European Journal of Operational Research 265 (2018) 169–177

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Stochastics and Statistics

# Dynamic allocation of stochastically-arriving flexible resources to random streams of objects with application to kidney cross-transplantation

Yael Perlman[a,*], Amir Elalouf[a], Uri Yechiali[b]

[a] Department of Management, Bar Ilan University, Ramat Gan 52900, Israel
[b] Department of Statistics and Operations Research, School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel

ARTICLE INFO

*Article history:*
Received 12 February 2017
Accepted 29 July 2017
Available online 4 August 2017

*Keywords:*
OR in health services
Dynamic allocation
Flexible-resource
Kidney transplantation

ABSTRACT

Two distinct random streams of discrete objects flow into a system and queue in two separate lines. Concurrently, two distinct types of resources arrive stochastically over time. Upon arrival, each resource unit is matched with a waiting object. One resource type is 'flexible' and can be allocated to either one of the object types. However, units of the other, non-flexible, resource type can be allocated only to units of one specific object type. The allocation probabilities are not fixed and may depend on both queue sizes of the two objects. If a resource unit is not allocated immediately, it is lost. The goal is to find an optimal state-dependent probabilistic dynamic allocation policy. We formulate the system as a two-dimensional Markov process, analyze its probabilistic behavior, and derive its performance measures. We then apply the model to the problem of kidney cross-transplantation and propose a new measure of system effectiveness, called Expected Value of Transplantation (EVT), based on the histocompatibility between kidneys and candidates. We further show that it is possible to balance the objectives of achieving equity in candidates' expected waiting times (EW) and maximizing EVT by equating the value of EW/EVT between the two groups.



## 1. Introduction

This paper studies a dynamic flexible-resource allocation problem that is encountered in numerous operational settings and is particularly salient in the context of live organ transplantation. We consider two random streams of discrete objects, denoted $S_B$ and $S_O$, which flow into a system and queue in two separate lines, denoted $Q_B$ and $Q_O$, respectively. In parallel, two distinct types of discrete resources, $R_B$ and $R_O$, arrive at the system stochastically over time. A unit of resource type $R_B$ can only be allocated to an $S_B$-type object, whereas a unit of resource type $R_O$ is 'flexible' and can be allocated to either one of the two object types: It is allocated to an $S_B$-type unit waiting in $Q_B$ (if the queue is not empty) with a probability that may depend on both queue sizes or, with the complementary probability, to a unit of an $S_O$-object waiting (if any) in $Q_O$. An arriving resource unit must be allocated to an object upon arrival, or it is lost. Thus, if a unit of resource $R_B$ arrives when there are no $S_B$-type objects in $Q_B$, it is discarded and disappears from the system, whereas an $R_O$-type resource is discarded upon arrival only when both queues are empty (i.e., $Q_B = Q_O = 0$). This paper seeks to identify an optimal state-dependent probabilistic dynamic allocation policy in the setup described.

Flexible-resource allocation models characterizing manufacturing systems that can produce multiple products simultaneously have been addressed in the literature (see, e.g., Buzacott & Shanthikumar, 1993; Sethi & Sethi, 1990; Perlman, 2013). Additional applications include telecommunication networks (Ross, 1995), in which incoming calls can be routed to multiple links. Similarly, in computer systems with multiple users and multiple servers, users can be dynamically routed to different servers, and computing capacity can be shared among different customers. Call centers are another important application of resource flexibility. As calls can vary by topic, urgency, duration, and level of difficulty, different call center operators may be trained to handle different subsets of call types (Koole & Mandelbaum, 2002; Shumsky, 2004). Aghari and Balcioglu (2009) performed extensive numerical studies based on realistic call center scenarios and showed that limited cross-training of operators (specifically, providing each agent with two additional skills) can considerably enhance the system's performance. Robbins and Harrison (2010) studied a call center queueing model with two customer types. They considered scenarios in which two separate teams were assigned to two different customer

* Corresponding author.
*E-mail address:* yael.perlman@biu.ac.il (Y. Perlman).

http://dx.doi.org/10.1016/j.ejor.2017.07.068
0377-2217/

types, or in which a single cross-trained team could serve both types. They showed that cross-training a small number of agents results in substantial benefits as opposed to relying on agents who can only serve one type of customer.

We provide and analyze a general model of the flexible-resource queueing system described above, and then focus on its application in the context of live kidney transplantation. Stanford, Lee, Chandok, and McAlister (2014) discussed the problem of allocating stochastically-arriving kidneys to random streams of transplantation candidates, who form separate queues according to their blood types. Their model considers two blood types—type O and type B—where type O kidneys can be given to any candidate, and type B kidneys can only be given to candidates with blood type B. The authors propose that it is possible to allocate a fixed fraction of blood type O kidneys to blood type B candidates such that the expected waiting times (EW) for transplantation for the two different types of candidates are the same. However, the fixed probability assumption overlooks the situation in which a type O kidney arrives while there are only type O candidates in the system, but no type B candidates. Our model addresses this issue by assuming that, in such a situation, an arriving type O kidney is allocated exclusively to a type O candidate. More generally, a key contribution of our paper is in letting the probability of cross transplantation depend dynamically on the actual number of B candidates present in the system, rather than assuming that the probability is fixed.

Human tissue cells contain antigens that are immunologically relevant to specific candidate and donor. The system of these antigens is known as the Human Leukocyte Antigen (HLA) system. HLA matching is one of the most important factors when deciding on kidney allocation. A review of the determinants of successful kidney transplantation is given in Bendersky and David (2016). Proper HLA matching decreases the risk of graft lost by about 40% (Takemoto, Port, Claas, & Duquesnoy, 2004). In applying our model to the kidney transplantation context, we propose a new means of measuring the effectiveness of the allocation system, beyond traditional metrics such as mean waiting times or mean queue sizes. This measure, which we refer to as the Expected Value of Transplantation (EVT), takes into account the extent to which the candidates and the kidneys they receive are compatible, i.e., matched in terms of their Human Leukocyte Antigen (HLA) groups. This measure constitutes another important contribution of our study. It is straightforward to generalize this measure to more traditional manufacturing systems, by attributing a value to each object-resource pair, representing the utility obtained from that specific pairing.

In a numerical analysis, we observe that long queues and long waiting times are associated with higher EVT values, as they increase the likelihood that an incoming kidney will find a well-matched candidate. On the other hand, long waiting times are expected to lead to deterioration in candidates' health, potentially culminating in death. Thus, we propose an additional measure of system effectiveness: the ratio of EW to EVT. The measure EW/EVT quantifies the rate of change in expected waiting time attributable to a change in EVT. Thus, this measure balances the two goals of achieving equitable waiting times and maximizing the overall quality of transplants. We show that only a small fraction of type O blood kidneys should be cross-transplanted to blood type B candidates in order to optimize the effectiveness of the system.

While the importance of having flexible resources or flexible servers in manufacturing and call center operations has long been recognized, in this paper we extend the scope of analysis by assuming that the number of available servers is not fixed but changes dynamically and randomly. Specifically, resources (e.g., kidneys, or servers) do not stay and wait for objects to arrive. Rather, individual resources arrive randomly over time, and each one must serve a waiting object (e.g., a transplantation candidate) as soon as it arrives; otherwise (i.e., if no appropriate objects are available), it is lost (i.e., disappears from the system).

The remainder of the paper is structured as follows. Section 2 presents the model formulation. In Section 3 we define and construct probability-generating functions (PGFs), and calculate the system's two-dimensional boundary probabilities, as well as its marginal state probabilities. In Section 4 we employ matrix geometric methods to further analyze the system and derive the system's stability condition and various performance measures. In Section 5 we formulate the EVT as a measure of the effectiveness of a system for the dynamic allocation of kidneys for (cross-) transplantation. In Section 6 we perform numerical analysis and conclude that only a small fraction of the flexible resource should be allocated to cross-transplantation in order to optimize the system's performance. Section 7 concludes the paper.

## 2. Model formulation

### 2.1. Problem description

Two Poisson streams of discrete objects, $S_B$ and $S_O$, flow into a system at rates of $\lambda_B$ and $\lambda_O$, respectively, and queue in two separate lines, $Q_B$ and $Q_O$. Concurrently, two types of discrete resources, $R_B$ and $R_O$, arrive stochastically with Poisson rates $\mu_B$ and $\mu_O$, respectively. All four processes are mutually independent. When a unit of resource $R_B$ arrives, it is allocated to an $S_B$ object waiting in $Q_B$. If $Q_B$ is empty, the unit is lost. However, an $R_O$-type resource can be allocated to an object of either type. The probability that a unit of $R_O$ is allocated to an $S_B$ object is assumed to depend both on the number of $S_B$ objects in $Q_B$ and on the number of $S_O$ objects in $Q_O$. If both queues are empty, the $R_O$ unit is lost.

Let $L_B$ and $L_O$ denote, respectively, the number of $S_B$ objects and the number of $S_O$ objects present in the system. It is assumed that the number of $S_B$ objects is bounded such that it cannot exceed a given value $N$. Let $P_{nm} = P(L_B = n, L_O = m)$, $n = 0, 1, \ldots, N$; $m = 0, 1, 2, \ldots$ denote the system's steady-state probabilities, and let $w_{nm}$ be the probability that an arriving $R_O$ resource is allocated to an $S_B$ object when the system is in state $(L_B = n, L_O = m)$. We assume that $w_{0m} = 0$ for all $m \geq 1$, since when $L_B = 0$, an $R_O$ resource is allocated only to an $S_O$ object. We further set $w_{no} = 1$ for $n = 1, \ldots, N$, since when no $S_O$ object is present, an arriving $R_O$ resource is allocated with probability 1 to an $S_B$ object (if present).

This process can be formulated as a two-dimensional continuous-time Markov process with a state transition-rate diagram as depicted in Fig. 1.

### 2.2. Balance equations

The set of balance equations for the system's state probabilities is constructed below.

For $n = 0$:

$$\begin{cases} m = 0 & P_{00}(\lambda_O + \lambda_B) = P_{01}\mu_O + P_{10}(\mu_O + \mu_B) \\ m \geq 1 & P_{0m}(\lambda_O + \lambda_B + \mu_O) = P_{0,m+1}\mu_O + P_{0,m-1}\lambda_O \\ & \quad + P_{1m}(w_{1m}\mu_O + \mu_B) \end{cases} \tag{1}$$

For $1 \leq n \leq N - 1$:

$$\begin{cases} m = 0 & P_{n0}(\lambda_O + \lambda_B + \mu_O + \mu_B) = P_{n-1,0}\lambda_A + P_{n,1}(1 - w_{n1})\mu_O \\ & \quad + P_{n+1,0}(\mu_O + \mu_B) \\ m \geq 1 & P_{nm}(\lambda_O + \lambda_B + \mu_O + \mu_B) = P_{n,m-1}\lambda_O + P_{n-1,m}\lambda_B \\ & \quad + P_{n,m+1}(1 - w_{n,m+1})\mu_O + P_{n+1,m}(w_{n+1,m}\mu_O + \mu_B) \end{cases} \tag{2}$$

For $n = N$:

$$\begin{cases} m = 0 & P_{N0}(\lambda_O + \mu_O + \mu_B) = P_{N,1}(1 - w_{N1})\mu_O + P_{N-1,0}\lambda_B \\ m \geq 1 & P_{Nm}(\lambda_O + \mu_O + \mu_B) \\ & \quad = P_{N,m-1}\lambda_O + P_{N,m+1}(1 - w_{N,m+1})\mu_O + P_{N-1,m}\lambda_B \end{cases} \tag{3}$$

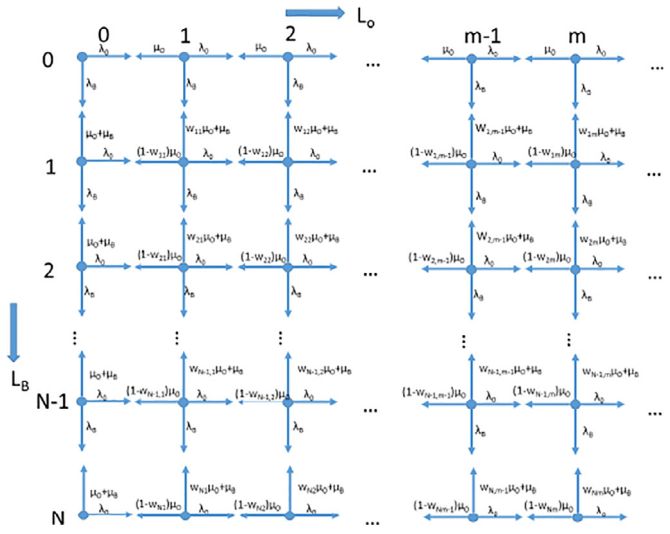



**Fig. 1.** Transition-rate diagram of $(L_B, L_O)$.

We note that the above set of equations cannot be solved analytically in its full general form, either by applying a generating functions method or by using a matrix geometric approach. Consequently, in what follows we relax the dependence of the probability $w_{nm}$ on both $n$ and $m$, and assume that the probability of allocating an $R_O$ resource to an $S_B$ object depends only on $L_B$, the number of $S_B$ objects present in the system. That is, we assume that $w_{n0} = 1$ for $n = 1, 2, ..., N$; $w_{0m} = 0$ for $m = 1, 2, ...$; and $w_{nm} = w_n \geq 0$ otherwise. In the next section we apply a generating functions method (see, e.g., Litvak and Yechiali, 2003; Perel and Yechiali, 2008; Perel and Yechiali, 2014) to analyze the system's probabilistic behavior.

## 3. Generating functions, boundary and marginal probabilities

Define, for each $0 \leq n \leq N$, the (partial) PGF

$$G_n(z) = \sum_{m=0}^{\infty} P_{nm} z^m \quad n = 0, 1, 2, ..., N.$$

Specifically, for $n = 0$, we multiply, for each $m$, the corresponding equation in (1) by $z^m$; by summing over all $m$ and rearranging terms, we obtain:

$$\left(\lambda_O(1-z) + \lambda_B + \mu_O\left(1 - \frac{1}{z}\right)\right)G_0(z) - (\mu_B + w_1\mu_O)G_1(z)$$
$$= \mu_O\left(1 - \frac{1}{z}\right)P_{00} + (1 - w_1)\mu_O P_{10} \equiv b_0(z) \tag{4}$$

Similarly, using the equations in (2), we obtain, for $1 \leq n \leq N-1$:

$$-\lambda_B G_{n-1}(z) + \left[\lambda_O(1-z) + \lambda_B + \mu_A + \mu_O\left(1 - \frac{1-w_n}{z}\right)\right]G_n(z)$$
$$-(\mu_A + w_{n+1}\mu_O)G_{n+1}(z)$$
$$= (1 - w_{n+1})\mu_O P_{n+1,0} - \mu_O(\frac{1-w_n}{z})P_{n0} \equiv b_n(z) \tag{5}$$

Finally, using the equations in (3), we obtain, for $n = N$:

$$-\lambda_B G_{N-1}(z) + \left[\lambda_O(1-z) + \mu_B + \mu_O\left(1 - \frac{1-w_N}{z}\right)\right]G_N(z)$$
$$= -\mu_O\frac{(1-w_N)}{z}P_{N0} \equiv b_N(z) \tag{6}$$

Let $d_0(z) = \lambda_O(1-z) + \lambda_B + \mu_O(1 - \frac{1}{z})$,and for $1 \leq n \leq N-1$, let $d_n(z) = \lambda_O(1-z) + \lambda_A + \mu_A + (1 - \frac{1-w_n}{z})\mu_O$.

Finally, let $d_N(z) = \lambda_O(1-z) + \mu_B + (1 - \frac{1-w_N}{z})\mu_O$.

Let $c_n = \mu_A + w_{n+1}\mu_O, \quad 0 \leq n \leq N$.

Eqs. (4)–(6) determine a set of linear equations for the unknown PGFs in the form $A(z)\vec{G}(z) = \vec{b}(z)$, where

$$A_{(N+1)\times(N+1)}(z)$$
$$= \begin{pmatrix} d_0(z) & -c_0 & 0 & \cdots & \cdots & \cdots & 0 \\ -\lambda_B & d_1(z) & -c_1 & 0 & \cdots & \cdots & 0 \\ 0 & -\lambda_B & d_2(z) & -c_2 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & -\lambda_B & d_{N-1}(z) & -c_{N-1} \\ 0 & \cdots & \cdots & \cdots & 0 & -\lambda_B & d_N(z) \end{pmatrix}$$

$$\vec{G}(z) = (G_0(z),\ G_1(z), \ldots, G_N(z))^t,$$

and

$$\vec{b}(z) = \begin{pmatrix} b_0(z) \\ \vdots \\ \vdots \\ b_n(z) \\ \vdots \\ \vdots \\ b_N(z) \end{pmatrix},$$

where $b_n(z)$, $n = 0,1,2,\ldots,N$, are given in Eqs. (4)–(6).

To obtain $G_n(z)$, we use Cramer's rule (see e.g. Perel & Yechiali, 2008). This leads to an expression of $G_n(z)$ in terms of the $N+1$ unknown probabilities $\{P_{n0},\ 0 \leq n \leq N\}$. In order to obtain the set $\{P_{n0}\}$, we need to find $N+1$ equations relating these $N+1$ unknowns. Since $G_n(z)$ is a PGF defined for all $|z| \leq 1$, each root of $|A(z)|$ in that interval is also a root of $|A_n(z)|$, $0 \leq n \leq N$. We claim that $|A(z)|$ has exactly $N+1$ roots in $0 < z < 1$. We will use those roots to obtain the $(N+1)$ probabilities $\{P_{n0}\}$.

**Theorem 1.** *The polynomial $|A(z)|$ has $2(N+1)$ roots, of which exactly $N+1$ are in $(0, 1)$, and the additional $N+1$ are in the open interval $(1, \infty)$.*

**Proof.** See Appendix. □

Denote by $z_{N+1,1}, z_{N+1,2} \ldots, z_{N+1,N}$ the $N+1$ roots of $|A(z)|$ in $(0, 1)$. The boundary probabilities $\{P_{n0},\ 0 \leq n \leq N\}$ are now calculated by using these $N+1$ roots and by solving the following set of $N+1$ equations: $|A_0(z_{N+1,1})| = 0$, $|A_1(z_{N+1,2})| = 0$, ..., $|A_N(z_{N+1,N+1})| = 0$, where the variables are the $N+1$ unknown boundary probabilities $P_{00}, P_{10}, ..., P_{N0}$. Then, when all boundary probabilities are calculated, the set of PGFs $\{G_n(z),\ n = 0, 1, 2, ..., N\}$ is completely determined. In addition, the marginal probabilities $P_{n\bullet}$ are given by

$$P_{n\bullet} = G_n(1) = \sum_{m=0} P_{nm} \qquad n = 0, 1, 2, ..., N. \tag{7}$$

With the aid of the boundary probabilities, we obtain the set of $N+1$ marginal probabilities $\{P_{n\bullet}$ by setting $z = 1$ in each of the $N$ equations corresponding to (4) and (5) to obtain

$$\lambda_B P_{n\bullet} = \mu_B P_{n+1,\bullet} + \mu_O P_{n+1,0} + w_{n+1}\mu_O(P_{n+1,\bullet} - P_{n+1,0}),$$
$$n = 0, 1, 2, ..., N-1 \tag{8}$$

Alternatively, Eq. (8) can be obtained by a horizontal cut between lines $n$ and $n+1$ in Fig. 1. Finally, together with the normalization equation $\sum_{n=0}^{N} P_{n\bullet} = 1$, the set $\{P_{n\bullet}\}$ is directly calculated.

The mean number of $S_B$-type objects in the system is given by

$$E[L_B] = \sum_{n=0}^{N} nP_{n\bullet}. \tag{9}$$

In a stable system the effective inflow of $S_B$-type objects is $\lambda_B(eff) = \lambda_B(1 - P_{N\bullet})$. Hence, by Little's law, the mean sojourn time of an $S_B$-type object is

$$E[W_B] = \frac{E[L_B]}{\lambda_B(1 - P_{N\bullet})}. \tag{10}$$

In the next section, we use matrix geometric analysis to obtain the mean number of $S_O$-type objects in the system.

## 4. Matrix geometric analysis

Arrange the set of system states as follows:

$$\{(00, 10, 20, ..., N0), (01, 11, 21, ..., N1), ...,$$
$$(0m, 1m, 2m, ..., Nm), ...\}.$$

Modifying the balance equations (1)–(3) for the case in which $w_{nm}$ depends only on $n$ — namely, $w_{n0} = 1$ for $n = 1, 2, ..., N$; $w_{0m} = 0$ for $m = 1, 2, ...$; and $w_{nm} = w_n \geq 0$ otherwise — then the generator matrix Q of the resulting level-dependent Quasi Birth and Death (QBD) process is given by

$$Q = \begin{pmatrix} B_0 & A_0 & 0 & 0 & \cdots & \cdots & \cdots \\ A_2 & A_1 & A_0 & 0 & \cdots & \cdots & \vdots \\ 0 & A_2 & A_1 & A_0 & \cdots & \cdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \cdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \cdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \cdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix},$$

where $B_0$, $A_0$, $A_1$ and $A_2$ are each an $(N+1)$-dimensional square matrix. We have:

$$B_0 = \begin{pmatrix} -(\lambda_O + \lambda_B) & \lambda_B & 0 & 0 & \cdots & \cdots & 0 \\ \mu_O + \mu_B & -(\lambda_O + \lambda_B + \mu_O + \mu_B) & \lambda_B & 0 & \cdots & \cdots & \vdots \\ 0 & \mu_O + \mu_B & -(\lambda_O + \lambda_B + \mu_O + \mu_B) & \lambda_B & \cdots & \cdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \cdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \cdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \cdots & \lambda_B \\ \vdots & \vdots & \vdots & \vdots & \vdots & \mu_O + \mu_B & -(\lambda_O + \mu_O + \mu_B) \end{pmatrix},$$

$$A_0 = \lambda_O I,$$

$$A_1 = \begin{pmatrix} -(\lambda_O + \lambda_B + \mu_O) & \lambda_B & 0 & \cdots & \cdots & 0 \\ w_1\mu_O + \mu_B & -(\lambda_O + \lambda_B + \mu_O + \mu_B) & \lambda_B & \cdots & \cdots & \vdots \\ 0 & w_2\mu_O + \mu_B & -(\lambda_O + \lambda_B + \mu_O + \mu_B) & \cdots & \cdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \cdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \cdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & -(\lambda_O + \lambda_B + \mu_O + \mu_B) & \lambda_B \\ \vdots & \vdots & \vdots & \vdots & w_N\mu_O + \mu_B & -(\lambda_O + \mu_O + \mu_B) \end{pmatrix}$$

and

$$A_2 = \begin{pmatrix} \mu_0 & 0 & 0 & 0 & \cdots & \cdots & \cdots \\ 0 & (1-w_1)\mu_0 & 0 & 0 & \cdots & \cdots & \vdots \\ 0 & 0 & (1-w_2)\mu_0 & 0 & \cdots & \cdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \cdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \cdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \cdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & (1-w_N)\mu_0 \end{pmatrix}.$$

Let $\vec{P_m} = (P_{0m}, P_{1m}, P_{2m}, ..., P_{Nm})$ be an $(N+1)$-dimensional row vector, $m = 1, 2, 3, ...$.

Also, set $\vec{P} = (\vec{P}_0, \vec{P}_1, \vec{P}_2, ..., \vec{P}_m, ...)$. The system-state probabilities are calculated by $\vec{P}\,Q = \vec{0}$ and $\vec{P}\,e^T = 1$, where $\vec{0}$ is a row vector of zeros, and $e^T$ is a column vector of ones. Then, the solution (Latouche & Ramaswami, 1999; Neuts, 1981) is given by $\vec{P}_0 B_0 + \vec{P}_1 A_2 = \vec{0}$ and $\vec{P}_m = \vec{P}_0 R^m$, $m = 1, 2, 3, ...$, where $\vec{P}_0[I - R]^{-1}e^T = 1$ and the rate matrix $R$ is the minimal nonnegative solution of the quadratic-matrix equation $A_0 + RA_1 + R^2A_2 = 0$.

The stability condition of the system is derived as follows:
Let $A = A_0 + A_1 + A_2$. Then,

$$A = \begin{pmatrix} -\lambda_B & \lambda_B & 0 & 0 & \cdots & \cdots & 0 \\ w_1\mu_O + \mu_B & -(\lambda_B + w_1\mu_O + \mu_B) & \lambda_B & 0 & \cdots & \cdots & \vdots \\ 0 & w_2\mu_O + \mu_B & -(\lambda_B + w_2\mu_O + \mu_B) & \lambda_B & \cdots & \cdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \cdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \cdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & -(\lambda_B + w_{N-1}\mu_O + \mu_B) & \lambda_B \\ \vdots & \vdots & \vdots & \vdots & \vdots & w_N\mu_O + \mu_B & -(w_N\mu_O + \mu_B) \end{pmatrix}.$$

The matrix $A$ expresses the infinitesimal generator matrix of an M/Mn/1/N queue with arrival rate $\lambda_n = \lambda_B, \quad n = 0, 1, 2, ..., N-1$, and a state-dependent service rate $\mu_n = w_n\mu_O + \mu_B, \ n = 1, 2, ..., N$. The stationary probability vector $\vec{\Pi} = (\Pi_0, \Pi_1, \Pi_2, ..., \Pi_N)$ of this queue is given by $\vec{\Pi}A = \vec{0}$ and $\vec{\Pi}\, e^T = \vec{1}$.

Let $\rho_i = \frac{\lambda_B}{w_i\mu_O + \mu_B}, \quad i = 1, 2, ..., N$. Then, $\Pi_n = (\prod_{i=1}^{n} \rho_i)\Pi_0, \quad n = 1, 2, ..., N$, with $\Pi_0^{-1} = 1 + \sum_{n=1}^{N} (\prod_{i=1}^{n} \rho_i)$. The stability condition of the system $Q$ is given by $\vec{\Pi}A_0e^T < \vec{\Pi}A_2e^T$ (Neuts, 1981).

Now, $\vec{\Pi}A_0e^T = \lambda_O$ and $\vec{\Pi}A_2e^T = \Pi_0\mu_O + \sum_{n=1}^{N} \Pi_N(1 - w_n)\mu_O = \Pi_0\mu_O[1 + \sum_{n=1}^{N} (1 - w_n)(\prod_{i=1}^{n} \rho_i)]$.

Thus, the stability condition becomes:

$$\frac{\lambda_O}{\mu_O} < \frac{1 + \sum_{n=1}^{N} (1 - w_n)\left(\prod_{i=1}^{n} \rho_i\right)}{1 + \sum_{n=1}^{N} \left(\prod_{i=1}^{n} \rho_i\right)}. \tag{11}$$

When $w_n = 0$ for all $n$, the system of $R_O$-type resources and $S_O$-type objects is an M/M/1 queue, and the stability condition reduces to $\frac{\lambda_O}{\mu_O} < 1$. When $w_n = 1$, then $\rho = \rho_i = \frac{\lambda_B}{\mu_O + \mu_B}, \quad i = 1, 2, ..., N$, and the stability condition is $\frac{\lambda_O}{\mu_O} < \frac{1-\rho}{1-\rho^{N+1}}$. This condition can be explained as follows: Consider an M/M/1/N queue with arrival rate $\lambda_B$ and a service rate $\mu_O + \mu_B$. Then, the fraction of time the queue is empty is $\frac{1-\rho}{1-\rho^{N+1}}$. This is a queue in which, as long as there are $S_B$ objects in the system, all resources (type $R_B$ and type $R_O$) are allocated to $S_B$ objects. $R_O$ resources are allocated to $S_O$ objects only when no $S_B$ objects are present. The proportion of time the latter event occurs is $\frac{1-\rho}{1-\rho^{N+1}}$. Thus, the rate of 'work' attributed to $S_O$ objects, namely, $\frac{\lambda_O}{\mu_O}$, cannot exceed this proportion of time.

The mean number of $S_O$ objects in the system is given by

$$E[L_O] = \sum_{m=0}^{\infty} m\vec{P_m}\, e^T = \vec{P_0} \sum_{m=1}^{\infty} mR^m\, e_{N+1}^T = \vec{P_0}\ R[I - R]^{-2}\ e^T. \tag{12}$$

By Little's law, the mean sojourn time of an $S_O$ object is

$$E[W_O] = \frac{E[L_O]}{\lambda_O}. \tag{13}$$

The mean waiting time for an arbitrary object is

$$E[W] = \frac{\lambda_B(1 - P_{N\bullet})}{\lambda_B(1 - P_{N\bullet}) + \lambda_O} E[W_B] + \frac{\lambda_O}{\lambda_B(1 - P_{N\bullet}) + \lambda_O} E[W_O]. \tag{14}$$

## 5. Application to kidney transplantation: a new measure – Expected Value of Transplantation

Several studies have addressed various aspects of the kidney allocation problem. David and Yechiali (1985) were among the first to model kidney allocation based on HLA considerations. They considered a single candidate and a stochastic stream of kidneys, where the decision whether to transplant an arriving kidney or reject it is based on the degree of histocompatibility between the candidate and the kidney. Those authors further extended the study of dynamic allocation process to parallel streams of candidates and offers (David and Yechiali, 1990) and to one attribute sequential assignment match process in discrete time (David and Yechiali, 1995). Zenios (1999) was the first to present a queueing model for transplant waiting times and, in a subsequent study, carried out simulations on data from kidney transplant waiting lists in the US (Zenios, Chertow, & Wein, 2000). Su and Zenios (2004) extended the kidney allocation problem to take into account the possibility that patients might refuse an available kidney in order to hold out for a higher-quality match. Bendersky and David (2016) recently studied a flexible single-candidate model for the kidney allocation problem based on a broad family of Gamma lifetime distributions. They obtained the optimal critical times of acceptance of offers of different qualities.

As discussed above (see also Stanford et al., 2014), our model assumes that kidneys corresponding to blood type O can serve multiple types of candidates, whereas, kidneys corresponding to blood type B can serve only one type of candidate (i.e., individuals with blood type B). The issue of cross-transplantation of kidneys corresponding to blood type O has given rise to the so-called "Blood Type O Problem", in which too many type O kidneys are cross-transplanted to compatible blood groups, thereby diminishing the supply of type O kidneys and causing notably longer waits for candidates with blood type O. Queueing models have recently begun to address this problem (see Drekic, Stanford, Woolford, & McAlister, 2015).

In what follows, we propose a new measure for evaluating the performance of a kidney transplantation queueing system. First, given that a kidney has been allocated to a particular queue, we propose that the specific waiting candidate who receives the kidney should be selected on the basis of a *best-fit rule*. In particular, we suggest that the kidney should be allocated to the candidate with the highest level of HLA match. We operationalize the HLA match as follows: When a kidney arrives and is allocated to the queue of a particular blood type, it is assigned a level of histocompatibility for each one of the waiting candidates, where there are $I$ possible levels. The kidney is then given to the candidate with the best fit, independently of his position in line, as defined below.

Let $H$ be a random variable denoting the number of mismatched HLA characteristics between a randomly-arriving kidney

and a random candidate. Let $f_i = P(H = i),\ i = 0, 1, 2, \ldots, I$, be the probability that a random candidate and a random kidney have $i$ mismatches; and let $F_i = P(H \le i)$, where $F_I = 1$. Let $X$ be a random variable denoting the 'transplantation value' between a random kidney and a random candidate. For example, X may denote the probability that the lifetime of the transplanted candidate will exceed a given number of years. The value of $X$ for $H = i$ mismatches is denoted by $x_i$, where, if $i < j$, then $x_i > x_j$. Consequently,

$$P(X = x_i) = P(H = i) = f_i, \text{ and } E[X] = \sum_{i=0}^{I} f_i x_i.$$

Suppose that $L_B = n \ge 1$, and that an arriving kidney is allocated to $Q_B$, the queue of candidates with blood type B. The $X$ values corresponding to the $n$ candidates waiting in $Q_B$ are denoted, respectively,$X_1, X_2, \ldots, X_n$; each of these variables is i.i.d. like $X$. Assuming that the kidney is allocated according to the *best-fit rule*, then the value of the allocation is $X^*_{(n)} = \max\{X_1, X_2, \ldots, X_n\}$. Now, denoting $\bar{F}_i = 1 - F_i$, we obtain:

$$E\left[X^*_{(n)}\right] = \underbrace{\left(1 - \bar{F}_0^n\right)}_{\substack{\text{The probability}\\ \text{that at least one}\\ \text{candidate has}\\ \text{zero mismatches}}} x_0 + \sum_{i=1}^{I} \underbrace{\left(\left(1 - \bar{F}_i^n\right) - \left(1 - \bar{F}_{i-1}^n\right)\right)}_{\substack{\text{The probability that the candidate with}\\ \text{thebest match has exactly i mismatches}}} x_i \ . \tag{15}$$

We define the EVT obtained from allocating a kidney (type B or type O) to a B candidate, $EVT_B$, as follows:

$$EVT_B = \sum_{n=0}^{N} P_{n\bullet} E\left[X^*_{(n)}\right], \tag{16}$$

where $P_{n\bullet}$ is given by Eq. (7), $E[X^*_{(0)}] = 0$, and when $L_B = 1$, $E[X^*_{(1)}] = E[X] = \sum_{i=0}^{I} f_i x_i$.

Similarly, when an arriving type O kidney is allocated to a candidate with blood type O, the EVT is given by

$$EVT_O = \sum_{m=0}^{\infty} P_{\bullet m} E\left[X^*_{(m)}\right], \tag{17}$$

where $P_{\bullet m} = \overrightarrow{P_m} \cdot e^T$.

Let $C_{nm} = w_n E[X^*_{(n)}] + (1 - w_n) E[X^*_{(m)}]$. Then, the overall EVT obtained from allocating a kidney according to the *best-fit rule* is given in the following theorem.

**Theorem 2.** $EVT_{best-fit} = \frac{\mu_B}{\mu_B+\mu_O} EVT_B + \frac{\mu_O}{\mu_B+\mu_O}\left[\sum_{m=1}^{\infty} P_{0m} E[X^*_{(m)}] + \sum_{n=1}^{N} P_{n0} E[X^*_{(n)}] + \sum_{n=1}^{N}\sum_{m=1}^{\infty} P_{nm} C_{nm}\right]$

**Proof.** The *best-fit* EVT equals the sum of two terms: (i) the product of the probability that a randomly-arriving kidney is of type B, i.e., $\frac{\mu_B}{\mu_B+\mu_O}$, and $EVT_B$; and (ii) the product of the complementary probability, $\frac{\mu_O}{\mu_B+\mu_O}$, and the weighted EVT, $\sum_{n=0}^{N}\sum_{m=0}^{\infty} P_{nm}(w_n E[X^*_{(n)}] + (1 - w_n) E[X^*_{(m)}])$, resulting from allocating an O-type kidney to either a B or an O candidate. The proof is completed by using the specific definition of $C_{nm}$ and since$E[X^*_{(0)}] = 0$. □

Conventional resource allocation methods are based on a first-come first-serve (FCFS) approach (see a detailed discussion and a list of limitations in Thekinen and Panchal, 2016). Under the FCFS rule, an incoming kidney that has been allocated to a given queue is assigned to the candidate at the front of the queue. Thus, the EVT of the entire system under the FCFS rule equals

$$EVT_{FCFS} = \frac{\mu_B}{\mu_B+\mu_O}(1 - P_{0\bullet})E[X] + \frac{\mu_O}{\mu_B+\mu_O}(1 - P_{00})E[X]. \tag{18}$$

This calculation is based on the assumption that a kidney of blood type B is allocated to a candidate with blood type B when there is at least one such candidate in the system, whereas a kidney of type O is allocated whenever at least one of the candidate queues is not empty. Clearly, $EVT_{best-fit} \ge EVT_{FCFS}$.

## 6. Numerical analysis: $w_n$ increases with $n$

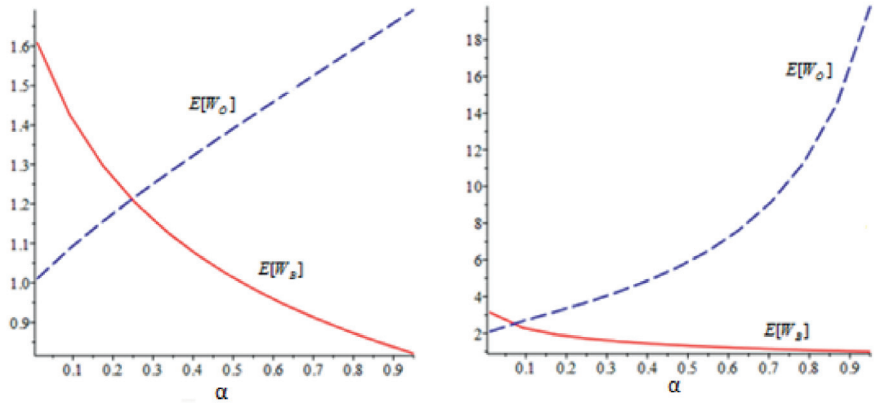
In what follows we carry out a numerical analysis in which we assume that the probability of allocating a kidney of blood type O to a candidate of blood type B increases with $n$, the number of candidates of blood type B who are waiting for transplant. Specifically, we define $w_n = \frac{\alpha n}{N}$ for $n = 1, 2, 3, \ldots, N$, where $0 \le \alpha \le 1$.

Since transplant waitlists are almost never empty, it is appropriate to assume that $\frac{\lambda_O}{\mu_O}$ is a value close to unity. Hence, we assume that $\lambda_O = 9$ and $\mu_O = 10$. Adopting the ratio of kidney availability rates presented in Stanford et al. (2014) for blood types B and O, respectively, we set $\lambda_B = \lambda_O \frac{9}{46}$, and $\mu_B = \mu_O \frac{9}{46}$. In addition, we set$N = 40$. For the EVT calculations we use the following data: $I = 4$, $f_i$= [0.0094, 0.0941, 0.3134, 0.4073, 0.1758], and $x_i$= [0.7, 0.62, 0.49, 0.47, 0.44] (see David & Yechiali, 1985).

Fig. 2 depicts the expected sojourn times of type O candidates and type B candidates— $E[W_O]$ and $E[W_B]$, respectively—as a function of $\alpha$ for two different values: $\lambda_O = 9$ and $\lambda_O = 9.5$ (where the value of $\lambda_B$ in each case is determined according to the ratio of kidney availability rates). Clearly, when $\alpha$ increases, more O kidneys are given to B candidates, causing higher values of $E[W_O]$. Note that when $\alpha$ equals zero, $E[W_O] = \frac{1}{\mu_O - \lambda_O}$, since the system for the O blood type becomes a regular M/M/1 queue. At the same time, $E[W_B]$ decreases as $\alpha$ increases. Denote by $\alpha^*$ the value of $\alpha$ that yields $E[W_O] = E[W_B]$. Stanford et al. (2014) assume that this point reflects equity and fairness in the allocation process. As depicted in Fig. 2, as $\lambda_O$ increases, the average number (as well as the mean waiting time) of O candidates increases, implying that $\alpha^*$decreases from $\alpha^* \cong 0.24$ (Fig. 2a) to $\alpha^* \cong 0.09$ (Fig. 2b). That is, the probability of allocating an O kidney to a B candidate decreases as $\lambda_O$ increases.

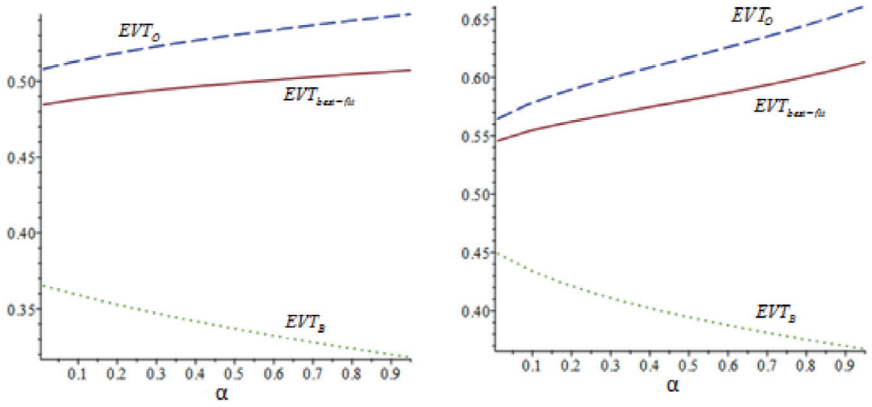
Thus, when the decision maker's objective is to equate the expected waiting times for the two types of candidates, the mean fraction of O kidneys cross-transplanted to B candidates is equal to $\bar{w}^* \equiv \sum_{n=1}^{N} P_{n\bullet} \frac{\alpha^* n}{N} = \frac{\alpha^*}{N} E[L_A]$. When $\lambda_O = 9$, then $\bar{w}^* \cong 0.013$, while when $\lambda_O = 9.5$, then $\bar{w}^* \cong 0.01$. Note that these small probabilities ensure that cross-transplantation is indeed a rare occurrence, thereby preventing the "blood type O problem" in which the average waiting time for transplantation candidates with blood type O is much higher than that of candidates with blood type B (Stanford et al., 2014).

Next, we consider a case in which the decision maker's objective is to maximize the EVT of the system. Fig. 3 depicts $EVT_O$, $EVT_B$ and $EVT_{best-fit}$ as functions of $\alpha$ for the two values of $\lambda_O$. Clearly, as $\lambda_O$ increases, all EVT values increase, since having more candidates in the system increases the probability of attaining a better transplantation fit (HLA match). In addition, when $\alpha$ increases, the probability of allocating an O kidney to a B candidate increases, resulting in higher values of $EVT_O$, since more O candidates accumulate in the queue, resulting in a higher probability of attaining a good fit. As depicted in Fig. 3, $EVT_{best-fit}$ also increases in $\alpha$.



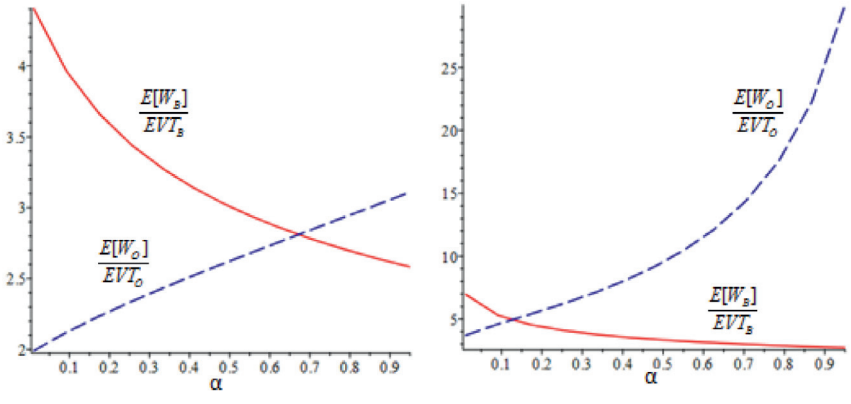

a: $\lambda_o = 9$ b: $\lambda_o = 9.5$

**Fig. 2.** $E[W_O]$ and $E[W_B]$as functions of $\alpha$ for a: $\lambda_O = 9$, and b: $\lambda_O = 9.5$.




a: $\lambda_o = 9$ b: $\lambda_o = 9.5$

**Fig. 3.** $EVT_{best-fit}$, $EVT_O$ and $EVT_B$ as functions of $\alpha$ for a: $\lambda_O = 9$, and b: $\lambda_O = 9.5$.




a: $\lambda_o = 9$ b: $\lambda_o = 9.5$

**Fig. 4.** $\frac{E[W_O]}{EVT_O}$ and $\frac{E[W_B]}{EVT_B}$ as a function of $\alpha$ for a: $\lambda_O = 9$, and b: $\lambda_O = 9.5$.

We further consider an additional objective measure that combines fairness and equity in candidates' waiting times with the benefit in terms of transplantation quality reflected in the EVT measure. Specifically, we propose the ratio [EW/EVT], which quantifies the rate of change in EW due to a change in EVT. We denote the value of $\alpha$ that equates the ratios $\frac{E[W_O]}{EVT_O} = \frac{E[W_B]}{EVT_B}$ by $\tilde{\alpha}$. Fig. 4a shows that $\tilde{\alpha} = 0.67$ for $\lambda_O = 9$, and Fig. 4b shows that $\tilde{\alpha} = 0.15$ for $\lambda_O = 9.5$. Let $\tilde{\tilde{w}} \equiv \sum_{n=1}^{N} P_{n\bullet} \frac{\tilde{\alpha} n}{N} = \frac{\tilde{\alpha}}{N} E[L_A]$. When $\lambda_O = 9$, then $\tilde{\tilde{w}} \cong 0.03$ and when $\lambda_O = 9.5$ then $\tilde{\tilde{w}} \cong 0.02$. That is, a relatively small fraction of O kidneys are cross-transplanted into B candidates.

Figs. 4 and 2 suggest that $\tilde{\alpha}$ is larger than $\alpha^*$. That is, taking into consideration the EVT measure by employing the ratio EW/EVT results in a higher probability of allocating an O kidney to a B candidate. When the objective is to equate the values of [EW/EVT] for the two queues, the waiting time of O candidates is longer than that under the policy that equates the values of EW; however, the value of $EVT_{best-fit}$ in the former case is greater, as shown in Fig. 3. That is, overall benefit to the system in terms of achieving a high value of EVT may come at a slight cost in terms of type O candidates' waiting time.

## 7. Conclusions

Resource flexibility can be beneficial in that it facilitates efficient resource utilization. Yet, it leads to the question: What is the optimal fraction of cross-allocation? To address this problem, we developed and analyzed a queuing model based on a dynamic approach, in which the probability of cross-allocation of a flexible resource depends on the number of objects waiting in the non-flexible queue. We applied our model to the context of kidney transplantation, in which candidates with blood type B can be allocated kidneys of either blood type B or type O, but candidates with blood type O can only receive type O kidneys. We proposed a measure, EVT, that takes into account the histocompatibility between kidneys and transplant candidates, as a new measure for evaluating the effectiveness of the kidney allocation system. We further suggest that it is possible to balance different objectives (namely, achieving equitable waiting times among different candidates while maximizing EVT) by striving to equate the value of the ratio EW/EVT between the two types of candidates. In a numerical analysis we show that, when the latter objective is used, only a small fraction of type O kidneys are ultimately allocated to type B candidates, and that type O candidates' average waiting time only slightly exceeds that of type B candidates.

## Appendix

Theorem 1 is proved using a cascade of supporting lemmas. To this end we define the following.

Let $q_0(z) = 1$. Define the determinants of the minors of the diagonal of the matrix $A(z)$, starting from the upper left corner, as follows:

$$q_1(z) = d_0(z) \tag{A1}$$

$$q_n(z) = d_{n-1}(z)q_{n-1}(z) - \lambda_B c_{n-2}q_{n-2}(z), \quad (n = 2, \ldots, N+1). \tag{A2}$$

Clearly, $q_0(z)$ has no roots. Next we find the roots of $q_1(z)$.

**Lemma 1.** *$q_1(z)$ has one root in (0, 1) and another root in (1,∞).*

**Proof.** Clearly, $z = 0$ is not a root of $q_1(z)$, so that $zd_0(z)$ is a quadratic function of z having two roots: $z_{1,1} = \frac{\lambda_O+\lambda_B+\mu_O-\sqrt{(\lambda_O+\lambda_B+\mu_O)^2-4\lambda_O\mu_O}}{2\lambda_O}$, and $y_{1,1} = \frac{\lambda_O+\lambda_B+\mu_O+\sqrt{(\lambda_O+\lambda_B+\mu_O)^2-4\lambda_O\mu_O}}{2\lambda_O}$. Now, since $\lambda_O\lambda_B > 0$, $z_{1,1} \in (0, 1)$ while $y_{1,1} > 1$ since $\mu_O > \lambda_O$. □

**Lemma 2.** *$q_n(z)$ and $q_{n+1}(z)$ have no common roots.*

**Proof.** By induction. Clearly, for $n = 0$ the claim is true. Assume the lemma holds for $n = k - 1 \le N$. If $z^*$ is a root of both $q_k(z)$ and of $q_{k+1}(z)$, then, by Eq. (A2) and since $c_n$ is not a function of z, $z^*$ is also a root of $q_{k-1}(z)$, which contradicts the induction's assumption. □

**Lemma 3.** *Given $z^*$ is a root of $q_{n-1}(z)$, then sign $(q_{n-2}(z^*)q_n(z^*))) = -1$.*

**Proof.** If $q_{n-1}(z^*) = 0$ by Eq. (A2), $q_n(z^*) = -\lambda_B c_{n-2}q_{n-2}(z^*)$, implying that $\text{sign}(q_{n-2}(z^*)q_n(z^*))) = -1$, as claimed. □

The following four lemmas follow from Eq. (A2).

**Lemma 4.** *$z^n q_n(z)$ is a polynomial of degree 2n for $0 \le n \le N + 1$.*

**Lemma 5.** $q_n(0^+) = (-1)^n \infty \quad 0 \le n \le N+1$

**Lemma 6.** $q_n(\infty) = (-1)^n \infty \quad 0 \le n \le N+1$

**Lemma 7.** $q_n(1) = (\lambda_B)^n \quad 0 \le n \le N+1$

**Lemma 8.** *For $1 \le n \le N+1$, $q_n(z)$ possesses exactly 2n roots from which n roots, denoted by $z_{n,1}, \ldots, z_{n,n}$, are in $(0, 1)$, and the other n roots, denoted by $y_{n,1}, \ldots, y_{n,n}$, are in $(1, \infty)$. In addition, $\text{sign}(q_{n-1}(z_{n,i})) = (-1)^{n-i}$ and $\text{sign}(q_{n-1}(y_{n,i})) = (-1)^{i-1}$ for $i = 1,\ldots,n$.*

**Proof.** By induction. For $n = 1$, by Lemma 1, $q_1(z)$ has exactly two roots, $0 < z_{1,1} < 1$ and $1 < y_{1,1} < \infty$. Since $q_0(z) = 1$ for all $z$, then evidently, $q_0(z_{1,1}) > 0$ and $q_0(y_{1,1}) > 0$.

For $n = 2$, by Lemma 5, $q_2(0^+) > 0$. By Lemma 7, $q_2(1) > 0$. Since $q_0(z) = 1 > 0$, then by Lemma 3, $q_2(z_{1,1}) < 0$ and $q_2(y_{1,1}) < 0$. By Lemma 6, $q_2(\infty) > 0$. Therefore, $q_2(z)$ has exactly 2 roots in (0,1) denoted by $z_{2,1}$ and $z_{2,2}$ and exactly 2 roots in (1,∞) denoted by $y_{2,1}$ and $y_{2,2}$. Hence, the claim is true for $n = 2$. For these roots, $0 < z_{2,1} < z_{1,1} < z_{2,2} < 1 < y_{2,1} < y_{1,1} < y_{2,2} < \infty$. In addition, $q_1(z) < 0$ in $(-\infty, z_{1,1})$ and in $(y_{1,1}, \infty)$, $q_1(z) > 0$ in $(z_{1,1}, y_{1,1})$. Therefore, $q_1(z_{2,1})\langle 0$, $q_1(z_{2,2}))0$, $q_1(y_{2,1}) > 0$, $q_1(y_{2,2}) < 0$, showing that

$\text{sign}(q_1(z_{n,i})) = (-1)^{n-i}$ and $\text{sign}(q_{n-1}(y_{n,i})) = (-1)^{i-1}$ holds for $n = 2$.

We assume that the Lemma holds for $n-1$ and prove for $n$.

By Lemma 5, $q_n(0^+) = (-1)^n \infty$ and by the induction assumption and Lemma 3 $\text{sign}(q_n(z_{n-1,i})) = (-1)^{n-i}$ for i= 1,...,n−1. Now, by Lemma 4, $q_n(1) > 0$. Therefore, $q_n(z)$ changes it's sign $n+1$ times in (0,1) and the roots satisfy $0 < z_{n,1} < z_{n-1,1} < z_{n,2}, \ldots, z_{n-1,n-1} < z_{n,n} < 1$. Since we know all the roots of $q_{n-1}(z)$ and $q_{n-1}(1) > 0$ and since $q_{n-1}(0^+) = (-1)^{n-1} \infty > 0$ then $\text{sign}(q_{n-1}(z_{n,i})) = (-1)^{n-i}$ $i = 1,\ldots,n$. Similarly, the same holds for the roots in (1,∞) and the proof is complete. □

## References


Ahghari, M., & Balcioğlu, B. (2009). Benefits of cross-training in a skill-based routing contact center with priority queues and impatient customers. *IIE Transactions, 41*, 524–536.

Bendersky, M., & David, I. (2016). Deciding kidney-offer admissibility dependent on patients' lifetime failure rate. *European Journal of Operational Research, 251*(2), 686–693.

Buzacott, J. A., & Shanthikumar, J. G. (1993). *Stochastic models of manufacturing systems*. Englewood Cliffs, NJ: Prentice Hall.

David, I., & Yechiali, U. (1985). A time-dependent stopping problem with application to live organ transplants. *Operations Research, 33*, 491–504.

David, I., & Yechiali, U. (1990). Sequential assignment match processes with arrivals of candidates and offers. *Probability in the Engineering and Informational Sciences, 4*, 413–430.

David, I., & Yechiali, U. (1995). One-attribute sequential assignment match processes in discrete time. *Operations Research, 43*, 879–884.

Drekic, S., Stanford, D. A., Woolford, D. G., & McAlister, V. C. (2015). A model for deceased-donor transplant queue waiting times. *Queueing Systems, 79*, 87–115.

Koole, G., & Mandelbaum, A. (2002). Queuing models of call centers, an introduction. *Annals of Operations Research, 113*, 41–59.

Latouche, G., & Ramaswami, V. (1999). *Introduction to matrix geometric methods in stochastic modeling*. Philadelphia, PA: SIAM ASA-SIAM Series on Statistics and Applied Probability.

Litvak, N., & Yechiali, U. (2003). Routing in queues with delayed information. *Queueing systems, 43*(1), 147–165.

Neuts, M. F. (1981). *Matrix-geometric solutions in stochastic models: An algorithmic approach*. Baltimore, MD: The Johns Hopkins University Press.

Perel, E., & Yechiali, U. (2008). Queues where customers of one queue act as servers of the other queue. *Queueing Systems, 60*(3), 271–288.

Perel, N., & Yechiali, U. (2014). The Israeli queue with infinite number of groups. *Probability in the Engineering and Informational Sciences, 28*(1), 1–19.
Perlman, Y. (2013). The effect of risk aversion on product family design under uncertain consumer segments. *International Journal of Production Research, 51*, 504–514.
Robbins, T., & Harrison, T. (2010). Cross training in call centers with uncertain arrivals and global service level agreements. *International Journal of Operations and Quantitative Management, 16*, 307–329.
Ross, K. W. (1995). *Multiservice loss models for broadband telecommunication networks*. London: Springer-Verlag.
Sethi, A. K., & Sethi, S. P. (1990). Flexibility in manufacturing: A survey. *The International Journal of Flexible Manufacturing Systems, 2*, 289–328.
Shumsky, R. A. (2004). Approximation and analysis of a call center with flexible and specialized servers. *OR Spectrum, 26*, 307–330.
Stanford, D. A., Lee, J. M., Chandok, N., & McAlister, V. (2014). A queuing model to address waiting time inconsistency in solid-organ transplantation. *Operations Research for Health Care, 3*(1), 40–45.
Su, X., & Zenios, S. A. (2004). Patient choice in kidney allocation: The role of the queueing discipline. *Manufacturing & Service Operations Management, 6*, 280–301.
Takemoto, S., Port, F. K., Claas, F. H., & Duquesnoy, R. J. (2004). HLA matching for kidney transplantation. *Human Immunology, 65*(12), 1489–1505.
Thekinen, J., & Panchal, J. H. (2016). Resource allocation in cloud-based design and manufacturing: A mechanism design approach. *Journal of Manufacturing Systems, 43*, 327–338.
Zenios, S. A. (1999). Modeling the transplant waiting list: A queueing model with reneging. *Queueing Systems, 31*, 239–251.
Zenios, S. A., Chertow, G. M., & Wein, L. M. (2000). Dynamic allocation of kidneys to candidates on the transplant waiting list. *Operations Research, 48*, 549–569.