R}$ as $f(\boldsymbol{x})=\mathit{\Gamma}_S(\boldsymbol{x})+C_W\mathit{\Gamma}_W(\boldsymbol{x})+C_D\mathit{\Gamma}_D(\boldsymbol{x})-C_I\sum\limits_{t=1}^{T}x_t$. To prove Proposition 1, we need to prove that for all $\boldsymbol{x}$, $\boldsymbol{x}+\boldsymbol{v}_i$, $\boldsymbol{x}+\boldsymbol{v}_j$ and $\boldsymbol{x}+\boldsymbol{v}_i+\boldsymbol{v}_j$ in $\mathbb{Z}_+^T$,

$$f(\boldsymbol{x}+\boldsymbol{v}_i)+f(\boldsymbol{x}+\boldsymbol{v}_j)\ge f(\boldsymbol{x})+f(\boldsymbol{x}+\boldsymbol{v}_i+\boldsymbol{v}_j) \tag{23}$$

whenever $\boldsymbol{v}_i,\ \boldsymbol{v}_j\in\boldsymbol{V}^\diamond$ and $\boldsymbol{v}_i\neq\boldsymbol{v}_j$ where $\boldsymbol{V}^\diamond$ is defined as (11). Lemma 7 below helps us prove (23).

Lemma 7. *Given a feasible schedule $\boldsymbol{x}\in\mathbb{Z}_+^T$ and random walk-ins $\boldsymbol{\beta}=(\beta_1,\beta_2,\ldots,\beta_T)$, let random variable y_t denote the number of patients waiting at the end of t, thus $y_{t+1}=(y_t+x_{t+1}+\beta_{t+1}-1)^+$. If $y_t^i+y_t^j\ge y_t+y_t^{ij}$ for any t (where no superscript indicates under schedule $\boldsymbol{x}$, superscript i indicates under schedule $\boldsymbol{x}+\boldsymbol{v}_i$, superscript j indicates under schedule $\boldsymbol{x}+\boldsymbol{v}_j$ and superscript ij indicates under schedule $\boldsymbol{x}+\boldsymbol{v}_i+\boldsymbol{v}_j$), then (23) holds for all feasible schedules.*

Proof of Lemma 7. Let s_t denote the number of scheduled patients waiting at the end of t, thus $s_{t+1}=(s_t+x_{t+1}-1)^+$. It is obvious that $\mathit{\Gamma}_S(\boldsymbol{x})=\sum_{t=1}^{T}s_t$, $\mathit{\Gamma}_S(\boldsymbol{x})+\mathit{\Gamma}_W(\boldsymbol{x})=\sum_{t=1}^{T}\mathbb{E}[y_t(\boldsymbol{x})]$ and $\mathit{\Gamma}_D(\boldsymbol{x})=\mathbb{E}[y_T(\boldsymbol{x})]+T$. So $f(\boldsymbol{x})$ is a summation of s_t, $\mathbb{E}[y_t]$ and $-x_t$ with non-negative weights (recall that $C_W\le 1$). If for any

t, we have $s_t^i + s_t^j \geq s_t + s_t^{ij}$, $\mathbb{E}[y_t^i] + \mathbb{E}[y_t^j] \geq \mathbb{E}[y_t] + \mathbb{E}[y_t^{ij}]$ and $-x_t^i - x_t^j \geq -x_t - x_t^{ij}$ (where the superscript indicates under the corresponding schedule), then the proof can be completed. Since $-x_t^i - x_t^j = -x_t - x_t^{ij}$ by definition, what we need to prove is $s_t^i + s_t^j \geq s_t + s_t^{ij}$ and $\mathbb{E}[y_t^i] + \mathbb{E}[y_t^j] \geq \mathbb{E}[y_t] + \mathbb{E}[y_t^{ij}]$. For the second one, by Lemma 6, we only need to prove $y_t^i + y_t^j \geq y_t + y_t^{ij}$. If we can prove $y_t^i + y_t^j \geq y_t + y_t^{ij}$, then we can prove $s_t^i + s_t^j \geq s_t + s_t^{ij}$ in the same way by replacing all y_t by s_t and letting $\beta_t = 0$. So if $y_t^i + y_t^j \geq y_t + y_t^{ij}$ holds for any t, then (23) holds for any $\boldsymbol{x}$ in $\mathbb{Z}_+^T$.

Now, we examine whether $y_t^i + y_t^j \geq y_t + y_t^{ij}$ holds case by case.

Case A: $\boldsymbol{v}_i = (-1, 0, 0, \dots, 0)$ and $\boldsymbol{v}_j = (1, -1, 0, \dots, 0)$

In this case, $\boldsymbol{x} + \boldsymbol{v}_i$ means that one patient in $t = 1$ is removed, $\boldsymbol{x} + \boldsymbol{v}_j$ means that one patient in $t = 2$ is moved to $t = 1$, $\boldsymbol{x} + \boldsymbol{v}_i + \boldsymbol{v}_j$ means that one patient in $t = 2$ is removed. So we must have $x_1 \geq 1$ and $x_2 \geq 1$.

Let y_1 denote the random number of patients waiting at the end of $t = 1$ under schedule $\boldsymbol{x}$, then, $y_1^i = (y_1 - 1)^+$, $y_1^j = y_1 + 1$ and $y_1^{ij} = y_1$. Then we have $y_1 - y_1^i \leq y_1^j - y_1^{ij}$ by Lemma 3. For $t = 2$, we have $y_2 = y_1 + x_2 + \beta_2 - 1$, $y_2^i = (y_1 - 1)^+ + x_2 + \beta_2 - 1$, $y_2^j = y_1 + x_2 + \beta_2 - 1$ and $y_2^{ij} = (y_1 + x_2 + \beta_2 - 1)^+$. Then $y_2^j = y_2$ and $y_2^i \geq y_2^{ij}$ by Lemma 3. Since the arrival process (including walk-ins and scheduled patients) stays same after $t = 2$, by Lemma 2, we have $y_t^j = y_t$ and $y_t^i \geq y_t^{ij}$ for any $t \geq 3$. So $y_t^i + y_t^j \geq y_t + y_t^{ij}$ for all t.

Case B: $\boldsymbol{v}_i = (-1, 0, 0, \dots, 0)$ and $\boldsymbol{v}_j = (0, \dots, 1, -1, 0, \dots)$

In this case, $\boldsymbol{x} + \boldsymbol{v}_i$ means that one patient in $t = 1$ is removed, $\boldsymbol{x} + \boldsymbol{v}_j$ means that one patient in $t = j + 1$ is moved to $t = j$, $\boldsymbol{x} + \boldsymbol{v}_i + \boldsymbol{v}_j$ means that one patient in $t = 1$ is removed and one patient in $t = j + 1$ is moved to $t = j$. So we must have $x_1 \geq 1$ and $x_{j+1} \geq 1$.

Let y_1 be the number of patients waiting at the end of $t = 1$ under schedule $\boldsymbol{x}$, then $y_1^i = (y_1 - 1)^+$, $y_1^j = y_1$ and $y_1^{ij} = (y_1 - 1)^+$. It is obvious that $y_1 - y_1^i \leq 1$. Then we know $y_t^i = y_t^{ij}$, $y_t^j = y_t$ and $y_t - y_t^i \leq 1$ for any $t \leq j - 1$ by Lemma 3. For $t = j$, $y_j = (y_{j-1} + x_j + \beta_j - 1)^+$, $y_j^i = (y_{j-1}^i + x_j + \beta_j - 1)^+$, $y_j^j = y_{j-1} + x_j + \beta_j$ and $y_j^{ij} = y_{j-1}^i + x_j + \beta_j$. Then we have $y_j - y_j^i \leq y_j^j - y_j^{ij}$ by Lemma 2. For $t = j + 1$, $y_{j+1} = (y_{j-1} + x_j + \beta_j - 1)^+ + x_{j+1} + \beta_{j+1} - 1$, $y_{j+1}^i = (y_{j-1}^i + x_j + \beta_j - 1)^+ + x_{j+1} + \beta_{j+1} - 1$, $y_{j+1}^j = (y_{j-1} + x_j + \beta_j + x_{j+1} + \beta_{j+1} - 2)^+$ and $y_{j+1}^{ij} = (y_{j-1}^i + x_j + \beta_j + x_{j+1} + \beta_{j+1} - 2)^+$. Then we have $0 \leq y_{j+1} - y_{j+1}^i \leq y_{j+1}^j - y_{j+1}^{ij}$ by Lemma 2, $y_{j+1}^j \leq y_{j+1}$, $y_{j+1}^{ij} \leq y_{j+1}^i$ and $y_{j+1} - y_{j+1}^{ij} \leq 1$ by Lemma 3. So for any $t > j + 1$, since the arrival process stays same, then $0 \leq y_t - y_t^i \leq y_t^j - y_t^{ij}$, $y_t^j \leq y_t$, $y_t^{ij} \leq y_t^i$ and $y_t - y_t^{ij} \leq 1$ by Lemma 5. Then we have $y_t^i + y_t^j \geq y_t + y_t^{ij}$ for all t.

Case C: $\boldsymbol{v}_i = (-1, 0, 0, \dots, 0)$ and $\boldsymbol{v}_j = (0, \dots, 0, 1)$

In this case, $\boldsymbol{x} + \boldsymbol{v}_i$ means that one patient in $t = 1$ is removed, $\boldsymbol{x} + \boldsymbol{v}_j$ means that one patient is added in $t = T$, $\boldsymbol{x} + \boldsymbol{v}_i + \boldsymbol{v}_j$ means that one patient in $t = 1$ is moved to $t = T$. So we must have $x_1 \geq 1$.

y_1 is the number of patients waiting at the end of $t = 1$ under schedule $\boldsymbol{x}$, then $y_1^i = (y_1 - 1)^+$, $y_1^j = y_1$ and $y_1^{ij} = (y_1 - 1)^+$. By Lemma 2, we have $y_t = y_t^j$, $y_t^i = y_t^{ij}$ and $y_t \leq y_t^i$ for any $t \leq T - 1$. For $t = T$, $y_T = (y_{T-1} + x_T + \beta_T - 1)^+$, $y_T^i = (y_{T-1}^i + x_T + \beta_T - 1)^+$, $y_T^j = y_{T-1} + x_T + \beta_T$ and $y_T^{ij} = y_{T-1}^i + x_T + \beta_T$, so $0 \leq y_T - y_T^i \leq y_T^j - y_T^{ij}$, $y_T \leq y_T^j$ and $y_T^i \leq y_T^{ij}$ by Lemma 2. Since there is no arrival after T, we have $0 \leq y_t - y_t^i \leq y_t^j - y_t^{ij}$, $y_t \leq y_t^j$ and $y_t^i \leq y_t^{ij}$ for any $t > T$, by Lemma 4. So $y_t^i + y_t^j \geq y_t + y_t^{ij}$ for all t.

Case D: $\boldsymbol{v}_i = (0, \dots, 1, -1, 0, \dots)$ and $\boldsymbol{v}_j = (0, \dots, 1, -1, 0, \dots)$ $(j = i + 1)$

In this case, $\boldsymbol{x}+\boldsymbol{v}_i$ means that one patient in $t=j$ is moved to $t=i$, $\boldsymbol{x}+\boldsymbol{v}_j$ means that one patient in $t=j+1$ is moved to $t=j$, $\boldsymbol{x}+\boldsymbol{v}_i+\boldsymbol{v}_j$ means that one patient in $t=j+1$ is moved to $t=i$. So we must have $x_j \geq 1$ and $x_{j+1} \geq 1$.

For $t<i$, there is no difference between these four schedules. For $t=i$, y_i^i is the number of patients waiting at the end of $t=i$ under schedule $\boldsymbol{x}^i$, then $y_i=(y_i^i-1)^+$, $y_i^j=(y_i^i-1)^+$ and $y_i^{ij}=y_i^i$, so $y_i=y_i^j$ and $y_i^i=y_i^{ij}$. For $t=j$, $y_j=(y_i^i-1)^+ + x_j+\beta_j-1$, $y_j^i=(y_i^i+x_j+\beta_j-2)^+$, $y_j^j=(y_i^i-1)^+ + x_j+\beta_j$ and $y_j^{ij}=y_i^i+x_j+\beta_j-1$. Since $y_j^j-y_j=1$, then $y_j^{ij}-y_j^i \leq y_j^j-y_j$ by Lemma 3. For $t=j+1$, $y_{j+1}=(y_i^i-1)^+ + x_j+\beta_j-1+x_{j+1}+\beta_{j+1}-1$, $y_{j+1}^i=(y_i^i+x_j+\beta_j-2)^+ + x_{j+1}+\beta_{j+1}-1$, $y_{j+1}^j=(y_i^i-1)^+ + x_j+\beta_j+x_{j+1}+\beta_{j+1}-2$ and $y_{j+1}^{ij}=(y_i^i+x_j+\beta_j-1+x_{j+1}+\beta_{j+1}-2)^+$. Then we have $y_{j+1}=y_{j+1}^j$ and $y_{j+1}^{ij} \leq y_{j+1}^i$ by Lemma 3. Since the arrival process after $j+1$ stays same, $y_t=y_t^j$ and $y_t^{ij} \leq y_t^i$ for any $t>j+1$, by Lemma 2. So $y_t^i+y_t^j \geq y_t+y_t^{ij}$ for all t.

Case E: $\boldsymbol{v}_i=(0,\dots,1,-1,0,\dots)$ and $\boldsymbol{v}_j=(0,\dots,1,-1,0,\dots)$ $(j>i+1)$

In this case, $\boldsymbol{x}+\boldsymbol{v}_i$ means that one patient in $t=i+1$ is moved to $t=i$, $\boldsymbol{x}+\boldsymbol{v}_j$ means that one patient in $t=j+1$ is moved to $t=j$, $\boldsymbol{x}+\boldsymbol{v}_i+\boldsymbol{v}_j$ means that one patient in $t=i+1$ is moved to $t=i$ and one patient in $t=j+1$ is moved to $t=j$. So we must have $x_{i+1} \geq 1$ and $x_{j+1} \geq 1$.

For $t<i$, there is no difference between these four schedules. For $t=i$, y_i^i is the number of patients waiting at the end of $t=i$ under schedule $\boldsymbol{x}^i$, then $y_i=(y_i^i-1)^+$, $y_i^j=(y_i^i-1)^+$ and $y_i^{ij}=y_i^i$, so $y_i=y_i^j$ and $y_i^i=y_i^{ij}$. For $t=i+1$, $y_{i+1}=(y_i^i-1)^+ + x_{i+1}+\beta_{i+1}-1$, $y_{i+1}^i=(y_i^i+x_{i+1}+\beta_{i+1}-2)^+$, $y_{i+1}^j=(y_i^i-1)^+ + x_{i+1}+\beta_{i+1}-1$ and $y_{i+1}^{ij}=(y_i^i+x_{i+1}+\beta_{i+1}-2)^+$. By Lemma 3, we have $y_t=y_t^j$, $y_t^i=y_t^{ij}$ and $0 \leq y_t-y_t^i \leq 1$ for $i<t \leq j-1$. For $t=j$, $y_j=(y_{j-1}+x_j+\beta_j-1)^+$, $y_j^i=(y_{j-1}^i+x_j+\beta_j-1)^+$, $y_j^j=y_{j-1}+x_j+\beta_j$ and $y_j^{ij}=y_j^i+x_j+\beta_j$. Then $y_j-y_j^i \leq y_j^j-y_j^{ij}$ by Lemma 2. For $t=j+1$, $y_{j+1}=(y_{j-1}+x_j+\beta_j-1)^+ + x_{j+1}+\beta_{j+1}-1$, $y_{j+1}^i=(y_{j-1}^i+x_j+\beta_j-1)^+ + x_{j+1}+\beta_{j+1}-1$, $y_{j+1}^j=(y_{j-1}+x_j+\beta_j+x_{j+1}+\beta_{j+1}-2)^+$ and $y_{j+1}^{ij}=(y_{j-1}^i+x_j+\beta_j+x_{j+1}+\beta_{j+1}-2)^+$. Since $0 \leq y_{j-1}-y_{j-1}^i \leq 1$, then $0 \leq y_{j+1}-y_{j+1}^i \leq y_{j+1}^j-y_{j+1}^{ij}$, $y_{j+1}^j \leq y_{j+1}$, $y_{j+1}^{ij} \leq y_{j+1}^i$ and $y_{j+1}-y_{j+1}^{ij} \leq 1$ by Lemma 2 and 3. Since the arrival process after $j+1$ stays same, then we have $0 \leq y_t-y_t^i \leq y_t^j-y_t^{ij}$, $y_t^j \leq y_t$, $y_t^{ij} \leq y_t^i$ and $y_t-y_t^{ij} \leq 1$ for any $t>j+1$ by Lemma 5. So $y_t^i+y_t^j \geq y_t+y_t^{ij}$ for all t.

Case F: $\boldsymbol{v}_i=(0,\dots,1,-1,0,\dots)$ and $\boldsymbol{v}_j=(0,\dots,0,1)$

In this case, $\boldsymbol{x}+\boldsymbol{v}_i$ means that one patient in $t=i+1$ is moved to $t=i$, $\boldsymbol{x}+\boldsymbol{v}_j$ means that one patient is added to $t=T$, $\boldsymbol{x}+\boldsymbol{v}_i+\boldsymbol{v}_j$ means that one patient in $t=i+1$ is moved to $t=i$ and one patient is added to $t=T$. So we must have $x_{i+1} \geq 1$.

For $t<i$, there is no difference between these four schedules. For $t=i$, y_i^i is the number of patients waiting at the end of $t=i$ under schedule $\boldsymbol{x}^i$, then $y_i=(y_i^i-1)^+$, $y_i^j=(y_i^i-1)^+$ and $y_i^{ij}=y_i^i$, so $y_i=y_i^j$ and $y_i^i=y_i^{ij}$. For $t=i+1$, $y_{i+1}=(y_i^i-1)^+ + x_{i+1}+\beta_{i+1}-1$, $y_{i+1}^i=(y_i^i+x_{i+1}+\beta_{i+1}-2)^+$, $y_{i+1}^j=(y_i^i-1)^+ + x_{i+1}+\beta_{i+1}-1$ and $y_{i+1}^{ij}=(y_i^i+x_{i+1}+\beta_{i+1}-2)^+$. By Lemma 3, we have $y_t=y_t^j$, $y_t^i=y_t^{ij}$ and $0 \leq y_t-y_t^i \leq 1$ for $i<t<T$. For $t=T$, $y_T=(y_{T-1}+x_T+\beta_T-1)^+$, $y_T^i=(y_{T-1}^i+x_T+\beta_T-1)^+$, $y_T^j=y_{T-1}+x_T+\beta_T$ and $y_T^{ij}=y_T^i+x_T+\beta_T$. Then $0 \leq y_T-y_T^i \leq y_T^j-y_T^{ij}$, $y_T \leq y_T^j$ and $y_T^i \leq y_T^{ij}$ by Lemma 2. There is no arrival after T, so we have $0 \leq y_t-y_t^i \leq y_t^j-y_t^{ij}$, $y_t \leq y_t^j$ and $y_t^i \leq y_t^{ij}$ for any $t>T$ by Lemma 4. So $y_t^i+y_t^j \geq y_t+y_t^{ij}$ for all t.

Case G: $\boldsymbol{v}_i=(0,\dots,0,1,-1)$ and $\boldsymbol{v}_j=(0,\dots,0,1)$

In this case, $\boldsymbol{x}+\boldsymbol{v}_i$ means that one patient in $t=T$ is moved to $t=T-1$, $\boldsymbol{x}+\boldsymbol{v}_j$ means that one patient is added to $t=T$, $\boldsymbol{x}+\boldsymbol{v}_i+\boldsymbol{v}_j$ means that one patient is added to $t=T-1$. So we must have $x_T \geq 1$.

For $t<T-1$, there is no difference between these four schedules. For $t=T-1$, y^i_{T-1} is the number of patients waiting at the end of $t=T-1$ under schedule $\boldsymbol{x}^i$, then $y_{T-1}=(y^i_{T-1}-1)^+$, $y^j_{T-1}=(y^i_{T-1}-1)^+$ and $y^{ij}_{T-1}=y^i_{T-1}$, so $y_{T-1}=y^j_{T-1}$ and $y^i_{T-1}=y^{ij}_{T-1}$. For $t=T$, $y_T=(y^i_{T-1}-1)^+ + x_T+\beta_T-1$, $y^i_T=(y^i_{T-1}+x_T+\beta_T-2)^+$, $y^j_T=(y^i_{T-1}-1)^+ + x_T+\beta_T$ and $y^{ij}_T=y^i_{T-1}+x_T+\beta_T-1$. Then we have $0\leq y_T-y^i_T\leq y^j_T-y^{ij}_T$, $y_T\leq y^j_T$ and $y^i_T\leq y^{ij}_T$ by Lemma 2. And there is no arrival after T, then $0\leq y_t-y^i_t\leq y^j_t-y^{ij}_t$, $y_t\leq y^j_t$ and $y^i_t\leq y^{ij}_t$ for any $t>T$ by Lemma 4. So $y^i_t+y^j_t\geq y_t+y^{ij}_t$ for all t.

Till now, we examine all cases and show that $y^i_t+y^j_t\geq y_t+y^{ij}_t$ holds for any t. Q.E.D.

By Lemma 7, we have (23) hold for any $\boldsymbol{x}$ in $\mathbb{Z}^T_+$, which means that $f(\boldsymbol{x})$ is multimodular on its domain. This completes the proof of Proposition 1.

C.2. Proof of Corollary 1

Corollary 2.2 in Altman et al. (2000) states that: if $\mathcal{A}$ is the set of all integer points in a convex set and a function f is multimodular in $\mathcal{A}$, then a local minimum of f in $\mathcal{A}$ is a global minimum in $\mathcal{A}$. By Corollary 2.2 in Altman et al. (2000) and Proposition 1, we prove the desired results.

C.3. Proof of Proposition 2

Given a schedule $\boldsymbol{x}$ which overbooks at some slots, i.e., $x_t>1$ for some t, let $f(\boldsymbol{x})$ denote its associated cost. Let t_1 be the largest t such that $x_t>1$. Let t_2 be the smallest t such that $t_2>t_1$ and $x_t=0$.

If $t_2\neq\phi$, construct another schedule $\boldsymbol{x}'$, such that $x'_t=x_t$ for $t\neq t_1$ and $t\neq t_2$, $x'_{t_1}=x_{t_1}-1$ and $x'_{t_2}=1$. It is easy to see that no walk-ins are served during $[t_1,t_2]$ since scheduled patients have higher priority and there are unserved scheduled patients in these slots for both two schedules. It means that the number of (scheduled) patients waiting at the end of slot t_2 does not change. So $\mathit{\Gamma}_D(\boldsymbol{x}')=\mathit{\Gamma}_D(\boldsymbol{x})$. It is straightforward that $\mathit{\Gamma}_S(\boldsymbol{x}')\leq\mathit{\Gamma}_S(\boldsymbol{x})$ since one scheduled patient is moved to t_2 from t_1, and $\mathit{\Gamma}_W(\boldsymbol{x}')=\mathit{\Gamma}_W(\boldsymbol{x})$ since this movement does not influence the wait time of walk-ins. And $\sum_{t=1}^T x'_t=\sum_{t=1}^T x_t$, then we finally have $f(\boldsymbol{x}')\leq f(\boldsymbol{x})$.

If $t_2=\phi$, construct another schedule $\boldsymbol{x}''$, such that $x''_t=x_t$ for $t\neq t_1$, $x'_{t_1}=x_{t_1}-1$. It is easy to see that no walk-in is served during $[t_1,T]$ since scheduled patients have higher priority and there are unserved scheduled patients in these slots for both two schedules. It means that the number of (scheduled) patients waiting at the end of slot T decreases by 1 . So $\mathit{\Gamma}_D(\boldsymbol{x}')=\mathit{\Gamma}_D(\boldsymbol{x})-1$. And it is straightforward that $\mathit{\Gamma}_S(\boldsymbol{x}')\leq\mathit{\Gamma}_S(\boldsymbol{x})$ and $\mathit{\Gamma}_W(\boldsymbol{x}')\leq\mathit{\Gamma}_W(\boldsymbol{x})$ since one scheduled patient is removed from t_1. $\sum_{t=1}^T x'_t=\sum_{t=1}^T x_t-1$. As $C_I+C_O=C_D$, we finally have $f(\boldsymbol{x}')\leq f(\boldsymbol{x})$.

$f(\boldsymbol{x}')\leq f(\boldsymbol{x})$ indicates that the cost will not increase if an occurrence of overbooking is eliminated.

This completes the proof.

C.4. Proof of Proposition 3

To show the multimodularity of the objective function in the model with no-show behavior, the key point is still to prove $f(\boldsymbol{x}+\boldsymbol{v}_i)+f(\boldsymbol{x}+\boldsymbol{v}_j)\geq f(\boldsymbol{x})+f(\boldsymbol{x}+\boldsymbol{v}_i+\boldsymbol{v}_j)$. Recall the proof of Proposition 1 in Section C.1, the idea here stays the same. The only difference is x_t, the number of scheduled patients in t, becomes a random variable α_t. For the slots such that $x_t = x_t^i = x_t^j = x_t^{ij}$, i.e., the number of scheduled patients does not change, we can treat α_t and β_t as one variable. For the slots when the number of scheduled patients changes, we can separate it into two parts: the patients who are still scheduled at t and the one who is added, removed or moved to other slots. The first part can be treat as β_t. As for the second part, we need to take the conditional probability of their show-up status into account. In the inequality $f(\boldsymbol{x}+\boldsymbol{v}_i)+f(\boldsymbol{x}+\boldsymbol{v}_j) \geq f(\boldsymbol{x})+f(\boldsymbol{x}+\boldsymbol{v}_i+\boldsymbol{v}_j)$, there are two changes in total (we call the patient who changes his slot in $\boldsymbol{x}+\boldsymbol{v}_i$ $(\boldsymbol{x}+\boldsymbol{v}_j)$ patient i (j), then in $\boldsymbol{x}+\boldsymbol{v}_i+\boldsymbol{v}_j$ both patient i and j change their slots). If patient i and j do not show up, then the four schedules are same; if patient i shows up and the other one does not show up, then $\boldsymbol{x} = \boldsymbol{x}+\boldsymbol{v}_j$ and $\boldsymbol{x}+\boldsymbol{v}_i = \boldsymbol{x}+\boldsymbol{v}_i+\boldsymbol{v}_j$ which means that the LHS equals to RHS; if both two patients show up, then the proof in Section C.1 can be directly applied here. This completes the proof.

C.5. Examples of Complex Optimal Schedules with Walk-ins

Example 1. Proposition 2 shows that there exists an optimal schedule without overbooking when only walk-ins are present. However, Table C8 gives an example that shows when no-shows are also present, the unique optimal schedule has to overbook.

Example 2. Robinson and Chen (2010) show that the optimal appointment schedule, if no walk-ins present, has a structural property of "No Holes". That is, if no patients are scheduled for an appointment slot, then none will be scheduled for any subsequent slots. However, we find that this property does *not* hold when walk-ins are present, e.g., we see a hole at $t = 3$ in Table C8.

Table C8 An Example of the Unique Optimal Schedule with No-shows.

t=1	t=2	t=3	t=4	t=5	t=6	t=7	t=8	t=9
3	2	0	1	2	1	1	0	0

Notes. $T = 9$, $C_W = 0.8$, $C_D = 25$, $C_I = 10$, $q^s = 0.4$, and the walk-in arrival process follows a Poisson process with arrival rate vector [0.6 0.4 0.7 0.6 0.01 0.4 0.1 0.5 0.2].

Table C9 Optimal Schedules With and Without No-shows.

No-show Probability	t=1	t=2	t=3	t=4	t=5
0	1	0	1	1	0
0.2	0	1	1	1	0

Notes. $T = 5$, $C_W = 0.8$, $C_D = 25$, $C_I = 11$, walk-ins in each period follow the Poisson distribution with arrival rates [1.5 0.3 0.1 0.1 0.1].

Example 3. When walk-ins are present, we cannot rely on the optimal solution of the problem without no-shows to narrow down the search space for the problem with no-shows. To be specific, let $\boldsymbol{x}^{w/o}$ be the optimal schedule with neither no-shows nor walk-ins and $\boldsymbol{x}^{w/}$ be the optimal schedule with no-shows but without walk-ins, then we can derive a necessary (and thus weaker) condition from the "No Holes" property in Robinson and Chen (2010):

$$\sum_{i=1}^{t} x_i^{w/} \geq \sum_{i=1}^{t} x_i^{w/o}, \ \forall t = 1, 2, \ldots, T. \tag{24}$$

We defer the derivation of (24) to the next subsection. Condition (24) suggests that, without walk-ins, when no-show rate is strictly positive, it is better to "front-load" the schedule. Intuitively, one may contend that

Condition (24) also holds when walk-ins are present. If Condition (24) indeed holds with walk-ins, then one may use the structural results obtained in Proposition 2 (for the system with walk-ins but without no-shows) to reduce the search space for the optimal schedule with both walk-ins and no-shows. However, Condition (24) does *not* hold when walk-ins are present. Table C9 shows an example in which the optimal schedule is more front-loaded when the no-show rate is zero. To give an explanation, consider moving the first scheduled patient from $t = 1$ to $t = 2$ in Table C9 while maintaining the positions of other scheduled patients. This movement does not affect the wait time of scheduled patients; it reduces the wait time (cost) of walk-ins, but increases the total duration (cost) of service. Whether to move the first scheduled patient to a later slot or not depends on the tradeoff between the latter two costs. As the no-show rate increases, there is less reduction in the wait time of walk-ins, but the increment of overall service duration is also smaller. Thus there may still be a net benefit of moving the first scheduled patient to a later slot.

C.6. Proof of Condition (24)

We use the following definition in our proof. The vector $\boldsymbol{a}$ is said to majorize the vector $\boldsymbol{b}$ (denoted $\boldsymbol{a} \succ \boldsymbol{b}$) if

$$\sum_{i=1}^{k} a_i \geq \sum_{i=1}^{k} b_i, \ \forall k = 1, 2, \ldots, n-1,$$

and

$$\sum_{i=1}^{n} a_i = \sum_{i=1}^{n} b_i.$$

Condition (24) can be equivalently written as $\boldsymbol{x}^{\boldsymbol{w}/} \succ \boldsymbol{x}^{\boldsymbol{w}/\boldsymbol{o}}$. We note that $\boldsymbol{x}^{\boldsymbol{w}/\boldsymbol{o}}$ has the following pattern:

$$\boldsymbol{x}^{\boldsymbol{w}/\boldsymbol{o}} = (1, 1, \ldots).$$

That is, we start from the first slot, and assign one patient to each subsequent slot and all remaining patients (if any) to slot T. The "No Hole" property in Robinson and Chen (2010) implies that empty slots, if any, will be placed towards the end in $\boldsymbol{x}^{\boldsymbol{w}/}$. In other words, compared to $\boldsymbol{x}^{\boldsymbol{w}/\boldsymbol{o}}$, $\boldsymbol{x}^{\boldsymbol{w}/}$ "pushes" patients scheduled towards the end upfront in the schedule, leading to a "larger" vector in terms of the majorization order. Therefore, we have $\boldsymbol{x}^{\boldsymbol{w}/} \succ \boldsymbol{x}^{\boldsymbol{w}/\boldsymbol{o}}$.

This completes the proof.

C.7. Lemma 8

LEMMA 8. *Let $\bar{\boldsymbol{x}}$ be the optimal schedule for the problem*

$$\min_{\boldsymbol{x} \in \mathbb{Z}_+^T} \ \mathbb{E}_{\omega^o}\Big[\Upsilon(\boldsymbol{x}, \omega^o) - C_I \sum_{t=1}^{T} \alpha_t(x_t, \omega^0)\Big] \tag{A}$$

where

$$\Upsilon(\boldsymbol{x}, \omega^o) = \Big\{ C_D y_T^s \Big| (17),\ (18) \Big\}.$$

Let $\bar{n} = \sum_{t=1}^{T} \bar{x}_t$, then $\bar{n} \geq n^*$.

Proof of Lemma 8. It is obvious that $(\bar{n}, 0, 0, \ldots, 0)$ is an optimal schedule for problem (A). Note that problem (**T1**) can be rewritten as

$$\min_{\boldsymbol{x} \in \mathbb{Z}_+^T} \ \mathbb{E}_{\omega^o}\Big[\Upsilon(\boldsymbol{x}, \omega^o) - C_I \sum_{t=1}^{T} \alpha_t(x_t, \omega^0)\Big] \tag{B}$$

where

$$\Upsilon(\boldsymbol{x},\omega^o)=\Big\{\sum_{t=1}^{\overline{T}} y_t^s + C_W\sum_{t=1}^{\overline{T}} y_t^w + C_D(y_T^s+y_T^w)\Big|(17),\ (18)\Big\}.$$

Consider two arbitrary solutions $\boldsymbol{x}'$ and $\boldsymbol{x}$ such that $\sum_{t=1}^T x_t=\bar{n}$, $\boldsymbol{x}'=\boldsymbol{x}+\boldsymbol{\delta}$ and $\boldsymbol{\delta}\geq 0$. Let $k=\sum_{t=1}^T\delta_t>0$. We shall show next that if we drop $\boldsymbol{\delta}$ from the schedule $\boldsymbol{x}+\boldsymbol{\delta}$ for problem (B), we will not be worse off. So $\bar{n}\geq n^*$, where n^* is the optimal total number of patients to schedule in problem (B).

For ease of discussion, we call patients in $\boldsymbol{x}$ the first $\bar{n}$ patients, and those in $\boldsymbol{\delta}$ the last k patients. Let Δ_A be the difference of $C_D\mathbb{E}[y_T^s]$ if $(\bar{n},0,0,\dots,0)+\boldsymbol{\delta}$ is applied in problem (A) rather than $(\bar{n},0,0,\dots,0)$, and let Δ_B be the difference of $C_D\mathbb{E}[y_T^s]$ if $\boldsymbol{x}+\boldsymbol{\delta}$ is applied in problem (B) rather than $\boldsymbol{x}$. Because the last k patients in $\boldsymbol{x}'$ will get served immediately after the first $\bar{n}$ patients finish their services, and obviously, schedule $(\bar{n},0,0,...,0)+\boldsymbol{\delta}$ is the one which can finish the first $\bar{n}$ patients most quickly, so $\Delta_A\leq\Delta_B$. Let Δ_α be the difference of $C_I\mathbb{E}[\sum_{t=1}^T\alpha_t(x_t)]$ if the total number of scheduled patients is $\bar{n}+k$ rather than $\bar{n}$. We have $\Delta_A\geq\Delta_\alpha$ since $(\bar{n},0,0,\dots,0)$ is optimal for problem (A). So we arrive at $\Delta_B\geq\Delta_\alpha$. Since $\mathbb{E}[y_t^s]$ and $\mathbb{E}[y_t^w]$ are nondecreasing in $\boldsymbol{x}$, then if we drop $\boldsymbol{\delta}$ from the schedule $\boldsymbol{x}+\boldsymbol{\delta}$ for problem (B), we will not be worse off. Thus, $\bar{n}\geq n^*$, completing the proof.

Indeed, this upper bound is constructed in a similar way to the one in Zacharias and Pinedo (2017) Theorem 1 (iii).

C.8. Proof of Proposition 4

We firstly prove that the equality between original formulation and the reformulated formulation, then we prove that for the reformulated version of equation (17) and (18), taking positive part can be translated to two linear inequalities.

It suffices to show that $\forall a_t,\ b_t,\ t=1,2,\dots,T$,

$$Pr(\alpha_t(x_t,\omega^o)=a_t,\beta_t(\omega^o)=b_t,\ \forall t=1,2,\dots,T)=Pr(\sum_{i=1}^{N_S}\gamma_{t,i}(\omega)z_{t,i}=a_t,\beta_t(\omega)=b_t,\ \forall t=1,2,\dots,T).$$

As events occurring in different slots are independent and walk-ins and scheduled patients are independent, we have

$$Pr(\alpha_t(x_t,\omega^o)=a_t,\beta_t(\omega^o)=b_t,\ \forall t=1,2,\dots,T)=\prod_{t=1}^T Pr(\alpha_t(x_t,\omega^o)=a_t)\cdot Pr(\beta_t(\omega^o)=b_t,\ \forall t=1,2,\dots,T).$$

Due to constraints of $z_{t,i}$ in (**T1-R**), we know that $\sum\limits_{i=1}^{N_S}\gamma_{t,i}(\omega)z_{t,i}$ and $\sum\limits_{i=1}^{N_S}\gamma_{s,i}(\omega)z_{s,i}$ have no overlapping terms, and thus are independent. It follows that

$$Pr(\sum_{i=1}^{N_S}\gamma_{t,i}(\omega)z_{t,i}=a_t,\beta_t(\omega)=b_t,\ \forall t=1,2,\dots,T)=\prod_{t=1}^T Pr(\sum_{i=1}^{N_S}\gamma_{t,i}(\omega)z_{t,i}=a_t)\cdot Pr(\beta_t(\omega)=b_t,\ \forall t=1,2,\dots,T).$$

For any t, we observe that

$$\begin{aligned}Pr(\sum_{i=1}^{N_S}\gamma_{t,i}(\omega)z_{t,i}=a_t)&=\binom{\sum_{i=1}^{N_S}z_{t,i}}{a_t}(q^s)^{a_t}(1-q^s)^{\sum_{i=1}^{N_S}z_{t,i}-a_t}=\binom{x_t}{a_t}(q^s)^{a_t}(1-q^s)^{x_t-a_t}\\&=Pr(\alpha_t(x_t,\omega^o)=a_t),\end{aligned}$$

where the second equality is resulted from our definition of the new decision variables $z_{t,i}$ that $x_t = \sum_{i=1}^{N_S} z_{t,i}$. So the reformulated formulation is equivalent to the original formulation.

Now we need to prove the second part. By definition, $\boldsymbol{y}_t$ ($\boldsymbol{y}_t^s$) is the number of waiting (scheduled) patients after time slot t. Then $y_{t-1} + \sum\limits_{i=1}^{N_S} \gamma_{t,i}(\omega) z_{t,i} + \beta_t(\omega) - y_t$ ($y_{t-1} - y_t$ when $t > T$) is the number of served patient at t. This number cannot be greater than 1. Then it must be 0 or 1 with integer constraints of y_t and y_t^s. Notice that the objective function is increasing in y_t and y_t^s, so the second stage problem in (**T1-R**) is choosing to serve a patient or not at each time to minimize the total weighted waiting patients. It is easy to see that the optimal solution must be serving a patient if there is any, which means that problem (**T1-R**) will give the same result as (**T1**).

C.9. Proof of Theorem 1

The result follows directly from the totally unimodularity of matrix $\boldsymbol{U}$ and Proposition 4.

C.10. Proof of Proposition 5

By strong duality, we have

$$\boldsymbol{v}(\boldsymbol{z}, \omega)^{tr}(\boldsymbol{M}(\omega)\boldsymbol{z} + \boldsymbol{d}(\omega)) = \Upsilon(\boldsymbol{z}, \omega), \ \forall \boldsymbol{z} \in \mathcal{Z}. \tag{25}$$

As (Dual) is a maximization problem, we know that

$$\boldsymbol{v}(\boldsymbol{z}', \omega)^{tr}(\boldsymbol{M}(\omega)\boldsymbol{z} + \boldsymbol{d}(\omega)) \le \Upsilon(\boldsymbol{z}, \omega), \ \forall \boldsymbol{z}' \in \mathcal{Z}. \tag{26}$$

Now, let $\Upsilon(\boldsymbol{z}') = \mathbb{E}_\omega[\Upsilon(\boldsymbol{z}', \omega)]$. Taking expectation of (26), adding $-h(\boldsymbol{z}')$ to both sides, we can see that for any given $\boldsymbol{z} \in \mathcal{Z}$,

$$\boldsymbol{a}(\boldsymbol{z}')\boldsymbol{z} + b(\boldsymbol{z}') - h(\boldsymbol{z}) \le \Upsilon(\boldsymbol{z}) - h(\boldsymbol{z}), \ \forall \boldsymbol{z}' \in \mathcal{Z}. \tag{27}$$

In particular, when $\boldsymbol{z}' = \boldsymbol{z}$, we have

$$\boldsymbol{a}(\boldsymbol{z})\boldsymbol{z} + b(\boldsymbol{z}) - h(\boldsymbol{z}) = \Upsilon(\boldsymbol{z}) - h(\boldsymbol{z}). \tag{28}$$

Let $\boldsymbol{z}^*$ be the optimal solution to (**T2**) with optimal value $\Upsilon(\boldsymbol{z}^*) - h(\boldsymbol{z}^*)$. It is clear that $(\boldsymbol{z}, u) = (\boldsymbol{z}^*, \Upsilon(\boldsymbol{z}^*) - h(\boldsymbol{z}^*))$ is a feasible solution to (**T2-D**). Now, let $(\boldsymbol{z}^{**}, u^{**})$ be the optimal solution to (**T2-D**), then there must be $u^{**} \le \Upsilon(\boldsymbol{z}^*) - h(\boldsymbol{z}^*)$. Suppose that $u^{**} < \Upsilon(\boldsymbol{z}^*) - h(\boldsymbol{z}^*)$. Then we have

$$\boldsymbol{a}(\boldsymbol{z}^{**})\boldsymbol{z}^{**} + b(\boldsymbol{z}^{**}) - h(\boldsymbol{z}^{**}) = \Upsilon(\boldsymbol{z}^{**}) - h(\boldsymbol{z}^{**}) \le u^{**} < \Upsilon(\boldsymbol{z}^*) - h(\boldsymbol{z}^*),$$

where the first equality is from (28) and the second inequality follows from (19). This contradicts with that $\boldsymbol{z}^*$ is the optimal solution to (**T2**), and therefore, there must be that $u^{**} = \Upsilon(\boldsymbol{z}^*) - h(\boldsymbol{z}^*)$.

This completes the proof.

C.11. Proof of Theorem 2

Following Proposition 5, we have two facts which provide theoretical support to our algorithmic approach. We use $(\boldsymbol{z}^*, u^*)$ to denote the optimal solution to (**T2-D**). Recall that $\boldsymbol{z}^*$ is also the optimal solution of (**T2**) and $u^* = \Upsilon(\boldsymbol{z}^*) - h(\boldsymbol{z}^*)$.

Fact 1. If $\overline{u} \ge u^*$ and a solution $\boldsymbol{z}^0$ does not satisfy $\boldsymbol{a}(\boldsymbol{z}')\boldsymbol{z}^0 + b(\boldsymbol{z}') - h(\boldsymbol{z}^0) \le \overline{u}$ for some $\boldsymbol{z}' \in \mathcal{Z}$, then $\boldsymbol{z}^0$ is not an optimal solution.

Fact 1 follows that $\varUpsilon(\boldsymbol{z}^*) - h(\boldsymbol{z}^*) \le \overline{u} < \boldsymbol{a}(\boldsymbol{z}')\boldsymbol{z}^0 + b(\boldsymbol{z}') - h(\boldsymbol{z}^0) \le \varUpsilon(\boldsymbol{z}^0) - h(\boldsymbol{z}^0)$.

Fact 2. If $\boldsymbol{z}^0$ is not an optimal solution, then its corresponding constraint $\boldsymbol{a}(\boldsymbol{z}^0)\boldsymbol{z} + b(\boldsymbol{z}^0) - h(\boldsymbol{z}) \le u$ in (**T2-D**) is redundant.

Fact 2 follows that $\boldsymbol{a}(\boldsymbol{z}^0)\boldsymbol{z}^* + b(\boldsymbol{z}^0) - h(\boldsymbol{z}^*) \le u^*$, indicating that $\boldsymbol{a}(\boldsymbol{z}^0)\boldsymbol{z} + b(\boldsymbol{z}^0) - h(\boldsymbol{z}) \le u$ is not active in the optimal condition.

Fact 1 suggests that we may eliminate a non-optimal solution $\boldsymbol{z}$ without fully solving $u(\boldsymbol{z})$. Fact 2 implies that if $\boldsymbol{z}^0$ is not optimal, then we do not need $\boldsymbol{a}(\boldsymbol{z}^0)$ and $b(\boldsymbol{z}^0)$. These are main ideas of Algorithm 1.

By Fact 1, a solution which does not satisfy a subset of (19) is not optimal. So in Algorithm 1, we get a solution which satisfies the current set of (19) after the inequality check in Line 6. By Fact 2, constraints associated with non-optimal solutions are redundant. Thus after Line 6, we only generate the constraints associated with the solution which satisfy the inequality (which is potentially optimal). The loop is not endless, because there are only a finite number of constraints that can be added. Once all neighbors of the current $\boldsymbol{z}^*$ are checked and no one can pass the condition check in Line 6 or improve $\overline{u}$, then the “while” loop is broken. We arrive at a local optimal solution. By Corollary 2, it is the global optimum.

D. Allocation Rule of the Heuristic

This allocation rule reserves holes by taking walk-ins into account, and it may also overbook in the early slots (i.e., front-loading) to reduce the impact of no-shows. It has two phases.

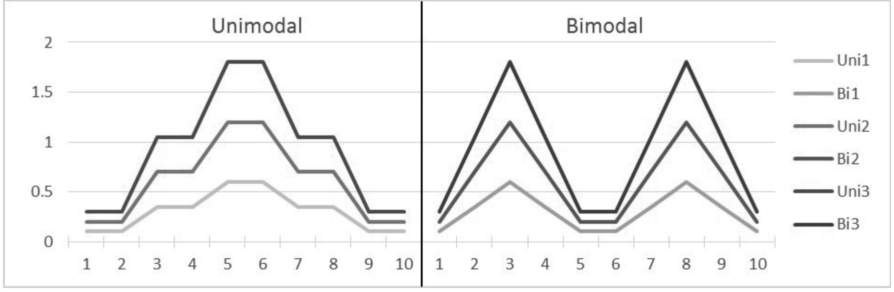
- First, we try to match "supply" and "demand" in each slot by calibrating the expected number of patients who arrive for service is each slot to be 1, adjusting for their waiting costs. Specifically, for each slot t, we calculate $EC_t = 1 - C_W \times \overline{\beta_t}$, which can be viewed as the remaining effective capacity in slot t for scheduled patients. (If $C_W = 0$, then walk-ins need not be counted and the whole slot t can be used by a scheduled patient. If $C_W = 1$, then the remaining effective capacity for scheduled patients is exactly $1 - \overline{\beta_t}$ because both walk-ins and scheduled patients are treated the same.) We then sequentially schedule $max(0, \lfloor \frac{1-C_W \times \overline{\beta_t}}{q^s} \rfloor)$ patients to slots $t = 1, 2, \ldots$. Note that if there are too many walk-ins in a slot (i.e., $\overline{\beta_t}$ is large) or their waiting cost rate C_W is high, we reserve holes.
- If we cannot exhaust allocating all n^h patients in the first phase, we consider front-loading. Specifically, we front-load up to $|\frac{1}{q^s}|$ patients to slots $1, 2, \ldots, T$ in a sequential manner. We repeat this process, if necessary, until all n^h patients are allocated.

Algorithm 2 Allocation Rule

```
$\overline{\beta_t}$ is the expected walk-ins in slot $t$, $q^s$ is the show-up probability
Initialize the solution $x_t \leftarrow 0$ for all $t$
Calculate the capacity of each slot by $EC_t \leftarrow 1 - C_W \times \overline{\beta_t}$
Let $n \leftarrow n^*$ be the number of unscheduled patients
for $t = 1 : T$ do
    Set the batch size $B_t \leftarrow \max(0, \lfloor \frac{EC_t}{q^s} \rfloor)$
    $x_t \leftarrow x_t + \min(n, B_t)$
    $n \leftarrow n - \min(n, B_t)$
end for
Set the batch size $B \leftarrow |\frac{1}{q^s}|$
while $n > 0$ do
    for $t = 1 : T$ do
        $x_t \leftarrow x_t + \min(n, B)$
        $n \leftarrow n - \min(n, B)$
    end for
end while
return $\boldsymbol{x} = (x_1, x_2, \ldots, x_T)$
```

E. Supplementary Figures and Tables of the Numerical Study

Figure E9 Expected Number of Walk-ins for $T = 10$.



Note. The expected number of walk-ins in each time slot for each scenario is specified as follows. Uni1: (0.1,0.1,0.35,0.35,0.6,0.6,0.35,0.35,0.1,0.1), Uni2 doubles Uni1, and Uni3 triples Uni1; Bi1: (0.1,0.35,0.6,0.35,0.1,0.1,0.35,0.6,0.35,0.1), Bi2 doubles Bi1, and Bi3 triples Bi1.

Table E10 Comparison of Computation Times (sec)

Parameters				Unimodal				Bimodal			
No-show Probability	C_I	C_D	C_W	Local Search	MILP	LS + LR	CGA	Local Search	MILP	LS + LR	CGA
0.5	5	15	0.5	180	1124	72	65	717	452	231	53
0.5	5	15	0.9	137	276	55	28	703	74	64	31
0.5	5	25	0.5	124	438	28	12	297	78	78	31
0.5	5	25	0.9	123	155	24	13	296	28	41	22
0.5	10	15	0.5	1421	62384	981	731	2521	4498	743	437
0.5	10	15	0.9	905	26573	761	731	2127	1482	383	228
0.5	10	25	0.5	158	4122	69	90	733	1634	194	118
0.5	10	25	0.9	188	1757	147	111	940	688	74	57
0.1	5	15	0.5	129	52	33	6	619	82	64	16
0.1	5	15	0.9	83	72	18	6	210	83	39	9
0.1	5	25	0.5	82	59	18	5	211	54	41	9
0.1	5	25	0.9	79	53	18	4	209	80	41	5
0.1	10	15	0.5	586	484	144	43	2159	188	197	46
0.1	10	15	0.9	337	826	174	38	915	90	129	28
0.1	10	25	0.5	219	88	111	24	984	122	114	16
0.1	10	25	0.9	131	121	33	11	593	94	69	22

Notes. (1) $T = 14$ and average walk-in rate is 0.6 per appointment slot. (2) Computation time includes the time of random sample generation and solving the model for the last three approaches. (3) All computations were conducted in a laptop computer equipped with Intel Core i7-5500U 2.40 GHz CPU, 8.00 GB RAM, 64-bit Windows 10 OS, R x64 3.3.1 and Gurobi 7.0.2.

Table E11 Computation Times of Solving General Problem Instances (sec)

Approach			CGA		MILP		
Walk-in Pattern			Correlated		Independent		Correlated
No-show Probability			$p=0.5$	$p=0.1$	$p_H=0.5$ $p_L=0.1$	$p=0.3$	$p_H=0.5$ $p_L=0.1$
C_I	C_D	C_W	Unimodal Walk-ins				
5	15	0.5	14	2	271	948	156
5	15	0.9	11	2	100	277	92
5	25	0.5	4	1	164	521	101
5	25	0.9	3	1	88	131	62
10	15	0.5	90	14	13832	5258	8675
10	15	0.9	24	6	3835	2355	1899
10	25	0.5	54	4	1881	1715	2031
10	25	0.9	22	5	742	1050	773
C_I	C_D	C_W	Bimodal Walk-ins				
5	15	0.5	34	2	533	735	134
5	15	0.9	20	2	179	213	88
5	25	0.5	8	2	181	346	79
5	25	0.9	7	2	84	167	61
10	15	0.5	271	25	9116	5117	1617
10	15	0.9	138	17	1303	1949	850
10	25	0.5	47	9	2073	1067	321
10	25	0.9	21	5	1261	890	312

Notes. (1) $T = 14$ and average walk-in rate is 0.6 per appointment slot. (2) In problem instances with correlated walk-ins, the correlation matrix is randomly generated and walk-ins in different time slots are assumed to follow a multivariate Poisson distribution. (3) In problem instances with homogeneous patient no-show probabilities, p denotes the no-show probability. (4) When patient no-show probabilities are heterogeneous, p_H and p_L denote the high and low patient no-show probabilities we consider, respectively. In these problem instances, we add a constraint that ensures the difference between the total numbers of patients with different no-show probabilities to schedule is at most 1, to avoid trivial solutions. (4) All computations were conducted in a desktop computer equipped with Intel Core i5-4590 3.30 GHz CPU, 4.00 GB RAM, 64-bit Windows 7 OS, R x64 3.3.1 and Gurobi 7.0.2.

Table E12 Computation Times of Solving Large-Scale Problems ($T = 30$)

Parameters				**Time to Optimum by CGA** (sec)	
No-show Probability	C_I	C_D	C_W	Unimodal	Bimodal
0.5	10	15	0.5	87654	86723
0.1	10	15	0.5	87378	164436

Notes. (1) We choose the parameter settings under which the computation takes the longest time in Table E10 above. (2) All computations were conducted in a laptop computer equipped with Intel Core i7-5500U 2.40 GHz CPU, 8.00 GB RAM, 64-bit Windows 10 OS, R x64 3.3.1 and Gurobi 7.0.2.

Table E13 Optimal Cost and Optimal Number of Scheduled Patients ($T = 10$)

No-show Probability	Cost Structure			No Walk-in		Uni1		Bi1		Uni2		Bi2		Uni3		Bi3	
	C_I	C_D	C_W	C^*	n^*	C^*	n^*	C^*	n^*	C^*	n^*	C^*	n^*	C^*	n^*	C^*	n^*
0.5	5	15	0.5	12.63	17	18.93	10	18.98	10	22.16	4	23.07	4	44.31	2	33.75	1
0.1	5	15	0.5	5.00	10	14.81	6	14.95	6	20.22	3	21.20	3	43.45	1	32.54	1
0.5	10	15	0.5	16.13	20	24.44	14	24.64	14	27.34	8	27.51	8	37.84	3	32.77	3
0.1	10	15	0.5	5.71	11	17.74	8	17.56	8	23.60	5	23.97	5	34.89	2	30.96	2
0.5	5	25	0.5	14.62	16	21.29	9	21.26	9	27.12	3	28.71	3	68.33	1	49.70	1
0.1	5	25	0.5	5.00	10	17.46	6	17.57	6	25.25	2	27.26	2	67.06	1	48.82	1
0.5	10	25	0.5	22.43	18	31.21	11	31.28	11	35.89	6	37.16	6	63.49	2	50.83	2
0.1	10	25	0.5	8.85	11	24.04	7	24.18	7	32.54	3	34.04	4	60.73	2	49.32	1
0.5	5	15	0.9	12.63	17	21.61	9	21.70	9	25.78	3	26.42	3	54.74	2	39.61	1
0.1	5	15	0.9	5.00	10	16.87	6	17.37	6	23.84	2	24.84	3	53.48	1	38.65	1
0.5	10	15	0.9	16.13	20	29.47	12	29.65	12	33.49	7	33.52	7	49.28	3	40.48	2
0.1	10	15	0.9	5.71	11	21.32	8	21.90	8	28.96	4	29.21	4	46.18	2	39.00	2
0.5	5	25	0.9	14.62	16	23.39	8	23.51	8	30.13	3	31.61	3	78.20	1	55.56	1
0.1	5	25	0.9	5.00	10	19.18	5	19.69	5	28.11	2	30.19	2	77.09	1	54.92	1
0.5	10	25	0.9	22.43	18	35.10	11	35.34	11	40.51	5	41.65	5	73.92	2	57.48	2
0.1	10	25	0.9	8.85	11	27.24	7	27.70	7	36.52	3	38.12	3	72.03	2	55.42	1

Notes. C^* is the cost of the optimal schedule; n^* represents the optimal number of scheduled patients; walk-in patterns Uni1, Bi1, Uni2, Bi2, Uni3 and Bi3 are specified in Figure E9.

Table E14 Comparison of the Cost and the Number of Scheduled Patients under the Heuristic and Optimal Solutions

No-show	Cost Structure			Uni1		Bi1		Uni2		Bi2		Uni3		Bi3	
Probability	c_I	c_D	c_W	$\frac{C^h}{C^*}-1$	n^h-n^*	$\frac{C^h}{C^*}-1$	n^h-n^*	$\frac{C^h}{C^*}-1$	n^h-n^*	$\frac{C^h}{C^*}-1$	n^h-n^*	$\frac{C^h}{C^*}-1$	n^h-n^*	$\frac{C^h}{C^*}-1$	n^h-n^*
$T=10$															
0.5	5	15	0.5	14%	1	14%	1	14%	1	6%	1	1%	-1	0%	0
0.1	5	15	0.5	18%	0	10%	0	10%	0	5%	0	0%	0	0%	0
0.5	10	15	0.5	14%	2	14%	2	11%	2	10%	2	17%	1	15%	1
0.1	10	15	0.5	17%	1	17%	1	10%	1	7%	1	0%	0	0%	0
0.5	5	25	0.5	21%	0	19%	0	4%	0	3%	0	0%	0	0%	0
0.1	5	25	0.5	28%	-1	20%	-1	12%	-1	6%	-1	0%	0	0%	0
0.5	10	25	0.5	10%	1	11%	1	20%	0	11%	0	4%	-1	3%	-1
0.1	10	25	0.5	12%	0	6%	0	12%	0	7%	-1	3%	-1	0%	0
0.5	5	15	0.9	7%	2	8%	2	12%	2	9%	2	0%	-1	0%	0
0.1	5	15	0.9	20%	0	9%	0	2%	1	5%	0	0%	0	0%	0
0.5	10	15	0.9	25%	4	24%	4	15%	3	16%	3	15%	1	14%	2
0.1	10	15	0.9	34%	1	29%	1	27%	2	23%	2	0%	0	0%	0
0.5	5	25	0.9	14%	1	16%	1	11%	0	5%	0	0%	0	0%	0
0.1	5	25	0.9	25%	0	15%	0	9%	-1	3%	-1	0%	0	0%	0
0.5	10	25	0.9	5%	1	6%	1	10%	1	7%	1	2%	-1	1%	-1
0.1	10	25	0.9	14%	0	7%	0	0%	0	3%	0	1%	-1	0%	0
$T=14$															
0.5	5	15	0.5	10%	2	11%	2	15%	1	7%	1	3%	-2	1%	-1
0.1	5	15	0.5	24%	0	14%	0	26%	0	20%	0	3%	-1	0%	0
0.5	10	15	0.5	20%	3	21%	3	17%	4	14%	4	12%	0	9%	1
0.1	10	15	0.5	22%	1	21%	1	15%	2	13%	2	0%	0	2%	1
0.5	5	25	0.5	20%	0	18%	1	7%	0	4%	0	1%	-1	0%	0
0.1	5	25	0.5	31%	-1	20%	-1	16%	0	7%	0	1%	-1	0%	0
0.5	10	25	0.5	7%	1	9%	1	24%	2	14%	1	8%	-3	4%	-2
0.1	10	25	0.5	15%	0	9%	0	19%	0	14%	0	8%	-1	2%	-1
0.5	5	15	0.9	11%	3	13%	3	16%	3	13%	2	1%	-1	0%	0
0.1	5	15	0.9	33%	1	24%	1	3%	1	10%	1	1%	-1	0%	0
0.5	10	15	0.9	37%	5	36%	5	23%	5	24%	6	9%	1	9%	2
0.1	10	15	0.9	42%	2	39%	2	38%	3	33%	3	1%	1	7%	1
0.5	5	25	0.9	13%	2	15%	2	9%	1	5%	1	1%	-1	0%	0
0.1	5	25	0.9	29%	-1	17%	-1	0%	0	3%	0	0%	0	0%	0
0.5	10	25	0.9	6%	2	7%	2	18%	3	15%	2	5%	-2	2%	-1
0.1	10	25	0.9	24%	1	17%	1	4%	1	7%	1	5%	-1	0%	0

Notes. C^* is the cost of the optimal schedule; n^* is the optimal number of scheduled patients; C^h is the cost of the heuristic schedule; n^h is the number of scheduled patients under the heuristic schedule; walk-in patterns Uni1, Bi1, Uni2, Bi2, Uni3 and Bi3 for $T=10$ are specified in Figure E9; their counterparts for $T=14$ are specified in Figure 2 of the main paper.

Table E15 **Changes in Objective Function Components by Adopting the Optimal Schedule for Provider KNI (Change Unit: Slots)**

$\Delta\Gamma_S$ $\Delta\Gamma_W$	Cost Structure: (C_I, C_D, C_W)															
$\Delta\Gamma_I$ $\Delta\Gamma_O$	(5,15,0.5)		(5,15,0.9)		(5,25,0.5)		(5,25,0.9)		(10,15,0.5)		(10,15,0.9)		(10,25,0.5)		(10,25,0.9)	
5/6/2011	1.20	4.57	1.20	4.84	1.20	5.42	1.36	5.79	-1.38	0.52	0.94	2.05	0.98	3.07	1.03	3.40
	-0.85	1.29	-0.87	1.27	-1.35	1.43	-1.35	1.42	0.35	0.56	-0.02	0.83	-0.36	1.13	-0.38	1.11
5/7/2011	56.15	20.85	56.15	21.12	56.15	21.70	56.32	22.07	53.57	16.80	55.90	18.33	55.94	19.35	55.99	19.68
	-1.42	8.40	-1.45	8.38	-1.92	8.54	-1.93	8.53	-0.23	7.67	-0.60	7.94	-0.94	8.24	-0.96	8.22
5/13/2011	9.67	11.61	9.67	11.87	9.67	12.46	9.83	12.83	7.08	7.56	9.41	9.08	9.45	10.11	9.50	10.44
	-1.17	3.52	-1.20	3.50	-1.68	3.66	-1.68	3.66	0.02	2.80	-0.35	3.06	-0.69	3.37	-0.71	3.35
5/14/2011	58.22	20.92	58.22	21.19	58.22	21.78	58.39	22.14	55.64	16.87	57.97	18.40	58.01	19.42	58.06	19.76
	-1.40	8.42	-1.43	8.40	-1.90	8.56	-1.91	8.56	-0.21	7.70	-0.58	7.96	-0.92	8.27	-0.94	8.25
5/20/2011	9.20	10.41	9.20	10.68	9.20	11.27	9.37	11.63	6.62	6.36	8.95	7.89	8.98	8.91	9.03	9.25
	-1.28	2.78	-1.31	2.75	-1.78	2.91	-1.79	2.91	-0.09	2.05	-0.46	2.32	-0.80	2.62	-0.82	2.60
5/21/2011	27.57	15.48	27.57	15.75	27.57	16.34	27.73	16.70	24.98	11.44	27.31	12.96	27.35	13.99	27.40	14.32
	-1.45	4.53	-1.48	4.51	-1.95	4.67	-1.96	4.66	-0.26	3.80	-0.63	4.07	-0.97	4.37	-0.99	4.35
5/27/2011	18.12	12.65	18.12	12.91	18.12	13.50	18.29	13.87	15.54	8.60	17.87	10.12	17.90	11.15	17.95	11.48
	-1.26	3.44	-1.29	3.41	-1.77	3.57	-1.77	3.57	-0.07	2.71	-0.44	2.98	-0.78	3.28	-0.80	3.26
5/28/2011	17.65	12.77	17.65	13.04	17.65	13.63	17.82	13.99	15.07	8.72	17.40	10.25	17.44	11.27	17.49	11.61
	-1.29	3.41	-1.31	3.38	-1.79	3.55	-1.80	3.54	-0.10	2.68	-0.47	2.95	-0.81	3.25	-0.83	3.23

Notes. (1) $T=8$ and the no-show rate is 0.36. (2) Rows represent clinic sessions in different days, and columns for different cost parameter settings. (3) Each cell contains four numbers, the upper left being the reduction of scheduled patients' wait time, the upper right being that of walk-ins' wait time, the lower left being that of provider idle time and the lower right being that of provider overtime. (4) The measurement unit is appointment slot.

Table E16 **Changes in Objective Function Components by Adopting the Heuristic Schedule for Providers GAR and KNI (Change Unit: Slots)**

$\Delta\Gamma_S$ $\Delta\Gamma_W$ $\Delta\Gamma_I$ $\Delta\Gamma_O$	Cost Structure: (C_I, C_D, C_W)															
	(5,15,0.5)		(5,15,0.9)		(5,25,0.5)		(5,25,0.9)		(10,15,0.5)		(10,15,0.9)		(10,25,0.5)		(10,25,0.9)	
GAR																
7/1/2011	-1.91	6.10	-1.91	6.10	-1.91	6.10	-1.91	6.10	-10.57	-8.40	-14.05	-8.69	-2.39	3.03	-4.88	-0.10
	-1.07	1.46	-1.07	1.46	-1.07	1.46	-1.07	1.46	1.33	0.48	1.27	0.43	-0.35	1.33	0.32	1.17
7/8/2011	0.08	2.18	0.08	2.18	0.08	2.18	0.08	2.18	-8.58	-12.32	-12.07	-12.61	-0.41	-0.89	-2.89	-4.02
	0.09	1.78	0.09	1.78	0.09	1.78	0.09	1.78	2.50	0.81	2.44	0.75	0.81	1.66	1.49	1.49
7/15/2011	-1.20	-2.61	-1.20	-2.61	-1.20	-2.61	-1.20	-2.61	-9.85	-17.11	-13.34	-17.41	-1.68	-5.68	-4.16	-8.81
	1.12	0.28	1.12	0.28	1.12	0.28	1.12	0.28	3.52	-0.70	3.47	-0.75	1.84	0.15	2.52	-0.01
7/22/2011	-1.18	8.21	-1.18	8.21	-1.18	8.21	-1.18	8.21	-9.84	-6.28	-13.32	-6.58	-1.66	5.14	-4.15	2.02
	-1.92	1.46	-1.92	1.46	-1.92	1.46	-1.92	1.46	0.48	0.48	0.43	0.43	-1.20	1.33	-0.52	1.17
7/29/2011	1.17	14.02	1.17	14.02	1.17	14.02	1.17	14.02	-7.49	-0.48	-10.97	-0.78	0.69	10.95	-1.80	7.82
	-1.43	2.78	-1.43	2.78	-1.43	2.78	-1.43	2.78	0.97	1.81	0.91	1.75	-0.71	2.66	-0.04	2.49
8/5/2011	1.42	0.40	1.42	0.40	1.42	0.40	1.42	0.40	-7.23	-14.10	-10.72	-14.40	0.94	-2.67	-1.54	-5.80
	0.73	2.42	0.73	2.42	0.73	2.42	0.73	2.42	3.13	1.44	3.08	1.39	1.45	2.29	2.13	2.13
KNI																
5/6/2011	-3.10	3.82	-1.07	5.06	0.07	6.24	0.07	6.24	-9.81	-0.47	-9.81	-0.47	-3.10	3.82	-3.10	3.82
	-0.80	1.34	-1.35	1.43	-1.91	1.50	-1.91	1.50	0.36	0.57	0.36	0.57	-0.80	1.34	-0.80	1.34
5/7/2011	51.85	20.10	53.88	21.34	55.03	22.52	55.03	22.52	45.15	15.81	45.15	15.81	51.85	20.10	51.85	20.10
	-1.38	8.45	-1.92	8.54	-2.49	8.61	-2.49	8.61	-0.22	7.68	-0.22	7.68	-1.38	8.45	-1.38	8.45
5/13/2011	5.37	10.86	7.39	12.10	8.54	13.28	8.54	13.28	-1.34	6.57	-1.34	6.57	5.37	10.86	5.37	10.86
	-1.13	3.57	-1.68	3.66	-2.24	3.74	-2.24	3.74	0.03	2.80	0.03	2.80	-1.13	3.57	-1.13	3.57
5/14/2011	53.92	20.18	55.95	21.42	57.10	22.59	57.10	22.59	47.22	15.88	47.22	15.88	53.92	20.18	53.92	20.18
	-1.36	8.47	-1.90	8.56	-2.47	8.64	-2.47	8.64	-0.20	7.70	-0.20	7.70	-1.36	8.47	-1.36	8.47
5/20/2011	4.90	9.67	6.93	10.91	8.08	12.08	8.08	12.08	-1.81	5.37	-1.81	5.37	4.90	9.67	4.90	9.67
	-1.24	2.82	-1.78	2.92	-2.35	2.99	-2.35	2.99	-0.08	2.06	-0.08	2.06	-1.24	2.82	-1.24	2.82
5/21/2011	23.27	14.74	25.29	15.98	26.44	17.15	26.44	17.15	16.56	10.45	16.56	10.45	23.27	14.74	23.27	14.74
	-1.41	4.58	-1.95	4.67	-2.52	4.74	-2.52	4.74	-0.25	3.81	-0.25	3.81	-1.41	4.58	-1.41	4.58
5/27/2011	13.82	11.90	15.85	13.14	17.00	14.32	17.00	14.32	7.12	7.61	7.12	7.61	13.82	11.90	13.82	11.90
	-1.22	3.48	-1.76	3.58	-2.33	3.65	-2.33	3.65	-0.06	2.72	-0.06	2.72	-1.22	3.48	-1.22	3.48
5/28/2011	13.35	12.03	15.38	13.27	16.53	14.44	16.53	14.44	6.65	7.74	6.65	7.74	13.35	12.03	13.35	12.03
	-1.24	3.45	-1.79	3.55	-2.36	3.62	-2.36	3.62	-0.09	2.69	-0.09	2.69	-1.24	3.45	-1.24	3.45

Notes. (1) For Provider GAR, $T = 12$ and the no-show rate is 0.16. (2) For Provider KNI, $T = 8$ and the no-show rate is 0.36. (2) Rows represent clinic sessions in different days, and columns for different cost parameter settings. (3) Each cell contains four numbers, the upper left being the reduction of scheduled patients' wait time, the upper right being that of walk-ins' wait time, the lower left being that of provider idle time and the lower right being that of provider overtime. (4) The measurement unit is appointment slot.

Table E17 Performance Improvement for Providers KNI and GAR by the Optimal Solution.

	$C_I = 5$				$C_I = 10$			
	$C_D = 15$		$C_D = 25$		$C_D = 15$		$C_D = 25$	
	$C_W = 0.5$	$C_W = 0.9$	$C_W = 0.5$	$C_W = 0.9$	$C_W = 0.5$	$C_W = 0.9$	$C_W = 0.5$	$C_W = 0.9$
Provider GAR								
Optimal C^*	22.69	26.65	27.80	30.94	26.65	32.99	36.04	41.29
Optimal n^*	5	5	4	4	7	7	6	6
7/1/11 $n = 7$	36%	35%	48%	48%	26%	21%	34%	31%
7/8/11 $n = 6$	50%	46%	58%	56%	47%	39%	50%	46%
7/15/11 $n = 3$	27%	21%	26%	23%	46%	36%	35%	29%
7/22/11 $n = 8$	32%	33%	46%	47%	11%	10%	26%	25%
7/29/11 $n = 9$	57%	57%	67%	66%	41%	39%	53%	51%
8/5/11 $n = 6$	59%	55%	67%	64%	56%	48%	59%	55%
Provider KNI								
Optimal C^*	13.26	15.41	15.75	17.58	14.67	18.34	21.23	24.06
Optimal n^*	6	6	5	5	9	8	7	7
5/6/11 $n = 10$	49%	49%	63%	63%	24%	21%	41%	40%
5/7/11 $n = 22$	90%	90%	93%	93%	82%	84%	87%	88%
5/13/11 $n = 14$	75%	75%	83%	82%	56%	57%	70%	70%
5/14/11 $n = 22$	90%	91%	93%	93%	84%	85%	88%	88%
5/20/11 $n = 13$	70%	71%	79%	79%	48%	50%	64%	64%
5/21/11 $n = 16$	81%	83%	87%	87%	68%	71%	77%	78%
5/27/11 $n = 14$	77%	77%	83%	83%	61%	62%	71%	72%
5/28/11 $n = 14$	75%	77%	82%	83%	57%	61%	70%	71%

Note: (1) For Provider GAR, $T = 12$ and the no-show rate is 0.16. (2) For Provider KNI, $T = 8$ and the no-show rate is 0.36. (3) Percentage improvement is evaluated as the percentage reduction in the expected total cost due to adopting the optimal schedule to replace the observed schedule.

Table E18 Performance Improvement for Providers KNI and GAR by the Heuristic Solution.

	$C_I=5$				$C_I=10$			
	$C_D=15$		$C_D=25$		$C_D=15$		$C_D=25$	
	$C_W=0.5$	$C_W=0.9$	$C_W=0.5$	$C_W=0.9$	$C_W=0.5$	$C_W=0.9$	$C_W=0.5$	$C_W=0.9$
Provider GAR								
Heuristic C^h	22.97	27.46	27.91	31.25	29.67	40.13	36.23	42.06
Heuristic n^h	5	5	4	4	9	9	6	6
7/1/2011 $n=7$	35%	33%	48%	47%	18%	4%	33%	30%
7/8/2011 $n=6$	49%	44%	58%	56%	41%	26%	50%	45%
7/15/2011 $n=3$	26%	18%	26%	22%	40%	22%	35%	27%
7/22/2011 $n=8$	31%	31%	46%	46%	1%	-10%	25%	24%
7/29/2011 $n=9$	56%	55%	67%	66%	34%	26%	52%	50%
8/5/2011 $n=6$	58%	53%	67%	64%	51%	37%	59%	54%
Provider KNI								
Heuristic C^h	15.85	20.85	18.23	21.86	20.01	26.52	27.76	32.51
Heuristic n^h	6	6	5	5	10	10	7	7
5/6/2011 $n=10$	39%	30%	57%	53%	-4%	-14%	23%	19%
5/7/2011 $n=22$	88%	87%	91%	91%	76%	77%	84%	84%
5/13/2011 $n=14$	70%	66%	80%	78%	41%	38%	61%	60%
5/14/2011 $n=22$	88%	87%	92%	91%	78%	78%	84%	84%
5/20/2011 $n=13$	64%	61%	76%	74%	29%	28%	53%	52%
5/21/2011 $n=16$	78%	77%	85%	84%	56%	58%	70%	71%
5/27/2011 $n=14$	72%	69%	80%	79%	47%	45%	62%	62%
5/28/2011 $n=14$	70%	68%	80%	79%	41%	43%	60%	61%

Notes. (1) For Provider GAR, $T=12$ and the no-show rate is 0.16. (2) For Provider KNI, $T=8$ and the no-show rate is 0.36. (3) Percentage improvement is evaluated as the percentage reduction in the expected total cost due to adopting the heuristic schedule to replace the observed schedule.

F. Dealing with Random Service Times

As discussed before, it is reasonable to assume deterministic service times in many practical contexts as providers can often adjust their time with patients depending on the progress of the day. It is, however, sometimes important to consider the variability in service times when scheduling patients. There are empirical evidences suggesting that provider service times may follow the exponential distribution (Kopach et al. 2007). Some previous literature, such as Kaandorp and Koole (2007), Hassin and Mendel (2008), has also considered exponentially distributed service times in their scheduling models. Our models can be extended to incorporate such random service times.

Specifically, suppose that the service times of patients are i.i.d. exponential random variables with mean ϕ. Then, the number of potential departures within a single slot of time (given that there are enough patients waiting) has a Poisson distribution with mean $1/\phi$. Let δ_t be this potential number of departures in slot t, then $Pr(\delta_t = i) = \frac{\phi^{-i}}{i!}e^{-1/\phi}$, $i = 0, 1, 2, \ldots$. Note that this distributional result is independent of time given the memoryless property of the exponential distribution.

Recall that (17) defines the relationship between y_t and y_{t-1}, where y_t is the total number of patients waiting at the end of t. By expanding the definition of the random scenario ω^o to include the uncertainty of random service times, we can redefine the relationship of y_t and y_{t-1} as follows.

$$y_t = \begin{cases} (y_{t-1} + \alpha_t(x_t, \omega^o) + \beta_t(\omega^o) - \delta_t(\omega^o))^+ & \text{for } 1 \le t \le T \text{ with } y_0 = 0, \\ (y_{t-1} - \delta_t(\omega^o))^+ & \text{for } T < t \le \overline{T}, \end{cases} \tag{29}$$

where $\delta_t(\omega^o)$ is the aforementioned number of potential maximum departures in slot t associated with scenario ω^o. The correctness of the recursive equation (29) follows from the memoryless property of the exponential distribution.

Similarly, we can redefine the relationships between y_t^s and y_{t-1}^s, where y_t^s is the total number of scheduled patients waiting at the end of t. Then, we can follow the same reformulation process above and use the same solution approach to solve this new problem. We note that, however, if the service time distribution is *not* exponential, the problem becomes much more complicated because we would need to record the service starting time of each patient in the system state. We leave this more general extension to future research.