Manufacturing & Service Operations Management

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

A Scenario-Based Approach for Operating Theater Scheduling Under Uncertainty

Nickolas K. Freeman, Sharif H. Melouk, John Mittenthal

MANUFACTURING & SERVICE OPERATIONS MANAGEMENT
Vol. 18, No. 2, Spring 2016, pp. 245–261
ISSN 1523-4614 (print) | ISSN 1526-5498 (online)
http://dx.doi.org/10.1287/msom.2015.0557

A Scenario-Based Approach for Operating Theater Scheduling Under Uncertainty

Nickolas K. Freeman
University of Houston, Houston, Texas 77204, nfreeman@bauer.uh.edu

Sharif H. Melouk, John Mittenthal
University of Alabama, Tuscaloosa, Alabama 35487, {smelouk@cba.ua.edu, jmittent@cba.ua.edu}

Elective operation scheduling significantly affects the financial health of a hospital and additional metrics important to stakeholders in an operating theater (OT) environment. In this research, we develop a novel scheduling formulation that explicitly considers the uncertainty in elective operation durations and also plans for potential randomly arriving urgent demands. Using a scenario-based modeling approach, the objective of this formulation is to maximize the expected profit associated with the OT schedule. Since the complexity of the problem is NP-hard, we develop a two-step, heuristic solution approach that allows us to solve practical-sized instances in reasonable time. Experimentation shows that incorporating uncertainty via scenarios increases profit and OT utilization when compared to deterministic scheduling methods. Moreover, explicitly considering urgent arrivals results in a significant reduction in the time that patients of this type wait to receive service, with little impact on other key metrics.

Keywords: healthcare management; math programming
History: Received: May 26, 2014; accepted: July 28, 2015. Published online in *Articles in Advance* November 6, 2015.

1. Introduction

Healthcare operations are of increasing interest in the literature. The survey papers of Cardoen et al. (2010a), Guerriero and Guido (2011), Gupta and Denton (2008), and Rais and Viana (2011) demonstrate that many recent works focus on a specific unit within hospitals, the operating theater. According to Guerriero and Guido (2011), the operating theater (OT) consists of multiple operating rooms (ORs) and is one of the most critical and expensive resources of a hospital. Specifically, recent estimates indicate the OT accounts for more than 40% of total hospital revenue as well as a similar proportion of hospitals' costs. Muñoz et al. (2010) project that total surgical expenditures in the United States will grow from \$572 billion in 2005 to \$912 billion (in 2005 dollars) in the year 2025. Thus, the impact of OT efficiency on the financial well-being of hospitals is likely to increase.

Cardoen et al. (2010a) suggest that financial objectives, although largely overlooked, may be the most general performance measure applicable to OT scheduling. In particular, a financial objective accounts for the conflicting objectives of OT stakeholders, i.e., hospital management, OT staff, and patients. Elective operations are a primary source of revenue for a hospital. Thus, hospital management desires OT schedules with both high utilization and low overtime usage. Both patients and hospital staff prefer reliable schedules. We account for these preferences in a financial objective that includes costs for deviations in the OT schedule and overtime. Although the stated preferences are often conflicting, an OT scheduling approach that reduces total cost or increases revenue provides additional capital that can facilitate improvements (e.g., additional capacity) that benefit all stakeholders. Regardless of the objective choice, OT scheduling is difficult because of uncertainty in the duration of operations and in the arrival of patients requiring urgent attention. Although the body of literature regarding OT scheduling is large, the survey of Cardoen et al. (2010a) identifies the need for increased effort in determining scheduling strategies that incorporate uncertain operation durations and account for potential randomly arriving urgent patients. We address these major issues in this work.

The surgical scheduling literature typically classifies operations as either elective or urgent. Elective operations may be planned, whereas urgent operations occur randomly and require prompt attention. Uncertainty exists with both operation types. For elective operations, the uncertainty typically resides in the time requirements for the operation. With respect to urgent operations, there is uncertainty in both the arrival time and the requirements of the operation. Typical strategies for accommodating urgent patients are (1) reserving capacity in dedicated operating rooms for urgent operations and (2) performing urgent operations in operating rooms that become available as elective operations finish. The medical literature suggests that the second policy is most common. The survey of Cardoen et al. (2010b) supports this claim, finding that, of 52 hospitals in Belgium, 85% of the respondents reported performing urgent operations in the first room that becomes available. Only 4% of the respondents indicated they intentionally reserve operating rooms for urgent patients.

Surgical suites typically operate following either an open or block scheduling policy. Under an open scheduling policy, schedules allocate elective operations to empty time slots in any capable operating room. Block scheduling policies designate time intervals in specific operating rooms for operations belonging to a particular surgical specialty or a particular surgeon. Only 4% of respondents to the survey of Cardoen et al. (2010b) report following an open scheduling policy, whereas the other 96% of the respondents report following some form of block scheduling.

This research studies the problem of developing a single day schedule for a set of elective operations in an OT. The schedule provides a plan for an upcoming day in the OT that specifies the assignment of available elective operations to specific operating rooms along with planned start times for scheduled elective operations. Each elective operation has uncertain setup, procedure, and cleanup times. We assume that sufficient historical data exist to determine appropriate distributions for the time requirements of each elective operation and that the total time requirements of the elective operations may exceed the time available in the OT. The OT follows a block scheduling policy, and the eligibility of elective operations for each operating room block is given. To accommodate the potential arrival of urgent patients, schedules include opportunities for performing urgent operations (referred to as break-in-moments (BIMs) following van Essen et al. 2012b) within a specified duration from arrival.

In practice, averages of historical data, or medical staff predictions, are commonly used to estimate the time requirements of elective operations (Tyler et al. 2003, Cardoen et al. 2010b, Chen et al. 2010). For a particular problem instance, we use the time requirement distributions for each elective operation to generate a set of scenarios. Each scenario contains a specific realization of the setup, procedure, and cleanup time requirements for all available elective operations. A scenario-based mixed integer programming model uses these scenarios to determine an assignment of elective operations to operating room blocks, along with planned start times, that maximizes the expected profit over the set of scenarios. We refer to the problem as the scenario-based elective surgery scheduling (SBESS) problem. Similar to van Essen et al. (2012a), Gupta and Denton (2008), and Tyler et al. (2003), we use simulation to assess schedules provided by the SBESS model.

The SBESS employs a financial objective that includes the revenue associated with performing elective operations as well as costs for overtime and tardiness of elective operations. We define the tardiness of an elective operation as the difference between the planned and actual start times of an elective operation. Although the SBESS does not directly measure OT utilization, including elective operation revenue in the objective results in a natural preference for OT schedules with high utilization. Experiments show that schedules found by the SBESS model result in higher expected profit and OT utilization than schedules found by deterministic solution approaches for OT scheduling that appear in the literature.

The remainder of the paper is organized as follows. In §2, we discuss related works in the literature. In §3, we present a scenario-based mixed integer programming formulation for the SBESS problem. In §4, we describe a two-step solution approach to solve practical-sized instances efficiently. Moreover, we show that the solution to a subproblem included in the two-step solution provides an upper bound for an instance of the SBESS problem. In §5, we present experiments focusing on the performance of the two-step solution approach, the performance of the upper bounding procedure, the effects of OT size, and the effects of incorporating break-in-moment requirements. We present major insights of the work in §6. Finally, in the appendix, we show that the complexity of the SBESS problem is NP-hard.

2. Literature Review

A large proportion of the existing research regarding elective operation scheduling does not consider uncertain elective operation time requirements or the possibility of urgent arrivals. Figure 1 displays the characteristics of several recent works incorporating at least one of these two issues. The majority of the included works only consider the assignment of elective operations to time periods or to specific operating room blocks. Denton et al. (2007) and Gul et al. (2011) address the sequencing of elective operations as well as the assignment of elective operation start times. The elective operation sequences are determined using heuristics including sequencing in order of increasing mean procedure time, decreasing mean procedure time, increasing variance of procedure time, and increasing coefficient of variation of procedure time. Both works suggest that sequencing elective operations in order of increasing mean procedure time performs well. The inclusion of break-in-moment requirements in the SBESS problem adds

Figure 1 Related Works

Reference	Uncertain EO durations?	UO considerations?	EO sequencing?	EO start time assignment?	Assigns EOs to:	Objective function contributors	Solution approach
Lamiri et al. (2008)		✓			Periods	Costs to perform EOs overtime	SP
Lamiri et al. (2009)		✓			Periods	Costs to perform EOs overtime	SP
Erdem et al. (2012)		✓			Periods	EO waiting time overtime resource costs UO deferral costs	MILP heuristics
Hans et al. (2008)	✓				Rooms/Blocks	OT utilization overtime	Heuristics
Shylo et al. (2012)	✓				Rooms/Blocks	OT utilization	MINLP
Addis et al. (2014)	✓				Rooms/Blocks	EO waiting time	MILP
Denton et al. (2007)	✓		✓	✓	Rooms/Blocks	EO tardiness	SP heuristics
Gul et al. (2011)	✓		✓	✓	Rooms/Blocks	EO tardiness overtime	Heuristics
Min and Yih (2010)	✓	✓			Rooms/Blocks	Costs to perform EOs overtime	SP
Rachuba and Werners (2014)	✓	✓			Periods	EO waiting time overtime	MILP
Wang et al. (2014)	✓	✓			Rooms/Blocks	Costs to perform EOs overtime	Heuristics
van Essen et al. (2012b)	✓	✓		✓	Rooms/Blocks	Time between BIMs	Heuristics
SBESS	✓	✓	✓	✓	Rooms/Blocks	Profit	MILP

Note. EO, elective operation; UO, urgent operation; SP, stochastic programming; MILP, mixed-integer linear programming; MINLP, mixed-integer nonlinear programming.

additional complexity to sequencing decisions that diminishes the performance of the described heuristic methods. The work of van Essen et al. (2012b) assumes that the sequence of elective operations is given and seeks to determine start times so that break-in-moments occur at regular intervals throughout the day.

Figure 1 shows that a diverse set of objectives for OT scheduling is considered in the existing literature. In particular, we observe waiting time, utilization, financial, and patient deferral objectives as defined in Cardoen et al. (2010a). The financial objective of the SBESS, along with the constraints, incorporates all of the mentioned objective categories when determining a single day schedule for an OT. Specifically, we consider the waiting time of both elective and urgent patients via tardiness costs and the inclusion of break-in-moment requirements, respectively. The inclusion of revenue and overtime costs for elective operations results in a preference for schedules with high utilization of the planned OT time. The consideration of revenue also provides an incentive to schedule as many elective operations as possible, lowering the number of deferrals.

Stochastic programming is a common modeling approach for OT scheduling problems with uncertain elective operation time requirements. In particular, several works successfully apply a variant of stochastic programming known as sample average approximation (SAA) in the context of OT scheduling (see Lamiri et al. 2008, 2009; Min and Yih 2010). The SAA approach approximates the expected value of the objective function for a stochastic problem using a sample average. Specifically, a set of scenarios is generated and these are assumed to be equally likely. The SAA approach simplifies the problem of assigning probabilities to scenarios when the uncertain parameters follow continuous distributions. Although an SAA approach considers all scenarios to be equally likely, a typical SAA implementation involves generating a set of scenarios large enough that low-probability scenarios have an appropriately low frequency of occurrence and impact on the problem objective. A key difference between OT scheduling problems where SAA is successfully applied and the SBESS is the inclusion of elective operation sequencing decisions. The inclusion of sequencing decisions in the SBESS results in the problem being NP-hard, regardless of the number of scenarios included. Denton et al. (2007) and Gul et al. (2011) also consider elective operation sequencing and start time assignment but must resort to heuristics to realize practical solution times. Our scenario-based modeling approach uses a relatively low number of scenarios to provide information regarding the shape of the underlying elective operation time requirement distributions. In contrast to an SAA approach, we approximate the likelihood of each scenario using a discretization of the elective operation time requirement distributions.

Although all aspects of the SBESS problem are extant in the literature, this research is the first to present a scenario-based model for (1) assigning, (2) sequencing, and (3) determining explicit start times for a set of available elective operations with uncertain time requirements while ensuring that opportunities exist for a potential arriving urgent demand to receive prompt attention. The scenario-based model for the SBESS problem may be solved exactly for small OT environments, i.e., three operating rooms or fewer. We develop a hierarchical, two-step solution approach for larger instances where the assignment of elective operations to operating room blocks occurs in the first step and sequencing and start time assignment occur in the second step.

3. Problem Statement

Our problem environment is based on a large regional hospital located in Alabama. The OT consists of K operating rooms, and, on any given day, the hospital faces the problem of assigning and sequencing up to N elective operations in the OT. We assume that the operating theater divides the total time for each room into B blocks where the starting and ending times for all operating room blocks are given. We denote the starting time of block b in operating room k as v_{bk} and the ending time of block b in operating room k as w_{bk}. The setup, procedure, and cleanup times for elective operations are uncertain. However, we assume that sufficient historical data, or expert knowledge, exist to generate S scenarios, each containing realizations of the uncertain setup, procedure, and cleanup times for the N elective operations.

The SBESS model includes a decision variable specifying whether an elective operation is planned, and if so, the position of the operation in the sequence of an eligible operating room and a planned start time. Based on the assignments and planned start times, additional decision variables specify actual start times for planned elective operations in the S scenarios. We require the actual start time to be equal to or later than the planned start time. Also, we require that the expected completion time for elective operations performed in an operating room block be earlier than the ending time of the block.

Our formulation seeks to maximize expected profit, where profit is defined as the difference between the revenue of scheduled elective operations and the costs of overtime and start time tardiness of elective operations. We assume that the costs of opening and staffing the operating rooms comprising the OT are sunk and do not include them in the SBESS objective. This assumption is consistent with the assumptions that the block assignments, and corresponding room opening and staffing decisions, are made in advance and known. We also assume that all operating rooms are identically equipped and that the setup, procedure, and cleanup time requirements for elective operations are the same in all rooms. However, our model can handle cases where elective operation time requirements depend on the assigned operating room by introducing an index for the operating room in the time requirement parameters. Following the $\alpha \mid \beta \mid \gamma$ convention commonly used in the scheduling literature (Graham et al. 1979), the problem can be classified as $\mathrm{P} \mid \mathrm{M}_j, p_j \sim F(\mu_j, \sigma_j) \mid \sum_j(\pi_j A_j - \xi_j f(\mathbb{E}[C_j]))$ where P indicates identical parallel machines (rooms), M_j indicates eligibility restrictions for job j (elective operation j), $p_j \sim F(\mu_j, \sigma_j)$ indicates that the processing time of each job follows a random distribution defined by the mean time requirement of job j (μ_j) and the standard deviation of the time requirement of job j (σ_j), and the expression $\sum_j(\pi_j A_j - \xi_j f(\mathbb{E}[C_j]))$ represents the objective as the difference between a weighted measure that depends on the assignment of job j and a weighted function that depends on the expected completion time of job j. Note that the eligibility restrictions arise from limitations imposed by the block schedule, which specifies the operating room blocks assigned to the surgical specialty necessary to perform each elective operation.

Individuals needing urgent care may arrive during the day and require access to one of the operating rooms. A managerial constraint for such urgent arrivals is that their expected waiting time for access to an operating room must be less than or equal to Δ. Such a break-in-moment exists in an operating room k, at time τ, if the procedure or cleanup for an elective operation is not taking place in the operating room at time τ. To model the break-in-moment requirements, we divide the total OT time into a set $\mathcal{T}$ of time intervals of length $\Delta/2$ and ensure that we may expect at least one opportunity to perform an urgent operation exists in each interval. Since we desire break-in-moments to occur throughout the planned OT day, we need $\lceil \bar{w}/(\Delta/2)\rceil = \lceil 2\bar{w}/\Delta\rceil$ time intervals, where $\bar{w}$ represents the maximum end time across all operating room blocks. We index the set $\mathcal{T}$ by $t = 0, \ldots, \lfloor 2\bar{w}/\Delta\rfloor$, with $t \in \mathcal{T}$ representing the interval from time $t(\Delta/2)$ to $(t+1)(\Delta/2)$. The final time interval of the day, $\lfloor 2\bar{w}/\Delta\rfloor$, starts at time $\lfloor 2\bar{w}/\Delta\rfloor(\Delta/2)$ and ends at time $\lceil 2\bar{w}/\Delta\rceil(\Delta/2) \geq \bar{w}$. Note that we are not including urgent operations in the schedule. Instead, we are ensuring that the plan can accommodate a potential urgent patient.

We assume that all operating rooms are capable of performing urgent operations, and preemption is allowed only during the setup of an elective operation. Since we prohibit urgent operations from interrupting an ongoing elective operation during the procedure and cleanup phases, we generate individual activity times and add them together in the SBESS model. We employ binary variables to determine the elective operation schedule impact on the occurrence of break-in-moments in each of the $\lfloor 2\bar{w}/\Delta\rfloor$ time intervals in each scenario. We do not model the use of shared hospital resources, e.g., recovery rooms and central sterilization units, since the demand on such areas is not completely determined by the OT schedule. Notation used in the SBESS formulation follows.

Sets

$\mathcal{N}$—set of elective operations $\quad \mathcal{N} = \{1, \ldots, N\}$
$\mathcal{K}$—set of operating rooms $\quad \mathcal{K} = \{1, \ldots, K\}$
$\mathcal{B}$—set of operating room blocks $\quad \mathcal{B} = \{1, \ldots, B\}$
$\mathcal{P}$—set of operating room block positions $\quad \mathcal{P} = \{1, \ldots, P\}$
$\mathcal{S}$—set of scenarios $\quad \mathcal{S} = \{1, \ldots, S\}$
$\mathcal{T}$—set of time intervals $\quad \mathcal{T} = \{0, \ldots, \lfloor 2\bar{w}/\Delta\rfloor\}$

Indices

m, n—elective operation indices $\quad m, n \in \mathcal{N}$
k—operating room index $\quad k \in \mathcal{K}$
b—block index $\quad b \in \mathcal{B}$
p—position index $\quad p \in \mathcal{P}$
s—scenario index $\quad s \in \mathcal{S}$
t—time interval index $\quad t \in \mathcal{T}$

Parameters

r_{ns}—setup time for elective operation n in scenario s
q_{ns}—procedure time plus cleanup time of elective operation n in scenario s
Δ—desired maximum time between the arrival and start of an urgent operation
π_n—revenue associated with performing elective operation n
α_k—cost of one time unit of overtime in operating room k
β_n—cost of delaying elective operation n by one time unit
v_{bk}—start time of block b in operating room k
w_{bk}—end time of block b in operating room k
$\bar{w}$—maximum end time across all operating room blocks, i.e., $\bar{w} = \max_{b\in\mathcal{B},\, k\in\mathcal{K}}(w_{bk})$

$$e_{nbk}\begin{cases}1 & \text{if operating room } k \text{ is configured to accommodate operation } n \text{ during block } b,\\ 0 & \text{otherwise.}\end{cases}$$

M—number strictly larger than $\bar{w}$

Decision Variables

$$A_{npbk} = \begin{cases}1 & \text{if elective operation } n \text{ is sequenced for position } p \text{ in block } b \text{ of operating room } k,\\ 0 & \text{otherwise.}\end{cases}$$

Y_n—planned start time of the setup for elective operation n
X_{ns}—actual start time of the setup for elective operation n in scenario s
O_k—expected overtime requirements for operating room k
Z_n—expected difference between the planned start time of operation n and the actual start time of the operation n (i.e., start time tardiness)

$$\delta_{nt} = \begin{cases}1 & \text{if the expected start time for operation } n\text{'s procedure is prior to time interval } t,\\ 0 & \text{otherwise;}\end{cases}$$

$$\gamma_{nt} = \begin{cases}1; & \text{if the expected completion time for operations } n\text{'s cleanup is after time interval } t \text{ elapses},\\ 0 & \text{otherwise;}\end{cases}$$

$$\Theta_{nt} = \begin{cases}1 & \text{if the expected duration of operation } n\text{'s procedure and/or cleanup occupies a room for the entire duration of time interval } t,\\ 0 & \text{otherwise.}\end{cases}$$

The SBESS formulation follows where $\mathbb{E}$ represents the expected value over the set of scenarios $\mathcal{S}$. Since the underlying time requirement distributions are likely to be continuous, considering a finite set of scenarios is necessary to achieve a tractable model:

$$\text{maximize} \sum_{n\in\mathcal{N}}\sum_{p\in\mathcal{P}}\sum_{b\in\mathcal{B}}\sum_{k\in\mathcal{K}}\pi_n A_{npbk} - \sum_{k\in\mathcal{K}}\alpha_k O_k - \sum_{n\in\mathcal{N}}\beta_n Z_n. \tag{1}$$

Assignment constraints:

$$\sum_{p\in\mathcal{P}}\sum_{b\in\mathcal{B}}\sum_{k\in\mathcal{K}} A_{npbk} \le 1, \quad \forall n\in\mathcal{N}; \tag{2}$$

$$\sum_{n\in\mathcal{N}} A_{npbk} \le 1, \quad \forall p\in\mathcal{P}, b\in\mathcal{B}, k\in\mathcal{K}; \tag{3}$$

$$\sum_{n\in\mathcal{N}} A_{npbk} \ge \sum_{n\in\mathcal{N}} A_{np+1bk}, \quad \forall p\in\mathcal{P}\backslash\{P\}, b\in\mathcal{B}, k\in\mathcal{K}; \tag{4}$$

$$A_{npbk} \le e_{nbk}, \quad \forall n\in\mathcal{N}, p\in\mathcal{P}, b\in\mathcal{B}, k\in\mathcal{K}; \tag{5}$$

$$\sum_{n\in\mathcal{N}}\sum_{p\in\mathcal{P}} A_{npbk}\mathbb{E}(r_{ns}+q_{ns}) \le w_{bk}-v_{bk}, \quad \forall b\in\mathcal{B}, k\in\mathcal{K}. \tag{6}$$

Start time constraints:

$$Y_n \ge \sum_{p\in\mathcal{P}} v_{bk}A_{npbk}, \quad \forall n\in\mathcal{N}, b\in\mathcal{B}, k\in\mathcal{K}; \tag{7}$$

$$\mathbb{E}(r_{ns}+q_{ns}) - M\left(1-\sum_{p\in\mathcal{P}} A_{npbk}\right) \le w_{bk}-Y_n, \quad \forall n\in\mathcal{N}, b\in\mathcal{B}, k\in\mathcal{K}; \tag{8}$$

$$Y_n \le X_{ns}, \quad \forall n\in\mathcal{N}, s\in\mathcal{S}; \tag{9}$$

$$X_{ns}+M(2-A_{mpbk}-A_{n(p+1)bk}) \ge X_{ms}+r_{ms}+q_{ms}, \quad \forall m,n\in\mathcal{N}, p\in\mathcal{P}\backslash\{P\}, b\in\mathcal{B}, k\in\mathcal{K}, s\in\mathcal{S}; \tag{10}$$

$$X_{ns}+M(2-A_{mp(b-1)k}-A_{n1bk}) \ge X_{ms}+r_{ms}+q_{ms}, \quad \forall m,n\in\mathcal{N}, p\in\mathcal{P}, b\in\mathcal{B}\backslash\{1\}, k\in\mathcal{K}, s\in\mathcal{S}. \tag{11}$$

Expected time to break-in-moment constraints:

$$\mathbb{E}(X_{ns}+r_{ns}) - M(1-\delta_{nt}) \le t\frac{\Delta}{2}, \quad \forall n\in\mathcal{N}, t\in\mathcal{T}; \tag{12}$$

$$\mathbb{E}(X_{ns}+r_{ns}) + M\delta_{nt} \ge t\frac{\Delta}{2}, \quad \forall n\in\mathcal{N}, t\in\mathcal{T}; \tag{13}$$

$$\mathbb{E}(X_{ns}+r_{ns}+q_{ns}) + M(1-\gamma_{nt}) \ge (t+1)\frac{\Delta}{2}, \quad \forall n\in\mathcal{N}, t\in\mathcal{T}; \tag{14}$$

$$\mathbb{E}(X_{ns}+r_{ns}+q_{ns}) - M\gamma_{nt} \le (t+1)\frac{\Delta}{2}, \quad \forall n\in\mathcal{N}, t\in\mathcal{T}; \tag{15}$$

$$\Theta_{nt} \le \delta_{nt}, \quad \forall n\in\mathcal{N}, t\in\mathcal{T}; \tag{16}$$

$$\Theta_{nt} \le \gamma_{nt}, \quad \forall n\in\mathcal{N}, t\in\mathcal{T}; \tag{17}$$

$$\Theta_{nt} \ge \delta_{nt}+\gamma_{nt}-1, \quad \forall n\in\mathcal{N}, t\in\mathcal{T}; \tag{18}$$

$$\sum_{n\in\mathcal{N}} \Theta_{nt} \le K-1, \quad \forall t\in\mathcal{T}. \tag{19}$$

Objective function variable definitions:

$$Z_n \ge \mathbb{E}(X_{ns}-Y_n), \quad \forall n\in\mathcal{N}; \tag{20}$$

$$O_k \ge \mathbb{E}(X_{ns}+r_{ns}+q_{ns}-w_{Bk}-(M+r_{ns}+q_{ns})(1-A_{npBk})), \quad \forall n\in\mathcal{N}, p\in\mathcal{P}, k\in\mathcal{K}. \tag{21}$$

Variable domains:

$$X_{ns} \ge 0, \quad \forall n\in\mathcal{N}, s\in\mathcal{S}; \tag{22}$$

$$O_k \ge 0, \quad \forall k\in\mathcal{K}; \tag{23}$$

$$A_{npbk} \in \{0,1\}, \quad \forall n\in\mathcal{N}, p\in\mathcal{P}, b\in\mathcal{B}, k\in\mathcal{K}; \tag{24}$$

$$\delta_{nt}, \gamma_{nt}, \Theta_{nt} \in \{0,1\}, \quad \forall n\in\mathcal{N}, t\in\mathcal{T}. \tag{25}$$

The objective function of our model, given by Equation (1), maximizes the difference between the revenue generated from scheduled elective operations and the expected costs of overtime and start time tardiness. Requirements for the A_{npbk} variables are defined by constraint sets (2)–(6). Constraint set (2) ensures that each elective operation is assigned to at most one position in the sequence of one operating room block. Constraint set (3) ensures that at most one elective operation is assigned to any position in an operating room block. Constraint set (4) ensures that elective operations are not assigned to a position of an operating room block sequence if no assignment is made for the previous position. Constraint set (5) prohibits assigning elective operations to ineligible operating room blocks. Ineligibility may be a result of the operating room lacking resources or equipment necessary to perform an elective operation, or the operation requiring a surgical specialty that is not assigned to the operating block. Constraint set (6) ensures that the expected time requirements for elective operations assigned to a particular operating room block do not exceed the time allotted for the block.

Constraint sets (7)–(11) define the elective operation start time decision variables. Constraint set (7) requires the planned start time for elective operations to be later than the start time of their assigned operating room block. Constraint set (8) requires the completion of expected total time requirements for elective operations before the end of their assigned operating room blocks. Constraint set (9) requires the actual start time for elective operations to be greater than their planned start time, in all scenarios. Constraint set (10) ensures that the actual start time of an elective operation in position $p+1$ of an operating room block sequence is greater than the completion time of the preceding operation, in each scenario. Constraint set (11) ensures that the actual start time of an elective operation in the first position of an operating room block sequence is greater than the completion time of all elective operations in the preceding block of the same operating room, in each scenario.

Constraint sets (12)–(19) capture the expected timing of the procedure and cleanup for elective operations and ensures that the expected time between break-in-moments is at most Δ. Constraint sets (12) and (13) define the δ_{nt} variables, and constraint sets (14) and (15) define the γ_{nt} variables. Together, constraint sets (16)–(18) ensure that Θ_{nt} takes a value of 1 if and only if $\delta_{nt} = \gamma_{nt} = 1$, indicating that the expected time for operation n's procedure takes place throughout the time interval t. Constraint set (19) requires that at least one break-in-moment is expected in all time intervals.

Constraint sets (20) and (21) define continuous variables used in the SBESS objective. Specifically, constraint set (20) requires the variable Z_n to be greater than or equal to the expected difference between the actual and planned start times of elective operation n. Together, constraint sets (7), (9), and (20) ensure that Z_n is nonnegative for all $n \in \mathcal{N}$. Constraint set (21) requires the variable O_k to be greater than or equal to the expected difference between the completion of all elective operations performed in the last operating room block and the end time of the final operating room block. Constraint sets (22) and (23) define lower bounds for the X_{ns} and O_k variables. Constraint sets (24) and (25) define the binary variables.

The SBESS formulation is also valid for an OT following an open scheduling policy. In such a case, a single block must be specified for each of the available operating rooms, and constraint sets (5) and (11) must be excluded. If the number of blocks varies among the set of operating rooms, the start and end times must be adjusted to ensure that no elective operations are assigned to invalid blocks. For example, consider an OT with two operating rooms where operating room one uses a single block and operating room two uses two blocks. In this case, B equals two and the start and end times of block one in operating room one can be set to zero to ensure that no elective operations are assigned to the first block. Note that, in the room using a single block, the full block time is assigned to the final block instead of the first. This is necessary to ensure the correct behavior of constraint set (21), which defines overtime based on the time requirements of elective operations assigned to block B.

4. Solution Approach

The complexity of the SBESS formulation leads to high computational time requirements. Thus, we develop a two-step solution approach to reduce the computational time required to determine an SBESS solution. The first step solves a multiple knapsack subproblem that determines the assignments of elective operations to a block of an eligible operating room. The multiple knapsack solution is used to generate a simplified SBESS problem that is solved in the second step to provide sequences of operations for each room in the OT. Details of the multiple knapsack subproblem are discussed in §4.1. Section 4.2 contains a discussion of the two-step solution approach. We discuss implementation of the described solution approach in §4.3.

4.1. Multiple Knapsack Subproblem

Ignoring break-in-moment restrictions and sequencing of elective operations, the overall problem reduces to assigning elective operations to operating room blocks to maximize the benefit of scheduled operations. This assignment problem is a multiple knapsack problem (see Lin (1998)), with operating blocks representing knapsacks that have a capacity equal to the corresponding block time. Elective operations represent items to be placed in one of the available knapsacks with a "size" equal to the sum of the expected setup plus procedure times. Addis et al. (2014), Lamiri et al. (2008, 2009), and Hans et al. (2008) address similar problems. The formulation for the multiple knapsack subproblem, referred to as MKPS, follows.

MKPS Formulation
Decision Variables

$$A'_{nbk} = \begin{cases} 1 & \text{if operation } n \text{ is assigned to block} \\ & b \text{ of operating room } k, \\ 0 & \text{otherwise.} \end{cases}$$

$$\text{maximize} \quad \sum_{n\in\mathcal{N}} \sum_{b\in\mathcal{B}} \sum_{k\in\mathcal{K}} A'_{nbk}\pi_n \tag{26}$$

$$\text{subject to} \quad \sum_{b\in\mathcal{B}} \sum_{k\in\mathcal{K}} A'_{nbk} \le 1, \quad \forall n\in\mathcal{N}; \tag{27}$$

$$\sum_{n\in\mathcal{N}} A'_{nbk}\mathbb{E}(r_{ns}+q_{ns}) \le w_{bk} - v_{bk}, \quad \forall b\in\mathcal{B}, k\in\mathcal{K}; \tag{28}$$

$$A'_{nbk} \le e_{nbk}, \quad \forall n\in\mathcal{N}, b\in\mathcal{B}, k\in\mathcal{K}; \tag{29}$$

$$A'_{nbk} \in \{0,1\}, \quad \forall n\in\mathcal{N}, b\in\mathcal{B}, k\in\mathcal{K}. \tag{30}$$

The objective function, given by Equation (26), seeks to maximize the total revenue of assigned elective operations. Constraint set (27) ensures that each elective operation is assigned to at most one block of one operating room. Constraint set (28) ensures that the expected setup plus procedure time requirements of all elective operations assigned to a particular operating room block do not exceed the amount of time allotted to the block. As in the SBESS formulation, the expectation is over the set $\mathcal{S}$. Constraint set (29) requires that operations only be assigned to eligible blocks. Constraint set (30) defines the domain of the decision variables.

Theorem 1. *The objective function value of the MKPS for a particular problem instance provides an upper bound on the corresponding SBESS objective value.*

Proof of Theorem 1. The MKPS problem is a simplification of the SBESS problem that does not consider the sequencing of elective operations and break-in-moment requirements. Therefore, all solutions to the SBESS formulation are feasible in the corresponding MKPS problem. Moreover, the exclusion of sequencing decisions in the MKPS problem prohibits the capture of any overtime or start time tardiness costs. Thus, the objective value of the MKPS associated with a particular problem instance is always greater than or equal to the corresponding SBESS objective function value.

4.2. Two-Step Solution Approach

Solving the MKPS provides a vector of elective operation to operating room block assignments. We use the elective operations assigned in the MKPS to create a reduced set of elective operations $\mathcal{N}'=\{n\in\mathcal{N}\colon A'_{nbk}=1\}$ and restrict the eligibility of elective operations in $\mathcal{N}'$ to their respective MKPS assignments. Additionally, we use the MKPS solution to determine the maximum size of the position set, P, and a latest position for each operating room block, P_{bk}, where there are no assignments for positions greater than P_{bk} in block b of operating room k. These modifications significantly reduce the number of binary decision variables and constraints, thus allowing large instances to be solved. Algorithm 1 shows the steps of the two-step approach. Algorithm 1 solves the MKPS in line 3 and uses the elective operation to operating room block assignments to update the set $\mathcal{N}'$, specify the corresponding values of e'_{nbk}, increment P_{bk}, and set P_{max} in lines 4–15 (i.e., step one). A reduced version of the SBESS is solved in line 16 (i.e., step two), using the modifications stated in lines 17–20.

Algorithm 1 (Two-Step Solution Approach for SBESS)

1: Input: SBESS data
2: Initialize: $\mathcal{N}'=\varnothing$, $P_{\max}=0$, $P_{bk}=0, \forall b\in\mathcal{B}, k\in K$, and $e'_{nbk}=0, \forall n\in\mathcal{N}, b\in\mathcal{B}, k\in\mathcal{K}$
3: Solve MKPS
4: **for** $n\in N$ **do**
5:   **for** $b\in\mathcal{B}$ **do**
6:     **for** $k\in K$ **do**
7:       **if** $A'_{nbk}=1$ **then**
8:         $\mathcal{N}'=\mathcal{N}'\cup n$, $e'_{nbk}=1$, & $P_{bk}=P_{bk}+1$
9:         **if** $P_{bk}>P_{\max}$ **then**
10:           $P_{\max}=P_{bk}$
11:         **end if**
12:       **end if**
13:     **end for**
14:   **end for**
15: **end for**
16: Solve SBESS with the following modifications:
17: – Replace the set $\mathcal{N}$ with $\mathcal{N}'$, and modify constraint domains accordingly
18: – Construct set $\mathcal{P}'=\{1,\ldots,P_{\max}\}$
19: – Modify constraint set (5) to $A_{npbk} \le e'_{nbk}$, $\forall n\in\mathcal{N}', p\in\mathcal{P}', b\in\mathcal{B}, k\in\mathcal{K}$
20: – Add the constraint set $A_{npbk}=0, \forall n\in\mathcal{N}', b\in\mathcal{B}, k\in\mathcal{K}, p=P_{bk}+1,\ldots,P_{\max}$

4.3. Implementation

With respect to implementation, all computational components of this research are coded using the commercial software Visual Studio 2012, CPLEX 12.6, and the open source Boost 1.56 libraries. The majority of hospitals capture data regarding the setup, procedure, and cleanup times for the OT (Surgical Directions 2014). Assuming that these data are maintained by a hospital, time requirement distributions may be fit. Even when historical data exist, determining accurate distributions may be difficult because of the potentially large number of variations in elective procedures. In such cases, the best case, average, and worst case estimates or methods for estimating location parameters of a lognormal distribution (see Spangler et al. 2004) may be used to fit approximate distributions. Since we combine the procedure and cleanup times in the SBESS formulation, it may be advantageous to combine the data for these two elements before fitting a distribution to capture any correlation. Once reasonable distributions are determined, a means to import the block schedule details and eligibility of elective operations must be established. The computational time requirements of the developed two-step approach, discussed in §5, allow decision makers to generate schedules for the next OT day once the current day finishes. Such use allows for unplanned occurrences during the current day (e.g., a canceled case) to be incorporated into the schedule for the following day.

The SBESS and MKPS objective functions favor scheduling elective operations that provide higher amounts of revenue. Thus, it is possible that elective operations with lower revenue potential frequently may be excluded from the resulting schedules. This issue is overcome by replacing π_n with a surrogate measure based on a combination of revenue and time on the waiting list.

5. Experimentation

We conduct methodological experiments to compare the performance of the two-step solution approach to the SBESS formulation. We also use simulation to perform comparative experiments among schedules found using scenario-based models to schedules found by deterministic scheduling methods. In addition to expected profit, we present OT utilization results for the comparative experiments to illustrate the impact of the scheduling methods on this important performance measure.

Section 5.1 describes the experimental design used in both sets of experiments. Results of the methodological and comparative experiments are presented in §§5.2 and 5.3, respectively. Section 5.4 presents the results of two additional sets of experiments. The first experiment considers the impact of OT size with respect to solution quality, computational time, and break-in-moment requirements. We observe that increasing the size of the OT causes a natural reduction in both the maximum and average time between break-in-moments. However, the second set of additional experiments shows that the inclusion of break-in-moment requirements reduces the waiting time for an urgent patient regardless of the OT size.

5.1. Experimental Setup

The hospital upon which our problem is based considers historical surgery time requirements to be confidential. Hence, we develop test problems from studies that exist in the literature. Specifically, we use procedure information from Gul et al. (2011), Min and Yih (2010), and Ferreira et al. (2008) to construct three test cases for experimentation. We exclude data for neurosurgery and cardiac procedures in Ferreira et al. (2008) and Min and Yih (2010) since these procedures typically require specialized rooms where urgent operations are not normally performed. In test cases with excluded operations, we adjust the relative frequency of the remaining procedure types accordingly. Studies published in the medical and academic literatures suggest that elective procedure time requirements are best approximated by a log-normal distribution (Hancock et al. 1988, Strum et al. 2000, Gul et al. 2011). We follow this established assumption.

Each test case differs significantly with respect to the expected time requirements for elective procedures and the amount of variability in elective procedure time requirements. The elective procedures defining test case 1 are relatively short with elective operations requiring more than one hour occurring with a frequency of less than 12%. On the other hand, the expected time requirements of elective procedures in test cases 2 and 3 are both longer than one hour. To assess the variability of the time requirements for each elective operation, we calculate the coefficient of variation (CV), which is the ratio of the standard deviation to the mean. The elective procedures composing test cases 1 and 2 all possess CV values less than 1, indicating low to moderate variability, whereas the procedures of test case 3 possess higher CV values, on average. Moreover, two of the six procedures in test case 3 have a CV value exceeding 1, indicating a high level of variability. We use factorial experiments

Table 1 Cost Parameters

	Case 1	Case 2	Case 3
Regular time			
Overhead cost	\$15/min	\$12/min	\$10/min
Staffing cost	\$15/min	\$18/min	\$20/min
	30/min	\$30/min	\$30/min
Overtime			
Overhead cost	\$15/min	\$12/min	\$10/min
Staffing cost	\$22.50/min	\$27/min	\$30/min
Total	\$37.50/min	\$39/min	\$40/min

to separate any influences due to the test cases in our analysis of the results.

Setup and cleanup times for all elective operations are generated from the uniform distribution with bounds of 12 and 20 minutes. We choose the setup and cleanup time distribution such that the minimum total time between the completion of one elective operation and the start of the following operation falls within the range of 25–40 minutes. Macario (2006) suggests tht this range is suitable for a hospital with average efficiency.

We base the costs used in our experiments on estimates from Macario (2010). In all experimental cases, we establish operating room time cost as \$30 per minute. Similar to Batun et al. (2011), we use the per minute operating room cost to approximate the per minute cost of start time tardiness. To determine overtime costs, we divide the operating room time cost into fixed overhead and staffing cost components. Macario (2010) suggests a reasonable fixed overhead cost is \$10 per minute of operating room time, with per minute staffing costs greater than or equal to overhead costs. We use the suggestion of Macario (2010) to create three cases for the SBESS cost parameters. Specifically, we vary the per minute staffing cost from \$15 to \$20 with fixed overhead costs making up the remainder of the operating room cost. We assume that the overtime staffing cost is 1.5 times normal staffing cost and calculate the per minute total overtime cost as the sum of the fixed overhead and overtime staffing costs. Table 1 details the three cost settings.

We consider three cases for specifying the revenue of elective operations. Specifically, the work of Macario (2010) finds that the charge for a two-hour elective procedure ranges from \$3,520 to \$9,647. We use the cheapest charge to specify a "low" setting of \$30 per minute and the most expensive charge to specify a "high" setting of \$80 per minute. A "medium" setting uses the average of the low and high charges, \$55 per minute. The revenue for a particular elective operation is set to the product of the applicable charge and the expected procedure time.

We include three deterministic models in our experiments. The first model, EV, mimics the common practice of using medical staff predictions or averages of historical data to estimate elective operation time requirements and uses the expected value of the corresponding time requirement distribution (Tyler et al. 2003, Cardoen et al. 2010b, Chen et al. 2010). The remaining two deterministic models, JH65 and JH70, mimic the "job hedging" method used in Gul et al. (2011). The job hedging method estimates the actual time requirements of elective operations using a predetermined percentile of the corresponding time requirement distribution. Gul et al. (2011) investigates the performance of job-hedging models using the 50th, 55th, 60th, 65th, 70th, 75th, 80th, and 85th percentiles for minimizing the sum of waiting and overtime costs in an OT environment. In their experiments, models using the 65th and 70th percentiles yielded the best solution in 77% of the instances considered.

We also consider three scenario-based models that differ by the number of scenarios. Specifically, we consider scenario-based models with 5, 10, and 15 scenarios. The first scenario of each scenario-based model uses the expected value for all time requirements to capture the centering of the associated distributions. The remaining scenarios are constructed by randomly sampling from the elective operation time requirement distributions. Thus, the remaining scenarios provide information regarding dispersion of the time requirements distributions. We use common random numbers, a variance reduction technique (Ross 2006), when generating scenarios. So, the first five scenarios in the 10-scenario model are identical to those comprising the 5-scenario model. Similarly, the first 10 scenarios in the 15-scenario model are identical to those comprising the 10-scenario model.

Factorial designs are used in both the methodological and comparative experiments. Table 2 provides descriptions for the different levels of the design factors. All operating rooms have a planned time of 540 minutes (i.e., 9 hours) that is divided into two blocks of 270 minutes. The value of M is set to 600 minutes throughout our experiments.

All instances are solved using the C++ API of CPLEX 12.6 on a Dell PC with an Intel Xeon 2.53 GHz processor and 64 GB of RAM. We determine probabilities for the scenario-based models using the random distribution capabilities of Boost version 1.56, a collection of C++ libraries for enhanced mathematical functionality. The Boost random library uses a discretization of (+/−0.5) to determine the probability that a continuous variable takes a specific value. Because of the small probability associated with scenarios, it may be necessary to scale the probability values to avoid numerical issues. Klotz and Newman (2013) describe methods to overcome such numerical issues when solving linear programs. The number of elective operations available to schedule (N) is specified by dividing the total OT time by the expected procedure time for the current test case and rounding up. A terminating condition of at most 30 minutes of CPU time is enforced in all experiments.

Table 2 **Factors and Levels for Designed Experiments**

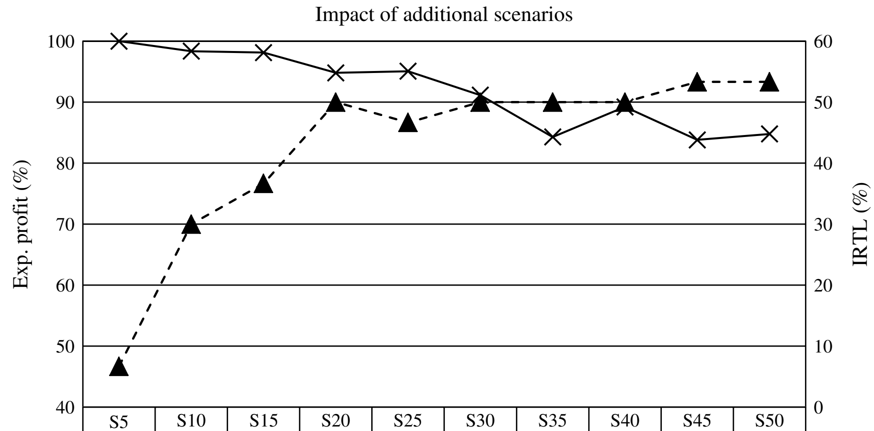
Factor	Levels	Interpretation
BIM	N	No BIM requirements
	Y	BIM requirements incorporated ($\Delta = 60$)
Costs	Low	$\beta_n = \$30.00, \forall n \in \mathcal{N}, \alpha_k = \$37.50, \forall k \in \mathcal{K}$
	Medium	$\beta_n = \$30.00, \forall n \in \mathcal{N}, \alpha_k = \$39.00, \forall k \in \mathcal{K}$
	High	$\beta_n = \$30.00, \forall n \in \mathcal{N}, \alpha_k = \$40.00, \forall k \in \mathcal{K}$
Model	EV	Deterministic model, using the mean for all time requirements
	JH65	Deterministic model, using the 65th percentile for all time requirements
	JH70	Deterministic model, using the 70th percentile for all time requirements
	S5	Scenario-based model, using 5 scenarios
	S10	Scenario-based model, adding 5 scenarios to those of S5
	S15	Scenario-based model, adding 5 scenarios to those of S10
Revenue	Low	$\pi_n = \$30 \times \mathbb{E}[r_{ns} + q_{ns}], \forall n \in \mathcal{N}$
	Medium	$\pi_n = \$55 \times \mathbb{E}[r_{ns} + q_{ns}], \forall n \in \mathcal{N}$
	High	$\pi_n = \$80 \times \mathbb{E}[r_{ns} + q_{ns}], \forall n \in \mathcal{N}$
Rooms	3	Operating theater consisting of three operating rooms
	6	Operating theater consisting of six operating rooms
Test case	1	Elective operation time requirements and frequency mimic Gul et al. (2011)
	2	Elective operation time requirements and frequency mimic Min and Yih (2010)
	3	Elective operation time requirements and frequency mimic Ferreira et al. (2008)

The choice to limit the number of scenarios to 15 is based on the results of a preliminary experiment that shows degrading expected profit as the number of scenarios increases. The factorial experiment includes all three test cases and considers OTs with three and six operating rooms. Revenue and cost parameters are set at the medium values previously described. Figure 2 summarizes the results of the experiment showing the relative expected profit (Exp. Profit) and the number of instances failing to solve to optimality within a 30-minute time limit (instances reaching time limit (IRTL)) as the number of scenarios increases from 5 to 50, in increments of 5.

5.2. Methodological Experiments

Theorem 1 showed that the objective function value of an optimal MKPS solution provides an upper bound on the optimal SBESS objective function value for the same instance. We use the described relationship to compare the performance of the SBESS formulation to the two-step approach (TS). We base our comparison on the results of the factorial design described in Table 3.

Figure 2 **Impact of Number of Scenarios**



	S5	S10	S15	S20	S25	S30	S35	S40	S45	S50
Exp. profit (%)	100	98	98	95	95	91	84	89	84	85
IRTL (%)	7	30	37	50	47	50	50	50	53	53

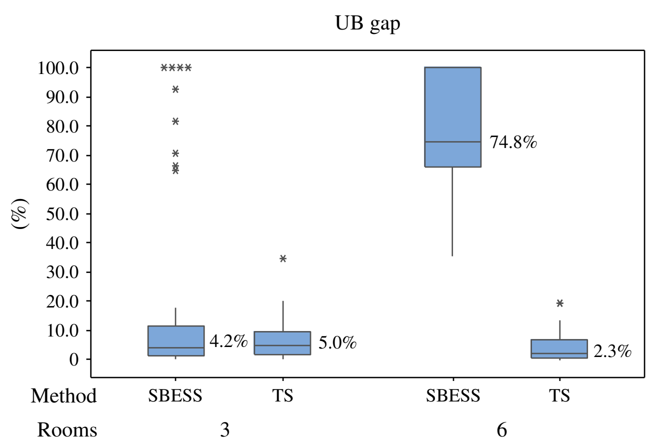
Notes. All expected profit values, provided on the left vertical axis, are expressed as a percent of the expected profit realized when only five scenarios are included. The IRTL values, provided on the right vertical axis, capture the number of instances where CPLEX is unable to prove optimality within a 30-minute time limit. The figure shows that expected profit generally decreases as additional scenarios are included because of an increase in computational complexity.

Figure 3 displays box plots and median values for the gap between (i) the MKPS upper bound and the SBESS objective function value, indicated by "SBESS," and (ii) the MKPS upper bound and the two-step solution approach, indicated by "TS," across all three-room OT instances and all six-room OT instances. Specifically, we calculate the gap measure as $((UB-\hat{Z})/UB)\times 100$ where UB is the optimal MKPS objective value and $\hat{Z}$ is either the SBESS objective value of the best solution found from solving the SBESS formulation or two-step solution approach within the 30-minute time limit. The results highlight the sensitivity of the SBESS formulation to the size of problem instances. In particular, when the OT consists of only three rooms, the majority of solutions found using only the SBESS formulation yield objective function values within 10% of the MKPS upper bound. Moreover, the median gap value is 0.8% lower for solutions found from solving the SBESS formulation in lieu of the two-step approach when the OT has three rooms. However, the performance of the SBESS formulation significantly deteriorates when the OT size is doubled to six rooms. In this case, the gap associated with all solutions exceeds 30% with a median of 74.8%. On the other hand, the performance of the two-step solution approach improves to provide a median gap of 2.3%. The improvement in the performance of the two-step solution approach as the OT size increases is due to easier satisfaction of the BIM requirements. The MKPS subproblem does not account for BIM requirements. Thus, the assignment of elective operations given by the MKPS may be infeasible for the SBESS problem. However, as the number of rooms in the OT increases, there is a natural increase in BIMs (as demonstrated in §5.4). Thus, as the OT size increases, it is less likely that an MKPS solution violates the BIM requirements and solutions found by the two-step solution approach more closely match the MKPS solution.

Table 3 **Factorial Design**

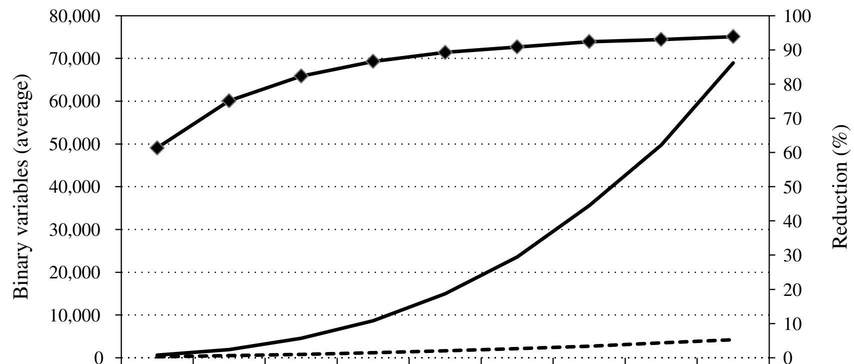
Factor	Levels
BIM	Y
Costs	Medium
Model	S5, S10, S15
Revenue	Medium
Rooms	3, 6
Test case	1, 2, 3

Figure 3 **(Color online) SBESS vs. Two-Step Solution Approach**



Note. The box plots measure the gap between the upper bound (UB) specified by the optimal MKPS objective function value and solutions found via the SBESS formulation (SBESS) and the two-step solution approach (TS).

Figure 4 **Binary Variable Reduction via the Two-Step Approach**



Rooms	2	3	4	5	6	7	8	9	10
Before MKPS reduction	618	1,906	4,551	8,685	14,954	23,564	35,552	49,773	68,946
After MKPS reduction	239	474	804	1,161	1,606	2,155	2,707	3,462	4,233
Reduction (%)	61	75	82	87	89	91	92	93	94

Rooms

The consistent performance of the two-step approach as the problem size increases is due to a significant reduction in binary variables. Figure 4 plots the average binary variables required to solve an SBESS instance before and after the reduction performed via Algorithm 1, along with the percentage of the reduction, as the size of the OT increases from 2 to 10 operating rooms. The significant reduction of binary variables reduces the computational difficulty of solving an SBESS instance and improves the quality of solutions that may be determined within fixed time limits.

5.3. Comparative Experiments

To assess the performance of solutions found via the two-step solution approach in an OT environment, we conduct a factorial experiment to compare the performance of the deterministic and scenario-based models. This experiment uses all factors and levels described in Table 2. Assessing the performance of schedules found by the various models in an OT requires several steps. Figure 5 summarizes the entire process from generating problem instances through the analysis of the simulation results.

First, we generate S scenarios consisting of independent realizations of the setup, procedure, and cleanup time requirements for all N elective operations. Each realization is sampled from the respective time requirement distribution associated with a particular elective operation. A set of scenarios along with the additional SBESS parameters constitutes a single problem instance. We generate five instances for each combination of the design factors given in Table 2, resulting in 3,240 instances. Next, we solve each instance using the two-step solution approach described in Algorithm 1. The solution returned by the two-step approach provides the assignment of elective operations to operating room blocks and the planned start time of each elective operation. We use a simulation model to capture (i) profit, (ii) OT utilization, (iii) maximum time between consecutive BIMs (Max BIM), and (iv) average time between consecutive BIMs (Avg BIM). The simulation model mimics the OT environment using setup, procedure, and cleanup times that follow the respective random distributions for each scheduled elective operation. We perform 1,000 replications of the simulation for each instance and store the average values for all performance measures. To capture the time between break-in-moments, we determine the minutes elapsing between all integer time points τ that are less than the total planned OT time, $\bar{w}$, and the earliest break-in-moment later than τ. An illustrative example is given in Figure 6.

Figure 7 provides 95% confidence intervals for the profit associated with each level of the model factor. The results are presented relative to the average profit of the EV model and suggest 95% confidence that the profit associated with schedules found via the scenario-based models is greater than or equal to the profit associated with schedules found via the deterministic models. Scenario-based model S5 achieves the highest profit values. Moreover, there is no overlap among the confidence intervals for model S5 and any other model. The performance of model S5 in comparison to the deterministic methods highlights the value of capturing the variability of time requirements via scenarios. We observe decreasing profit as the number of scenarios increases. Figure 9, which

Figure 5 **Procedure for Comparative Experiments**

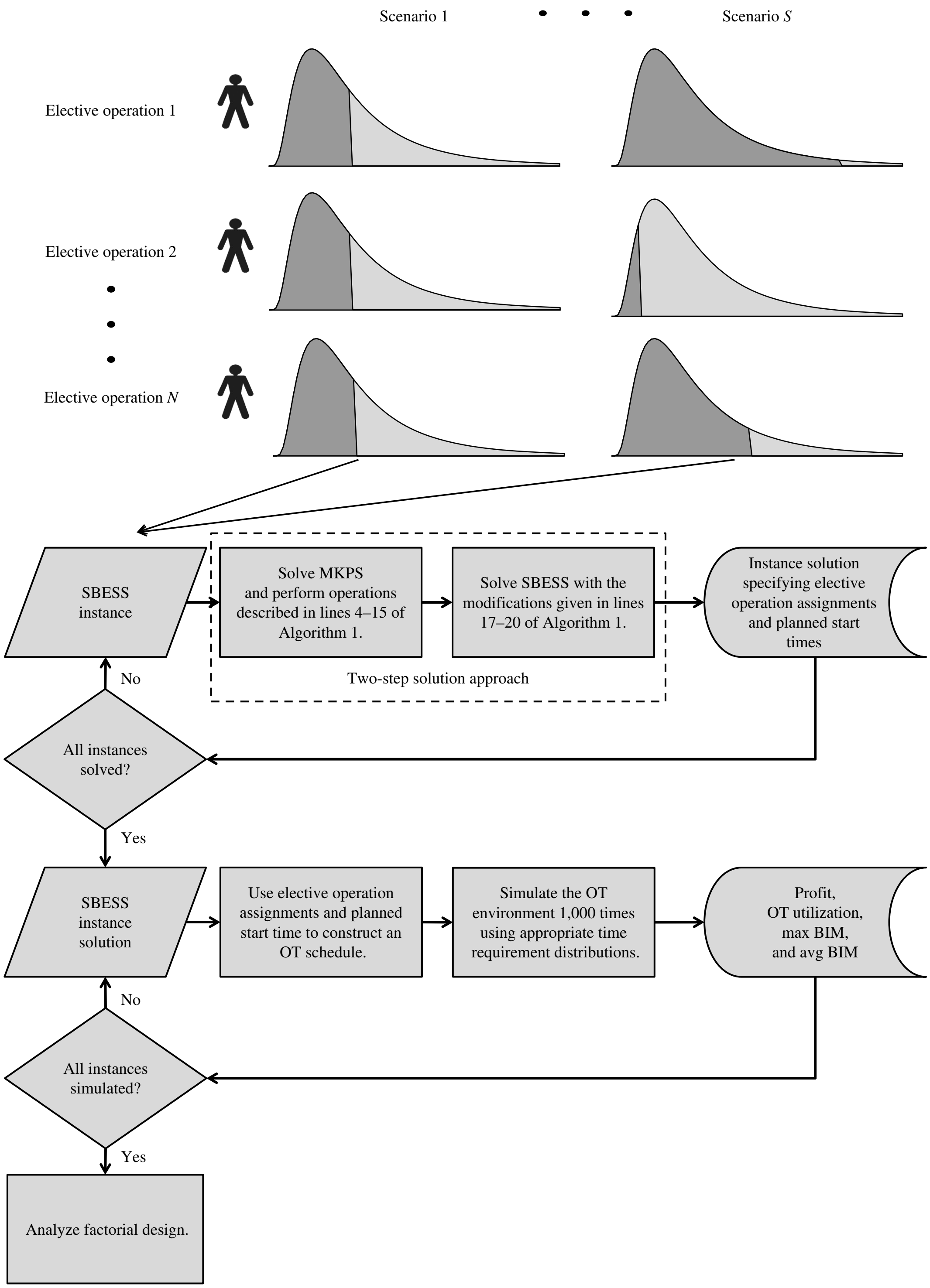
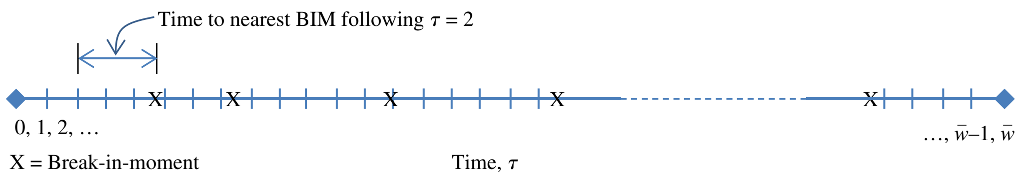


Figure 6 **(Color online) Capturing Time to BIM**



Note. This figure shows that the time to the next BIM from time τ is determined by the difference between the time the next BIM occurs and τ.

is discussed later, suggests that the described phenomenon is partly due to an increase in the computational complexity of the problem as the number of scenarios increases.

Figure 8 provides 95% confidence intervals for the OT utilization associated with each level of the model factor. As in the profit results, we observe that schedules found by model S5 attain higher OT utilization than all other model levels. Thus, from Figures 7 and 8, we may infer that incorporating the variability of time requirements via scenarios results in schedules that make better use of available OT time and increase profit. These benefits seem to lessen as the number of scenarios increases.

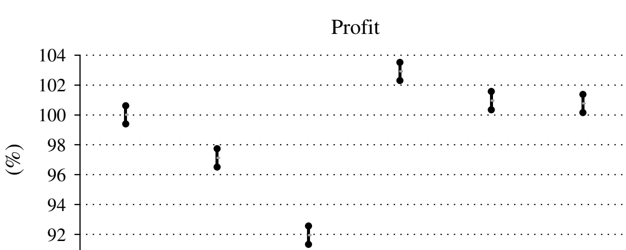
Tyler et al. (2003) investigates the trade-offs among OT utilization, overtime requirements, and tardiness of elective operations in an "ideal" OT environment. Specifically, the work simulates an OT environment where a given set of elective operations is to be performed. The elective operation time requirements follow identical distributions with relatively short mean values (i.e., less than one hour) and low variability (i.e., CV values equal to 0.3). The simulation suggests that achieving OT utilization greater than 85% results in significant increases in start time tardiness and overtime. Our results confirm this observation with all models achieving average OT utilization levels less than 85%. However, model S5 provides schedules within 2.5% of the 85% threshold, on average, although the degree of uncertainty considered is much greater than the ideal setting of Tyler et al. (2003).

Figure 9 provides 95% confidence intervals for the computational time requirements and the percent of instances requiring at least 30 minutes in the left and right panels, respectively. The figure shows that scenario-based models require more computational effort to solve and that an increasing number of problem instances reach the 1,800-second limit as the number of scenarios increases. This increasing computational burden enables the S5 model to outperform models S10 and S15.

5.4. Additional Experiments

We now consider the impact of OT size on profit, OT utilization, average time between consecutive BIMs, maximum time between consecutive BIMs, and CPU time in seconds. We use a radar graph containing spokes for each of the stated performance measures where the data value along a spoke indicates the magnitude of the observed effect for a specific level of OT size, relative to the maximum observed effect across all OT size levels. A line connects data values corresponding to the same OT size level, producing a star-like appearance. The described radar graph is given in Figure 10. The graph highlights the scalability of

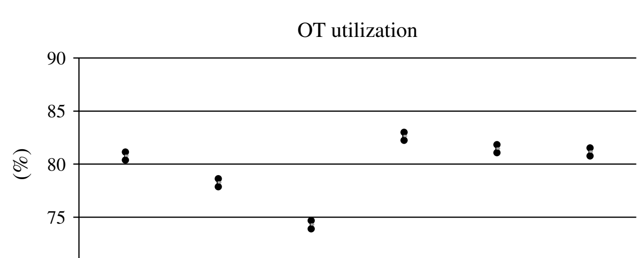
Figure 7 **Relative Profit for Model Levels**



	EV	JH65	JH70	S5	S10	S15
High (%)	100.6	97.7	92.6	103.5	101.6	101.4
Avg (%)	100.0	97.1	91.9	102.9	100.9	100.8
Low (%)	99.4	96.5	91.3	102.3	100.3	100.1

Note. All profit values are expressed as percent of the average profit observed for the EV model.

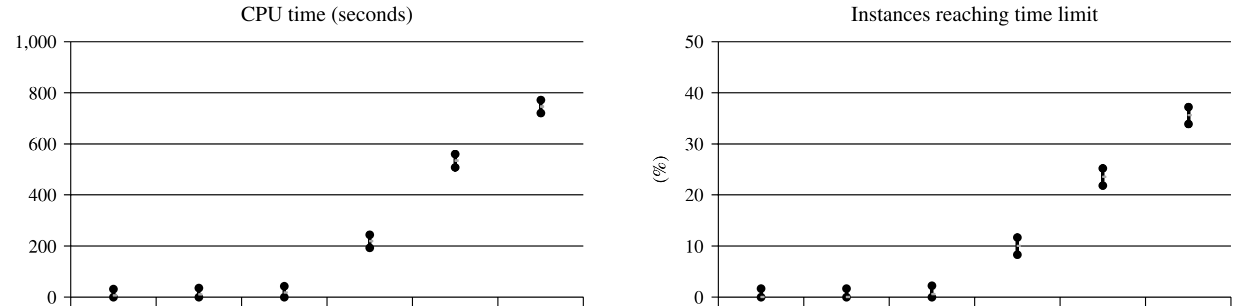
Figure 8 **OT Utilization for Model Levels**



	EV	JH65	JH70	S5	S10	S15
High (%)	81.1	78.6	74.7	83.0	81.8	81.5
Avg (%)	80.8	78.3	74.3	82.6	81.5	81.2
Low (%)	80.4	77.9	73.9	82.3	81.1	80.8

Note. Similar to Guerriero and Guido (2011), we define OT utilization as the total time spent performing elective operations during the planned time for operating rooms, including setup and cleanup, divided by the sum of the planned operating time for all operating rooms.

Figure 9 **Computational Performance for Model Levels**



	EV	JH65	JH70	S5	S10	S15
High	31.7	35.7	43.0	244.1	559.1	771.3
Avg	5.8	9.8	17.1	218.2	533.2	745.4
Low	0.0	0.0	0.0	192.4	507.4	719.6

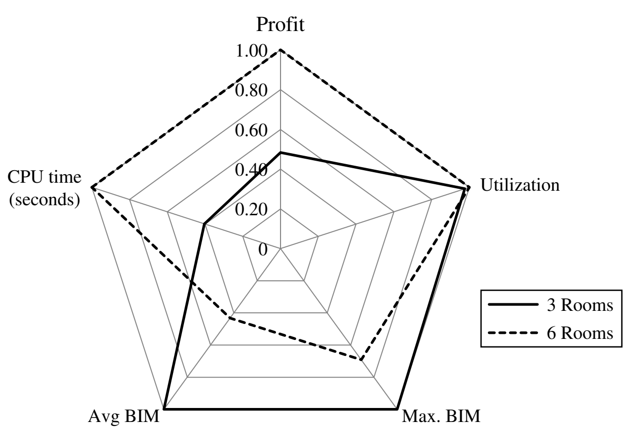
	EV	JH65	JH70	S5	S10	S15
High (%)	1.7	1.7	2.2	11.7	25.2	37.2
Avg (%)	0.0	0.0	0.6	10.0	23.5	35.6
Low (%)	0.0	0.0	0.0	8.3	21.8	33.9

Note. The solution time for a particular instance reaching the time limit signifies that CPLEX is unable to prove that the incumbent solution is optimal within 30 minutes (1,800 seconds) of CPU time.

the two-step solution approach, which achieves a proportional increase in profit as the OT size doubles, with little, if any, deterioration in OT utilization. With respect to BIM requirements, both the maximum and average time between consecutive BIMs decreases in the larger OT. Since a BIM occurs during the setup of elective operations, when the cleanup for an elective operation is complete, and during any idle time between elective operations, more break-in-moments will naturally exist in a larger OT. Although we see no deterioration in the quality of schedules determined by the two-step solution approach as the OT size increases, computational time requirements increase.

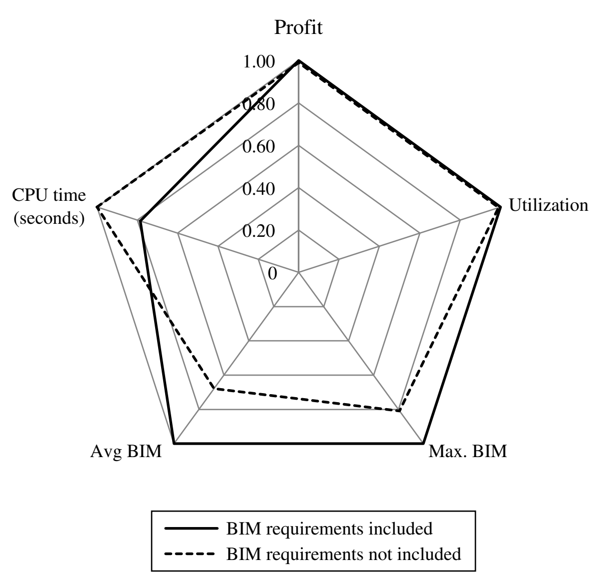
Figure 10 shows that the maximum and average time between consecutive BIMs naturally decreases as the OT size increases. However, incorporating BIM requirements improves both the average and maximum time an urgent patient waits, regardless of the OT size. This observation is illustrated in Figure 11, which shows that the average and maximum time between consecutive BIMs reduces by approximately 25% and 20%, respectively, when BIM requirements are included. Moreover, the inclusion of BIM requirements has little effect on either profit or OT utilization; however, computational time requirements increase by approximately 25%. Ultimately, the results show that the inclusion of BIM requirements when scheduling for an OT allow the hospital to provide quicker service to urgent patients with little impact on key OT performance indicators. This characteristic of a hospital is especially desirable from the perspective of emergency response providers.

Figure 10 **Impact of OT Size**



Notes. All values are expressed relative to the maximum value for a particular performance measure. For example, when the OT consists of three operating rooms, the expected profit is approximately 0.5, or 50%, of the profit observed when the OT consists of six operating rooms.

Figure 11 **Impact of BIM Requirements**



6. Conclusion

This research presents a novel, scenario-based method that effectively incorporates the uncertain time requirements of elective operations into operating theater scheduling decisions and accounts for potentially arriving urgent patients. Our two-step solution approach offers a significant reduction in the number of binary assignment variables associated with an SBESS instance and thus allows the attainment of high-quality schedules within practical time frames.

Experimental results show that the scenario-based models provide schedules with higher, and statistically significant, levels of profit and OT utilization. The stated benefits are obtained by including only five scenarios with the first scenario using the mean value for all time requirements. Moreover, accounting for urgent operations via break-in-moment requirements significantly reduces the maximum and average times an urgent patient waits to receive care, with little deterioration in profit and OT utilization. The ability to provide quicker service to urgent patients is desirable from the perspective of both patients and emergency medical service providers.

Schedules found by the two-step solution approach provide a plan for a single day in an OT. However, hospitals often consider a multiday horizon when developing an OT schedule. Thus, future work focuses on evolving the two-step approach to develop multiday schedules. The multiday scheduling procedure must account for disruptive events, i.e., the arrival of an urgent patient, that require modification of the initial schedule. Thus, there is also a need for the development of an adaptive scheduling component that monitors deviations from the initial plan and makes adjustments when significant disruptions occur. Bean et al. (1991) demonstrate the utility of such an adaptive scheduling approach in an automotive manufacturing environment. Another area of future research is the application of a modified SBESS formulation to address other scheduling environments subject to uncertainty.

Acknowledgments

The authors thank the associate editor and two anonymous reviewers for their comments and suggestions. Their inputs resulted in significant improvements to the practicality and exposition of this paper.

Appendix. SBESS Complexity

We show that the SBESS problem is NP-hard through a reduction of a traveling salesman problem (TSP) with variable arc costs to an SBESS instance. Consider a tour-planning problem among a set of $\mathcal{N}$ cities where arc travel times are uncertain. Specifically, there is a set of S scenarios, $\mathcal{S}$, each containing different realizations for the travel time of each arc. Assume that the travel time associated with the arc connecting an origin city n to a destination city m in scenario s is independent of the destination. Under this assumption, we denote the travel time from an origin n to any destination in scenario s as τ_n^s. In addition, assume that the sum of the arc travel times for all cities is equal to a constant D in each scenario (i.e., $\sum_{n\in\mathcal{N}}\tau_n^s = D, \forall s\in\mathcal{S}$), and that the duration of the tour must not exceed D. The objective is to determine a sequence, visiting each city once, and an earliest arrival time for cities in each position of the sequence such that the expected sum of the differences between the earliest arrival time and the actual arrival time at each city is minimized. Specifically, we seek to determine the position p that each city should be visited and the earliest arrival time for cities in all tour positions. We assume that the tour starts and ends at a home city 0, with negligible travel time between the start node and any city. Letting A_{np} be a binary decision variable taking a value of 1 if city n is in position p of the tour sequence and letting Y_p represent the earliest arrival time at the city in position p of the tour, an expression for the objective function we wish to minimize follows where $\mathbb{E}$ represents the expected value over the set of scenarios $\mathcal{S}$. The city in position 1 is omitted in the outer sum since it is visited from the start node with a travel time of 0.

$$\sum_{p=2}^{N}\mathbb{E}\left(\sum_{\hat{p}=1}^{p-1}\sum_{n=1}^{N}\tau_n^s A_{n\hat{p}} - Y_p\right)$$

Although we assume that the travel time from an origin city n to any destination city is independent of the destination, the objective we define is sequence dependent. A mixed integer programming (MIP) formulation for the problem follows. In the formulation, we seek to assign each of the N cities to a position $p=1,\ldots,N$ in the tour sequence as well as an earliest arrival time to city n. We allow the index for the tour position p to take on values of 0 and $N+1$, and require that the start node is assigned to both of these positions (i.e., position 0 of the tour sequence is the home city, positions $1\ldots N$ define the sequence among the cities, and at position $N+1$ the tour returns to the home).

$$\text{Minimize} \quad \sum_{p=2}^{N}\mathbb{E}\left(\sum_{\hat{p}=1}^{p-1}\sum_{n=1}^{N}\tau_n^s A_{n\hat{p}} - Y_p\right) \tag{31}$$

$$\text{Subject to} \quad \sum_{n=1}^{N} A_{np} = 1, \quad p=1,\ldots,N; \tag{32}$$

$$\sum_{p=1}^{N} A_{np} = 1, \quad n=1,\ldots,N; \tag{33}$$

$$Y_p \le \sum_{\hat{p}=1}^{p-1}\sum_{n=1}^{N}\tau_n^s A_{n\hat{p}}, \quad p=2,\ldots,N,\ s\in\mathcal{S}; \tag{34}$$

$$A_{00} = 1; \tag{35}$$

$$A_{0(N+1)} = 1; \tag{36}$$

$$A_{np} \in \{0,1\}, \quad n,p=1,\ldots,N; \tag{37}$$

$$c_p \ge 0, \quad p=2,\ldots,N. \tag{38}$$

The objective function seeks to minimize the expected difference between the earliest arrival time and the actual arrival time for cities assigned to all positions of the tour. Note that we only need to include the costs associated with cities in position $2\ldots P$ because the travel time from the home city to the first city and the travel time from the city in position N back to the home city take no travel time. Constraint set (32) ensures that one city is assigned to each tour position. Constraint set (33) ensures that each position $p=1,\ldots,N$ is assigned a city. Together, constraint sets (32) and (33) serve to eliminate subtours. Constraint set (34) requires the value of Y_p to be less than or equal to the earliest possible arrival time to a city in position p of the tour sequence, across all scenarios. Constraints (35) and (36) require the tour to begin and end at the home city 0. Constraint sets (37) and (38) define the domain of the assignment variables.

Although the form is unconventional, the described problem is a TSP with variable arc costs. Moreover, the described problem directly corresponds to a special case SBESS instance consisting of a single operating room block. In this instance, elective operations are represented by cities, the maximum allowable time between the arrival and start of an urgent demand is large (i.e., $\Delta\to\infty$), the time allotted for the operating room block is D, and the elective operations to sequence within the block are given. The uncertain travel times correspond to the sum of the setup, procedure, and cleanup times for the elective operations (i.e., $\tau_n^s = r_{ns}+q_{ns}$, $n=1,\ldots,N$, $s=1,\ldots,S$). Moreover, the sum of these times equals D in all scenarios. There is no revenue associated with performing elective operations, and the only costs arise as a result of the tardiness of elective operations (i.e., the difference between actual arrival times and earliest arrival times). A solution of the described TSP does provide a solution for the described SBESS instance. However, since the TSP is NP-hard for the case where arc costs are explicitly defined, the version we describe with variable arc costs is also NP-hard. Moreover, since we can construct a TSP equivalent to an SBESS instance, the SBESS is also NP-hard.

References

Addis B, Carello G, Grosso A, Tanfani E (2014) A rolling horizon framework for the operating rooms planning under uncertain surgery duration. Working paper, Universite' de Lorraine, Nancy, France.

Batun S, Denton B, Huschka T, Schaefer A (2011) Operating room pooling and parallel surgery processing under uncertainty. *INFORMS J. Comput.* 23(2):220–237.

Bean JC, Birge J, Mittenthal J, Noon C (1991) Matchup scheduling with multiple resources, release dates and disruptions. *Oper. Res.* 39(3):470–483.

Cardoen B, Demeulemeester E, Beliën J (2010a) Innovative applications of O.R.: Operating room planning and scheduling: A literature review. *Eur. J. Oper. Res.* 201(3):921–932.

Cardoen B, Demeulemeester E, Van Der Hoeven J (2010b) On the use of planning models in the operating theatre: Results of a survey in Flanders. *Internat. J. Health Planning Management* 25(4):400–414.

Chen CK, Lin C, Hou TH, Wang SH, Lin HM (2010) A study of operating room scheduling that integrates multiple quantitative and qualitative objectives. *J. Nursing Res.* 18(1):62–74.

Denton B, Viapiano J, Vogl A (2007) Optimization of surgery sequencing and scheduling decisions under uncertainty. *Health Care Management Sci.* 10(1):13–24.

Erdem E, Qu X, Shi J (2012) Rescheduling of elective patients upon the arrival of emergency patients. *Decision Support Systems* 54(1):551–563.

Ferreira R, Coelli F, Pereira W, Almeida R (2008) Optimizing patient flow in a large hospital surgical centre by means of discrete-event computer simulation models. *J. Evaluation Clinical Practice* 14(6):1031–1037.

Graham RL, Lawler EL, Lenstra JK, Rinnooy Kan AHG (1979) Optimization and approximation in deterministic sequencing and scheduling: A survey. *Ann. Discrete Math.* 5:287–326.

Guerriero F, Guido R (2011) Operational research in the management of the operating theatre: A survey. *Health Care Management Sci.* 14(1):89–114.

Gul S, Denton B, Fowler J, Huschka T (2011) Bi-criteria scheduling of surgical services for an outpatient procedure center. *Production Oper. Management* 20(3):406–417.

Gupta D, Denton B (2008) Appointment scheduling in health care: Challenges and opportunities. *IIE Trans.* 40(9):800–819.

Hancock WM, Walter PF, More RA, Glick ND (1988) Operating room scheduling data base analysis for scheduling. *J. Medical Systems* 12(6):397–409.

Hans E, Wullink G, van Houdenhoven M, Kazemier G (2008) Robust surgery loading. *Eur. J. Oper. Res.* 185(3):1038–1050.

Klotz E, Newman A (2013) Practical guidelines for solving difficult linear programs. *Surveys Oper. Res. Management Sci.* 18(12): 1–17.

Lamiri M, Grimaud F, Xie X (2009) Optimization methods for a stochastic surgery planning problem. *Internat. J. Production Econom.* 120(2):400–410.

Lamiri M, Xie X, Dolgui A, Grimaud F (2008) A stochastic model for operating room planning with elective and emergency demand for surgery. *Eur. J. Oper. Res.* 185(3):1026–1037.

Lin E (1998) A bibliographical survey on some well-known non-standard knapsack problems. *Inform. Systems Oper. Res.* 36(4):274–317.

Macario A (2006) Are your hospital operating rooms "efficient"?: A scoring system with eight performance indicators. *Anesthesiology* 105(2):237–240.

Macario A (2010) What does one minute of operating room time cost? *J. Clinical Anesthesia* 22(4):233–236.

Min D, Yih Y (2010) Scheduling elective surgery under uncertainty and downstream capacity constraints. *Eur. J. Oper. Res.* 206(3):642–652.

Muñoz E, Muñoz W, Wise L (2010) National and surgical health care expenditures, 2005–2025. *Ann. Surgery* 251(2):195–200.

Rachuba S, Werners B (2014) Robust approach for scheduling in hospitals using multiple objectives. *J. Oper. Res. Soc.* 65(4): 546–556.

Rais A, Viana A (2011) Operations research in healthcare: A survey. *Internat. Trans. Oper. Res.* 18(1):1–31.

Ross SM (2006) *Simulation (Statistical Modeling and Decision Science)* (Elsevier Academic Press, Burlington, MA).

Shylo O, Prokopyev O, Schaefer A (2012) Stochastic operating room scheduling for high-volume specialties under block booking. *INFORMS J. Comput.* 25(4):682–692.

Spangler W, Strum D, Vargas L, May J (2004) Estimating procedure times for surgeries by determining location parameters for the lognormal model. *Health Care Management Sci.* 7(2):97–104.

Strum D, May J, Vargas L (2000) Modeling the uncertainty of surgical procedure times—comparison of log-normal and normal models. *Anesthesiology* 92(4):1160–1167.

Surgical Directions (2014) Systematic approach to scheduling smoothes out the daily wrinkles. *OR Manager* 30(6):21–23.

Tyler D, Pasquariello C, Chen C (2003) Determining optimum operating room utilization. *Anesthesia Analgesia* 96(4):1114–1121.

van Essen J, Hurink J, Hartholt W, van den Akker B (2012a) Decision support system for the operating room rescheduling problem. *Health Care Management Sci.* 15(4):355–372.

van Essen JT, Hans EW, Hurink JL, Oversberg A (2012b) Minimizing the waiting time for emergency surgery. *Oper. Res. Health Care* 1(2–3):34–44.

Wang Y, Tang J, Fung R (2014) A column-generation-based heuristic algorithm for solving operating theater planning problem under stochastic demand and surgery cancellation risk. *Internat. J. Production Econom.* 158:28–36.