

Singapore Management University
Institutional Knowledge at Singapore Management University

Research Collection Lee Kong Chian School Of
Business

Lee Kong Chian School of Business

8-2008

Dynamic Allocation of Airline Check-in Counters: A Queueing Optimisation Approach

Mahmut PARLAR
McMaster University, Canada

Sharafali MOOSA
Singapore Management University, sharafalim@smu.edu.sg
DOI: <https://doi.org/10.1287/mnsc.1070.0842>

Follow this and additional works at: https://ink.library.smu.edu.sg/lkcsb_research

Part of the [Operations and Supply Chain Management Commons](#), and the [Operations Research, Systems Engineering and Industrial Engineering Commons](#)

Citation

PARLAR, Mahmut and MOOSA, Sharafali. Dynamic Allocation of Airline Check-in Counters: A Queueing Optimisation Approach. (2008). *Management Science*. 54, (8), 1410-1424. Research Collection Lee Kong Chian School Of Business.
Available at: https://ink.library.smu.edu.sg/lkcsb_research/881

This Journal Article is brought to you for free and open access by the Lee Kong Chian School of Business at Institutional Knowledge at Singapore Management University. It has been accepted for inclusion in Research Collection Lee Kong Chian School Of Business by an authorized administrator of Institutional Knowledge at Singapore Management University. For more information, please email libIR@smu.edu.sg.



Management Science

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

Dynamic Allocation of Airline Check-In Counters: A Queueing Optimization Approach

Mahmut Parlar, Moosa Sharafali,

To cite this article:

Mahmut Parlar, Moosa Sharafali, (2008) Dynamic Allocation of Airline Check-In Counters: A Queueing Optimization Approach. Management Science 54(8):1410-1424. <http://dx.doi.org/10.1287/mnsc.1070.0842>

Full terms and conditions of use: <http://pubsonline.informs.org/page/terms-and-conditions>

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact permissions@informs.org.

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

Copyright © 2008, INFORMS

Please scroll down for article—it is on subsequent pages



INFORMS is the largest professional society in the world for professionals in the fields of operations research, management science, and analytics.

For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

Dynamic Allocation of Airline Check-In Counters: A Queueing Optimization Approach

Mahmut Parlar

DeGroote School of Business, McMaster University, Hamilton, Ontario L8S 4M4, Canada,
parlar@mcmaster.ca

Moosa Sharafali

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

This paper was motivated by an observation in an international airport with regard to allocation of resources for check-in counters. In an exclusive check-in counter system, each flight has a dedicated number of counters that will be open until at least a half-hour before the scheduled departure of that flight. Currently, in many of the airports around the world, the decision to open or close check-in counters is done on an ad hoc basis by human schedulers. In doing so, the schedulers are almost always forced to perform a balancing act in meeting the quality of service stipulated by the airport authority vis-à-vis the optimal allocation of the resources to the counters. There appear to be very few academic and application papers in counter management, and most of those that have looked into this problem have resorted to simulation to study the queue characteristics. Ours is the first paper to show that for a specific flight, this complicated problem is amenable to analytical treatment. We first propose a multicounter queueing model with a special type of arrival process reflecting reality from the population of passengers booked for the flight. Most importantly, we derive the time-dependent operating characteristics to the queueing process under a specified time-window constraint. Then a stochastic dynamic programming model is formulated to determine the optimal numbers of counters to open over the time window specified. A numerical example is provided to illustrate the model solution and gain managerial insights.

Key words: queues; transient results; dynamic programming; applications; transportation; scheduling

History: Accepted by Michael Fu, stochastic models and simulation; received September 20, 2006. This paper was with the authors 2½ months for 2 revisions. Published online in *Articles in Advance* May 14, 2008.

1. Introduction

Globally, the recent years have witnessed a dramatic surge in demand for air transportation. Notwithstanding the post-9/11 impact, factors like globalization and the measures taken by governments in deregulating air traffic have borne fruit in the form of increased passenger load factor. The advent of low-cost carriers, like ValueAir and Air Asia in Asia-Pacific, Jet Star in Australia, and similar airlines in other regions, bear testimony to this. According to Boeing's Current Market Outlook 2007,¹ worldwide passenger traffic growth will average 4.5% up to the year 2026. In the developing economies of the Asia-Pacific region, this figure will be around 5.5%, and in China alone the growth is expected to be 8.1%. A similar trend is also expected in cargo traffic.

Although increase in demand for air transportation is a promising sign for airlines, civil aviation

authorities, terminal operators, and ancillary service providers, it also raises a whole range of challenges due to the increased congestion that will result at the terminals. Even though extra capacity will be added to profitable routes, airline terminal handling nonetheless still remains confined to within the terminals. Also, increased capacity only leads to greater congestion on the runway of airports, and hence the need to ensure timely aircraft takeoffs. Typically, any delay in scheduled flight departures will cascade down in further upheavals in downstream flight itineraries for the rest of the day. As such, these operational challenges require immediate attention and swift resolution. The imperative is to clear passengers checking into a terminal counter within a specified time window, without having to incur unnecessary additional counter openings or compromises on customer service. In the worst case, terminal operators have to open or even "borrow" additional counters from nearby counters meant for other scheduled flights to ease the surge in demand and to meet the

¹ Accessed November 11, 2007, <http://www.boeing.com/commercial/cmo/highlights.html>.

deadline set by the airport authority. These are usually intended on a discretionary basis to meet the upsurge in the economy-class passengers who form the bulk of passengers and are without any prior check-in done. Clearly, costwise, this knee-jerk reaction is not optimal because it incurs additional cost and strain in staffing, counterspace rental, equipment, and operations planning.

The motivation for the work presented in this paper arose from an observation by the authors in an international airport on the balancing act the airport operator had to perform in meeting the performance standards set by the airport authority while at the same time maintaining a judicious use of the valuable resources. For example, the most important concern for the aviation authority and the ground service provider is the ability to consistently maintain high-quality schedules for the daily check-in counter management, given the variations in the skill and experience level of each human scheduler. The check-in counter allocation schedule (or the delay experienced) has a very important impact on the overall image of the airport and airlines as perceived by the passengers. An efficient schedule simply translates to shorter queues, less congestion at the counters, faster check-in, more time to spend at the duty-free section, and, consequently, less harried passengers.

This paper presents a model based on the evaluation of the operational practices of international airports. Although the trend in airports now is to use paperless tickets and self check-in kiosks, this is not so in many of the airports around the world, and especially in the high-growth Asia-Pacific region. Currently, the decision to open or close check-in counters is done on an ad hoc “gut-feel” basis by the counter supervisor in most of Asia (presumably similar to those practices found in the retail sector). Hence, this research is motivated for the exact purpose of developing a counter allocation management system to help predict the resource requirements at international airports. The thrust of such a system is to assign valuable resources efficiently to meet business demands without compromising service standards.

This research should be approached in two phases. The first phase is the queueing and statistical analysis phase to help determine the optimal number of counters needed for each flight over time to minimize a certain expected cost function (while implicitly achieving a desired level of customer service). The second phase is the optimization phase required to schedule and assign counters to flights, subject to meeting the various airport- and airline-specific constraints. This paper attempts to model and analyze the first phase only. For this phase of work, we first obtain the operational characteristics of the basic

queueing model and then use stochastic dynamic programming to determine the number of counters to allocate to each scheduled flight for the duration of time required. The basic model proposed is a multicounter single queue with the arrival process occurring according to a “passenger show-up process” from the finite population of passengers holding confirmed bookings. In stochastic processes, this arrival process is referred to as a death process (see Bhat 1984, p. 208; Feller 1968, p. 478). In the most generous case, this population is no more than 550 passengers on the Boeing 747-400, the capacity of the largest passenger plane today. We highlight that the multicounter model considered in this paper will also be useful for a multicounter system with the assumption of independent arrivals to each counter.

Our main contribution to the research literature in this area is in deriving the time-dependent solution to the queueing process. Lee’s (1966) very interesting anecdotal description of his experiences in modeling this problem provides detailed insights into the difficulties encountered in modeling and analysis. Hence, we feel that our work in this paper is an important progress in this area. Another significant contribution of our work is the dual purpose our results can serve. Our main thrust here is to minimize the operating cost for the service provider through the use of the dynamic programming model analyzed. However, if the service provider’s primary objective is to meet the quality of service (QoS) mandated by the airport authority, our results can still be used to analyze the resource requirements to meet the QoS. Because our model is analytical and more realistic, the optimization results are on firmer ground than those based on simulation.

The rest of this paper is organized as follows. Section 2 presents the relevant literature for the area of study. The model and assumptions are presented in §3. We perform the analysis of the multicounter model in §4, where we seek to determine the time-dependent operating characteristics of a stochastic process operating under a specified time-window constraint. In §5, we discuss a more general model with Erlang service times and compare it to the standard model with exponential service times using the properties of the time-to-absorption random variable for the queueing process. In §6, we present a statistical procedure for estimating the parameters of the pure-death process where the show-up rate is time dependent. In §7, our objective is to determine the optimal number of counters to open over time. We do this by developing a stochastic dynamic programming model that provides some managerial insights. Suggestions

for future research and some concluding remarks on this piece of applied work are highlighted in the final section. All appendices are provided online in the e-companion.²

2. Literature Review

The literature in this area is scant, and there has been a host of methods used to model the real-life problem. The earliest work in this area is by Lee (1966) who provides a very interesting account of a similar exercise carried out over a period of time for an airport in the United Kingdom, where the passenger arrival stream was assumed to be Poisson, and an $M/M/s$ queue was used to model the check-in counter system. A recent work on this problem for the Hong Kong Airport by Chun and Mak (1999) also assumes a Poisson arrival stream at the counter and beta distributed service times. The work still employed only simulation to determine the number of counters to open for each flight, but markedly differed from the earlier works (see, e.g., Haeme et al. 1988), all of which simulated the overall situation in an airport terminal building after a schedule had been developed. Further, one can only find predominantly application papers, which rely mainly on simulation to present a *prima facie* case. For example, Bitauld et al. (1997) describe the IBM and Air Canada joint effort called "Journey Management." They apply the simulation technique to capture the impact of new technologies in managing congestion in airports. They highlight IBM's initiative in developing a library of building blocks and templates, called the IBM Journey Management Library, for use with simulation tools.

SmartAirport³ is a comprehensive software suite developed jointly by IBM and Ascent Technology Inc. This integrated application suite can merge the planning and real-time management of multiple airline activities, such as gate management and real-time personnel allocation, making the best use of limited and costly assets. On the counter allocation aspect, it claims to use probabilistic models to analyze the problem, but no information on the type of models used is made available.

Other recent papers on Asian airports, and related to our work here, include that of Littler and Whitaker (1997). They have provided a procedure to estimate the staffing requirements to meet a preset processing time target, using stochastic simulation of passenger arrivals into the terminal of an airport in New Zealand. More recently, Park and Ahn (2003) revisited

the problem of passenger arrivals at the Gimpo airport in Korea. They have argued that utilizing check-in counters efficiently is key for operators (the space owners) and for airlines (the renters). They derive the optimal assignment for check-in operations making use of passenger arrival distribution patterns to determine the most appropriate number of check-in counters. Using a different methodology, Yan et al. (2004) provide an integer programming model to assign common-use check-in counters for Taipei's International Airport. However, due to the large problem size and complexity, they had to resort to a heuristic method to solve the model.

Very recently van Dijk and van der Sluis (2006) considered a check-in counter management problem that is similar in spirit to ours. They highlight the fact that one of the important features of this problem for any single flight is the finite calling population size. Hence, they resort to "terminating simulation" to identify the number of counters to open, unlike almost all the other works, which use nonterminating simulation. In our paper, we too incorporate this realistic characteristic in our model, but instead of using simulation, we demonstrate the analytical tractability of the model.

From the foregoing, it is clear that there has not been much academic research progress, particularly in terms of the structure of the problem and gaining analytical insights. Hence, this paper is an attempt to fill this gap in the literature. Most importantly, time-dependent solution is what is required for this problem, and we have demonstrated that it is possible to obtain closed-form formulas. Use of these formulas will enhance the quality of the solution for such real-life problems. As noted in the introduction, in this paper we use stochastic dynamic programming and build a basic model for allocating counters to a single flight, which would then be extended in the future to capture the comprehensive check-in allocation problem for allocating counters to multiple flights, subject to capacity limitations.

3. Multicounter Single-Line Check-In Counter System

There are two kinds of check-in counter systems used in airports. They are the common-use system and the exclusive-use system. The common-use system consists of a long continuum of counters. Passengers, irrespective of their flight, can check in at any of those counters. Obviously, the common-use system yields higher utilization of the resources. As opposed to this, the exclusive-use system is a set of counters dedicated to every flight. They are opened just a few hours before the scheduled departure of a flight and closed about a half-hour before the departure.

² An electronic companion to this paper is available as part of the online version that can be found at <http://mansci.journal.informs.org/>.

³ Accessed November 11, 2007, <http://www.ascent.com/resource-management-airports.html>.

In the airport terminal the authors visited, the system in use is an exclusive-use system. Further, the terminal has been designed for the use of an island-type check-in facility. In an island-type system, the counters are arranged in a U-shape around a single conveyor. Although this provides some flexibility, the current layout of the islands does not permit the change to a common-use check-in counter system. Therefore, in this paper the focus is on the exclusive-use system.

The problem analyzed in this paper has some resemblance to the management of call centers and supermarket check-out counters. We refer the reader to the website <http://www.math.vu.nl/obp/callcenters/>⁴ for a wealth of information on call center management, and to Koltai and Kalló (2007) for supermarket check-out counter management. However, the striking features in our model that do not exist in the above two areas are (i) the calling population size is finite, and (ii) the system will never reach the steady state.

The four primary features that stand out in typical check-in counter queues at the terminal are as follows: (i) The system in use is an exclusive-use system, with multiple counters and a single queue of passengers. These counters will service only those passengers booked for that flight. (ii) The finite value of the number of confirmed passengers (i.e., calling population size) for each flight is known a priori. This implies that the calling population size for these counters will decrease with time; that is, such a system will never reach the steady state. Therefore, a transient solution will be needed to effectively understand and manage the queue. (iii) Counters typically open three (or four) hours before the onset of the scheduled flight departure and have to close 30 minutes before the boarding gate closes irrespective of whether or not all passengers show up at the counters. Those who failed to show up are classified as “no-shows,” in the air travel parlance, and (iv) on observing the growing queue, the counter staff always tends to speed up service, making the service time depend on the state of the system. Hadidi (1969) was the first to consider such service rates. In fact, our observation suggests that the counter staff tend to spend more time with passengers when the queue is relatively short or empty. At the same time, when the queue grows to unwieldy levels within the traffic island (area around the counter), some other unallocated staff or supervisor goes to the help of the staff manning the overly busy counter, thereby increasing the effective service rate for the particular counter.

Note that in a common-use system where passengers of any flight at any time can check in at any

of the counters (as opposed to the exclusive-use system described in feature (i) above), the calling population will be, effectively, infinity, and so the system is expected to reach the steady state. It is possible that the arrival process may be Poisson due to the superposition of several renewal arrival processes. For such a system, infinite population queueing results can be used. However, our model, which reflects reality in the airport visited, clearly does not admit a steady-state solution.

Queueing models with finite population size have already been studied in the context of machine repair. Again, the feature that distinguishes our model from this literature is that in the machine repair problem, a repaired machine returns to the pool after repair, whereas a passenger never returns to the pool in our model, but proceeds to take the flight. It should be noted again that this is the sole reason that the check-in counter system we model will never reach the steady state. The other area of study where a similar stochastic process input is considered is the stochastic epidemic or, more generally, the stochastic population models. However, it should be noted that these models are more akin to infinite server queueing models where waiting never occurs, whereas in the airport check-in counter scenario waiting occurs and so is very relevant.

Let us add here that Hadidi (1969), who was the first to consider such state-dependent service rates in a single-server system, refers to it as “a potentially infinite capacity system.” Hadidi and Conolly (1969) comment further on the analogy of the state-dependent single-server system to the $M/M/\infty$ system by indicating that even though the single server increases his rate of service when an arrival occurs, the problem of waiting time is still relevant.

One additional comment is needed on the assumption of exponential service-time distribution. Because the counters are open only for a finite length of time (e.g., $T = 3$ hours), the following natural question may arise: Isn't a truncated exponential (truncated at T) more suitable than a distribution with infinite support because the counters are open only for a duration of T hours? However, in reality, there may be no-shows, and so there will be passengers who might arrive beyond T , also. Further, such an assumption would also destroy the Markovian nature of the process and make the model intractable.

We assume that arrivals form a “death process” from the population of N travelers booked on the flight. Formally, a death process is defined as follows: Consider a population of size N . For some reason, the members of the population get removed from the population at random times. This phenomenon of removal is called a death. The random lifetime of a

⁴ Accessed November 11, 2007.

member is assumed to follow an exponential distribution with mean λ^{-1} . Now, the stochastic process $A(t)$, which is the number of removals until time t , is a death process where each death corresponds to a removal from the population (and, for our problem, showing up at the airport). Given that at time 0 a total of m individuals have been removed (already showed up), the conditional probability of $A(t)$ is given as (Bhat 1984, p. 208)

$$\Pr\{A(t) = i \mid A(0) = m\} = \binom{N-m}{i-m} (1 - e^{-\lambda t})^{i-m} (e^{-\lambda t})^{(N-i)}. \quad (1)$$

This is clearly the binomial distribution where $1 - e^{-\lambda t}$ is the probability that an individual will “die” before t , i.e., will show up at the airport before t . Thus, the expected value of $A(t)$ is $E[A(t) \mid A(0) = m] = (N - m)(1 - e^{-\lambda t})$.

We now see that an arrival at a check-in counter for a flight is a removal from the population of passengers booked for that flight. Henceforth, we will use the term “passenger show-up process” instead of the term “death process.”

Let us assume the show-up rate to be λ . As highlighted above in the discussion of the features of check-in counter queues [feature (iv)], it is appropriate that we consider a state-dependent service rate. Hence, we assume a special type of state-dependent service according to which the effective instantaneous service rate depends both on the number of passengers in the check-in counter system and also on the number of counters open at that instant. This means that if at time t there are $c \equiv c(m, n)$ counters open and there are k passengers in the system (including the passengers undergoing service at the counters), then the instantaneous service rate at any counter is $ck\mu$, i.e., the probability that a departure occurs in $(t, t + \Delta t)$ is $ck\mu\Delta t + o(\Delta t)$. This is almost always true in the check-in counter system because new counters will be opened only if they are absolutely needed. Further, if any counter is idle, then that would signal the closure of that counter. This is so in reality, and also in the dynamic programming model to be considered in the sequel, where the system will be observed in short time intervals to decide whether to add more counters or close down idle counters.

The number of counters to open, $c \equiv c(m, n)$, is a decision variable in this problem. We analyze the queue behavior for the multi-check-in-counter system. Specifically, we use the notation in Table 1 listed in alphabetical order.

We also define $P_{m,n}^c(i, j, t) = \Pr\{A(t) = i, S(t) = j \mid A(0) = m, S(0) = n\}$ as the transient probability that at time t , i have arrived and j have been served given that at time 0, m had already arrived and n

Table 1 List of Notation in Alphabetical Order

Symbol	Description
$A(t) = m$	Number of passengers that have arrived by time t equals m
$c \equiv c(m, n)$	Number of check-in counters that are open (decision variable)
C_s	The unit cost of operating a counter [\$/counter-time]
C_w	The unit cost of making a passenger wait [\$/passenger-time]
h	Variable cost for each arrived passenger not cleared check-in by time T [\$/passenger]
λ	The passenger show-up rate [passenger/time]
μ	The service rate [passenger/time]
N	Number of passengers booked for the flight
T	Duration of time the check-in counter system will be open
τ	Time to absorption from state $(0, 0)$ to state (N, N)
$P_{m,n}^c(i, j, t)$	Transient probability of finding the system in state $(A(t), S(t)) = (i, j)$ at time t given that the system was in state (m, n) at time 0 and c counters were open
$\Pi_{m,n}(u, v, t)$	The p.g.f. of $P_{m,n}^c(i, j, t)$
$\hat{\Pi}_{m,n}(u, t)$	The p.g.f. of $Q_{m,n}(k, t)$
$Q_{m,n}(k, t)$	Transient probability of finding k passengers in the system at time t , i.e., $Y(t) = k$, given that the system was in state (m, n) at time 0
$S(t) = n$	Number of passengers that have been served by time t equals n
$Y(t) = A(t) - S(t)$	Number of passengers in the system at time t

had been served ($i = m, \dots, N$ and $j = n, \dots, i$) and $c \equiv c(m, n)$ counters had been opened. The probability generating function (p.g.f.) of $P_{m,n}^c(i, j, t)$ is defined as $\Pi_{m,n}(u, v, t) = \sum_{i=m}^N \sum_{j=n}^i P_{m,n}^c(i, j, t) u^i v^j$.

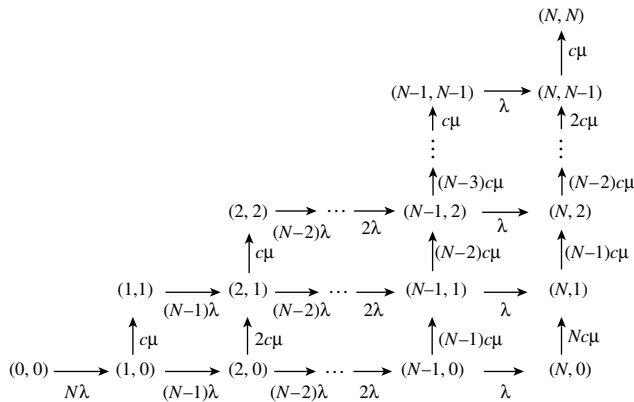
Before we present an analysis of the transient solution of $P_{m,n}^c(i, j, t)$, we would like to point out a useful property of the $(A(t), S(t))$ process: Assume that the process is in state $(0, 0)$ when the counter(s) open at time $t = 0$. Eventually, the process is absorbed in state (N, N) after all passengers arrive and all are served. If we define τ as the time to absorption, then τ is the phase-type random variable associated with the continuous-time Markov chain (CTMC) with $\hat{N} = \frac{1}{2}(N + 1)(N + 2)$ states $(0, 0), (1, 0), \dots, (N, 0) \mid (1, 1), \dots, (N, 1) \mid (2, 2), \dots, (N, 2) \mid \dots \mid (N, N)$, as depicted in Figure 1. Thus, the infinitesimal generator matrix for the resulting CTMC can be written as

$$\mathbf{Q}_{\hat{N} \times \hat{N}} = \left[\begin{array}{c|c} \mathbf{S} & \mathbf{S}_0 \\ \hline \mathbf{0} & 0 \end{array} \right],$$

where \mathbf{S} is the $(\hat{N} - 1) \times (\hat{N} - 1)$ subgenerator matrix and \mathbf{S}_0 is an $(\hat{N} - 1) \times 1$ column vector.

If we further define the $1 \times \hat{N}$ row vector $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_{\hat{N}})$ as the vector of initial probabilities, then for our case $\boldsymbol{\alpha} = (1, 0, \dots, 0)$, assuming the process

Figure 1 Transition Rate Diagram of the Airline Check-In Counter with $c \geq 1$ Counters and N Confirmed Passengers



Notes. The state vector is (m, n) , where m is the number of passengers who have already arrived and n is the number of passengers who have already been served. Both the arrival rates and service rates are state dependent. When the state vector is (m, n) , the arrival rate is $\lambda_{m,n} = (N - m)\lambda$ and the service rate is $\mu_{m,n} = c(m - n)\mu$.

starts at the first state $(0, 0)$. With these definitions, the cumulative distribution function $F_\tau(t) = \Pr(\tau \leq t)$ of the time until absorption is found as (see Neuts 1981, p. 45)

$$F_\tau(t) = 1 - \alpha \exp(\mathbf{S}t)\mathbf{e}, \quad (2)$$

where $\mathbf{e} = (1, \dots, 1)'$ and $\exp(\mathbf{S}t)$ is the matrix exponential function. It is also possible to show (Neuts 1981, p. 46) that the r th moment of τ is

$$E(\tau^r) = (-1)^r r! \alpha (\mathbf{S}^{-1})^r \mathbf{e}, \quad (3)$$

where $(\mathbf{S}^{-1})^r$ is the r th power of the inverse of the \mathbf{S} matrix. Thus, the mean $E(\tau)$ and the variance $\text{Var}(\tau) = E(\tau^2) - [E(\tau)]^2$ can be easily computed after inverting the $(\hat{N} - 1) \times (\hat{N} - 1)$ matrix \mathbf{S} .

As we will see later in §5, the observation that the absorption time from $(0, 0)$ to (N, N) is a phase-type random variable [denoted by $PH(\alpha, \mathbf{S})$] can be useful in the analysis of more general cases where the lifetimes and/or service times are not exponential. For example, by defining either or both of these random variables as Erlang (rather than exponential), the fundamental structure of the process does not change, and we can still calculate the distribution function $F_\tau(t)$ and the moments $E(\tau^r)$, albeit at the expense of dealing with a much larger state space (depending on the number of stages used in the Erlang random variables).

4. Transient Analysis of the System

It is now easy to see that the stochastic process $\{(A(t), S(t)): t \geq 0\}$ is a Markov process. In this section our goal is to find an exact expression for the

time-dependent (transient) probability $P_{m,n}^c(i, j, t)$ for a given (m, n) with $m \leq i \leq N$, $n \leq j \leq i$, and $t \geq 0$. To that end, we establish a system of differential equations in terms of $P_{m,n}^c(i, j, t)$, whose solution would produce the required transient probabilities.

4.1. Transient Distribution of the Process $(A(t), S(t))$

To calculate the conditional probabilities $P_{m,n}^c(i, j, t) = \Pr\{A(t) = i, S(t) = j \mid A(0) = m, S(0) = n\}$, we develop a system of differential equations in terms of the unknown function $P_{m,n}^c(i, j, t)$ for fixed (given) values of (m, n) at time 0.

Case I. $i = m$ and $j = n$.

This implies that there has been neither an arrival nor a departure in the interval $(0, t)$. Hence, the relevant difference equation for $P_{m,n}^c(m, n, \cdot)$ over the infinitesimal interval $(t, t + \Delta t)$ is $P_{m,n}^c(m, n, t + \Delta t) = [1 - (N - m)\lambda\Delta t - (m - n)c\mu\Delta t]P_{m,n}^c(m, n, t) + o(\Delta t)$. Note here that to be in state (m, n) at time $t + \Delta t$ after starting at the same state (m, n) at time t is equivalent to saying that there must have been no arrivals and no departures in the short time interval Δt . There are still $(N - m)$ customers at time t who still have not arrived (with individual arrival rates of λ), and there are already $(m - n)$ customers in the system waiting to be served (by c open counters, each with an individual service rate of μ , totalling $c\mu$). Thus, the probability of no arrivals and no departures in Δt is $[1 - (N - m)\lambda\Delta t - (m - n)c\mu\Delta t + o(\Delta t)]$. Rearranging terms, dividing both sides by Δt and letting $\Delta t \rightarrow 0$ gives, $dP_{m,n}^c(m, n, t)/dt = -[(N - m)\lambda + (m - n)c\mu]P_{m,n}^c(m, n, t)$.

For the remaining cases, we apply the standard steps of rearranging terms, dividing both sides by Δt , and letting $\Delta t \rightarrow 0$, and obtain the following:

Case II. $i = m + 1, m + 2, \dots, N$ and $j = n$.

For this case, more passengers have arrived, but none have left between 0 and t . Thus,

$$dP_{m,n}^c(i, n, t)/dt = -[(N - i)\lambda + (i - n)c\mu]P_{m,n}^c(i, n, t) + (N - i + 1)\lambda P_{m,n}^c(i - 1, n, t).$$

Case III. $i = m + 1, m + 2, \dots, N$ and $j = n + 1, n + 2, \dots, i - 1$.

Here, more passengers have arrived and some have also left between 0 and t , but there still remain some passengers who have not been served. This gives $dP_{m,n}^c(i, j, t)/dt = -[(N - i)\lambda + (i - j)c\mu]P_{m,n}^c(i, j, t) + (N - i + 1)\lambda P_{m,n}^c(i - 1, j, t) + (i - j + 1)c\mu P_{m,n}^c(i, j - 1, t)$.

Case IV. $i = m, m + 1, m + 2, \dots, N$ and $j = i$.

In this case, all passengers who have arrived have left, and there is no one at time t . Thus,

$$dP_{m,n}^c(i, i, t)/dt = -(N - i)\lambda P_{m,n}^c(i, i, t) + c\mu P_{m,n}^c(i, i - 1, t).$$

Case V. $i = m$ and $j = n + 1, n + 2, \dots, i - 1$.

Finally, in this case no more passengers have arrived, some have left, but there still remain some in the system. For this case we have

$$\begin{aligned} dP_{m,n}^c(m, j, t)/dt &= -[(N - m)\lambda + (m - j)c\mu]P_{m,n}^c(m, j, t) \\ &\quad + (m - j + 1)c\mu P_{m,n}^c(m, j - 1, t). \end{aligned}$$

Multiplying both sides of the (i, j) th differential equation $dP_{m,n}^c(i, j, t)/dt$ by $u^i v^j$ and summing over i and j , the above system of differential equations transforms into the following partial differential equation (PDE) in terms of the p.g.f. $\Pi \equiv \Pi_{m,n}(u, v, t) = \sum_{i=m}^N \sum_{j=n}^i P_{m,n}^c(i, j, t) u^i v^j$ with the initial condition $\Pi_{m,n}(u, v, 0) = u^m v^n$:

$$\begin{aligned} \frac{\partial \Pi}{\partial t} + u[\lambda(u - 1) - c\mu(v - 1)] \frac{\partial \Pi}{\partial u} + c\mu v(v - 1) \frac{\partial \Pi}{\partial v} \\ = N\lambda(u - 1)\Pi. \end{aligned} \tag{4}$$

The next theorem gives the exact solution of the PDE for the unknown function $\Pi_{m,n}(u, v, t)$:

THEOREM 1. *The solution of the PDE in (4) is*

$$\begin{aligned} \Pi_{m,n}(u, v, t) = u^m v^n \{ e^{-\lambda t} + [\alpha(c, t) + \beta(c, t)v]u \}^{N-m} \\ \cdot [e^{-c\mu t} + (1 - e^{-c\mu t})v]^{m-n}, \end{aligned} \tag{5}$$

for $0 \leq m \leq N$ and $0 \leq n \leq m$, where

$$\alpha(c, t) \equiv \frac{\lambda}{\lambda - c\mu} (e^{-c\mu t} - e^{-\lambda t}), \quad c \geq 1, t \geq 0, \tag{6}$$

$$\begin{aligned} \beta(c, t) \equiv 1 + \frac{c\mu e^{-\lambda t} - \lambda e^{-c\mu t}}{\lambda - c\mu} \\ = (1 - e^{-\lambda t}) - \alpha(c, t), \quad c \geq 1, t \geq 0, \end{aligned} \tag{7}$$

with $\lim_{\lambda \rightarrow c\mu} \alpha(c, t) = c\mu t e^{-c\mu t}$ and $\lim_{\lambda \rightarrow c\mu} \beta(c, t) = e^{-c\mu t} (e^{c\mu t} - 1 - c\mu t)$.

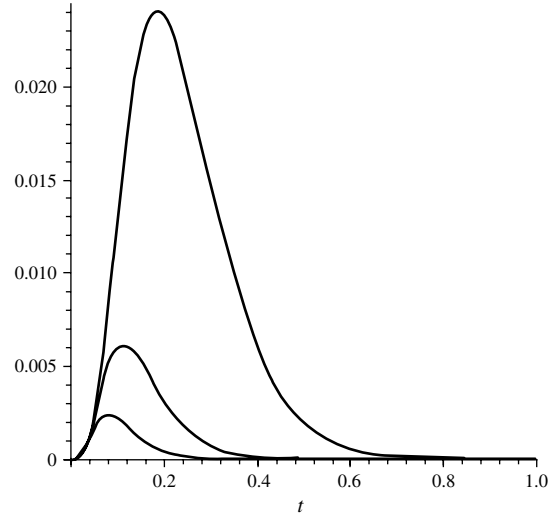
PROOF. See Online Appendix A. \square

Because the exact form of the p.g.f. $\Pi_{m,n}(u, v, t)$ is available, it can now be expanded to find the exact form of the distribution of the process $\{A(t), S(t)\}$. This result is provided in the next theorem.

THEOREM 2. *The transient probability distribution $P_{m,n}^c(i, j, t) = \Pr\{A(t) = i, S(t) = j \mid A(0) = m, S(0) = n\}$ of $\{A(t), S(t)\}$ for $m \leq i \leq N$ and $n \leq j \leq i$, is given as*

$$\begin{aligned} P_{m,n}^c(i, j, t) &= \binom{N-m}{i-m} e^{-(N-i)\lambda t} \\ &\cdot \sum_{r=0}^{\max(i-m, m-n)} \binom{i-m}{r} [\alpha(c, t)]^{i-m-r} [\beta(c, t)]^r \\ &\times \binom{m-n}{j-n-r} e^{-(m-j+r)c\mu t} (1 - e^{-c\mu t})^{j-n-r}, \end{aligned} \tag{8}$$

Figure 2 Transient Probability $P_{m,n}^c(i, j, t)$ for $N = 10, c = 1, 2, 3$ with $(m, n) = (4, 2), (i, j) = (7, 3)$, and $(\lambda, \mu) = (1.5, 5)$; the Graph with $c = 1$ Is the Highest One



with the convention that a binomial coefficient $\binom{a}{b} = 0$, if either $b > a$ or $b < 0$.

PROOF. See Online Appendix B. \square

Note that the transient probability $P_{m,n}^c(i, j, t)$ is a function of the decision variable c . This result will be useful when we formulate the functional equations of DP to determine the optimal number of counters to open. See Figure 2 for the graph of $P_{m,n}^c(i, j, t)$ for $N = 10, c = 1, 2, 3, (m, n) = (4, 2), (i, j) = (7, 3)$, and $(\lambda, \mu) = (1.5, 5)$. From the figure we note that the transition probability as a function of t is unimodal.

4.2. Transient Distribution of the Number in the System $Y(t)$

To compute the expected cost incurred for the passengers waiting to be served, we need to find the transient distribution of the number in the system $Y(t) = A(t) - S(t)$. Because the p.g.f. $\Pi_{m,n}(u, v, t) = E[u^{A(t)} v^{S(t)} \mid A(0) = m, S(0) = n]$ of the vector stochastic process $\{A(t), S(t)\}$ is available, the p.g.f. $\hat{\Pi}_{m,n}(u, t)$ of $Y(t)$ is obtained simply as

$$\begin{aligned} \hat{\Pi}_{m,n}(u, t) &= E[u^{A(t)-S(t)} \mid A(0) = m, S(0) = n] = \Pi_{m,n}(u, u^{-1}, t) \\ &= \{ [1 - \alpha(c, t)] + \alpha(c, t)u \}^{N-m} [(1 - e^{-c\mu t}) + e^{-c\mu t}u]^{m-n}. \end{aligned}$$

The next theorem gives the exact form of the distribution $Q_{m,n}(k, t) = \Pr\{Y(t) = k \mid A(0) = m, S(0) = n\}$ of the $Y(t)$ process where we define

$$\begin{aligned} \gamma_{m,n}(i, k, t) &\equiv [\alpha(c, t)]^{k-i} [1 - \alpha(c, t)]^{N-m-(k-i)} e^{-ic\mu t} [1 - e^{-c\mu t}]^{m-n-i}. \end{aligned}$$

THEOREM 3. *The distribution of $Y(t)$ is found as follows:*

For $m \leq (N + n)/2$, we have

$$Q_{m,n}(k, t) = \begin{cases} \sum_{i=0}^k \binom{N-m}{k-i} \binom{m-n}{i} \gamma_{m,n}(i, k, t), & \text{for } k \leq m-n-1 \\ \sum_{i=0}^{m-n} \binom{N-m}{k-i} \binom{m-n}{i} \gamma_{m,n}(i, k, t), & \text{for } m-n \leq k \leq N-m \\ \sum_{i=k-(N-m)}^{m-n} \binom{N-m}{k-i} \binom{m-n}{i} \gamma_{m,n}(i, k, t), & \text{for } N-m+1 \leq k \leq N-n, \end{cases}$$

and for $m > (N + n)/2$, we obtain

$$Q_{m,n}(k, t) = \begin{cases} \sum_{i=0}^k \binom{N-m}{k-i} \binom{m-n}{i} \gamma_{m,n}(i, k, t), & \text{for } k \leq N-m-1 \\ \sum_{i=0}^{N-m} \binom{N-m}{k-i} \binom{m-n}{i} \gamma_{m,n}(i, k, t), & \text{for } N-m \leq k \leq m-n \\ \sum_{i=k-(N-m)}^{N-m} \binom{N-m}{k-i} \binom{m-n}{i} \gamma_{m,n}(i, k, t), & \text{for } m-n+1 \leq k \leq N-n. \end{cases}$$

PROOF. Follows from the observation that the distribution of $Y(t)$ is the convolution of two binomial distributions. \square

4.3. Expected Number of Passengers in the System

Because the p.g.f. of $Y(t)$ is available, it is now easy to derive all the moments of $Y(t)$. For example, we can obtain the expected number in the system given that at time 0 a total of $m - n$ passengers are in the system as $E[Y(t) | Y(0) = m - n] = (d/du) \hat{\Pi}_{m,n}(u, t)|_{u=1} = (N - m)\alpha(c, t) + (m - n)e^{-c\mu t}$.

The expected number in the system is a function of several parameters, but let us concentrate on two important parameters (c, t) and write the expected number as $f(c, t) \equiv E[Y(t) | Y(0) = m - n]$. The following lemma and the corollary provide some insights into the behavior of $f(c, t)$ as a function of time t when the number of counters is fixed.

LEMMA 1. *For a fixed $c = \bar{c}$, the $\alpha(\bar{c}, t)$ function is unimodal.*

PROOF. See Online Appendix C. \square

COROLLARY 1. *Consider a fixed $c = \bar{c}$. If $\lambda(N - m) > b(m - n)$, the expected number in the system $f(t) \equiv f(\bar{c}, t)$ is a unimodal function of time t with a unique maximizer. Otherwise, $f(t)$ is monotone decreasing in time t .*

PROOF. See Online Appendix D. \square

We see that $\lambda(N - m)$ is the instantaneous effective arrival rate at time t and $b(m - n)$ is the instantaneous departure rate at time t . Therefore, when $\lambda(N - m) > b(m - n)$, the instantaneous arrival rate exceeds the instantaneous departure rate. Hence, Corollary 1 implies that in such a situation we can expect an accumulation of passengers (a “bunching up”) at some later time. Otherwise, the expected number in the system will decline as time progresses.

The following lemma and the corollary provide insights into the behavior of $f(c, t)$ as a function of c when the time t is fixed where we assume, for analytical ease, that c is a continuous variable.

LEMMA 2. *For a fixed $t = \bar{t}$ and $\lambda \neq c\mu$, the $\alpha(c, \bar{t})$ function is a monotone-decreasing function of c .*

PROOF. See Online Appendix E. \square

Note that in the analysis presented in Lemma 2, we examined the behavior of $\alpha(c)$ as a function c , assuming that the parameters (λ, μ) are fixed and $t = \bar{t}$. We do not consider here the limiting case where $\lambda \rightarrow c\mu$ because it is not possible to fix λ and vary c at the same time.

COROLLARY 2. *Consider a fixed $t = \bar{t}$ and $\lambda \neq c\mu$. The expected number in the system $f(c) \equiv f(c, \bar{t})$ is a monotone-decreasing function of the number of counters c .*

PROOF. See Online Appendix F. \square

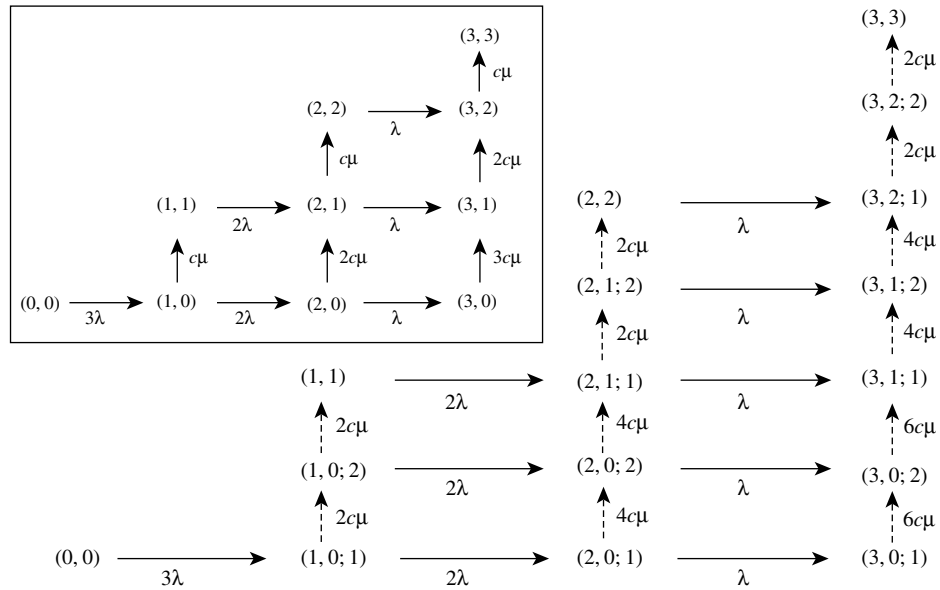
If there is a service-level requirement that the expected number in the system should not exceed a given value, say, η , at time $t = \bar{t}$, then the unique value of c that satisfies this requirement can be computed by solving the transcendental equation $f(c, \bar{t}) = \eta$.

5. A More General Process with Erlangian Service Times

In our previous discussion we have always assumed that the lifetimes and the service times are both exponential, and with this assumption we derived analytic expressions for quantities of interest (the most important being the transient solution of the conditional probability $P_{m,n}^c(i, j, t)$). In this section we present a brief discussion of a more general case where the service times are Erlang random variables (r.v.) with k stages. (The Erlang r.v. with parameters (k, μ) is constructed as the sum of k i.i.d. exponential random variables, each with rate $k\mu$ and mean $1/(k\mu)$.) The generality we introduce can also be extended to the case where the lifetimes are also Erlang, but we do not pursue this direction in the present work.

Consider first a simple example with $N = 3$, where all service times are still exponential as in the original problem. For this case, the transition rate diagram given in Figure 1 for N confirmed passengers reduces to the diagram in the upper-left corner of Figure 3.

Figure 3 Transition Rate Diagrams for $N = 3$



Notes. The smaller diagram in the upper-left corner corresponds to the simple case where all lifetimes and service times are exponential. The larger diagram at the lower-right corner corresponds to the more general case where the service times are Erlang with two stages (but with the same mean as the exponential).

With $N = 3$, the corresponding CTMC possesses a total of $\hat{N} = \frac{1}{2}(4)(5) = 10$ states for which the 9×9 subgenerator matrix S is obtained as

$$S_{9 \times 9} = \begin{matrix} (0,0) \\ (1,0) \\ (2,0) \\ (3,0) \\ (1,1) \\ (2,1) \\ (3,1) \\ (2,2) \\ (3,2) \end{matrix} \begin{bmatrix} -3\lambda & 3\lambda & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -(2\lambda + c\mu) & 2\lambda & 0 & c\mu & 0 & 0 & 0 & 0 \\ 0 & 0 & -(\lambda + 2c\mu) & \lambda & 0 & 2c\mu & 0 & 0 & 0 \\ 0 & 0 & 0 & -3c\mu & 3c\mu & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -2\lambda & 2\lambda & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -(\lambda + c\mu) & \lambda & c\mu & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -2c\mu & 0 & 2c\mu \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\lambda & \lambda \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -c\mu \end{bmatrix}$$

For general (λ, μ, c) values, we were able to invert the S matrix symbolically and obtain the mean of τ as

$$E(\tau) = \frac{22\lambda^4 + 101c\mu\lambda^3 + 143c^2\mu^2\lambda^2 + 89c^3\mu^3\lambda + 22c^4\mu^4}{6c\lambda\mu(2\lambda + c\mu)(\lambda + 2c\mu)(\lambda + c\mu)} \tag{9}$$

We have also obtained a closed-form expression for the variance, but due to its complicated form, we do not display it here.

If we now assume, for example, that $(\lambda, \mu, c) = (1, 5, 1)$, we find the mean and the standard deviation of τ from (3) as $E(\tau) = 2.09$ and $\sqrt{\text{Var}(\tau)} = 1.169$, respectively. (The mean of τ can be computed more easily from (9).) We also find the exact distribution, $F_\tau(t)$, of the time-to-absorption random variable τ from (2) as

$$F_\tau(t) = 1 + 0.0003e^{-15t} + 0.273e^{-11t} - 0.616e^{-10t} + 1.757e^{-7t} - 1.458e^{-6t} - 0.187e^{-5t} - 1.813e^{-3t} + 4.927e^{-2t} - 3.883e^{-t},$$

which is increasing in t , $F_\tau(0) = 0$, and $\lim_{t \rightarrow \infty} F_\tau(t) = 1$. Note that $F_\tau(1) = 0.144$, $F_\tau(2) = 0.560$, $F_\tau(3) = 0.818$, and $F_\tau(4) = 0.930$; that is, even though the mean time to absorption is 2.09 (hours), there is a 0.144 probability that actual time may be shorter than one (hour). If the counter opens three hours before the flight, then the probability is only 0.818 that all passengers will arrive and will be served within three hours.

We now pose the following questions: (i) What if the service times were in fact all Erlang random variables with several stages, rather than exponential as discussed above (both exponential and Erlang service times having the same means)? (ii) How does this change the problem's structure, and what are the percentage errors for the mean and the standard deviation of the time-to-absorption r.v. τ when the service times are assumed (possibly incorrectly) as exponential rather than Erlang?

To answer these questions on a limited scale, first assume that the service times are all Erlang with two stages. Because each stage is exponentially distributed, this assumption results in the transition rate diagram that is in the lower-right corner of Figure 3. Each service time is now represented by the sum of two i.i.d. exponentials; hence, the state space enlarges and we obtain a CTMC with a total of 16 states. For example, state $(1, 0; 1)$ represents the first stage of the service time of one passenger who has already arrived, but whose service time is not yet completed. Similarly, $(1, 0; 2)$ represents the second stage of the same passenger's service time, which has not yet been completed. Naturally, this idea can be used to incorporate Erlang random variables for any lifetime

and/or service time, but at the expense of quickly enlarging the state space. It is important to note that because the enlarged state space still continues to be the state space of a CTMC, it is possible, in principle, to develop the system of differential equations as was done in §4 for the Erlang service-time case and compute the transient probabilities—at least numerically (which is beyond the scope of the present paper). For this model, again using $(\lambda, \mu, c) = (1, 5, 1)$ as the base values of the parameters, we find the mean and standard deviation of the absorption time r.v. τ as $E(\tau) = 2.049$ and $\sqrt{\text{Var}(\tau)} = 1.163$, respectively. The exact distribution, $G_\tau(t)$, of τ is also computed from (2) as

$$G_\tau(t) = 1 + 0.04te^{-30t} + 0.01e^{-30t} - 14.35te^{-21t} - 34.40e^{-21t} - 20.74te^{-20t} + 34.25e^{-20t} + 180.04te^{-12t} + 627.78e^{-12t} + 1,074.07te^{-11t} + 358.12e^{-11t} + 358.97te^{-10t} - 985.78e^{-10t} - 1.79e^{-3t} + 4.50e^{-2t} - 3.70e^{-t},$$

which is increasing in t , and $G_\tau(0) = 0$, $\lim_{t \rightarrow \infty} G_\tau(t) = 1$ with $G_\tau(1) = 0.157$, $G_\tau(2) = 0.576$, $G_\tau(3) = 0.826$, and $G_\tau(4) = 0.933$. When we compare these results to the $F_\tau(t)$ values given above, we observe that if indeed the service times were Erlangian, then using the exponential model would result in more “pessimistic” completion (i.e., absorption) times because $F_\tau(t) < G_\tau(t)$ for all t considered. (In fact, for the case under consideration, the same property holds true for all $t \geq 0$.)

To answer the second question, we have conducted some experiments and calculated the mean and the standard deviation of the absorption time r.v. τ for different values of the parameters (λ, μ, c) for both the exponential and Erlang service times. Starting with the base values of $(\lambda, \mu, c) = (1, 5, 1)$, we kept two of these parameters constant and varied the third one to obtain the results in Table 2. For example, we calculate the error for the expected value $E(\tau)$ for the base case as $100 \cdot [(2.090 - 2.049)/2.049]\% = 2.00\%$.

These results show that percentage errors increase and they can be as large as 10% (or more) as λ approaches μ , i.e., when congestion increases. On the other hand, as c increases, errors for both the mean and the standard deviation diminish and approach zero. Even though this is a very limited experiment, the movement and magnitude of these errors are likely to be mirrored in the performance of an optimization model that determines the time-dependent allocation of the check-in counters. Thus, if the management predicts a highly congested system, it may be worthwhile to use a more accurate model that may involve Erlangian service (and/or lifetime) distributions.

In this section we considered a generalization where the service times at each state depicted in

Table 2 Errors Resulting from Using the Exponential Rather Than the Erlang Service Times in the Calculation of the Mean and Standard Deviation of the Absorption Time Random Variable τ

λ	μ	c	$E(\tau)$			$\sqrt{\text{Var}(\tau)}$		
			Exponential	Erlang	Error (%)	Exponential	Erlang	Error (%)
1	5	1	2.090	2.049	2.00	1.169	1.163	0.51
2	5	1	1.205	1.148	4.96	0.605	0.580	4.31
3	5	1	0.919	0.857	7.23	0.433	0.391	10.74
1	2	1	2.583	2.431	6.25	1.250	1.164	7.38
1	3	1	2.300	2.211	4.02	1.191	1.160	2.67
1	4	1	2.166	2.108	2.75	1.175	1.162	1.11
1	5	2	1.950	1.937	0.67	1.1655	1.1656	0.001
1	5	3	1.907	1.901	0.31	1.1657	1.1661	0.034
1	5	4	1.887	1.884	0.16	1.1660	1.1663	0.025

Figure 1 was assumed an Erlang r.v. with two stages giving rise to the transition rate diagram in Figure 3 for the case of $N = 3$ confirmed passengers. Naturally, the service times can be generalized even further (at the expense of increasing the number of states of the CTMC) by using an arbitrary number of stages for each Erlang r.v. It may even be possible to assume that the lifetimes are Erlang with any number of stages, say, l . However, such a generalization would probably require a computational approach (rather than an analytic approach as used in this paper) due to the resulting high dimensionality.

6. Estimation of Parameters

So far we have just analyzed only a basic model of a check-in counter system with the assumption that the arrival process is a death process with constant death or “show-up” rate. However, in real operations the arrival process is actually nonstationary, i.e., the show-up rate is actually time dependent. In the airport the authors visited, each flight had its own idiosyncrasies with respect to arrivals and service. It should also be so in other airports around the world. For some flights, passengers tend to arrive very early, and for some others—especially for very early morning flights—they tend to arrive late. Thus, the pattern of arrivals varies widely across flights (see, for example, Chun and Mak 1999). For the dynamic programming model to be analyzed in the next section, the nonstationary show-up rates are needed as inputs. Therefore, for each flight we would need to know the passenger arrival profile. Using this profile, we can divide the interval T into smaller subintervals over which the show-up rate λ is constant. Thus, we need to estimate the show-up rates for each of these subintervals. For our work, we use the procedure given in Basawa and Prakasa Rao (1980)

(see also Keiding 1975), which we briefly describe below:

Consider an arbitrary subinterval k of length T_k and observe the show-ups continuously at $t \in [\tau_0, \tau_0 + T_k]$, where $\tau_0 \equiv \tau_{k0}$ is the start of the subinterval. Assume that at the start of this subinterval there is a total of $x_0 \equiv x_{k0}$ passengers who have not yet arrived at the airport. Let $\tau_1 < \tau_2 < \dots < \tau_n$ be the epochs during this subinterval where $n \equiv n_k$ transitions (deaths, or removals) occur.

We form the likelihood function (that is, the joint density function of the sample observations) as

$$L_k(\lambda_k) = [x_0 \lambda_k \exp(-x_0 \lambda_k (\tau_1 - \tau_0))] \cdot [x_1 \lambda_k \exp(-x_1 \lambda_k (\tau_2 - \tau_1))] \cdots [x_{n-1} \lambda_k \exp(-x_{n-1} \lambda_k (\tau_n - \tau_{n-1}))] \cdot [\exp(-x_n \lambda_k (T_k - \tau_n))].$$

Here, $x_{i+1} = x_i - 1$ is the number of passengers not yet arrived at τ_{i+1} (for $i = 0, 1, \dots, n-1$) with $x_0 \equiv x_{k0}$, and the last term $\exp(-x_n \lambda_k (T_k - \tau_n))$ represents the contribution of the last interval $T_k - \tau_n$ during which no deaths occur. Defining $V_{T_k} \equiv \sum_{i=0}^{n-1} x_i (\tau_{i+1} - \tau_i) + x_n (T_k - \tau_n)$ as the total time lived before “death,” we have $L_k(\lambda_k) = \lambda_k^n \exp(-\lambda_k V_{T_k}) \prod_{i=0}^{n-1} x_i$. The usual steps of maximizing the log-likelihood function $\log L_k(\lambda_k)$ gives the maximum-likelihood estimator of the parameter as $\hat{\lambda}_k = n_k / V_{T_k}$.

In the dynamic programming model we consider in the next section, we assume that the arrival times of passengers, which are used to estimate the arrival rates, have been observed. To illustrate, consider a simple example of a small airplane with $N = 15$ booked passengers. We take $T = 3$ hours and assume that the historic passenger arrival profile reveals that the λ s are constant over intervals of one-hour duration. Hence, to estimate the λ s, the management observes and groups the arrival instants for every $\bar{T} = 1$ hour. Suppose the arrival times $[\tau_1 < \dots < \tau_n]$ of passengers are observed as in the third column of the following table. Using $\hat{\lambda}_k = n_k / V_{T_k}$, the estimates $\hat{\lambda}$ are computed as in the last column of the table.

Period k	No. of passengers		Arrival times [$\tau_1 < \dots < \tau_n$] (hrs.)	$\hat{\lambda}_k$
	arrived (n_k)			
1	4		[0.32, 0.34, 0.42, 0.47]	0.31
2	6		[1.15, 1.46, 1.47, 1.58, 1.93, 1.96]	0.69
3	5		[2.11, 2.44, 2.57, 2.71, 2.87]	1.83

7. The Dynamic Programming Model

Having obtained the exact expressions for the transient probabilities $P_{m,n}^c(i, j, t)$ and the expected number of passengers in the system $E[Y(t)]$, we can optimize the system by deciding on the number of counters to open. We highlight here that our work is

the first paper to consider optimization of the check-in counter system based on cost, because almost all the papers use only the service-level approach. Although service-level constraints can be included in our model, we have chosen optimization based on cost only because most of the time cost is the main concern.

7.1. Total Expected Cost in a Subinterval

We assume that the management will be observing the system periodically, say, every 20 minutes, and will decide to increase or decrease the number of counters to keep open in order to minimize an expected total cost defined as the sum of the cost of waiting passengers plus the cost of keeping the counters open. In the airport the authors visited, the service provider has supervisory staff who continuously monitor the counter congestion and acts to minimize the congestion by opening more counters. The service provider also had commissioned some mobile counters to ease the congestion at the counters.

To find the optimal solution, we will construct a stochastic dynamic programming model. To that end, consider a subinterval $[t_k, t_{k+1}]$. Assume that at time t_k , we observe a total of m_k passengers to have arrived and n_k to have departed. Let c_k be the number of counters to keep open at t_k . The total expected wait during this subinterval in *passenger-hours* is

$$W_k(c_k) = E \left[\int_{t_k}^{t_{k+1}} Y(s) ds \mid A(t_k) = m_k, S(t_k) = n_k \right] \\ = \int_{t_k}^{t_{k+1}} E[Y(s) \mid A(t_k) = m_k, S(t_k) = n_k] ds \\ = \int_{t_k}^{t_{k+1}} [(N - m_k)\alpha(c_k, s) + (m_k - n_k)e^{-c_k \mu s}] ds \\ = \left[\frac{\lambda(N - m_k)}{\lambda - c_k \mu} + (m_k - n_k) \right] \frac{(e^{-c_k \mu t_k} - e^{-c_k \mu t_{k+1}})}{c_k \mu} \\ + \frac{(N - m_k)}{\lambda - c_k \mu} (e^{-\lambda t_{k+1}} - e^{-\lambda t_k}).$$

If c_k counters are open during the subinterval $[t_k, t_{k+1}]$, then we incur a (deterministic) service cost of $C_s(t_{k+1} - t_k)c_k$ where C_s is the unit cost of operating a counter. Defining C_w as the unit cost of making a passenger wait, the total expected cost incurred during the subinterval $[t_k, t_{k+1}]$ with c_k counters are open is $g_k(c_k) = g_k(c_k, m_k, n_k) = C_w W_k(c_k) + C_s(t_{k+1} - t_k)c_k$, which consists of the cost of waiting passengers (first term) and the cost of operating the counters (second term).

7.2. Functional Equations

To determine the optimal time-dependent policy for the number of counters to keep open, we now formulate a stochastic dynamic programming model. We

Downloaded from informs.org by [202.161.43.77] on 06 January 2016, at 05:57. For personal use only, all rights reserved.

first divide the total time T into K subintervals of equal length, i.e., $[t_1, t_2], [t_2, t_3], \dots, [t_k, t_{k+1}]$ where $t_1 = 0$ and $t_{K+1} = T$ and write the DP functional equation as follows:

Define $V_k(m, n)$ as the minimum expected cost-to-go from time t_k , i.e., the beginning of the subinterval $[t_k, t_{k+1}]$, to the final time T using the optimal policy when m passengers have arrived and n have been served at time t_k . This gives, for $k = 1, \dots, K$,

$$V_k(m, n) = \min_{\substack{c_k^{\min} \leq c_k \leq c_k^{\max} \\ c_k(N, N)=0, V_k(N, N)=0}} \left[g_k(c_k) + \sum_{i=m}^N \sum_{j=n}^i P_{m,n}^c(i, j, t_{k+1} - t_k) V_{k+1}(i, j) \right], \quad (10)$$

where c_k^{\min} is the minimum number of counters that must be kept open before the flight (usually one), and c_k^{\max} is the maximum number of counters that can be made available for the flight. The conditional transient probability $P_{m,n}^c(i, j, t_{k+1} - t_k)$ appearing in (10) is calculated from (8).

Note that the minimization is performed with $c_k(N, N) = 0$ and $V_k(N, N) = 0$, because we assume that if all customers have arrived and all have been served at some time t_k , then we close all counters and set the value function to zero because no costs will be incurred after t_k .

To establish the boundary condition, we assume that there is a unit variable cost of h resulting from uncompleted passenger service. In other words, for each passenger who has arrived but whose service is not yet completed by the time the flight is to take off, the system incurs a cost of h dollars. Thus, the boundary condition is written as $V_{K+1}(m, n) = h(m - n)$.

7.3. Numerical Example

Consider now an example with the following data:

T	K	N	c_k^{\min}	c_k^{\max}	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\lambda}_3$	μ	C_w	C_s	h
1	3	10	1	5	0.58	1.60	2.74	5	40	60	20

Thus, in this example, a maximum of $c_k^{\max} = 5$ counters are available for $T = 1$ hour before the flight, and the management decides $K = 3$ times (i.e., every 20 minutes) on the number of counters to keep open, with the stipulation that there must be at least $c_k^{\min} = 1$ counter open at all times. Thus, the decisions must be made at epochs $t_1 = 0$ hours, $t_2 = \frac{1}{3}$ hours, and $t_3 = \frac{2}{3}$ hours to determine the optimal number of counters to open— $c_1^*(m, n)$, $c_2^*(m, n)$, and $c_3^*(m, n)$, during $[0, \frac{1}{3}]$, $[\frac{1}{3}, \frac{2}{3}]$, and $[\frac{2}{3}, 1]$, respectively, if the system is found in state (m, n) at t_k , $k = 1, 2, 3$. The estimates of the arrival-rate parameters $(\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3)$ are obtained using the procedure described in §6.

To obtain our cost estimates, we performed an Internet search and used values for (C_w, C_s, h) that are very close to these estimates available in the public domain. We point out that our intention here is just to highlight the usability of our model rather than the use of very realistic estimates as input into our models.

7.3.1. Passenger Delay Cost C_w . To estimate this cost we consulted three sources: European Organization for the Safety of Air Navigation (EOSAN 2005), Federal Aviation Administration (2003), and Smith (2004). Whereas EOSAN’s estimates range from €38 to €49 per hour,⁵ Smith gives an estimate of €18. The FAA-APO estimate for personal travel is \$23.30, for business travel is \$40.10, and for all purposes it is \$28.60 per hour. We have used $C_w = \$40$, which is very close to the estimate recommended by Federal Aviation Administration for business class passengers.

7.3.2. Check-In Counter Operating Cost C_s . There does not appear to be much information available on this cost because we could locate only one reference, Aéroport International Strasbourg (2006), which provides costs, among others, for the following two cases: (i) For rental of check-in counters, the cost is estimated as €11.68 per hour and, (ii) for services of airport personnel the estimates range from €25 per hour (for a maintenance agent) to €70 per hour (for a project manager). In the international airport we visited, it is the service provider who also provides the counter staff. This is unlike other airports, where the check-in counters are managed by the airline’s staff. Therefore, in our work the check-in counter operating cost includes the service of a highly qualified worker. In addition to these cost components there are other overhead charges such as use of the software, telephones, washrooms, etc. In all, we have used an estimate of $C_s = \$60$ in our numerical example.

7.3.3. Aircraft Delay Cost h . Here, h is the cost incurred by the airline for every arrived passenger who has not cleared check-in by the time the counter is expected to close. This will most likely result in a delay to the departure of the aircraft. In the airlines operations literature there have been many studies on airline schedule disruptions (see, for example, Shavell 2000). However, these papers consider mainly delays and diversions due to major reasons like weather, etc. Some of the cost components in the delay cost are (Shavell 2000) additional fuel, crew time, maintenance, passenger costs like meals and accommodation, and/or payments to other airlines. Only one document, EOSAN (2005), provides cost estimates due to delays in air transportation. It provides a

⁵ On November 11, 2007, the exchange rate was €1 = \$1.46.

Table 3 Optimal Number of Counters to Open $c_k^*(m, n)$ for Any State (m, n) at $k = 1, 2, 3$

c_1^*, c_2^*, c_3^*	n											
	0	1	2	3	4	5	6	7	8	9	10	
$m = 0$	1, 1, 4											
$m = 1$	1, 1, 4	1, 1, 4										
$m = 2$	1, 1, 4	1, 1, 4	1, 1, 4									
$m = 3$	1, 1, 3	1, 1, 3	1, 1, 3	1, 1, 3								
$m = 4$	1, 1, 3	1, 1, 3	1, 1, 3	1, 1, 3	1, 1, 3							
$m = 5$	1, 1, 3	1, 1, 3	1, 1, 3	1, 1, 3	1, 1, 3	1, 1, 3						
$m = 6$	1, 1, 3	1, 1, 3	1, 1, 3	1, 1, 3	1, 1, 3	1, 1, 3	1, 1, 3					
$m = 7$	2, 2, 3	1, 1, 3	1, 1, 3	1, 1, 3	1, 1, 3	1, 1, 3	1, 1, 2	1, 1, 2				
$m = 8$	2, 2, 3	2, 2, 3	1, 1, 3	1, 1, 3	1, 1, 3	1, 1, 2	1, 1, 2	1, 1, 2	1, 1, 2			
$m = 9$	2, 2, 3	2, 2, 3	2, 2, 3	2, 2, 3	1, 1, 2	1, 1, 2	1, 1, 2	1, 1, 2	1, 1, 2	1, 1, 1		
$m = 10$	2, 2, 3	2, 2, 3	2, 2, 3	2, 2, 3	2, 2, 3	2, 2, 2	2, 2, 2	1, 1, 2	1, 1, 2	1, 1, 1	0, 0, 0	

list of sources and their corresponding delay costs per minute. However, for our purposes, we need the delay cost per passenger which can be obtained by dividing the average duration of delay by the number of passengers still to be served. In our numerical example, we used an estimate of $h = \$20$ per passenger.

Now, implementing the DP algorithm with the above parameter values, we find the optimal policy as in Table 3, which gives the optimal number of counters to keep open during the one-hour interval, i.e., $c_k^*(m, n)$ for any combination of the state variables (m, n) at $t_k, k = 1, 2, 3$, respectively:

Note in Table 3 the interesting (and expected) monotonicity in the optimal policy: For any fixed value of (\bar{m}, \bar{n}) , the number of counters to open is nondecreasing as we approach the time to takeoff. In other words, we have for fixed $(\bar{m}, \bar{n}), c_k(\bar{m}, \bar{n}) \leq c_{k+1}(\bar{m}, \bar{n}), k = 1, 2, \dots, K - 1$. We also note that for fixed (\bar{k}, \bar{n}) , the number of counters to keep open is also nondecreasing in the number of passengers who have already arrived, m , i.e., $c_k(m, \bar{n}) \leq c_k(m + 1, \bar{n}), m = 0, \dots, N - 1$. Finally, for fixed (\bar{k}, \bar{m}) , the number of counters to open is nonincreasing in the number of passengers who have been serviced, i.e., $c_k(\bar{m}, n) \geq c_k(\bar{m}, n + 1), n = 0, \dots, N - 1$. This is also an intuitive result—if there are fewer passengers who are yet to arrive, one should keep open fewer counters.

It is interesting to note that in this example $c_1^* = c_2^* = 1$ when $m - n$ is small, e.g., $m - n = 1$ or 2. This means that it is optimal to open only one counter in the early stages ($k = 1, 2$) when there are few passengers in the system. However, when we reach the last subinterval ($k = 3$), we find that c_3^* may exceed one even if $m - n$ is small. This happens when m itself is small (i.e., very few passengers have arrived so far), so that a large number of passengers are expected to arrive in the last stage (subinterval). Thus, it makes sense to open more than one counter in the last stage to accommodate the large influx of passengers that will arrive just before the plane takes off.

In Table 4 we find the values assumed by the value function $V_k(m, n)$ for any $(m, n), k = 1, 2, 3$.

As in the optimal policy, we also observe in Table 4 that the value functions are monotonic. For example, for a fixed (\bar{k}, \bar{n}) , the value function is decreasing in m , that is, $V_k(m, \bar{n}) > V_k(m + 1, \bar{n}), m = 0, \dots, N - 1$. Similarly, for a fixed (\bar{k}, \bar{m}) , the value function is also decreasing in m , i.e., $V_k(\bar{m}, n) > V_k(\bar{m}, n + 1), n = 0, \dots, N - 1$. We should note that these intuitive monotonicity properties are also observed in other problems with more passengers (larger values of N) and more frequent observations (larger values of K).

8. Conclusion

In this paper we examine the problem of optimal dynamic assignment of check-in counters for a flight with a given number of booked passengers. Because the counters are kept open for a period of only a few hours before the flight, we compute the transient probabilities of the queueing process. Using these probabilities, we calculate some important operating characteristics of the system including the expected number of passengers in system and the probability of an empty system at any time t . Because the arrival rate of the passengers is usually time dependent, we present a procedure to estimate these rates over different subintervals. We then develop a stochastic dynamic programming model to find the optimal dynamic assignment of the counters to minimize a suitable expected cost function. The paper concludes with the discussion of a numerical example and the resulting monotonicity properties of the optimal policy and the corresponding value function.

As mentioned in the introduction, in this paper we have analyzed only the single-flight problem. The natural next step is to apply this analysis for all the flights scheduled to depart during a day or during a particular period of time to identify the number of counters and the number of counter clerks to allocate to each flight. Then, using the constraint on the total

Downloaded from informs.org by [202.161.43.77] on 06 January 2016, at 05:57. For personal use only, all rights reserved.

Table 4 Value Function $V_k(m, n)$ for Any State (m, n) at $k = 1, 2, 3$

V_1, V_2, V_3	n											
	0	1	2	3	4	5	6	7	8	9	10	
$m = 0$	172, 157 145											
$m = 1$	171, 151 139	164, 149 138										
$m = 2$	170, 144 133	163, 143 132	156, 141 132									
$m = 3$	168, 137 126	162, 136 125	155, 135 124	172, 133 123								
$m = 4$	167, 131 117	160, 129 116	153, 128 116	146, 126 115	139, 125 114							
$m = 5$	164, 123 109	157, 122 108	150, 120 107	143, 118 106	137, 117 106	130, 115 105						
$m = 6$	160, 115 100	154, 113 99	147, 112 99	140, 110 98	133, 108 97	126, 106 97	119, 104 96					
$m = 7$	154, 106 92	149, 104 91	142, 102 90	135, 101 90	128, 99 89	121, 97 88	114, 94 86	107, 92 82				
$m = 8$	143, 96 83	139, 94 82	135, 92 82	128, 90 81	121, 88 80	113, 86 79	106, 83 75	99, 81 72	92, 78 68			
$m = 9$	126, 80 75	122, 79 74	117, 78 73	113, 77 73	108, 75 72	101, 72 68	93, 69 65	85, 65 61	78, 61 57	70, 57 47		
$m = 10$	87, 53 66	83, 52 66	78, 51 65	73, 49 64	68, 48 61	64, 47 57	59, 45 54	53, 42 50	43, 36 47	31, 28 39	0, 0 0	

number of counters in the airport and the number of counter clerks available for each period, the resource allocation problem for the entire airport for a specific periods can be solved as shown in Chun (1995) and van Dijk and van der Sluis (2006). We consider this step to be pretty straightforward. Another extension is to consider other types of counter topology, like common-use counters. We conjecture this to be a much simpler problem because it is possible here to assume the arrival process to be Poisson because the arrivals are superpositions of several point processes.

9. Electronic Companion

An electronic companion to this paper is available as part of the online version that can be found at <http://mansci.journal.informs.org/>.

Acknowledgments

The authors thank two anonymous referees and an associate editor for their insightful comments, which helped to improve this paper. They also acknowledge with thanks the discussions they had with Mark Goh, Brian Rodrigues, and Chung-Piaw Teo during the early stages of this research. Their research was supported by the Natural Sciences and Engineering Research Council of Canada. Research was conducted while the second author was at the Department of Management and Marketing, Faculty of Economics and Commerce, University of Melbourne, Australia.

References

Aéroport International Strasbourg. 2006. Concession from the chamber of commerce and industry of Strasbourg and the Lower Rhine: Practical user guide. Accessed July 26, 2007, http://www.strasbourg.aeroport.fr/pdf/tarifs/tarifs_aero_2006_GB.pdf.

Basawa, I. V., B. L. S. Prakasa Rao. 1980. *Statistical Inference for Stochastic Processes*. Academic Press, London.

Bhat, U. N. 1984. *Elements of Applied Stochastic Processes*, 2nd ed. John Wiley, New York.

Bitauld, P., K. Burch, S. El-Taji, E. Fanucchi, M. Montevecchi, J. Ohlsson, A. Palella, R. Rushmeier, J. Snowdon. 1997. Journey management. *OR/MS Today* 24(5) 30–33.

Chun, H. W. 1995. Solving check-in counter constraints with ILOG SOLVER. *Proc. 1st ILOG Solver and ILOG Schedule Internat. Users Meeting, Abbaye des Vaux de Cernay, France*.

Chun, H. W., W. T. R. Mak. 1999. Intelligent resource simulation for an airport check-in counter allocation system. *IEEE Trans. Systems, Man and Cybernetics—Part C: Appl. Rev.* 29(3) 325–335.

European Organization for the Safety of Air Navigation. 2005. Standard inputs for Eurocontrol cost benefit analyses. Accessed July 27, 2007, <http://www.eurocontrol.int/eatm/gallery/content/public/library/CBA-standard-values.pdf>.

Federal Aviation Administration. 2003. Treatment of values of passenger time in economic analysis. Technical report, APO Bulletin APO-03-1, U.S. Department of Transportation, Washington, D.C.

Feller, W. 1968. *An Introduction to Probability Theory and Its Applications*, 3rd ed. John Wiley, New York.

Downloaded from informs.org by [202.161.43.77] on 06 January 2016, at 05:57. For personal use only, all rights reserved.

- Hadidi, N. 1969. On the service time distribution and the waiting time process of a potentially infinite capacity queueing system. *J. Appl. Probab.* **6** 594–603.
- Hadidi, N., B. W. Conolly. 1969. On the improvement of the operational characteristics of single server queues by the use of a queue length dependent service mechanism. *Appl. Statist.* **18**(3) 229–240.
- Haeme, R. A., J. L. Huttinger, R. W. Shore. 1988. Airline performance modeling to support schedule development: An application case study. M. Abrams, P. Haigh, J. Comfort, eds. *Proc. 1988 Winter Simulation Conf.*, AIEE, San Diego, 800–806.
- Keiding, N. 1975. Maximum likelihood estimation in the birth-death process. *Ann. Statist.* **3** 363–372.
- Koltai, T., N. Kalló. 2007. Quantitative analysis of waiting time in express lines at cash desks. *5th ANZAM Oper. Management Sympos.*, Melbourne, Australia, 1–12.
- Lee, A. M. 1966. *Applied Queueing Theory*. MacMillan, London.
- Littler, R. A., D. A. Whitaker. 1997. Estimating staffing requirements at an airport terminal. *J. Oper. Res. Soc.* **48**(2) 124–131.
- Neuts, M. F. 1981. *Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach*. Dover, New York.
- Park, Y., S. B. Ahn. 2003. Optimal assignment for check-in counters based on passenger arrival behaviour at an airport. *Transportation Planning Tech.* **26**(5) 397–416.
- Shavell, Z. A. 2000. The effect of schedule disruption on the economics of airline operations. *3rd USA/Europe Air Traffic Management R&D Seminar*. MITRE Corporation, Napoli.
- Smith, P. 2004. Barriers to marginal social cost pricing in the air transport sector—A guide for the non-economist. Accessed July 27, 2007, http://www.imprint-eu.org/public/Papers/IMPRINT%20final_phil%20smith.doc.
- van Dijk, N. M., E. van der Sluis. 2006. Check-in computation and optimization by simulation and IP in combination. *Eur. J. Oper. Res.* **171** 1152–1168.
- Yan, S., C. H. Tang, M. Chen. 2004. A model and a solution algorithm for airport common use check-in counter assignments. *Trans. Res. Part A* **38** 101–125.