

SINGLE SERVER QUEUES WITH ARCH-TYPE DEPENDENCE

JASON S. ROTH AND N. U. PRABHU

ABSTRACT. In this paper, we propose a class of single server queueing models in which the interarrival times or service times exhibit an ARCH-type (autoregressive conditionally heteroskedastic) dependence. The main feature of this dependence is the volatility that depends on the sample functions of the underlying process. We denote this property as QARCH. To illustrate the analysis involved, we investigate the system M/QARCH/1, and derive upper and lower bounds for its mean queue length in steady state expressed in terms of the mean queue length in M/G/1 with the same first two moments of the service time.

1. INTRODUCTION

The basic assumption underlying the GI/G/1 queue is that the interarrival times $\{u_n, n \geq 1\}$ and the service times $\{v_n, n \geq 1\}$ are two independent sequences of independent and identically distributed (IID) random variables. Attempts have been made to relax this assumption, namely, to introduce dependence within the terms of these sequences or between them. It is the nature of the problem that a single consolidated theory that incorporates all of these dependencies cannot be expected to emerge from these attempts. For this reason, although the literature in this area is fairly extensive, the only results obtained are for models with a specific degree of dependence. In broad terms, the problem is to model $\{(u_n, v_n), n \geq 1\}$ as a bivariate time series and apply to it the techniques of time series analysis. In this paper, we cite only relevant work that deals with $\{u_n\}$ or $\{v_n\}$. The earliest paper in this area is due to Finch (1963) who considered a queue with an arbitrary sequence $\{u_n\}$ and a sequence $\{v_n\}$ of IID random variables with Erlangian density (including the special case of exponential density). A further investigation of this model was carried out by Finch and Pearce (1965) in the case where $\{u_n\}$ is a moving average process and $\{v_n\}$ is an IID sequence with exponential density. Pearce (1965) extended these results to the system with batch arrivals. Pearce (1967) also considered the queue where $\{v_n\}$ is a moving average process. In all of this work, the main object of study is the queue length, and the approach is analytical.

Jacobs and Lewis (1977) constructed a sequence of dependent exponential random variables as follows. Let $\{e_n, n \geq 1\}$ be a sequence of IID random variables with density $\lambda e^{-\lambda x}$, and let $\{J_n, n \geq 1\}$ and $\{K_n, n \geq 1\}$ be two independent sequences of IID Bernoulli random variables. For $n \geq 1$, put

$$X_n = \beta e_n + J_n A_{n-1}$$

where

$$A_n = \rho A_{n-1} + K_n e_n.$$

The sequence $\{X_n\}$ is called EARMA (exponential mixed moving average autoregressive) process of order (1,1). If A_0 is chosen to be independent of all the other random variables and with density $\lambda e^{-\lambda x}$, then it turns out that $\{X_n\}$ is a stationary sequence with marginal density $\lambda e^{-\lambda x}$ and positive autocorrelation. Jacobs (1978) investigated a closed queueing network with two nodes, where the service times at the first node are given by an EARMA process that is independent of

AMS 1991 subject classifications. Primary 90B15; secondary 60K250.

Key words and phrases. correlated interarrival and service times, QARCH, queue length process, recurrence relations, stationary time series, volatility.

the service times at the second node, which are IID with the exponential density. Jacobs (1980) also considered single server queues in which $\{(u_n, v_n)\}$ is a cross-correlated EARMA process, namely, the interarrival times and service times are EARMA processes constructed as above with a common $\{A_n\}$ sequence.

Finally, Resnick and Samorodnitsky (1997) have investigated the GI/M/1 queue in which the interarrival time sequence $\{u_n\}$ exhibits long range dependence (where the current state has a strong dependency on the remote past).

In this paper, we formulate an ARCH (autoregressive conditionally heteroskedastic) model for the service time sequence $\{v_n\}$. The ARCH concept is due to Engle (1982) and has been the subject of recent research in time series analysis. The main feature of an ARCH model is the volatility that depends on the sample functions of the underlying process. Our model for $\{v_n\}$ is described by the recurrence relation

$$(1) \quad v_n = (\alpha_0 + \alpha_1 v_{n-1} + \alpha_2 v_{n-2} + \cdots + \alpha_p v_{n-p}) e_n$$

where $\alpha_0, \alpha_1, \dots, \alpha_p$ ($n \geq p+1 \geq 2$) are nonnegative constants and $\{e_n, n \geq p\}$ is an IID sequence of positive random variables with $E(e_n) = 1$ and $E(e_n^2) = c^2$. In addition, we assume that e_n is independent of $v_{n-1}, v_{n-2}, \dots, v_{n-p}$. For given $v_{n-1}, v_{n-2}, \dots, v_{n-p}$, (1) shows that v_n has mean

$$(2) \quad \alpha_0 + \alpha_1 v_{n-1} + \alpha_2 v_{n-2} + \cdots + \alpha_p v_{n-p}$$

and variance

$$(3) \quad (\alpha_0 + \alpha_1 v_{n-1} + \alpha_2 v_{n-2} + \cdots + \alpha_p v_{n-p})^2 (c^2 - 1).$$

Thus, the conditional variance of v_n , given $v_{n-1}, v_{n-2}, \dots, v_{n-p}$ can be high or low depending on the magnitudes of $\alpha_i v_{n-i}$ ($i = 1, 2, \dots, p$). This conditional heteroskedasticity is what characterizes ARCH models such as (1).

We can rewrite (1) as

$$(4) \quad V_n = A_n V_{n-1} + B_n \quad (n \geq p)$$

where

$$(5) \quad V_n = \begin{pmatrix} v_n \\ v_{n-1} \\ \vdots \\ v_{n-p+1} \end{pmatrix}, \quad A_n = \begin{pmatrix} \alpha_1 e_n & \alpha_2 e_n & \cdots & \alpha_{p-1} e_n & \alpha_p e_n \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}, \quad B_n = \begin{pmatrix} \alpha_0 e_n \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Here B_n is a p -dimensional random vector and A_n is a $p \times p$ random matrix. Clearly, $\{(A_n, B_n), n \geq 1\}$ is a sequence of IID matrices such that (A_n, B_n) is independent of V_{n-1} . The difference equation (4) is of the type studied by Furstenberg and Kesten (1960) and Kesten (1973). We restrict ourselves to the case $p = 1$, in which case (1) reduces to

$$(6) \quad v_n = (\alpha_0 + \alpha_1 v_{n-1}) e_n \quad (n \geq 1).$$

We shall refer to this sequence $\{v_n\}$ as QARCH (queueing ARCH). In Figures 1 and 2 below, we have plotted sample paths of this sequence for the following special cases.

1. e_n has exponential density.
2. e_n has the distribution

$$(7) \quad P\{e_n = a\} = \begin{cases} p & \text{if } a = 1 - \beta, \\ 1 - 2p & \text{if } a = 1, \\ p & \text{if } a = 1 + \beta, \end{cases}$$

where $0 < p < 1$ and $0 < \beta < 1$. For certain choices of p and β , e_n models severe jumps that occur infrequently.

We consider a single server queue with Poisson arrivals at a rate λ ($0 < \lambda < \infty$) and QARCH service times; the appropriate notation for this queue is M/QARCH/1. Within this framework, other queues of interest are QARCH/M/1, GI/QARCH/1, etc.

In Section 2, we derive the properties of the sequence $\{v_n\}$. It turns out that under appropriate constraints on the parameters α_0, α_1, c^2 and with a proper choice of $v_0, \{v_n\}$ is a stationary time series. Our analysis follows the main outlines of the investigation of a somewhat more general model carried out by Embrechts et al. (1997). However, our approach is more elegant and can also be used to simplify these authors' proofs in some details. In Section 3, we consider the system M/QARCH/1. As in the classical system M/G/1, the epochs of departure provide a regenerative set for the queue length. However, we have to expand the state description to (Q_n, v_n) , where v_n is the service time of the departing customer. (In M/G/1, the marginal chain $\{Q_n\}$ is also Markovian, which is not the case here.) While we are unable to derive the steady state distribution of the process $\{(Q_n, v_n)\}$, we establish some bounds for the mean queue length.

In Section 4, simulation studies are carried out to support the conclusions of section 3 and also provide additional insights into the behavior of the queue length process.

2. PROPERTIES OF $\{v_n\}$

For a given v_0 , it is obvious that the recurrence relations (6) can be solved to yield an expression for v_n in terms of e_1, e_2, \dots, e_n . This expression will be the starting point of our discussion. We have the following.

Theorem 1. *For a given v_0 , the equations (6) have the unique solution*

$$(8) \quad v_n = \alpha_0 \sum_{j=0}^{n-1} \alpha_1^j e_n e_{n-1} \cdots e_{n-j} + \alpha_1^n v_0 e_1 e_2 \cdots e_n \quad (n \geq 1).$$

More generally, we have

$$(9) \quad v_{m+n} = \alpha_0 \sum_{j=0}^{n-1} \alpha_1^j e_{m+n} e_{m+n-1} \cdots e_{m+n-j} + \alpha_1^n v_m e_{m+1} e_{m+2} \cdots e_{m+n} \quad (m \geq 0, n \geq 1).$$

Proof. It suffices to prove (9) for $m \geq 0$ and $n \geq 1$. The result (8) will then follow for $m = 0$. For $m \geq 0, n = 1$, the right hand side of (9) equals

$$\alpha_0 e_{m+1} + \alpha_1 v_m e_{m+1} = (\alpha_0 + \alpha_1 v_m) e_{m+1} = v_{m+1}$$

by definition. This proves (9) for $m \geq 0, n = 1$. Assume that (9) holds for $m \geq 0$ and up to some n . Then

$$\begin{aligned} v_{m+n+1} &= (\alpha_0 + \alpha_1 v_{m+n}) e_{m+n+1} \\ &= \alpha_0 e_{m+n+1} \\ &\quad + \alpha_1 e_{m+n+1} \left(\alpha_0 \sum_{j=0}^{n-1} \alpha_1^j e_{m+n} e_{m+n-1} \cdots e_{m+n-j} + \alpha_1^n v_m e_{m+1} e_{m+2} \cdots e_{m+n} \right) \\ &= \alpha_0 e_{m+n+1} + \alpha_0 \sum_{j=0}^{n-1} \alpha_1^{j+1} e_{m+n+1} e_{m+n} \cdots e_{m+n-j} + \alpha_1^{n+1} v_m e_{m+1} e_{m+2} \cdots e_{m+n+1} \\ &= \alpha_0 \sum_{j=0}^n \alpha_1^j e_{m+n+1} e_{m+n} \cdots e_{m+n+1-j} + \alpha_1^{n+1} v_m e_{m+1} e_{m+2} \cdots e_{m+n+1} \end{aligned}$$

which agrees with the right hand side of (9) for $n + 1$. The proof is completed by induction. \square

We are interested in the limit behavior of v_n as $n \rightarrow \infty$. The existence of such a possible limit is suggested by the fact that the distribution of the first term on the right hand side of (8) is invariant with respect to the permutation

$$(e_1, e_2, \dots, e_n) \rightarrow (e_n, e_{n-1}, \dots, e_1).$$

Therefore,

$$(10) \quad v_n \stackrel{d}{=} \alpha_0 \sum_{j=0}^{n-1} \alpha_1^j e_1 e_2 \cdots e_{j+1} + \alpha_1^n v_0 e_1 e_2 \cdots e_n$$

where the convergence as $n \rightarrow \infty$ of the expression on the right hand side depends on whether its second term converges to 0 as $n \rightarrow \infty$.

Theorem 2. *Let the random variables e_n be such that $\log e_n$ has a finite mean, and let $\log \alpha_1 + \log e_1 < 0$. Then, as $n \rightarrow \infty$, v_n converges in distribution to the random variable v_∞ , where*

$$(11) \quad v_\infty = \alpha_0 \sum_{j=0}^{\infty} \alpha_1^j e_1 e_2 \cdots e_{j+1}$$

independently of v_0 .

Proof. By the strong law of large numbers,

$$\frac{1}{n} \sum_{j=1}^n \log(\alpha_1 e_j) \rightarrow \log(\alpha_1 e_1) < 0 \text{ as } n \rightarrow \infty$$

with probability one. Therefore,

$$\sum_{j=1}^n \log(\alpha_1 e_j) \rightarrow -\infty$$

with probability one, and so also

$$\alpha_1^n e_1 e_2 \cdots e_n \rightarrow 0.$$

□

We note the special case where the e_n have density e^{-x} . We have

$$E(\log e_n) = \int_0^{\infty} e^{-x} \log x dx = -\gamma$$

where γ is Euler's constant ($=0.5772157$).

Letting $n \rightarrow \infty$ in (6), we find that v_∞ satisfies the distributional equation

$$(12) \quad v_\infty \stackrel{d}{=} (\alpha_0 + \alpha_1 v_\infty) e_1$$

where v_∞ is independent of e_1 . If we choose $v_0 \stackrel{d}{=} v_\infty$, then from (6), we obtain

$$v_1 \stackrel{d}{=} (\alpha_0 + \alpha_1 v_\infty) e_1 \stackrel{d}{=} v_\infty$$

and by induction $v_n \stackrel{d}{=} v_\infty$ for all $n \geq 1$. Thus, we may call v_∞ a “stationary” solution of (12). The uniqueness of the solution is guaranteed by the following.

Theorem 3. *The random variable v_∞ defined by (11) is the unique solution of (12).*

Proof. We have already proved that v_∞ is a solution of (12). Suppose v'_∞ is a second solution. Let $\{v'_n\}$ be the sequence obtained from the recurrence relations (6) with $v_0 \stackrel{d}{=} v'_\infty$. Then, as before, $v'_n \stackrel{d}{=} v'_\infty$. From (10), we obtain

$$v'_\infty \stackrel{d}{=} \alpha_0 \sum_{j=0}^{n-1} \alpha_1^j e_1 e_2 \cdots e_{j+1} + \alpha_1^n v'_\infty e_1 e_2 \cdots e_n$$

for $n \geq 1$. As $n \rightarrow \infty$, this gives $v'_\infty \stackrel{d}{=} v_\infty$ as was to be proved. \square

Equation (11) represents v_∞ as a weighted sum of products of IID random variables. As far as we are aware, this representation does not lead to a standard probability distribution. However, the moments of v_∞ can be easily derived (under appropriate conditions on α_1) either from (11) or else from (12).

The most important result for our purpose is the following.

Theorem 4. *Let $\mathbb{E}(e_n^2) = c^2 (\geq 1)$ and assume that $\alpha_1 < c^{-1}$. Choose $v_0 \stackrel{d}{=} v_\infty$. Then the sequence $\{v_n\}$ is stationary in the strict sense, with*

- a. $\mathbb{E}(v_n) = \frac{\alpha_0}{1-\alpha_1}$
- b. $\text{Var}(v_n) = \left(\frac{\alpha_0}{1-\alpha_1}\right)^2 \frac{c^2-1}{1-c^2\alpha_1^2}$
- c. $\text{Corr}(v_m, v_{m+n}) = \alpha_1^n$.

Proof. As we have already seen, the choice of $v_0 \stackrel{d}{=} v_\infty$ leads to $v_n \stackrel{d}{=} v_\infty$. Now, given v_m , the random variables $e_{m+1}, e_{m+2}, \dots, e_{m+n}$ are independent of v_m and have the same distribution as e_1, e_2, \dots, e_n . Therefore, from (9), we obtain

$$(13) \quad v_{m+n} \stackrel{d}{=} \alpha_0 \sum_{j=0}^{n-1} \alpha_1^j e_n e_{n-1} \cdots e_{n-j} + \alpha_1^n v_m e_1 e_2 \cdots e_n.$$

Since $v_m \stackrel{d}{=} v_\infty$, this gives

$$(v_m, v_{m+n}) \stackrel{d}{=} (v_\infty, v_n)$$

in view of (8). More generally,

$$(v_{m+n_1}, v_{m+n_2}, \dots, v_{m+n_n}) \stackrel{d}{=} (v_\infty, v_{n_2-n_1}, v_{n_3-n_2}, \dots, v_{n_n-n_{n-1}})$$

which proves the strict sense stationarity. To derive a and b, we use (12). Thus,

$$\mathbb{E}(v_\infty) = \alpha_0 + \alpha_1 \mathbb{E}(v_\infty)$$

and

$$\mathbb{E}(v_\infty^2) = (\alpha_0^2 + 2\alpha_0\alpha_1 \mathbb{E}(v_\infty) + \alpha_1^2 \mathbb{E}(v_\infty^2))c^2$$

which lead to a and b. To prove c, we use (9). Thus,

$$\begin{aligned} \mathbb{E}(v_m v_{m+n}) &= \mathbb{E}(v_m) \alpha_0 \frac{1-\alpha_1^n}{1-\alpha_1} + \alpha_1^n \mathbb{E}(v_m^2) \\ &= \left(\frac{\alpha_0}{1-\alpha_1}\right)^2 + \alpha_1^n \text{Var}(v_m) \end{aligned}$$

leading to c. \square

3. THE QUEUE LENGTH PROCESS IN M/QARCH/1

Let $D_0 = 0, D_1, D_2, \dots$ be the successive epochs of departure and denote by v_n the service time of the customer leaving at D_n . Also, let Q_n be the queue length (the number of customers waiting to be served plus the one at the counter, if any) at time D_n^+ . Then we have the recurrence relations

$$(14) \quad Q_{n+1} = Q_n + A(v_{n+1}) - 1_{Q_n > 0}$$

$$(15) \quad v_{n+1} = (\alpha_0 + \alpha_1 v_n) e_{n+1}$$

for $n \geq 0$, where $A(t)$ is the number of arrivals during a time interval of length t . From (14)–(15), it follows that $\{(Q_n, v_n), n \geq 0\}$ is a time-homogeneous Markov process on the state space $\{0, 1, 2, \dots\} \times [0, \infty)$. We shall call this the queue length process. From (14)–(15), we find that for $0 < z < 1, \theta > 0$,

$$\mathbb{E}(z^{Q_{n+1}} e^{-\theta v_{n+1}} | Q_n, v_n) = z^{Q_n-1} e^{-(\theta+\lambda-\lambda z)(\alpha_0+\alpha_1 v_n)e_{n+1}} 1_{Q_n > 0} + e^{-(\theta+\lambda-\lambda z)(\alpha_0+\alpha_1 v_n)e_{n+1}} 1_{Q_n=0}.$$

Therefore,

$$\begin{aligned} z \mathbb{E}(z^{Q_{n+1}} e^{-\theta v_{n+1}}) &= \mathbb{E}(z^{Q_n} e^{-(\theta+\lambda-\lambda z)(\alpha_0+\alpha_1 v_n)e_{n+1}}) \\ &\quad - (1-z) \mathbb{E}(z^{Q_n} e^{-(\theta+\lambda-\lambda z)(\alpha_0+\alpha_1 v_n)e_{n+1}}; Q_n = 0). \end{aligned}$$

We shall assume that as $n \rightarrow \infty$,

$$(16) \quad (Q_n, v_n) \xrightarrow{d} (Q, v'_\infty).$$

Then,

$$(17) \quad \mathbb{E}(z^Q e^{(\theta+\lambda-\lambda z)(\alpha_0+\alpha_1 v'_\infty)e_1} - z^{Q+1} e^{-\theta v'_\infty}) = (1-z) \mathbb{E}(e^{-(\theta+\lambda-\lambda z)(\alpha_0+\alpha_1 v'_\infty)e_1}; Q = 0).$$

For fixed $\theta > 0$, letting $z \rightarrow 1$ in (17), we obtain

$$(18) \quad \mathbb{E}(e^{-\theta(\alpha_0+\alpha_1 v'_\infty)e_1}) = \mathbb{E}(e^{-\theta v'_\infty}).$$

In view of Theorem 3, we conclude that the marginal distribution of v'_∞ is the same as that of v_∞ as defined by (11). Accordingly, we shall denote the limit in (16) as (Q, v_∞) , although in the joint distribution of (Q, v'_∞) , the variable v'_∞ behaves differently.

Writing (17) as

$$\mathbb{E}(e^{-(\theta+\lambda-\lambda z)(\alpha_0+\alpha_1 v_\infty)e_1}; Q = 0) = \frac{\mathbb{E}(z^Q e^{-(\theta+\lambda-\lambda z)(\alpha_0+\alpha_1 v_\infty)e_1}) - \mathbb{E}(z^{Q+1} e^{-Q v_\infty})}{1-z}$$

and letting $z \rightarrow 1$, we obtain

$$(19) \quad \begin{aligned} \mathbb{E}(e^{-\theta(\alpha_0+\alpha_1 v_\infty)e_1}; Q = 0) &= \mathbb{E}((Q+1)e^{-\theta v_\infty}) - \mathbb{E}(Q e^{-\theta(\alpha_0+\alpha_1 v_\infty)e_1}) \\ &\quad - \lambda \mathbb{E}(e^{-\theta(\alpha_0+\alpha_1 v_\infty)e_1} (\alpha_0 + \alpha_1 v_\infty) e_1). \end{aligned}$$

Letting $\theta \rightarrow 0$ in this, we obtain

$$(20) \quad \mathbb{P}\{Q = 0\} = 1 - \rho$$

where $\rho = \lambda \mathbb{E}(v_\infty)$ and $0 < \rho < 1$ is a necessary condition for the convergence in (16). Differentiating (19) with respect to θ and letting $\theta \rightarrow 0$, we obtain

$$(21) \quad \mathbb{E}(Q(\alpha_0 + \alpha_1 v_\infty)) + \mathbb{E}((\alpha_0 + \alpha_1 v_\infty); Q = 0) - \mathbb{E}(Q v_\infty) = \mathbb{E}(v_\infty) - \lambda \mathbb{E}(v_\infty^2).$$

Finally, taking the second derivative of (17) with respect to z and letting $z \rightarrow 1$ and $\theta \rightarrow 0$, we find that

$$(22) \quad \lambda \mathbb{E}(Q(\alpha_0 + \alpha_1 v_\infty)) + \lambda \mathbb{E}((\alpha_0 + \alpha_1 v_\infty); Q = 0) + \frac{\lambda^2}{2} \mathbb{E}(v_\infty^2) = \mathbb{E}(Q).$$

Although the results (21) and (22) do not yield in general an explicit result for $E(Q)$, some important conclusions can be drawn.

Theorem 5. *In the queue $M/QARCH/1$, the queue length process $\{Q_n, v_n\}$ converges in distribution as $n \rightarrow \infty$ only if $\rho < 1$, where $\rho = \lambda E(v_\infty)$. In this case, we have the following.*

a. *In the system with deterministic service times ($e_n \equiv 1$),*

$$(23) \quad E(Q) = \rho + \frac{1}{2} \frac{\rho^2}{1 - \rho}.$$

b. *More generally,*

$$(24) \quad \frac{\alpha_1}{1 - \alpha_1} \frac{d\rho^2 - \rho}{1 - \rho} < E(Q) - \left(\rho + \frac{1}{2} \frac{\lambda^2 E(v_\infty^2)}{1 - \rho} \right) < \frac{\alpha_1}{1 - \alpha_1} \frac{d\rho^2}{1 - \rho},$$

where

$$(25) \quad d = \frac{c^2(1 - \alpha_1^2)}{1 - c^2\alpha_1^2} \geq c^2 \geq 1.$$

Proof. First we prove a. If $e_n \equiv 1$, then the equation $v_\infty = (\alpha_0 + \alpha_1 v_\infty)e_1$ has the unique solution

$$v_\infty = \frac{\alpha_0}{1 - \alpha_1}.$$

Using this in (22), we obtain

$$\rho E(Q) + \rho(1 - \rho) + \frac{1}{2} \lambda^2 E(v_\infty^2) = E(Q)$$

which simplifies to (23).

Now we prove b. From (21) and (22), we find that

$$E(Q) - \rho + \lambda^2 E(v_\infty^2) = \lambda E(Qv_\infty) + \frac{1}{2} \lambda^2 E(v_\infty^2).$$

This can be written as

$$(1 - \rho) E(Q) = \rho + \frac{1}{2} \lambda^2 E(v_\infty^2) + \lambda(\text{Cov}(Q, v_\infty) - \lambda \text{Var}(v_\infty)).$$

The desired inequalities (24) follow from the following result. □

Theorem 6. *In the $M/QARCH/1$ system in steady state, we have*

$$(26) \quad \frac{\alpha_1}{1 - \alpha_1} (d\rho^2 - \rho) < \lambda(\text{Cov}(Q, v_\infty) - \lambda \text{Var}(v_\infty)) < \frac{\alpha_1}{1 - \alpha_1} d\rho^2.$$

Proof. We simplify (21) as

$$(27) \quad (1 - \alpha_1) \lambda(\text{Cov}(Q, v_\infty) - \lambda \text{Var}(v_\infty)) = \alpha_1 \lambda^2 E(v_\infty) - \alpha_1 \lambda E(v_\infty; Q > 0).$$

The calculations of Section 2 show that $\lambda^2 E(v_\infty^2) = d\rho^2$. Also,

$$(28) \quad 0 < \lambda E(v_\infty; Q > 0) < \rho$$

since $0 < \rho < 1$. From (27) and (28), we obtain (26). □

Remark 1. *In the system $M/G/1$ with Poisson arrivals at rate λ and the service times having the same mean and second moment as v_∞ , the mean queue length is given by*

$$\rho + \frac{1}{2} \frac{\lambda^2 E(v_\infty^2)}{1 - \rho}.$$

Theorem 5b establishes bounds for $E(Q)$ in $M/QARCH/1$ in terms of the mean queue length in $M/G/1$.

Remark 2. *Proceeding as in the M/QARCH/1 system, but noting that v_{n+1} does not depend on Q_n or v_n , we can prove that*

$$(29) \quad \text{Cov}(Q, v_\infty) = \lambda \text{Var}(v_\infty),$$

a result that does not seem to have been noticed in the literature. The inequalities (26) establish bounds for $\text{Cov}(Q, v_\infty)$ in M/QARCH/1 in terms of the corresponding quantity in M/G/1.

Remark 3. *The bounds in (24) depend on α_1 . In particular, if $\alpha_1 = 0$, then $v_n = \alpha_0 e_n$ and our system reduces to M/G/1 with service times $\alpha_0 e_n$ ($n \geq 1$). As is to be expected, the mean queue length reduces to the known expression for this M/G/1.*

Remark 4. *For steady state sojourn times S (waiting time plus service time) of customers, we have the formula $E(Q) = \lambda E(S)$. Therefore, the result (24) provides bounds for $E(S)$.*

Remark 5. *We do not claim that the bounds in (24) and (26) are as tight as possible. To demonstrate the behavior of these bounds, we evaluated (24) in the special case where the e_n have the exponential density and $\alpha_0 = 1$ (without loss of generality). These results are displayed in Figures 3 and 4 below.*

4. SIMULATION RESULTS

To verify results from previous sections and gain new insights about the behavior of queues with arch-type dependence, we simulated an M/QARCH/1 system in which e_n is exponentially distributed. We also simulated a corresponding system (M/G/1) as a point of comparison. To make a “fair” comparison, we chose the “G” distribution to be v_∞ (IID) for the QARCH process. With this selection, the two systems have the same marginal distributions, and we can attribute differences in queueing performance to volatility. In addition, we used the same set of interarrival times for the M/QARCH/1 and M/G/1 systems.

4.1. Generating IID from v_∞ . Equation (11) expresses v_∞ as a sum of weighted products of standard exponentials. As stated above, a closed form representation for its density or cdf cannot be derived. This poses a problem from the simulation perspective. Most standard techniques for generating IID samples from a distribution require the knowledge of the distribution’s density and/or cdf.

In this case, we have a recursive representation for the series of QARCH random variables. The most direct (although a not very efficient) way to generate IID samples from v_∞ is to run the recursion and store every h th observation (where h is chosen to be sufficiently large). We used $h = 25$ so that consecutive realizations were essentially uncorrelated ($\text{Corr}(\cdot) = \alpha_1^{25}$).

4.2. M/QARCH/1. We simulated this system with customers having arrival rates

$$\lambda \in \left\{ \frac{1}{20}, \frac{2}{20}, \dots, \frac{9}{20} \right\}$$

and service times having the QARCH parameters $\alpha_0 = 1$ and $\alpha_1 = 1/2$. This implies that the long-run mean service time is 2 and the long-run variance is 8. The set of traffic intensities are

$$\rho \in \left\{ \frac{1}{10}, \frac{2}{10}, \dots, \frac{9}{10} \right\}.$$

Both the M/QARCH/1 and the M/G/1 systems were simulated for a “warm-up period” and then the length of time until the 10,000th customer’s service completion. The results are presented below.

In Figure 7, the mean queue lengths are plotted for both systems. The theoretical bounds for $E(Q)$ in the M/QARCH/1 system are also plotted. Clearly, the customers in the queue with QARCH service times see longer lines; the difference is maximized for larger traffic intensities.

Time series plots of the queue lengths in the two systems where the traffic intensity is 0.9 are shown in Figures 5 and 6. Long lines persist longer in the M/QARCH/1 system. In that system, long lines persist since large service requests are followed by more large service requests. In the M/G/1 system, long lines are quickly attended to since consecutive service times are independent.

In Figure 8, we present a comparison of the histograms of the marginal distributions of queue lengths seen by arrivals where the traffic intensity is 0.9. The queue length distribution in the M/QARCH/1 system has much fatter tails than that in the M/G/1 system.

REFERENCES

- P. EMBRECHTS, C. KLUPPELBERG and T. MIKOSCH (1997): *Modelling Extremal Events*. Springer-Verlag, New York.
- R. F. ENGLE (1982): Autoregressive Conditional Heteroskedastic Models with Estimates of the Variance of United Kingdom Inflation. *Econometrica* 50:987–1007.
- P. D. FINCH (1963): The Single Server Queueing System with Non-Recurrent Input-Process and Erlang Service Time. *JAMS* 3:220–236.
- P. D. FINCH and C. PEARCE (1965): A Second Look at a Queueing System with Moving Average Input Process. *JAMS* 5:100–106.
- H. FURSTENBERG and H. KESTEN (1960): Products of Random Matrices. *Ann. Math. Statist.* 31:457–469.
- P. A. JACOBS (1978): A Cyclic Queueing Network with Dependent Exponential Service Times. *J. Appl. Prob.* 15:573–589.
- P. A. JACOBS (1980): Heavy Traffic Results for Single-Server Queues with Dependent (EARMA) Service and Interarrival Times. *Adv. Appl. Prob.* 12:517–529.
- P. A. JACOBS and P. A. W. LEWIS (1977): A Mixed Autoregressive-Moving Average Exponential Sequence and Point Process (EARMA(1,1)). *Adv. Appl. Prob.* 9:87–104.
- H. KESTEN (1973): Random Difference Equations and Renewal Theory for Products of Random Matrices. *Acta. Math.* 131:207–248.
- C. PEARCE (1965): A Queueing System with Moving Average Input Process and Batch Arrivals. *JAMS* 5:434–442.
- C. PEARCE (1967): Queues with Moving Average Service Times. *JAMS* 4:553–570.
- S. RESNICK and G. SAMORODNITSKY (1997): Performance Decay in a Single Server Exponential Queueing Model with Long Range Dependence. *Oper. Res.* 45:235–243.

SCHOOL OF OPERATIONS RESEARCH AND INDUSTRIAL ENGINEERING, CORNELL UNIVERSITY, ITHACA, NY 14853
E-mail address: jsr14@cornell.edu

SCHOOL OF OPERATIONS RESEARCH AND INDUSTRIAL ENGINEERING, CORNELL UNIVERSITY, ITHACA, NY 14853
E-mail address: questa@orie.cornell.edu

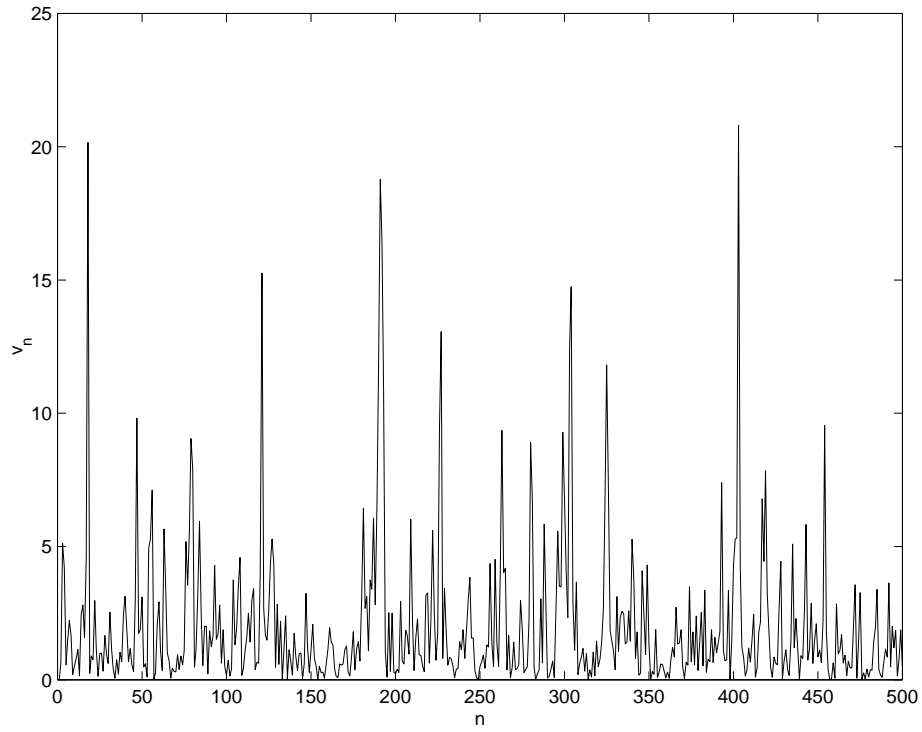


FIGURE 1. Sample Path of v_n where e_n has the Exponential Density ($\alpha_0 = 1, \alpha_1 = 0.5$)

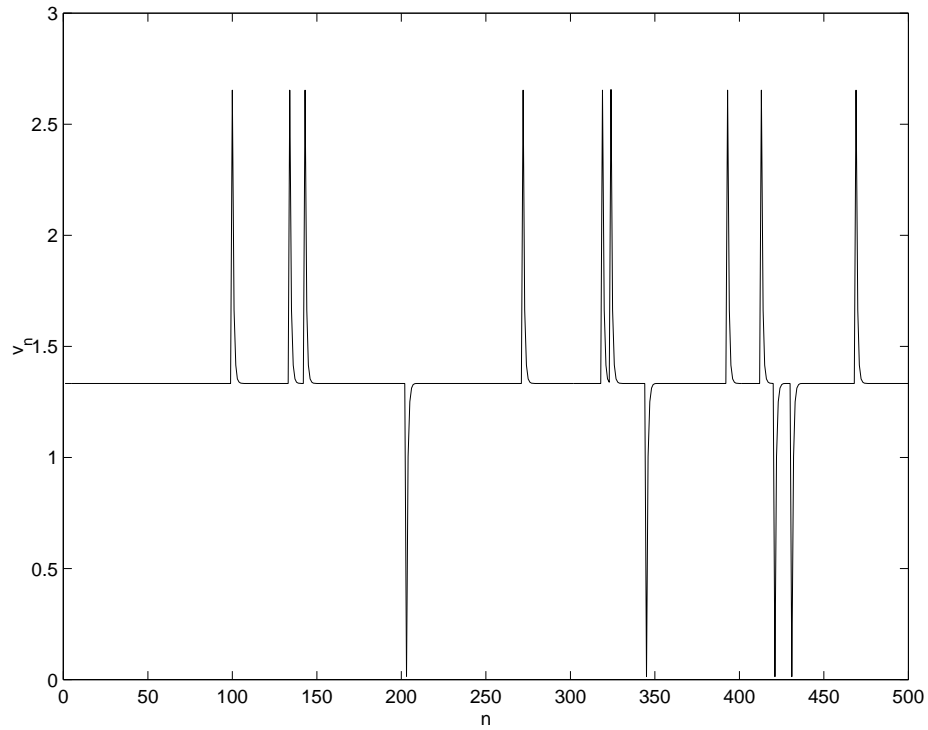
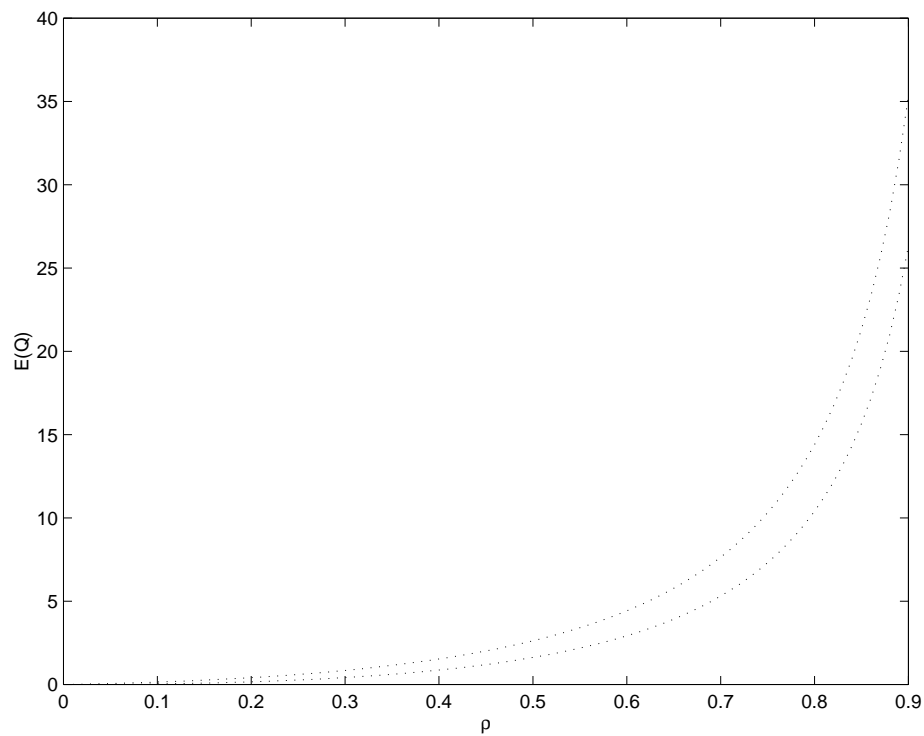
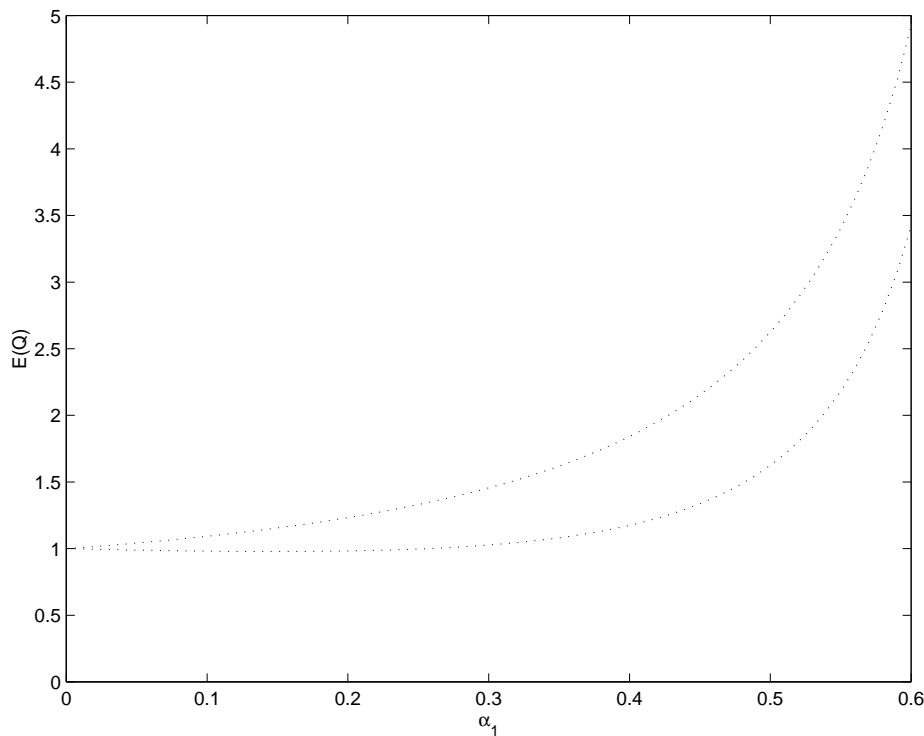


FIGURE 2. Sample Path of v_n where e_n has the Distribution Defined in (7) ($\alpha_0 = 1, \alpha_1 = 0.25, p = 0.01, \beta = 0.99$)

FIGURE 3. Bounds for $E(Q)$ as a Function of ρ ($\alpha_0 = 1$, $\alpha_1 = 0.5$)FIGURE 4. Bounds for $E(Q)$ as a Function of α_1 ($\alpha_0 = 1$, $\rho = 0.5$)

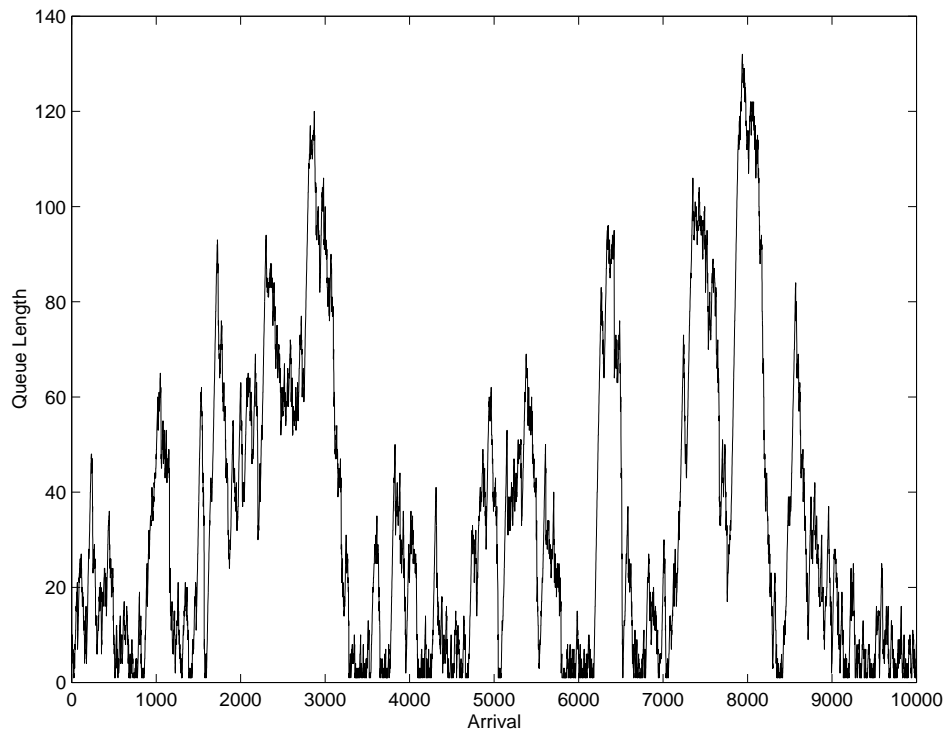


FIGURE 5. Queue Lengths Seen by Arrivals for M/QARCH/1 ($\rho = 0.9$)

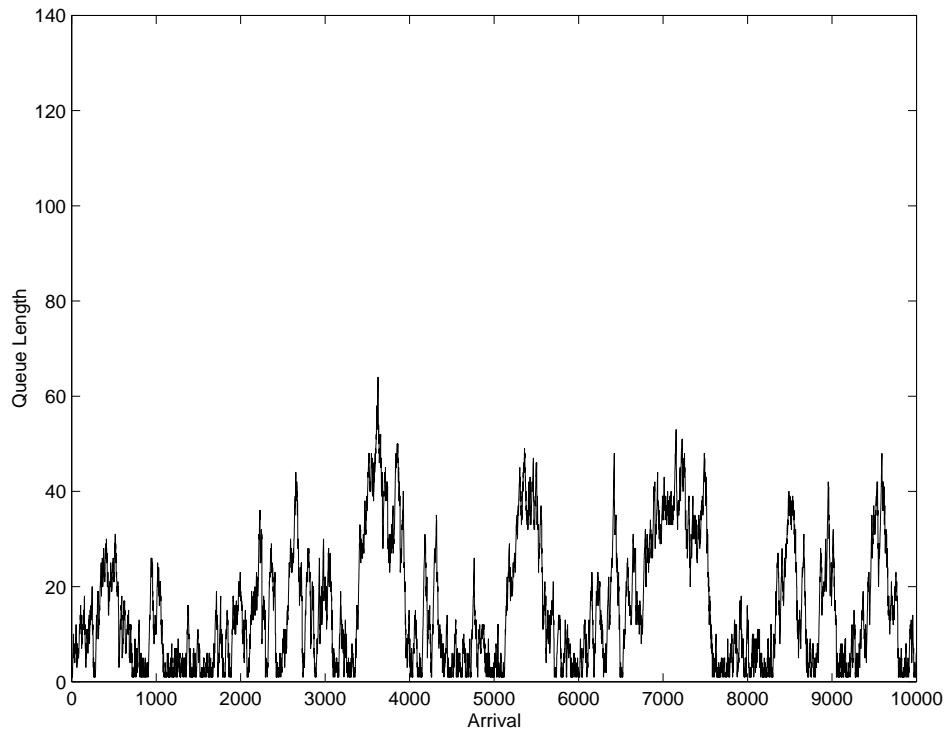
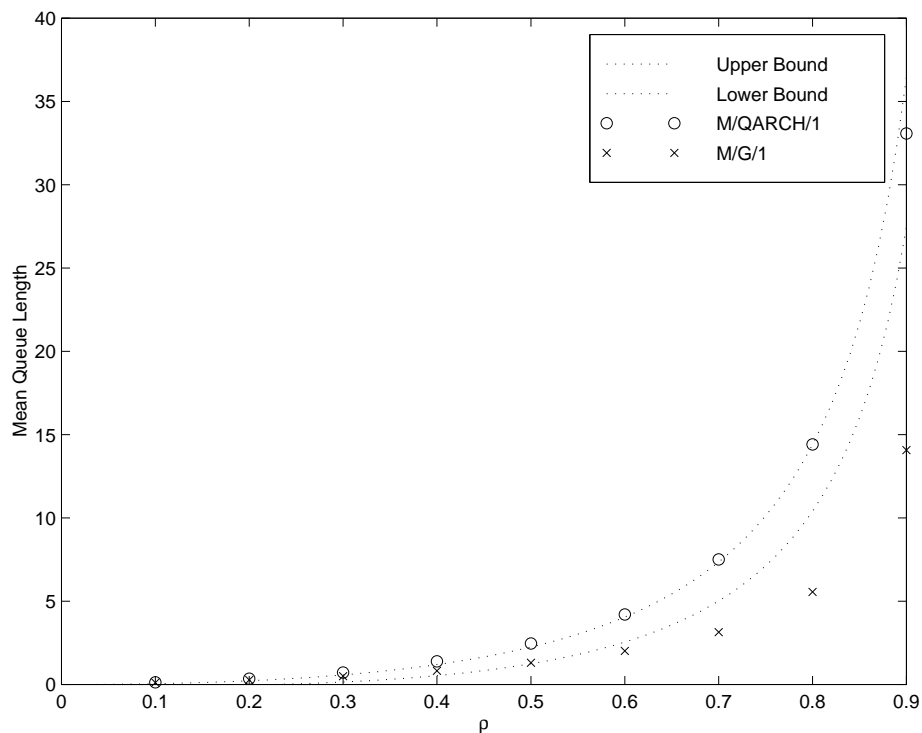
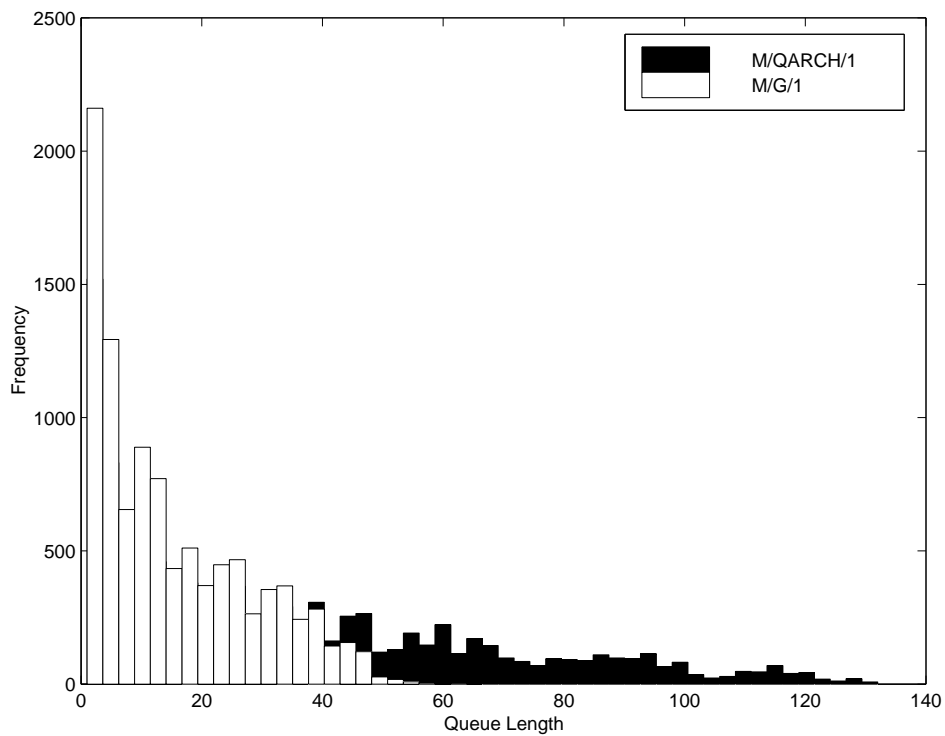


FIGURE 6. Queue Lengths Seen by Arrivals for M/G/1 ($\rho = 0.9$)

FIGURE 7. Mean Queue Length as a Function of ρ FIGURE 8. Queue Length Histogram Comparison — M/QARCH/1 vs. M/G/1 ($\rho = 0.9$)