IIE Transactions
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/uiie20

THE EFFECT OF CORRELATED ARRIVALS ON QUEUES

B. EDDY PATUWO [a], RALPH L. DISNEY [b] & DONALD C. McNICKLE [c]

[a] Department of Administrative Sciences, Kent State University, Kent, OH, 44241-0001

[b] Department of Industrial Engineering, Texas A&M University, College Station, TX, 77843-3131

[c] Department of Operations Research, University of Canterbury, Christchurch, New Zealand

Published online: 30 May 2007.

To cite this article: B. EDDY PATUWO, RALPH L. DISNEY & DONALD C. McNICKLE (1993) THE EFFECT OF CORRELATED ARRIVALS ON QUEUES, IIE Transactions, 25:3, 105-110, DOI: 10.1080/07408179308964296

To link to this article: http://dx.doi.org/10.1080/07408179308964296

THE EFFECT OF CORRELATED ARRIVALS ON QUEUES

B. EDDY PATUWO
Department of Administrative Sciences, Kent State University, Kent, OH 44241-0001

RALPH L. DISNEY
Department of Industrial Engineering, Texas A&M University, College Station, TX 77843-3131

DONALD C. McNICKLE
Department of Operations Research, University of Canterbury, Christchurch, New Zealand

Using Markov renewal arrival processes, a study of the effect of serial correlations in the arrival process on the mean queueing performance has been done. We show that positive serial correlations may have major impact on the mean queue lengths (and consequently on other performance measures).

■ Queueing theory has been used to model nearly every aspect of the U.S. economy from timber harvesting (see C. C. Hassler, *et al.* [6]) to flexible manufacturing systems (see Buzacott and Yao, *et al.* [1]). Indeed, not only does one find analytic results of abundance but one is impressed with the large number of computer simulation packages addressing queueing and other congestion problems. *Queueing Systems: Theory and Applications* (QUESTA, [14]) has devoted an issue to various applications of the theory.

Throughout most of the history of queueing, various independence conditions have been made. For example, it is common to see studies of queues with GI-arrivals (i.e., interarrival times are i.i.d. sequences of random variables).

Yet there seem to be scant results in the literature as to whether or not non-renewal arrival processes have any major impact on queueing properties. The purpose of this paper is to investigate this question in some detail.

The dependence among successive interarrival times has been evident in the studies of packet multiplexers for voice and data (Sriram and Whitt [18]). For an extensive simulation study of the effects of autocorrelations in both arrival as well as service times, see the report by Livny, *et al.* [9]. They also discuss the issue of applications in computer systems design.

We consider a queue with Markov renewal arrival processes to an exponential server. Queue disciplines are all FCFS, though for the results reported here that is not of consequence. Queue capacity is infinite and we are concerned with steady state performance measures. We call our system an MR/M/1 queue. There are other papers on this topic available (e.g., Cinlar [2], Neuts [10], [11]). But, to the best of our knowledge, this is the first numerical study of these queues attempting to gain insight into the effect of the dependency in the arrival process.

Markov renewal arrival processes occur in studies introducing "types," or priorities, or any one of a number of other means used to distinguish one arrival from another. It can be noted that the output process from M/GI/1/L for $L < \infty$ queues is a Markov renewal process and in the case $GI = E_k$ this process can have positive, negative or zero correlation depending on ϱ, k, L. Thus, if this output is used as input to a second queue, one can observe examples of the phenomenon considered in this paper. For details see Disney and Kiessler ([3], Chapter 6). The new idea here is that interarrival times can be correlated.

The result presented by Tin [19] caught our attention. He showed that the difference in the mean queue lengths between correlated and uncorrelated arrival queues can differ by a factor of 9. We will show later that worse cases than those reported by Tin can occur.

Earlier studies of correlated arrival queues are found in Gopinath and Morrison [5], Latouche [7], [8], and Runnenburg [15], [16]. Recently, the importance of the dependent arrival queues in the area of data and voice communication has been the subject of a series of papers published by Whitt and colleagues (see, for example, Sriram and Whitt [18], and Fendick, Saksena, and Whitt [4]).

Our contribution in this area is to the systematic study of the effect of the serial correlation of the arrival process on the mean queue length. A longer report has investigated these effects on standard deviations of the queue length and the caudal characteristic (a measure of the tail probabilities of the queue length distribution; see Neuts [12]). As might be expected (because these other measures also depend on the behavior of the queue

0740-817X/93/$3.00x.00

length process), there are added effects to those we report here, but in the cases we have considered the added effects are small relative to the results we report herein. Therefore, we do not pursue those topics here. Interested readers can find a more complete discussion in Patuwo [13]. (It should be noted that some of the matrices involved in our later computations are badly conditioned. Thus, in this reference, one needs to add double precision arithmetic to the program given therein. We have done that in preparing this paper.)

There are basically two queue length processes one could examine in a study such as ours: the queue length process embedded at arrival epochs, and the continuous time queue length process. Because of identities (21), (22), (23) which follow, one can expect similar behavior in either. The continuous time process strikes us as more meaningful to application and this is the one we attack. The embedded process would be more relevant to waiting time analysis, but as noted above we do not discuss that topic here. However, the embedded process is a means to our end, thus, a discussion of it preceeds our main results.

The Arrival Process

Let T^a be the arrival time process with elements T_n^a, $n = 0,1,2,\ldots$. We take $T_0^a = 0$ a.s. and let D^a be the interarrival time process with elements $D_{n+1}^a = T_{n+1}^a - T_n^a$. Let Z_n be the type of the n-th arrival and Z^a be the arrival-type process with state space $\{1,2\}$. Furthermore, let the semi-Markov kernel of the two-state Markov renewal arrival process (Z^a, D^a) be:

$$\mathbf{A}(t) = \begin{bmatrix} aF_1(t) & (1-a)F_2(t) \\ (1-b)F_1(t) & bF_2(t) \end{bmatrix}. \tag{1}$$

Here $F_i(t)$, $i = 1,2$, are Erlang c.d.f.'s with parameter λ_i and k_i where k_i is an integer, and a and b are constants between 0 and 1.

Consequently, the transition probability matrix of the underlying Markov chain Z^a *is:*

$$\mathbf{A} = \begin{bmatrix} a & (1-a) \\ (1-b) & b \end{bmatrix} \tag{2}$$

and the stationary probability vector of Z^a is $\pi = (\pi_1, \pi_2)$, where $\pi_1 = (1-b)/(1-\xi)$, $\pi_2 = (1-a)/(1-\xi)$, and $\xi = a + b - 1$. Notice that ξ is the subdominant eigenvalue of the stochastic matrix $\mathbf{A}$.

The marginal stationary interarrival time has the mixture of Erlang c.d.f.:

$$Pr[D_n^a \le t] = \pi_1 F_1(t) + \pi_2 F_2(t). \tag{3}$$

Letting $m_i = 1/\lambda_i$, $i = 1,2$,

$$E[D_n^a] = \pi_1 m_1 + \pi_2 m_2, \tag{4}$$

$$Var[D_n^a] = ((\pi_1 m_1^2 + \pi_2 m_2^2)/k) + \pi_1\pi_2(m_1-m_2)^2. \tag{5}$$

The lag-r serial correlation of the interarrival times turns out to be (see Patuwo [13]):

$$corr(r) = \pi_1\pi_2(m_1-m_2)^2\,\xi^r/Var[D_n^a], \tag{6}$$

and the coefficient of skewness γ is:

$$\gamma = \left\{\frac{(k+1)(k+2)}{k^2}(\pi_1 m_1^3 - \pi_2 m_2^3) - 3E[D_n^a]\,Var[D_n^a] - E[D_n^a]^3\right\}/\{Var[D_n^a]\}^{3/2}. \tag{7}$$

Let *corr* denote the lag-1 correlation coefficient of the interarrival times. Then the serial correlation (6) is characterized by the lag-1 correlation and the rate ξ:

$$corr(r) = corr \cdot \xi^{r-1}. \tag{8}$$

When $\xi > 0$, the serial correlation is strictly positive, and when $\xi < 0$, the serial correlation alternates in sign.

If $\xi = a + b - 1 = 0$ then the interarrival times are i.i.d. or the T^a process is a renewal process. This can be shown easily using one of the equivalence conditions given in Simon and Disney [17] or Disney and Kiessler [3].

A number of authors have pointed out that especially when the traffic intensity is high it is the correlations of all lags that influence the queue characteristics, not only the lag-1 correlation. From Sriram and Whitt [18], for example, the index of dispersion for intervals (IDI) is:

$$C_j^2 = \frac{var(D_n^a) + 2\sum_{r=1}^{j-1}\left(1-\frac{r}{j}\right)cov(D_n^a, D_{n+r}^a)}{E[D_n^a]^2} = \frac{var(D_n^a)}{E[D_n^a]^2}\left(1+2\sum_{r=1}^{j-1}\left(1-\frac{r}{j}\right)corr(r)\right),$$

a normalized variance taking account of some of the correlation in the arrival process. For the class of Markov renewal processes we are considering,

$$C_j^2 = scv\left(1+2\,corr\sum_{r=1}^{j-1}\left(1-\frac{r}{j}\right)\xi^{r-1}\right), \tag{9}$$

thus

$$C_\infty^2 = scv\left(1+\frac{2\,corr}{1-\xi}\right). \tag{10}$$

The Queueing Properties

We embed the queue process N^a at arrival times and call N_n^a the queue length (the number of customers in the system) at the n-th arrival epoch. In general, the queue length process of this Markov renewal queue is not Markov. But the process (N^a, Z^a, T^a) is Markov renewal and the joint process (N^a, Z^a) is the underlying Markov chain on $\{0,1,2,\ldots\} \times \{1,2\}$.

The transition matrix for the Markov chain (N^a, Z^a) has the familiar lower Hessenberg form of the GI/M/1 queue, though the elements are now 2 by 2 block matrices. We will call the elements on the n-th diagonal of the transition matrix $\mathbf{A}_n$ where $n = i - j + 1$, $i = 0,1,2,\ldots$, $j = 1,2,\ldots,i+1$.

What we have is a simple example of a queue with phase-type arrivals, (see Neuts [10]), so if the traffic intensity $\varrho = 1/(\mu . E[D_n^a]) < 1$, the limiting queue length probabilities are given by:

$$\lim_{n\to\infty} P[N_n^a = i] = \pi(\mathbf{I}-\mathbf{R})\mathbf{R}^i\mathbf{e}, \tag{11}$$

where $\mathbf{e}$ is the column vector of ones and $\mathbf{R}$ is the minimal nonnegative solution to the matrix polynomial equation:

$$\mathbf{R} = \sum_{n=0}^{\infty} \mathbf{R}^n \mathbf{A}_n, \tag{12}$$

and

$$\mathbf{A}_n = \int_0^{\infty} \frac{e^{-\mu t}(\mu t)^n}{n!} d\mathbf{A}(t). \tag{13}$$

In the numerical results to follow, we have used a truncated version of (12) in all of our computations.

In the next section we give the mean queue length embedded at arrival times. Then we exhibit the continuous time queue length process and tie together the embedded and the continuous time results.

Expected Queue Length Seen by Arriving Customers

Let L^a and L_q^a be the expected number of customers in the system and the expected number of customers waiting in the queue seen by arriving customers, respectively. Then:

$$L^a = E[N^a] = \pi(\mathbf{I}-\mathbf{R})^{-1}\mathbf{R}\mathbf{e}, \tag{14}$$

$$L_q^a = \pi(\mathbf{I}-\mathbf{R})^{-1}\mathbf{R}^2\mathbf{e}, \tag{15}$$

$$Var[N^a] = \pi(\mathbf{I}-\mathbf{R})^{-2}\mathbf{R}(\mathbf{I}+\mathbf{R})\mathbf{e}-(L^a)^2. \tag{16}$$

The Continuous Time Queue Length Process

Let $N(t)$ be the number of customers in the system at time t. Then from Neuts [10]:

$$\lim_{t\to\infty} P[N(t)=0] = 1 - \varrho, \tag{17}$$

$$\lim_{t\to\infty} P[N(t)=i] = \varrho\pi(\mathbf{I}-\mathbf{R})^{i-1}\mathbf{e},\ i = 1,2,\ldots \tag{18}$$

Let L^t and L_q^t be the expected number of customers in the system and the expected number of customers waiting in the queue at arbitrary time. Then

$$L^t = \varrho\pi(\mathbf{I}-\mathbf{R})^{-1}\mathbf{e}, \tag{19}$$

$$L_q^t = \varrho\pi(\mathbf{I}-\mathbf{R})^{-1}\mathbf{R}\mathbf{e} = \varrho\mathbf{L}^a. \tag{20}$$

The following 3 identities tie together the Little-like results for the embedded and the continuous time mean queue length:

$$L^t = L_q^a + \varrho, \tag{21}$$

$$L_q^t = \varrho L^a, \tag{22}$$

$$L^t = \varrho(L^a+1). \tag{23}$$

Results

(a) There are 5 parameters of the arrival process that can be used in the investigation of the mean queue length L^t of the MR/M/1 queue. They are: the mean, the variance, the coefficient of skewness, the lag-1 correlation coefficient, and the rate, ξ.

In order to fix as many parameters as possible we consider here only arrival processes whose mean, variance, and coefficient of skewness are the same as those of frequently used theoretical arrival processes: Poisson, Erlang-2, and hyperexponential. The remaining parameter, ξ, is fixed by selecting the smallest possible value of the number of stages, k_i, in the Erlang distributions $F_1(t)$ and $F_2(t)$. Extensive investigations of alternative parameter settings can be found in Patuwo [13]. In particular, it is shown there that the shapes of the L^t plots do not depend strongly on the value of k_i. It should be emphasized that this does not mean that the marginal distributions of the interarrival times are Erlang-2, Poisson, or hyperexponential, merely that they have matching first three moments.

The results are plotted against values of the lag-1 correlation, *corr*, in Figures 1 to 3. They show uniform increases in L^t with increasing values of *corr*.

Taking Figure 2, "Poisson-like Arrivals," (i.e., scv = 1, γ = 2), some values for the mean number in the system are:

	Corr. 0.0	0.1	0.4	0.65
ϱ=0.5	1.0	1.0046	1.2509	3.504
ϱ=0.7	2.3333	2.4877	3.7568	36.20
ϱ=0.9	9.0	0.9421	17.013	131.92

The values for a correlation of zero are calculated from the usual formula for the mean number in an M/M/1 queue. Thus we see increases in the mean number in the system, varying upward from about one order of magnitude as the correlation coefficient increases

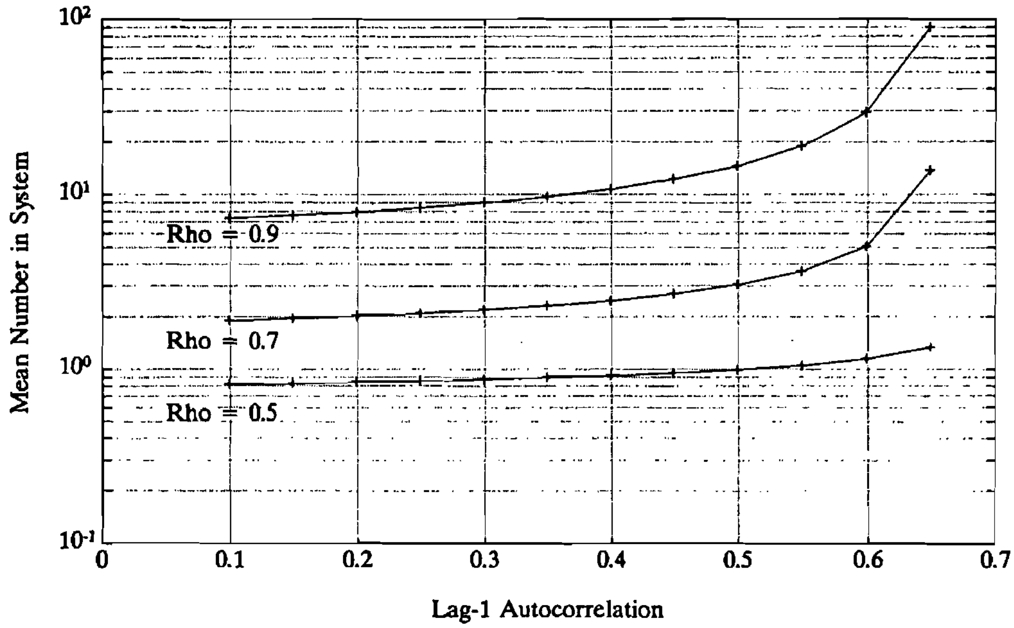

Figure 1. Erlang-2-like arrivals

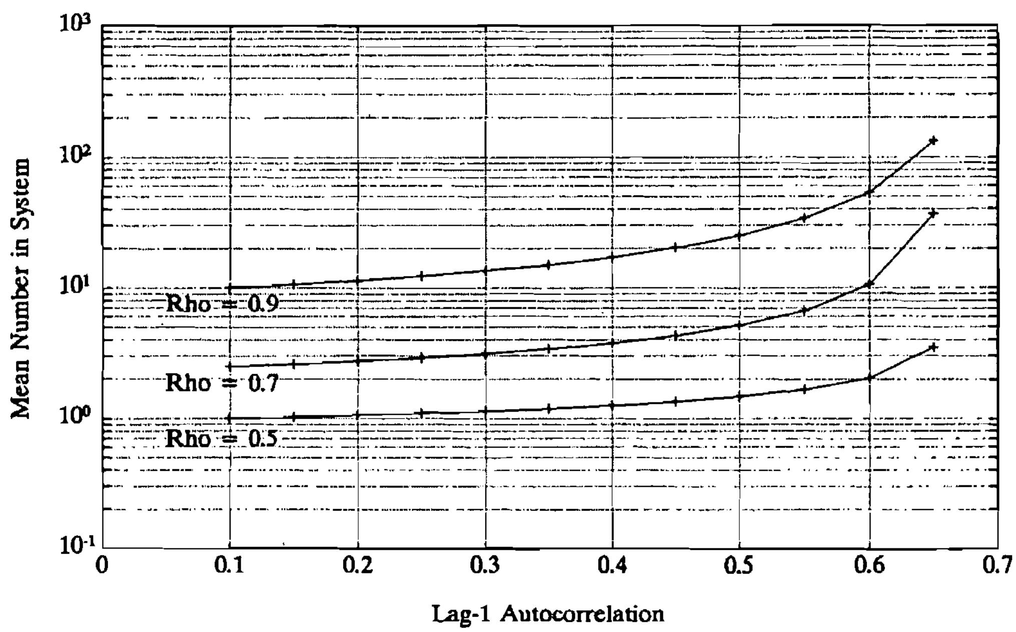

Figure 2. Poisson-like arrivals

to 0.65. Even small correlations of up to 0.4 can cause the mean number in the system to nearly double. We note that if sample sizes are small, correlations of this size can be quite hard to detect in an observed arrival process.

Another point worth noting is the strong family resemblance among Figures 1-3. While the effect of correlation is slightly more pronounced for the less variable (Erlang-2-like) arrival process, the effects are much the same across the three systems. Thus there appears to be a consistent effect due to correlation, which is largely independent of the form of the marginal distribution of the arrival process.

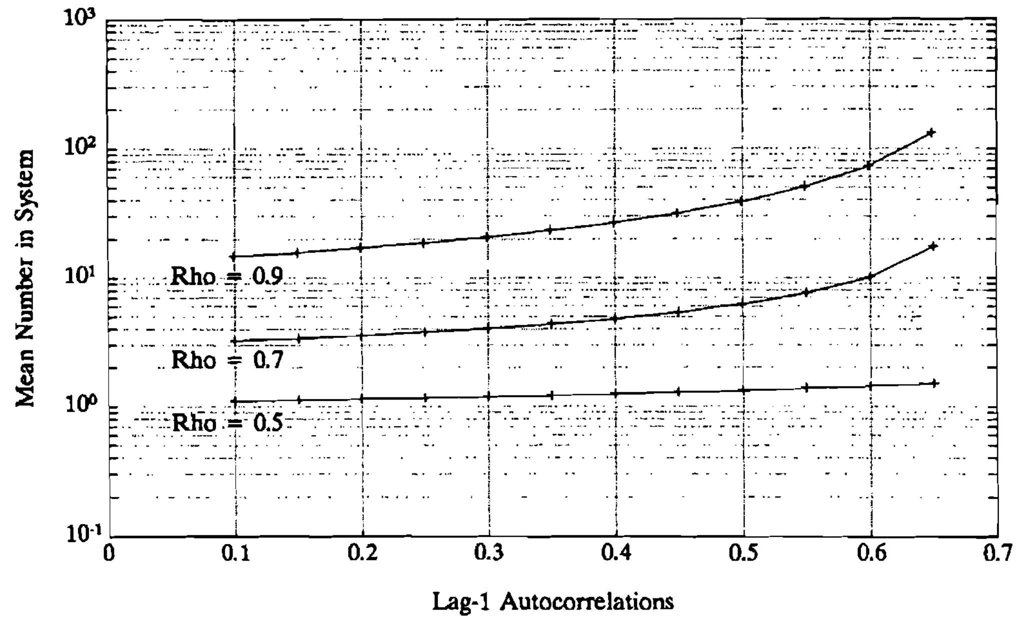

Figure 3. Hyperexponential-like arrivals

(b) Our purpose in part (a) above is to exhibit the effects of lag-1 correlation. However, it is known that, especially when traffic is heavy, the correlations of all lags are important, and that in fact the index of dispersion for intervals (IDI) is an appropriate measure of this effect. (See Sriram and Whitt [18]). In Figure 4 values of the mean number in the system produced by the "Poisson-like" arrival process have been plotted against the IDI values, calculated from (10).

The striking thing is the almost perfect linear response for traffic intensities greater than 0.6. It should be noted, however, that in practice it is often not easy to estimate

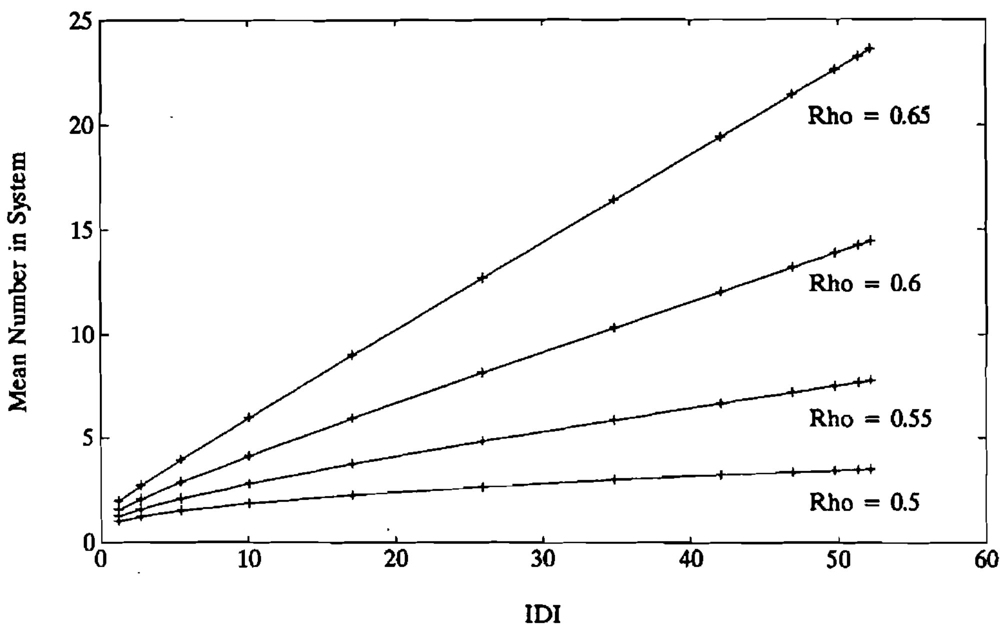

Figure 4. Poisson-like arrivals, Lt vs IDI

the IDI from a small set of observations. Also, for Markov renewal processes with more than two states the lag-r correlations do not have the simple geometric form (8). Hence calculating the IDI is much more difficult and depends on the complete eigenstructure of A. Thus we do not pursue this further in this paper.

Conclusions

Our investigation indicates that any queueing study should consider arrival processes with more dependence structure than has been considered for most of the history of queueing. That is, the renewal assumptions of GI or G queueing are providing results that for the user as well as for the producer may be seriously flawed. The point is that the results we have found appear to be consequences of correlation and not of distributional assumptions. Furthermore, approximations that replace correlated (e.g., Markov renewal) with renewal arrival processes may be quite misleading even if all one seeks is a mean queue length.

Acknowledgements

This paper was written while the third author was visiting the Industrial Engineering Department of Texas A&M University; the support of the Department and Prof. G. Kemble Bennett, Department Head, is gratefully acknowledged. The work in this paper was supported in part under NSF Grant DDM-8913658.

REFERENCES

[1] Buzacott, J. A., and Yao, D. D., "On Queueing Network Models of Flexible Manufacturing Systems," *QUESTA*, 1, 5-27 (1986).

[2] Cinlar, E., "Queues with Semi-Markovian Arrivals," *J. Appl. Probab.*, 4, 365-379 (1967).

[3] Disney, R. L., and Kiessler, P. C., *Traffic Process in Queueing Networks: A Markov Renewal Approach*, Johns Hopkins Univ. Press, Baltimore (1987).

[4] Fendick, K. W., Saksena, V. R., and Whitt, W., "Dependence in Packet Queues," *IEEE Trans. Commun.*, 37, 1173-83 (1989).

[5] Gopinath, B., and Morrison, J. A., "Single Server Queues with Correlated Inputs," *Computer Performance*, K. M. Chandy and M. Reiser, eds., North-Holland Publ. Co., 264-277 (1977).

[6] Hassler, C. C., Disney, R. L., and Sinclair, S. A., "A Discrete State, Continuous Parameter Markov Approach to Timber Harvesting System's Analysis," *Forest Sci.*, 34, 276-291 (1988).

[7] Latouche, G., "On a Markovian Queue with Weakly Correlated Interarrival Times," *J. Appl. Probab.*, 18, 190-223 (1981).

[8] Latouche, G., "An Exponential Semi-Markov Process with Applications to Queueing Theory," *Stochastic Models*, 1, 137-170 (185).

[9] Livny, M., Melamed, B., and Tsiolis, A. K., "The Impact of Autocorrelation on Queueing Systems," *Management Science*, 39, 322-339 (1993).

[10] Neuts, M. F., "Markov Chains with Applications in Queueing Theory, Which Have a Matrix-Geometric Invariant Probability Vector," *Adv. in Appl. Probab.*, 10, 185-212 (1978).

[11] Neuts, M. F., *Matrix Geometric Solutions in Stochastic Models*, Johns Hopkins University Press, Baltimore (1981).

[12] Neuts, M. F., "The Caudal Characteristic Curve of Queues," *Adv. in Appl. Probab*, 18, 221-254 (1986).

[13] Patuwo, B. E., "The Effects of the Dependency in the Markov Renewal Arrival Process on the Various Performance Measures of an Exponential Server Queues," Ph.D. thesis, Department of Industrial Engineering and Operations Research, Virginia Polytechnic Institute and State University (1989).

[14] *Queueing Systems: Theory and Applications*, Vol. 7, #1 (October, 1990).

[15] Runnenburg, J. Th., "An Example Illustrating the Possibilities of Renewal Theory and Waiting Line Theory for Markov Dependent Arrival Intervals," *Nederl. Akad. Wetensch. Pro. Ser.*, A, 64, 560-576 (1961).

[16] Runnenburg, J. Th., "Some Numerical Results on Waiting Time Distributions for Dependent Arrival-Intervals," *Statist. Nederlandica*, 12, 19-29 (1962).

[17] Simon, B., and Disney, R. L., "Some Conditions for the Equivalence Between Markov Renewal Processes and Renewal Processes, *New Zealand J. of Oper. Res.*, 12, 19-29 (1983).

[18] Sriram, K., and Whitt, W., "Characterizing Superposition Arrival Processes in Packet Multiplexers for Voice and Data," *IEEE J. on Selected Areas in Comm.*, SAC-4, 833-846 (1986).

[19] Tin, P., "A Queueing System with Markov Dependent Arrivals, *J. Appl. Probab.*, 22, 668-677 (1985).

B. Eddy Patuwo is an Assistant Professor in the Department of Administrative Sciences at Kent State University. He received a Ph.D. in Industrial Engineering and Operations research from Virginia Polytechnic Institute and State University. His research interests are in queueing theory and production/inventory systems.

Ralph L. Disney is Brockett Professor of Industrial Engineering at Texas A&M University, a Fellow of IIE, and has served in many capacities for that organization as well as other professional societies. He is the author of many articles and several books, the latest of which is *Traffic Processes in Queueing Network*, The Johns Hopkins University Press (1987), with Peter Kiessler. His interests are primarily in random processes and applications.

Don McNickle is a Senior Lecturer in the Operations Research Department, University of Canterbury, Christchurch, New Zealand. His research interests are in networks of queues, and problems in applied probability and statistics.

Received June 1990; revised March 1991. Handled by the Department of Engineering Statistics and Applied Probability.