The Annals of Applied Probability
1997, Vol. 7, No. 2, 326–339

STOCHASTIC MONOTONICITY FOR STATIONARY RECURRENCE TIMES OF FIRST PASSAGE HEIGHTS

By Eric V. Denardo, Eugene A. Feinberg[1] and Offer Kella[2]

Yale University, SUNY, Stony Brook and Hebrew University of Jerusalem

This paper deals with first passage heights of sums of renewal sequences, random walks, and Lévy processes. We prove that the joint age and excess (and therefore, the current life) stationary distributions of these heights are stochastically increasing (in the usual first-order sense) in the passage levels. As a preliminary tool, which is also of independent interest, a new decomposition of the stationary excess distribution, as a convolution of two other distributions, is developed. As a consequence of these results, certain monotonicity results are concluded for ratios involving convex functions. This paper is motivated by problems related to control of queues with removable servers which model single-machine produce-to-order manufacturing systems. Applications to these problems are provided.

1. Introduction. As is usual in renewal theory, S_k denotes the sum of k i.i.d. nonnegative random variables. Let $n(t)$ denote the smallest value of k for which S_k exceeds t. One of our results is that the stationary excess distribution of $S_{n(t)}$ is stochastically increasing (in the usual, first-order sense) in t. In this paper we prove this and stronger results, generalize them to a random walk with positive drift and, further, to a Lévy process, and discuss their applications. In addition, we establish a new decomposition of the stationary excess distribution of S_ν where ν is any finite mean stopping time. Theorem 2.1 expresses that distribution as the convolution of two simpler ones.

As is well known, $n(t)$ is a stopping time. Denote by ν another stopping time. Our basic result, Theorem 2.2, states that the stationary current life as well as the joint age and excess distributions of $S_{\nu \wedge n(t)}$ are stochastically increasing in t. The same theorem establishes the simpler (and possibly known) result that the stationary current life and joint age and excess distributions of S_k are stochastically increasing in k. However, Example 5.2 shows that even the excess distributions of $S_{k \wedge n(t)}$ need not be stochastically

Received May 1995; revised July 1996.

[1]Partially supported by NSF Grant DMI-95-00746.

[2]Supported in part by Grant 92-00035 of the U.S.–Israel Binational Science Foundation and by Grant 372/93-1 of the Israel Science Foundation.

AMS 1991 *subject classifications*. Primary 60K05; secondary 60G40, 60K25.

Key words and phrases. Age, excess life, current life, renewal sequence, stochastic order, $M^X/G/1$ queue, N-policy, Lévy process, convex.

increasing in k. It is evident (cf. Lemma 2.2) that stochastic monotonicity of a stationary current life distribution implies the same property of the corresponding excess distribution. But the converse is false (cf. Example 5.5).

The results announced above are renewal theoretic results, but they spring from and are applied to optimization of queues. Let us indicate how. A single-machine produce-to-order manufacturing system can be modeled as an $M^X/G/1$ queue, that is, as a single-server queue with a general service time distribution and exponential interarrival times of batches of customers. For the N-policy of Yadin and Naor (1963), this server is switched off when the system becomes empty and is switched on when the number of waiting customers is at least N. A natural question is whether the average waiting time W_N experienced by a customer increases with N. Lee and Srinivasan (1989) provided a formula for W_N, but it was not clear from this formula whether W_N is increasing or not. Their formula led us to observe that the monotonicity of W_N would follow directly from the monotonicity of the means of the stationary excess distributions of $\{S_{n(N)}|N \in \mathscr{N}\}$ and thus would follow from the stochastic monotonicity (in N) of these distributions. In the same paper, Lee and Srinivasan provided a formula for C_N, the average cost per unit time incurred by an N-policy, and they conjectured that C_N is quasi-convex (*unimodal* in their terminology). We prove this conjecture (Corollary 3.1) by observing that it follows from the decomposition presented in Theorem 2.1.

Section 2 presents our basic results on stochastic monotonicity. Section 3 relates these results to optimization of queues. Section 4 generalizes the results in Section 2 to random walks and Lévy processes. That section also contains results on the monotonicity of ratios involving convex functions. Section 5 deals with counterexamples to other directions of generalization. Collectively, these counterexamples suggest that the hypotheses in Sections 2 and 4 are the right ones to impose.

2. Setup and basic results. In what follows $\mathscr{R}^+$ is the set of nonnegative reals, $\mathscr{R}$ is the set of all reals, and $\mathscr{N}$ is the set of positive integers. Denote $a \wedge b \equiv \min(a, b)$, $f(x-) = \lim_{y \uparrow x} f(y)$, $f(x+) = \lim_{y \downarrow x} f(y)$. *Almost surely, without loss of generality, stochastically increasing and Laplace–Stieltjes transform* are abbreviated to a.s., w.l.o.g., SI and LST, respectively. Throughout, an empty sum is defined to be zero (i.e., $\sum_{i=a}^{b} c_i \equiv 0$ when $a > b$).

On a filtered probability space $(\Omega, \mathscr{F}, P)$ with filtration $\{\mathscr{F}_i | i \geq 0\}$, let $\{X_i | i \in \mathscr{N}\}$ be an adapted sequence ($X_i \in \mathscr{F}_i$, $i \in \mathscr{N}$) of identically distributed finite mean nonnegative random variables with $P[X_1 = 0] < 1$, such that X_i is independent of $\mathscr{F}_{i-1}$ for every $i \in \mathscr{N}$. In particular $\{X_i | i \in \mathscr{N}\}$ is a renewal sequence, that is, i.i.d. For convenience, let X be independent of $\mathscr{F}_\infty \equiv \bigvee_{i=0}^{\infty} \mathscr{F}_i$ (the smallest σ-field containing $\bigcup_{i=0}^{\infty} \mathscr{F}_i$) and be distributed like X_1. We define this filtration in order to allow the accommodation of stopping times that are determined by more than just our renewal sequence or even independent of it altogether. Henceforth, whenever we mention *stopping time* we mean a stopping time with respect to the above filtration which is not a.s. zero. Also, in this paper, a *random time* will mean a nonnegative integer-valued random variable (not necessarily a stopping time) which is not a.s. zero. Unless explicitly assumed otherwise, stopping and random times are allowed to be infinite with positive probability or even a.s.

We recall that a function of multiple variables is nondecreasing if it is nondecreasing in each variable. As is customary, we will use $\le_{\rm st}$ to denote the usual stochastic ordering. That is, for random vectors V and W, $V \le_{\rm st} W$ if $Eg(V) \le Eg(W)$ for every nondecreasing g. If V and W are random variables, an equivalent definition is $P[V > t] \le P[W > t]$ for every t. Throughout this paper, when we stochastically compare random vectors, we do not assume that they live on the same probability space. Rather, we are only concerned with their distributions. Although this might normally interfere with stochastic process notation, we find this convenient and we hope that it does not cause confusion. When we write $\{V(\lambda)|\lambda \in \Lambda\}$ is SI, where $\Lambda = \mathscr{R}^+, \mathscr{N}$, it means that $V(\lambda_1) \le_{\rm st} V(\lambda_2)$ for any $\lambda_1 < \lambda_2$, where $\lambda_1, \lambda_2 \in \Lambda$.

Set $S_0 = 0$, $S_n = \sum_{i=1}^n X_i$, $n(t) = \inf\{n|S_n > t\}$. As is well known, $n(t)$ is a stopping time with respect to the renewal sequence (hence, with respect to our filtration) and $n(t) > i$ if and only if $S_i \le t$. It is also well known that $En(t) < \infty$ for all $t \in \mathscr{R}^+$. Hence, $E\nu \wedge n(t) < \infty$ for any random time ν and $t \in \mathscr{R}^+$. Since $n(\cdot)$ is a right continuous process, we have by dominated convergence [$\nu \wedge n(t+h) \le \nu \wedge n(t+1)$ for $h \le 1$] that $E\nu \wedge n(t)$ is a right continuous nondecreasing function. Also we observe that $E\nu \wedge n(t) = 0$ for $t < 0$ and $E\nu \wedge n(\infty) = E\nu$ (finite or infinite).

For a given random variable Y with $EY < \infty$, we denote by (Y^a, Y^e) a random vector having the joint stationary *age* and *excess* (resp.) distribution associated with Y. Also we let $Y^c = Y^a + Y^e$ have the stationary *current* or *total* life distribution. It is well known that Y^a and Y^e are identically distributed, that $P[Y^a > s, Y^e > t] = P[Y^e > s + t]$ [e.g., Karlin and Taylor (1975), pages 193–195] and that

$$
\begin{aligned}
P[Y^c \le t] &= \frac{EY 1_{\{Y \le t\}}}{EY}, \\
P[Y^e \le t] &= \frac{EY \wedge t}{EY} = \frac{1}{EY}\int_0^t P(Y > y)\,dy.
\end{aligned}
\tag{2.1}
$$

In particular, for any Borel measurable g,

$$
\begin{aligned}
Eg(Y^c) &= \frac{EYg(Y)}{EY}, \\
Eg(Y^e) &= \frac{E\int_0^Y g(y)\,dy}{EY} = \frac{\int_0^\infty g(y)P[Y > y]\,dy}{EY},
\end{aligned}
\tag{2.2}
$$

provided the right-hand sides are well defined. The following is well known. The simple short proof is provided for completeness.

LEMMA 2.1. *Let* $U \sim \mathit{Uniform}(0,1)$ *and* Y^c *be independent. Then* $(UY^c, [1-U]Y^c)$ *and* (Y^a, Y^e) *are identically distributed.*

PROOF. Recalling that $P[Y^a > s, Y^e > t] = P[Y^e > s + t]$, we have

$$
\begin{aligned}
P[UY^c > s, (1-U)Y^c > t] &= P[s/Y^c < U < 1 - t/Y^c] \\
&= E[1-(s+t)/Y^c]^+ = \frac{EY[1-(s+t)/Y]^+}{EY} \\
&= \frac{E[Y-(s+t)]^+}{EY} = 1 - \frac{EY \wedge (s+t)}{EY} \\
&= P[Y^e > s+t].
\end{aligned}
\tag{2.3}
$$

where $a^+ = \max(a, 0)$. □

The following lemma implies that, for a family $\{V(t) | t \in \mathscr{R}\}$ of nonnegative random variables, $\{V^c(t) | t \in \mathscr{R}\}$ is SI if and only if $\{(V^a(t), V^e(t)) | t \in \mathscr{R}\}$ is SI.

LEMMA 2.2. *Let* Y *and* Z *be two nonnegative random variables. Then* $Y^c \le_{\mathrm{st}} Z^c$ *if and only if* $(Y^a, Y^e) \le_{\mathrm{st}} (Z^a, Z^e)$.

PROOF. Let $(Y^a, Y^e) \le_{\mathrm{st}} (Z^a, Z^e)$. For any nondecreasing function $g(x)$ on $\mathscr{R}$, the function $f(y, z) = g(y + z)$ is nondecreasing. Thus, $Eg(Y^c) = Ef(Y^a, Y^e) \le Ef(Z^a, Z^e) = Eg(Z^c)$.

Now let $Y^c \le_{\mathrm{st}} Z^c$. If $f(y, z)$ is a nondecreasing function, then $g(x) = f(ux, (1-u)x)$ is nondecreasing in x for any $u \in [0, 1]$. Let $f(y, z)$ be a nonnegative nondecreasing function. We apply Lemma 2.1 and get $Ef(Y^a, Y^e) = Ef(UY^c, (1-U)Y^c) = EE[f(UY^c, (1-U)Y^c) | U] \le EE[f(UZ^c, (1-U)Z^c) | U] = Ef(UZ^c, (1-U)Z^c) = Ef(Z^a, Z^e)$. □

REMARK 2.1. In view of Lemma 2.2, a natural question is whether $Y^e \le_{\mathrm{st}} Z^e$ implies $Y^c \le_{\mathrm{st}} Z^c$. Example 5.5 shows that the answer to this question is negative. However, $Y^e \le_{\mathrm{st}} Z^e$ does imply a result which is weaker than $Y^c \le_{\mathrm{st}} Z^c$. In light of Lemma 2.1 and $P[Y^a > s, Y^e > t] = P[Y^e > s + t]$, we have that $P[Y^a > s, Y^e > t] \le P[Z^a > s, Z^e > t]$. This inequality defines a two-dimensional stochastic order which is weaker than $\le_{\mathrm{st}}$; see Stoyan (1983) and Marshall and Olkin (1979).

For every random time ν with $E\nu < \infty$ we let S_ν^* be a random variable having the distribution

$$
P[S_\nu^* \le s] = \frac{E\nu \wedge n(s)}{E\nu}. \tag{2.4}
$$

Recalling that $\nu \wedge n(\cdot)$ is nondecreasing and right continuous with $E\nu \wedge n(0-) = 0$ and $E\nu \wedge n(\infty) = E\nu$, the right-hand side of (2.4) indeed represents a well-defined distribution function. The next lemma gives an alternative description of the distribution of S_ν^*.

LEMMA 2.3. *For every random time ν, with $E\nu < \infty$, and every bounded Borel measurable g,*

$$Eg(S_\nu^*) = \frac{E\sum_{i=1}^{\nu} g(S_{i-1})}{E\nu}. \tag{2.5}$$

PROOF. We have that

$$\begin{aligned} E\sum_{i=1}^{\nu} 1_{\{S_{i-1}\le s\}} &= E\sum_{i=1}^{\nu\wedge n(s)} 1 = E\nu\wedge n(s) \\ &= E\nu P[S_\nu^* \le s] = E\nu E 1_{\{S_\nu^*\le s\}}. \end{aligned} \tag{2.6}$$

Since (2.6) holds for any real s, by standard measure theoretic arguments the result extends to any bounded Borel measurable g. □

Let ν be a stopping time. Then,

$$\begin{aligned} P[S_\nu^c \le s] &= \frac{ES_\nu 1_{\{S_\nu\le s\}}}{ES_\nu}, \qquad P[S_\nu^e \le s] = \frac{ES_\nu \wedge s}{ES_\nu}, \\ P[S_\nu^* \le s] &= \frac{ES_\nu \wedge S_{n(s)}}{ES_\nu}, \end{aligned} \tag{2.7}$$

where the first two identities follow from (2.1), while the third is implied by $ES_\nu \wedge S_{n(s)} = ES_{\nu\wedge n(s)} = EXE\nu \wedge n(s)$ (the last equality is from Wald's identity) and by (2.4). From (2.7) we obtain that $S_\nu^* \le_{\mathrm{st}} S_\nu^e \le_{\mathrm{st}} S_\nu^c$, where the second ordering is well known. The following theorem introduces a new relationship between S_ν^e and S_ν^*. This is the decomposition property for S_ν^e to which we refer in the abstract and introduction.

THEOREM 2.1. *For a given finite mean stopping time ν, let S_ν^* and X^e be independent. Then S_ν^e and $S_\nu^* + X^e$ are identically distributed.*

PROOF. As ν is a stopping time, both the event $\{\nu \ge i\} = \{\nu > i-1\}$ and S_{i-1} are in $\mathscr{F}_{i-1}$, hence jointly independent of X_i. Therefore, for $\alpha > 0$

$$\begin{aligned} 1 - E\exp(-\alpha S_\nu) &= E\sum_{i=1}^{\nu} (1-\exp(-\alpha X_i))\exp(-\alpha S_{i-1}) \\ &= \sum_{i=1}^{\infty} E(1-\exp(-\alpha X_i)) E\exp(-\alpha S_{i-1}) 1_{\{\nu\ge i\}} \\ &= (1 - E\exp(-\alpha X)) E\sum_{i=1}^{\nu} \exp(-\alpha S_{i-1}). \end{aligned} \tag{2.8}$$

It is well known [it also follows from integration of (2.2) in parts] that for a finite mean positive random variable Y, the LST of the stationary excess time is given by

$$E\exp(-\alpha Y^e) = \frac{1 - E\exp(-\alpha Y)}{\alpha EY} \tag{2.9}$$

for $\alpha > 0$. Therefore, dividing both sides of (2.8) by $\alpha ES_\nu = \alpha EXE\nu$ (Wald's identity) and recalling Lemma 2.3 (with $g(\cdot) = e^{-\alpha \cdot}$), we have that

$$E\exp(-\alpha S_\nu^e) = E\exp(-\alpha X^e)E\exp(-\alpha S_\nu^*) \tag{2.10}$$

and the proof is complete. □

The following lemma prepares for the corollary which follows it. Note that this result still holds if all we would require is that $\{X_i | i \in \mathscr{N}\}$ is an arbitrary sequence of random variables having the property that $S_k \to \infty$ as $k \to \infty$ and $En(t) < \infty$ for all t.

Lemma 2.4. *The families $\{S_k^* | k \in \mathscr{N}\}$ and $\{S_{\nu \wedge n(t)}^* | t \in \mathscr{R}^+\}$, with any random time ν, are SI.*

Proof. We first observe that

$$P[S_k^* \leq s] = \frac{Ek \wedge n(s)}{k} = E\left[\frac{n(s)}{k} \wedge 1\right], \tag{2.11}$$

which is nonincreasing in k. This implies the stochastic monotonicity of $\{S_k^* | k \in \mathscr{N}\}$. Since $E\nu \wedge n(t) \wedge n(s) = [E\nu \wedge n(t)] \wedge [E\nu \wedge n(s)]$ we have that the distribution function of $S_{\nu \wedge n(t)}^*$ is given by $[E\nu \wedge n(\cdot)/E\nu \wedge n(t)] \wedge 1$ which is nonincreasing, hence SI, in t. □

As an immediate corollary of Theorem 2.1 and Lemma 2.4 we obtain the following. We note that a substantially stronger result (Theorem 2.2) will be given shortly.

Corollary 2.1. *The families $\{S_k^e | k \in \mathscr{N}\}$ and $\{S_{\nu \wedge n(t)}^e | t \in \mathscr{R}^+\}$, with any stopping time ν, are SI.*

Corollary 2.1 makes it natural to question why we only consider stopping times which are either deterministic (k) or of the form $\nu \wedge n(t)$ [$n(t) = \infty \wedge n(t)$ being a special case]. Perhaps $S_{\nu_1}^e \leq_{\text{st}} S_{\nu_2}^e$ for any two stopping times with $\nu_1 \leq \nu_2$? More generally, is it not the case that if $Y_1 \leq Y_2$ then $Y_1^e \leq_{\text{st}} Y_2^e$? Could we weaken the assumption that $\{X_i | i \in \mathscr{N}\}$ are i.i.d.? The answers to all of these questions are negative in general. These questions and counterexamples are discussed in Section 5. We now proceed to strengthen the monotonicity results described in Corollary 2.1.

Theorem 2.2. *The families $\{S_k^c | k \in \mathscr{N}\}$ and $\{S_{\nu \wedge n(t)}^c | t \in \mathscr{R}^+\}$, with any stopping time ν, are SI and therefore $\{(S_k^a, S_k^e) | k \in \mathscr{N}\}$ and $\{(S_{\nu \wedge n(t)}^a, S_{\nu \wedge n(t)}^e) | t \in \mathscr{R}^+\}$ are SI as well.*

Proof. Since $\{X_i 1_{\{S_k > y\}} | 1 \leq i \leq k\}$ are identically distributed, $i = 1, \dots, k$, then by (2.2),

$$P[S_k^c > y] = \frac{ES_k 1_{\{S_k > y\}}}{ES_k} = \frac{EX_1 1_{\{S_k > y\}}}{EX}. \tag{2.12}$$

Therefore, the right-hand side is clearly nondecreasing in k. To argue that $S^c_{\nu \wedge n(t)}$ is SI, we observe that for $t \le y$ we have that $S_i 1_{\{S_i > y\}} = 0$ for $i < n(t)$. Also S_{i-1} and $\{\nu \wedge n(t) \ge i\}$ are jointly independent of X_i. Thus for $t \le y$,

$$
\begin{aligned}
ES_{\nu \wedge n(t)} 1_{\{S_{\nu \wedge n(t)} > y\}} &= E \sum_{i=1}^{\nu \wedge n(t)} S_i 1_{\{S_i > y\}} \\
&= \sum_{i=1}^{\infty} E(S_{i-1} + X_i) 1_{\{S_{i-1} + X_i > y\}} 1_{\{\nu \wedge n(t) \ge i\}} \\
&= \sum_{i=1}^{\infty} E(S_{i-1} + X) 1_{\{S_{i-1} + X > y\}} 1_{\{\nu \wedge n(t) \ge i\}} \qquad (2.13)\\
&= E \sum_{i=1}^{\nu \wedge n(t)} (S_{i-1} + X) 1_{\{S_{i-1} + X > y\}} \\
&= E[\nu \wedge n(t)] E(S^*_{\nu \wedge n(t)} + X) 1_{\{S^*_{\nu \wedge n(t)} + X > y\}},
\end{aligned}
$$

where it is emphasized that $S^*_{\nu \wedge n(t)}$ and X are taken to be independent [we observe that the last equality in (2.13) follows from Lemma 2.3]. For every $t \ge y$, $S_{\nu \wedge n(t)} \le y$ if and only if $S_\nu \le y$, and $S_{\nu \wedge n(t)} = S_\nu$ on $\{S_\nu \le y\}$. Therefore, for $t \ge y$,

$$
\begin{aligned}
ES_{\nu \wedge n(t)} 1_{\{S_{\nu \wedge n(t)} > y\}} &= ES_{\nu \wedge n(t)} - ES_\nu 1_{\{S_\nu \le y\}} \qquad (2.14)\\
&= E[\nu \wedge n(t)] EX - ES_\nu 1_{\{S_\nu \le y\}}.
\end{aligned}
$$

Including y in the range in (2.13) as well as in (2.14) is done in order to insure that the monotonicity holds over the entire nonnegative real line. Combining (2.1), (2.13) and (2.14) we have that

$$
\begin{aligned}
P[S^c_{\nu \wedge n(t)} > y] EX &= \frac{ES_{\nu \wedge n(t)} 1_{\{S_{\nu \wedge n(t)} > y\}}}{ES_{\nu \wedge n(t)}} EX \\
&= \frac{ES_{\nu \wedge n(t)} 1_{\{S_{\nu \wedge n(t)} > y\}}}{E\nu \wedge n(t)} \qquad (2.15)\\
&= \begin{cases} E(S^*_{\nu \wedge n(t)} + X) 1_{\{S^*_{\nu \wedge n(t)} + X > y\}}, & t \le y, \\ EX - ES_\nu 1_{\{S_\nu \le y\}} / E\nu \wedge n(t), & t \ge y, \end{cases}
\end{aligned}
$$

once again, with the understanding that in (2.15), X and $S^*_{\nu \wedge n(t)}$ are considered independent. Clearly $g(\cdot) = E(\cdot + X) 1_{\{\cdot + X > y\}}$ is nondecreasing. Hence $P[S^c_{\nu \wedge n(\cdot)} > y]$ is nondecreasing on $[0, y]$ as, by Lemma 2.4, $\{S^*_{\nu \wedge n(t)} | t \in \mathscr{R}^+\}$ is SI. On $[y, \infty)$, $P[S^c_{\nu \wedge n(\cdot)} > y]$ is clearly nondecreasing as well. The last two statements of the theorem follow from Lemma 2.2. □

REMARK 2.2. If we change the assumption of i.i.d. to exchangeable, then (2.12) is still valid. Hence, $\{S^c_k | k \in \mathscr{N}\}$ as well as $\{(S^a_k, S^e_k) | k \in \mathscr{N}\}$ will also be SI under this assumption.

3. Control of $M^X/G/1$ queues with removable servers. In this section, we consider a stochastic model of a single-machine produce-to-order manufacturing system. We suppose that orders appear in i.i.d. batches according to a stationary Poisson process and that production times are i.i.d. Thus, the model is an $M^X/G/1$ queue. When the server completes a job or has no work to do, it might pay to switch the server off or to divert it to another purpose, which leads to the study of a queue with a "removable" server.

The optimization problem that presents itself is to minimize the aggregate cost per unit time, whose components are a set-up cost incurred each time the server is turned on, a linear operating cost per unit time while the server is on and a linear holding cost per unit time that each customer waits for service. The basic result is that average cost per unit time is minimized by an N-policy: turn the server on when the number of waiting customers is at least N, and turn it off when the system becomes empty. Lee and Srinivasan (1989) derived formulas for operating characteristics of an N-policy and constructed an algorithm that computes the best N-policy. See also Lee, Lee and Chae (1994) for further results in this direction. Federgruen and So (1991) proved the optimality of an N-policy over all policies, stationary or nonstationary.

One important issue in the $M^X/G/1$ model is whether the average waiting time W_N in a queue controlled by an N-policy is increasing in N. Another issue, conjectured on pages 717 and 718 in Lee and Srinivasan (1989), is whether the average cost C_N per unit time incurred by an N-policy is quasi-convex (*unimodal*, in their terminology) in N. Both issues are resolved in Corollary 3.1.

COROLLARY 3.1. *Consider an $M^X/G/1$ queue with a removable server. The expected customer waiting time W_N under an N-policy is nondecreasing in N, and the expected cost per unit time C_N is quasi-convex in N.*

PROOF. First we show the monotonicity of W_N. Using our notation, Theorem 4.2 in Lee and Srinivasan (1989) shows that $W_N = AES^e_{n(N-1)} + B$, where A and B are constants and $A > 0$. This equation can also be obtained from a general decomposition of a queue with a removable server [see Fuhrmann and Cooper (1985), Shanthikumar (1988) and Kella and Whitt (1991)]. Thus, W_N is nondecreasing in N if and only if $ES^e_{n(N)}$ has the same property. The latter is immediate from Corollary 2.1. Even so, we provide a simple direct proof. In fact, this proof is what motivated us to consider the substantially more general statements presented in Theorems 2.1 and 2.2.

In this section, X is integer valued. Without loss of generality, assume $X \in \mathscr{N}$. With $p_0 = 1$, $a_0 = EX^2$ and $b_0 = 2EX$, set $p_N = P[S_{n(N-1)} = N]$, $a_N = ES^2_{n(N)}$ and $b_N = 2ES_{n(N)}$ for $N \in \mathscr{N}$. Note that since $X \in \mathscr{N}$, either $S_{n(N-1)} = N$ or $S_{n(N)} = S_{n(N-1)}$. Thus, $S_{n(N)}$ is distributed like $S_{n(N-1)} + X1_{\{S_{n(N-1)}=N\}}$, X being independent of $S_{n(N-1)}$. This gives the recursions $b_N = b_{N-1} + 2EXp_N = 2EX\sum_{i=0}^N p_i$ and $a_N = a_{N-1} + (2NEX + EX^2)p_N =$

$EX^2 \sum_{i=0}^{N} p_i + 2EX \sum_{i=0}^{N} ip_i$. Since $ES^e_{n(N)} = a_N/b_N$,

$$ES^e_{n(N)} = \frac{EX^2}{2EX} + \frac{\sum_{i=0}^{N} ip_i}{\sum_{i=0}^{N} p_i} \tag{3.1}$$

and, by induction, the right-hand side is nondecreasing in N. Thus, W_N is nondecreasing in N.

We note that $EX^2/(2EX) = EX^e$, so that in view of Theorem 2.1, $ES^*_{n(N)} = \sum_{i=0}^{N} ip_i / \sum_{i=0}^{N} p_i$. In particular we identify $En(N) = \sum_{i=0}^{N} p_i$ and $E \sum_{i=1}^{n(N)} S_{i-1} = \sum_{i=0}^{N} ip_i$ (see Lemma 2.3), which is easy to verify directly.

For the second result, formulas (6.1)–(6.4) in Lee and Srinivasan (1989) express $C_N = a(ES_{n(N-1)})^{-1} + bES^e_{n(N-1)} + c$, where a, b, and c are constants and $a, b > 0$. Without loss of generality, we set $b = 1$, so that (3.1) gives

$$C_{N+1} - C_N = g_N f_N, \tag{3.2}$$

where

$$g_N = \frac{p_N}{\sum_{i=0}^{N-1} p_i \sum_{i=0}^{N} p_i} \tag{3.3}$$

is nonnegative and

$$f_N = \sum_{i=0}^{N-1} (N-i) p_i - \frac{a}{EX} \tag{3.4}$$

is increasing. Therefore, if $C_{N+1} > C_N$ for some N then $f_m > f_N > 0$ and $C_{m+1} \geq C_m$ for any $m \geq N$. So, C_N is quasi-convex. □

Although C_N is quasi-convex, it need not be convex; if $X = 2$ a.s., then $C_{2N-1} = C_{2N}$ for all $N \in \mathscr{N}$, in which case C_N cannot be convex.

Related work on the simpler $M/G/1$ queue (no batch arrivals) with a removable server includes Yadin and Naor (1963), Heyman (1968), Sobel (1969), Bell (1971), Hofri (1986), Kella (1989) and Altman and Nain (1993). Sobel (1969) showed that for any stationary policy for a $GI/G/1$ queue, there is an N-policy with the same or better average cost per unit time, but his method does not apply to batch arrivals.

Heyman (1968) noted that it can be difficult to estimate waiting costs, in which case it becomes natural to consider a pair of criteria, one being average operating cost per unit time and the other being the average waiting time. Rigorous analysis of this bicriterion problem for an $M/G/1$ queue was done by Feinberg and Kim (1996). In an effort to extend their work to $M^X/G/1$ queues, the following question immediately arises: is W_N increasing in N? This question motivated our research, and Corollary 3.1 answers it. Kim (1995) has used the results of this paper in a study of bicriterion optimization of an $M^X/G/1$ queue with a removable server.

4. Extensions and consequences. In this section we will consider generalizations of Theorem 2.2. To begin, do parts of Theorem 2.2 hold when $P[X < 0]$ is positive? Clearly, we cannot expect even $\{S_k | k \in \mathscr{N}\}$ itself to be SI when X can be negative, let alone $\{S_k^e | k \in \mathscr{N}\}$. However, as for other aspects of Theorem 2.2, the answer is indeed positive and is given as follows.

THEOREM 4.1. *Let* $\{X_i | i \in \mathscr{N}\}$ *be i.i.d. (not necessarily nonnegative) with* $E|X_1| < \infty$ *and* $EX_1 > 0$. *Then* $\{(S_{n(t)}^a, S_{n(t)}^e) | t \in \mathscr{R}^+\}$ *and* $\{S_{n(t)}^c | t \in \mathscr{R}^+\}$ *are SI.*

PROOF. With $T_0 = 0$ let $T_n = \inf\{S_i | S_i > T_{N-1}\}$ be the consecutive strictly increasing ladder heights associated with the random walk $\{S_i | i \geq 0\}$. Then $T_i - T_{i-1}$ are positive i.i.d. random variables having a finite mean [see, e.g., Chung (1974), pages 281 and 284, Theorems 8.4.4 and 8.4.7] and it is clear that $S_{n(t)} = T_{m(t)}$ where $m(t) = \inf\{i | T_i > t\}$. Therefore, the result follows from Theorem 2.2. □

Theorem 4.1 implies that part of Theorem 2.2 holds for positive mean random walks. A continuous time process which is continuous in probability and has stationary independent increments is called a Lévy process. This process is often viewed as a continuous time analogue of a random walk. Does Theorem 4.1 extend to Lévy processes? Preparing for the answer, let $\{Z_t | t \in \mathscr{R}^+\}$ be a càdlàg (strong Markov) version of a Lévy process satisfying $E|Z_1| < \infty$. Let $\{\mathscr{F}_t | t \in \mathscr{R}^+\}$ be a standard (right continuous, augmented) filtration, such that $Z_t - Z_s$ is independent of $\mathscr{F}_s$ for every $s < t$. As before, one possible such filtration is (the augmentation of) the one generated by $\{Z_t | t \in \mathscr{R}^+\}$. In the following theorem, when we say *stopping time*, we mean a stopping time (not a.s. zero) with respect to this filtration. Let $\sigma(t) = \inf\{s | Z_s > t\}$.

THEOREM 4.2. *If* $E|Z_1| < \infty$ *and* $EZ_1 > 0$, *then* $\{(Z_{\sigma(t)}^a, Z_{\sigma(t)}^e) | t \in \mathscr{R}^+\}$ *and* $\{Z_{\sigma(t)}^c | t \in \mathscr{R}^+\}$ *are SI. If in addition* $\{Z_t | t \in \mathscr{R}^+\}$ *is nondecreasing (a subordinator) then* $\{(Z_t^a, Z_t^e) | t \in \mathscr{R}^+\}$ *and* $\{Z_t^c | t \in \mathscr{R}^+\}$ *are SI and, for any stopping time* τ, $\{(Z_{\tau \wedge \sigma(t)}^a, Z_{\tau \wedge \sigma(t)}^e) | t \in \mathscr{R}^+\}$ *and* $\{Z_{\tau \wedge \sigma(t)}^c | t \in \mathscr{R}^+\}$ *are SI.*

PROOF. Consider $S_{k,n} = Z_{n2^{-k}}$. Then for every $k \in \mathscr{N}$, $\{S_{k,n} - S_{k,n-1} | n \in \mathscr{N}\}$ are i.i.d. having positive and finite mean. Setting $n_k(t) = \inf\{n | n \in \mathscr{N}, S_{k,n} > t\}$, we have by Theorem 4.1 that $\{(S_{k,n_k(t)}^a, S_{k,n_k(t)}^a) | t \in \mathscr{R}^+\}$ is SI. In particular $S_{k,n_k(t)}$ has a finite mean. Clearly $S_{k,n_k(t)}$ is nonincreasing in k; hence, by right continuity of $\{Z_t | t \in \mathscr{R}^+\}$, $S_{k,n_k(t)} \downarrow Z_{\sigma(t)}$ as $k \to \infty$. Therefore, by dominated convergence $ES_{k,n_k(t)} 1_{\{S_{k,n_k(t)} > y\}} \downarrow EZ_{\sigma(t)} 1_{\{Z_{\sigma(t)} > y\}}$ for any continuity point y of the distribution of $Z_{\sigma(t)}$. Similarly, we also have that $ES_{k,n_k(t)} \downarrow EZ_{\sigma(t)}$ as $k \to \infty$. Therefore, $S_{k,n_k(t)}^c$ converges in distribution to $Z_{\sigma(t)}^c$, which implies that $\{Z_{\sigma(t)}^c | t \in \mathscr{R}^+\}$; hence $\{(Z_{\sigma(t)}^a, Z_{\sigma(t)}^e) | t \in \mathscr{R}^+\}$ is SI [see Stoyan (1983), page 6, Proposition 1.2.3].

The proofs for $\{(Z_t^a, Z_t^e) | t \in \mathscr{R}^+\}$ and $\{(Z_{\tau \wedge \sigma(t)}^a, Z_{\tau \wedge \sigma(t)}^e) | t \in \mathscr{R}^+\}$, when $\{Z_t | t \in \mathscr{R}^+\}$ is a subordinator, are identical only that we apply Theorem 2.2 directly rather than Theorem 4.1; therefore, for the sake of brevity, they are omitted. The only potential complication is what the approximating stopping times should be for every approximating grid. However this is simple and standard, as we can approximate τ by $\nu_k = \sum_{n=1}^{\infty} n2^{-k} 1_{\{\tau \in [(n-1)2^{-k}, n2^{-k})\}}$ and observe that $\nu_k \downarrow \tau$, where in particular ν_k is a stopping time with respect to the filtration $\{\mathscr{F}_{n2^{-k}} | n \geq 0\}$. □

REMARK 4.1. Theorem 4.2 generalizes Theorems 2.2 and 4.1, which follow from it if we take $\{Z_t | t \in \mathscr{R}^+\}$ to be a compound Poisson process with jumps distributed like X, and restrict the stopping times to occur at points of jumps. On the other hand, Theorem 2.2 is more specific; it rests on the i.i.d. random variables $\{X_i | i \in \mathscr{N}\}$ which provide insight into what is going on and why.

In Section 3 we describe the initial question motivated by a queueing application. Namely, is $ES^2_{n(k)}/ES_{n(k)}$ nondecreasing when X is positive integer valued? Since $f(x) = x^2$ is convex with $f(0) = f(0+) = 0$ and $ES_{n(k)} = EXEn(k)$, a seemingly different generalization of this property is given by the following two corollaries.

COROLLARY 4.1. *Let f be convex on $(0, \infty)$ with $f(0) \leq f(0+) \leq 0$. If $E|X| < \infty$ and $EX > 0$, then $Ef(S_{n(t)})/En(t)$ is nondecreasing in t. If in addition $X \geq 0$ a.s., then $Ef(S_k)/k$ is nondecreasing in k and $Ef(S_{\nu \wedge n(t)})/E\nu \wedge n(t)$ is nondecreasing in t for any stopping time ν.*

COROLLARY 4.2. *Let f be convex on $(0, \infty)$ with $f(0) \leq f(0+) \leq 0$. If $\{Z_t | t \in \mathscr{R}^+\}$ is a càdlàg Lévy process with $E|Z_1| < \infty$ and $EZ_1 > 0$ then $Ef(Z_{\sigma(t)})/E\sigma(t)$ is nondecreasing in t. If in addition $\{Z_t | t \in \mathscr{R}^+\}$ is nondecreasing, then $Ef(Z_t)/t$ and $Ef(Z_{\tau \wedge \sigma(t)})/E\tau \wedge \sigma(t)$ are nondecreasing in t for any stopping time τ.*

PROOF OF COROLLARIES 4.1 AND 4.2. If f is convex on $(0, \infty)$, then it has a nondecreasing density f'. Hence, for any nonnegative random variable Y,

$$Ef(Y) - f(0)P[Y = 0] - f(0+)P[Y > 0] = E\int_0^Y f'(y)\,dy \tag{4.1}$$

(both sides being finite and equal or both infinite). If $EY < \infty$ then from (2.2) the right-hand side is $EYEf'(Y^e)$, so that

$$\frac{Ef(Y)}{EY} = f(0+)\frac{1}{EY} - [f(0+) - f(0)]\frac{P[Y = 0]}{EY} + Ef'(Y^e). \tag{4.2}$$

The result follows upon replacing Y by S_x, $x = n(t), k, \nu \wedge n(t)$ or by Z_x, $x = \sigma(t), t, \tau \wedge \sigma(t)$, applying Corollary 2.1 and Theorem 4.1 or Theorem 4.2, respectively, and noting that both $1/EY$ and $P[Y = 0]/EY$ become nonincreasing when Y is replaced by any nondecreasing process. □

5. Counterexamples. As discussed in the paragraph following Corollary 2.1, in this section we will explore certain natural directions in which one could try to generalize the results of Section 2, which were not discussed in Section 4. We show that these directions are fruitless. Here we will only consider the stationary excess time. Clearly a negative result regarding the stationary excess distribution trivially carries over to the joint distribution of the stationary age and access, and hence also to the stationary current life distribution. Under the conditions of Theorem 2.1 this would also imply a negative result for S^*.

EXAMPLE 5.1. When Y is deterministic, it is clear that $Y^c = Y$ and Y^e is uniformly distributed on $(0, Y)$ (Lemma 2.1). Hence if $Y(t)$ is an arbitrary (deterministic) increasing function, then so is $Y^c(t)$, and $Y^e(t)$ is clearly SI. Can it be the case that $Y^e(t)$ is SI for any nondecreasing process $Y(t)$? The answer is no. To see this, take Y_1 to be geometrically distributed (that is, $P[Y_1 = n] = (1-p)^n p$, $n = 0, 1, \dots$) and let $Y_2 = Y_1 + a$ where $0 < a < 1$. Then clearly $Y_1 < Y_2$. As $2EY^e = EY^2/EY$, it suffices to show that $EY_1^2/EY_1 > EY_2^2/EY_2$ in order to obtain a contradiction. It is easy to check that, with $q = 1 - p$,

$$\begin{aligned} \frac{EY_1^2}{EY_1} &= \frac{1+q}{p}, \\ \frac{EY_2^2}{EY_2} &= \frac{1+q}{p} - \frac{pa(1-a)}{q+pa}. \end{aligned} \tag{5.1}$$

EXAMPLE 5.2. Since both $\{k | k \in \mathscr{N}\}$ and $\{\nu \wedge n(t) | t \in \mathscr{R}^+\}$ in Theorem 2.2 are nondecreasing families of stopping times, is it the case that under the same assumptions $S_{\nu_1}^e \leq_{\text{st}} S_{\nu_2}^e$ for any two finite mean stopping times $\nu_1 \leq \nu_2$? If this is too much to hope for, then perhaps it is possible to verify the far more modest conjecture that $\{S_{k \wedge n(t)}^e | k \in \mathscr{N}\}$ is SI? Again, the answer is no. Let $S(t) = S_{n(t)}$ and $S_k(t) = S_{k \wedge n(t)}$. To obtain a counterexample, assume that $P[X = 1] = P[X = 8] = 1/2$. Then $P[S_2(8) = 2] = P[S_2(8) = 16] = P[S_3(8) = 16] = 1/4$, $P[S_2(8) = 9] = P[S_3(8) = 9] = 1/2$ and $P[S_3(8) = 3] = P[S_3(8) = 10] = 1/8$. Hence,

$$\frac{ES_2^2(8)}{ES_2(8)} = 11 + \frac{2}{3} + \frac{1}{18} > 11 + \frac{2}{3} = \frac{ES_3^2(8)}{ES_3(8)}. \tag{5.2}$$

EXAMPLE 5.3. In view of Remark 2.2, an immediate question is whether for an exchangeable sequence $\{S_{\nu \wedge n(t)}^e | t \in \mathscr{R}^+\}$, or even $\{S_{n(t)}^e | t \in \mathscr{R}^+\}$, is SI in t. Here too the answer is no. In particular, let $X = X_1 = X_2 = \cdots$ with $P[X = 1] = P[X = 3] = 1/2$. Then $P[S(3) = 4] = P[S(3) = 6] = 1/2$ as well as $P[S(4) = 5] = P[S(4) = 6] = 1/2$ [recall that $S_{n(t)}$ is strictly greater than t, rather than just greater than or equal to]. In particular, it is easy to check that

$$P[S^e(3) > 5] = \tfrac{1}{10} > \tfrac{1}{11} = P[S^e(4) > 5]. \tag{5.3}$$

EXAMPLE 5.4. Assuming that we insist on the independence of $\{X_i | i \in \mathscr{N}\}$, could we do away with the identical distribution assumption? Of course, if ν

and all the X_i's are deterministic, then so is $S_{\nu \wedge n(t)}$ and hence the results would clearly hold. However, once we introduce randomness, the results do not hold in general. For example assume that all random variables are exponentially distributed, with $EX_1 = 1$ and $EX_i = 1/2$ for $i \geq 2$. Then $S_{n(t)} = t + \gamma(t)$ where, by the memoryless property, $\gamma(t)$ is a $(e^{-t}, 1 - e^{-t})$ mixture of X_1 and X_2. With this in mind, both $ES^2_{n(t)}$ and $ES_{n(t)}$ can be easily computed. Simple differentiation gives that $ES^2_{n(t)}/ES_{n(t)}$ is strictly decreasing on $[0, t_0)$ where $t_0(t_0 + 1) = e^{-t_0}$. To check that ES^2_k/ES_k is not necessarily increasing, take any independent X_1 and X_2 with $EX_1 = EX_2 = 1$ and $\mathrm{Var}(X_1) > \mathrm{Var}(X_2) + 2$, in which case $ES^2_1/ES_1 > ES^2_2/ES_2$.

Lemma 2.2 states that $Y^c \leq_{\mathrm{st}} Z^c$ is equivalent to $(Y^a, Y^e) \leq_{\mathrm{st}} (Z^a, Z^e)$. The following example shows that just $Y^e \leq_{\mathrm{st}} Z^e$ does not imply $Y^c \leq_{\mathrm{st}} Z^c$.

EXAMPLE 5.5. Consider two random variables Y and Z with $P[Z = 1] = 2/3$, $P[Z = 2] = 1/3$ and $P[Y = 2/3] = 6/11$, $P[Y = 4/3] = 3/11$, $P[Y = 2] = 2/11$. Then $P[Z^c = 1] = P[Z^c = 2] = 1/2$ and $P[Y^c = 2/3] = P[Y^c = 4/3] = P[Y^c = 2] = 1/3$. We observe that $Y^c \leq_{\mathrm{st}} Z^c$ does not hold. We apply Lemma 2.1 and compute the distribution functions of X^e and Z^e:

$$P[Z^e \leq t] = \begin{cases} 0, & \text{if } t \leq 0; \\ \dfrac{3}{4}t, & \text{if } 0 < t \leq 1; \\ \dfrac{1}{2} + \dfrac{t}{4}, & \text{if } 1 < t \leq 2; \\ 1, & \text{if } t > 2; \end{cases} \tag{5.4}$$

and

$$P[Y^e \leq t] = \begin{cases} 0, & \text{if } t \leq 0; \\ \dfrac{11}{12}t, & \text{if } 0 < t \leq \dfrac{2}{3}; \\ \dfrac{1}{3} + \dfrac{5}{12}t, & \text{if } \dfrac{2}{3} < t \leq \dfrac{4}{3}; \\ \dfrac{2}{3} + \dfrac{t}{6}, & \text{if } \dfrac{4}{3} < t \leq 2; \\ 1, & \text{if } t > 2. \end{cases} \tag{5.5}$$

It is easy to see that $Y^e \leq_{\mathrm{st}} Z^e$.

Acknowledgment. We thank Joe Chang for useful discussions.

REFERENCES

ALTMAN, E. and NAIN, P. (1993). Optimal control of an $M/G/1$ queue with vacations. *IEEE Trans. Automat. Control* **38** 1766–1775.

BELL, C. (1971). Characterization and computation of optimal policies for operating an $M/G/1$ queueing system with removable server. *Oper. Res.* **19** 208–218.

CHUNG, K. L. (1974). *A Course in Probability Theory*, 2nd ed. Academic Press, New York.

FEDERGRUEN, A. and SO, K. C. (1991). Optimality of threshold policies in single server queueing systems with server vacations. *Adv. in Appl. Probab.* **23** 388–405.

FEINBERG, E. A. and KIM, D. J. (1996). Bicriterion optimization of an $M/G/1$ queue with a removable server. *Probability in the Engineering and Informational Sciences* **10** 57–73.

FUHRMANN, S. and COOPER, R. B. (1985). Stochastic decompositions in an $M/G/1$ queue with generalized vacations. *Oper. Res.* **33** 1117–1129.

HEYMAN, D. P. (1968). Optimal operating policies for $M/G/1$ queueing systems. *Oper. Res.* **16** 362–383.

HOFRI, M. (1986). Queueing systems with a procrastinating server. In *Proceedings of the Joint Conference on Computer Performance Modeling*, *Measurement and Evaluation*, *Raleigh*, *NC* 245–253.

KARLIN, S. and TAYLOR, H. M. (1975). *A First Course in Stochastic Processes*, 2nd ed. Academic Press, New York.

KELLA, O. (1989). The threshold policy in the $M/G/1$ queue with server vacations. *Nav. Res. Logist.* **36** 111–123.

KELLA, O. and WHITT, W. (1991). Queues with server vacations and Lévy processes with secondary jump input. *Ann. Appl. Probab.* **1** 104–117.

KIM, D. J. (1995). Bicriterion optimization of manufacturing systems. Ph.D. dissertation, Dept. Applied Mathematics and Statistics, SUNY, Stony Brook.

LEE, H. W., LEE, S. S. and CHAE, K. C. (1994). Operating characteristics of $M^x/G/1$ queue with N-policy. *Queueing Systems* **15** 387–399.

LEE, H. S. and SRINIVASAN, M. (1989). Control policies for the $M/G/1$ queueing system. *Management Sci.* **35** 708–721.

MARSHALL, A. W. and OLKIN, I. (1979). *Inequalities: Theory of Majorization and Its Applications*. Academic Press, New York.

SHANTHIKUMAR, J. G. (1988). On stochastic decomposition in $M/G/1$ type queues with generalized server vacations. *Oper. Res.* **36** 69–83.

SOBEL, M. J. (1969). Optimal average cost policy for a queue with start-up and shut-down costs. *Oper. Res.* **17** 145–162.

STOYAN, D. (1983). *Comparison Methods for Queues and Other Related Stochastic Models*. Wiley, New York.

YADIN, M. and NAOR, P. (1963). Queueing systems with a removable service station. *Oper. Res. Quart.* **14** 393–405.

E. V. DENARDO
CENTER FOR SYSTEMS SCIENCES
YALE UNIVERSITY
NEW HAVEN, CONNECTICUT 06520
E-MAIL: denardo@yalevm.cis.yale.edu

E. A. FEINBERG
HARRIMAN SCHOOL OF
MANAGEMENT AND POLICY
SUNY
STONY BROOK, NEW YORK 11794-3775
E-MAIL: efeinber@fac.har.sunysb.edu

O. KELLA
DEPARTMENT OF STATISTICS
HEBREW UNIVERSITY OF JERUSALEM
MOUNT SCOPUS, JERUSALEM 91905
ISRAEL
E-MAIL: mskella@olive.mscc.huji.ac.il