Probability in the Engineering and Informational Sciences, **16**, 2002, 513–522. Printed in the U.S.A.

WHAT IS A UNIT OF CAPACITY WORTH?

B. A. Chiera and P. G. Taylor
Department of Applied Mathematics
University of Adelaide
Adelaide, South Australia
E-mail: {bchiera,ptaylor}@maths.adelaide.edu.au

Consider a finite-capacity telecommunications link to which connection requests arrive in a Poisson process. Each connection carried on the link earns a certain amount of revenue for the link's manager. Now, assume that the manager is offered the opportunity to buy or sell a unit of the link's allocated capacity. Assuming that the manager has a knowledge of the current number of connections on the link, we demonstrate a method of calculating the buying and selling prices.

1. INTRODUCTION

There are many situations in which we might wish to place a value on the amount of capacity that is allocated to a telecommunications link. An obvious case occurs when the manager of the link has the option to purchase extra capacity, which will allow more customers to be carried, or to sell capacity, which will restrict the number of customers that can be carried. The question arises as to how much the manager should pay or charge for the resources that can be bought or sold.

Perhaps less obviously, the issue of the value of capacity is crucial to a class of schemes which can be used for the management of complex telecommunications networks. Consider the situation where origin–destination pairs are allocated logical paths (such as virtual paths in ATM networks or label-switched paths in multiprotocol label switching, MPLS, networks) that share physical resources. The number of connections that can be carried on such a logical path depends on some (perhaps notional) amount of capacity that is allocated to the path. If logical paths are allowed to transfer capacity between themselves, then we have precisely a situation where it makes sense to allocate a value to capacity. Even though logical paths may not be

 0269-9648/02 $12.50

https://doi.org/10.1017/S0269964802164084 Published online by Cambridge University Press

purchasing capacity in any real sense, efficient management schemes can be designed by assuming that they operate as if they are in such a "capacity market."

In this short note, we consider the problem of placing a value on capacity in an $M/M/C/C$ loss system; that is, a C-server loss system where each of the C servers represents the amount of capacity needed to carry one connection on a link. The link can buy or sell capacity so that C is increased to $C+1$ or decreased to $C-1$.

Previous work in this area, including that of Lanning, Massey, Rider, and Wang [6] and the references therein, has studied the prices that should be charged to customers in a dynamic loss system. This work is the closest in spirit to ours. However, the perspective of [6] is that of an Internet billing system where arrival rates are user cost dependent. Fulp and Reeves [3] concentrated more heavily on multimarket scenarios, where the price of bandwidth is determined on the basis of current and future usage.

Other proposed models include WALRAS (see Wellman [7]) which computes prices for bandwidth trading in a market-oriented environment by use of the tatonnement process in a competitive equilibrium. This model was set up as a producer–consumer problem which requires the simultaneous solution of three constrained linear programming (LP) problems. WALRAS has since been used in later research, in particular by Yamaki, Wellman, and Ishida [8], where bandwidth is traded at prices computed to reflect current and future capacity requirements. The main disadvantage of the WALRAS model is that the time for computation can exceed the time in which the underlying market changes. This problem was overcome, to some degree, in Yamaki, Yamaauchi, and Ishida [9], where the authors demonstrated that it is possible to interrupt the WALRAS calculation after some "adequate" time span and use the interrupted prices. The model is then restarted with the updated network information, only to be interrupted at some further "adequate" point in time.

In this article, we present an alternative approach to the pricing of capacity. First, we present a scheme which, given certain network parameters, will compute the expected lost revenue due to blocking, conditional on the system starting in a given state. We then translate this expected lost revenue into both a buying price and selling price for capacity in the system, again relying on knowledge of the state at time zero.

The article is organized as follows. In Section 2, we define a model to compute the expected lost revenue based on the current state of the system. In Section 3, a method for the translation of the expected loss in revenue into buying and selling prices is given. Our conclusions are given in Section 4.

2. MODEL AND ANALYSIS

As is well known (see, e.g., [5]), the $M/M/C/C$ loss system can be modeled by a continuous-time Markov chain with state space $\{0,1,\ldots,C\}$ and transition rates of the form

$$q_{n,n+1} = \begin{cases} \lambda, & 0 \le n < C \\ 0, & n = C, \end{cases} \tag{1}$$

$$q_{n,n-1} = n\mu, \quad n \le C, \tag{2}$$

https://doi.org/10.1017/S0269964802164084 Published online by Cambridge University Press

where λ and μ^{-1} are positive constants denoting the arrival rate and mean holding time of connections, respectively. Arrivals to a fully occupied system are lost. Let θ denote the expected revenue generated per accepted arrival to the system. We set up a model to compute the expected loss in revenue, conditional on a knowledge of the current number of customers in the system, as follows.

Let $R_n(t)$ denote the expected revenue lost in the interval $[0, t]$ given that there are n connections at time 0 and let $R_n(t|x)$ be the same quantity conditional on the fact that the first time that the network departs from state n is x. Because the link loses revenue at rate $\theta\lambda$ when C connections are present and not at all when less than C connections are present, we see that

$$R_n(t|x) = \begin{cases} 0, & n < C, t < x \\ \theta\lambda t, & n = C, \quad t < x \\ \dfrac{n\mu}{\lambda + n\mu} R_{n-1}(t-x) + \dfrac{\lambda}{\lambda + n\mu} R_{n+1}(t-x), & n < C, t \geq x \\ \theta\lambda x + R_{C-1}(t-x), & n = C, t \geq x. \end{cases} \tag{3}$$

When there are n connections on the link, let F_n be the distribution of the time until the first transition. Then,

$$R_n(t) = \int_0^\infty R_n(t|x)\, dF_n(x). \tag{4}$$

Due to the Markovian nature of the model, F_n is exponential with parameter $\lambda + n\mu$ when $n < C$ and exponential with parameter $C\mu$ when the link is full. Substituting (3) into (4), we see that there are three cases that we need to consider:

Case 1: $n = 0$. In this case, $dF_0(x) = \lambda e^{-\lambda x}\, dx$. From (3), we have

$$\begin{aligned} R_0(t) &= \int_0^t R_0(t|x)\, dF_0(x) \\ &= \int_0^t R_1(t-x)\lambda e^{-\lambda x}\, dx. \end{aligned} \tag{5}$$

Case 2: $0 < n < C$. In this case, $dF_n(x) = (\lambda + n\mu)e^{-(\lambda+n\mu)x}\, dx$, and from (3), $R_n(t)$ becomes

$$\begin{aligned} R_n(t) &= \int_0^t R_n(t|x)\, dF_n(x) \\ &= \int_0^t [n\mu R_{n-1}(t-x) + \lambda R_{n+1}(t-x)]e^{-(\lambda+n\mu)x}\, dx. \end{aligned} \tag{6}$$

Case 3: $n = C$. Finally, in the third case, $dF_C(x) = (C\mu)e^{-(C\mu)x}\,dx$ and

$$R_C(t) = \int_0^t R_C(t|x)\,dF_C(x) + \int_t^\infty R_C(t|x)\,dF_C(x)$$
$$= C\mu \int_0^t [R_{C-1}(t-x) + \theta\lambda x]e^{-(C\mu)x}\,dx$$
$$+ C\mu \int_t^\infty \theta\lambda t e^{-(C\mu)x}\,dx,$$

which is equivalent to

$$R_C(t) = C\mu \int_0^t R_{C-1}(t-x)e^{-C\mu x}\,dx + \frac{\theta\lambda}{C\mu}(1 - e^{-C\mu t}). \quad \textbf{(7)}$$

Taking the Laplace transforms of (5)–(7), we see that $\tilde{R}_n(s) = \int_0^\infty e^{-st}R_n(t)\,dt$ satisfies the system of equations

$$\tilde{R}_0(s) = \frac{\lambda}{s+\lambda}\tilde{R}_1(s), \quad \textbf{(8)}$$

$$\tilde{R}_n(s) = \frac{\lambda}{s+n\mu+\lambda}\tilde{R}_{n+1}(s) + \frac{n\mu}{s+n\mu+\lambda}\tilde{R}_{n-1}(s), \quad 0 < n < C, \quad \textbf{(9)}$$

$$\tilde{R}_C(s) = \frac{1}{s+C\mu}\left(C\mu\tilde{R}_{C-1}(s) + \frac{\theta\lambda}{s}\right). \quad \textbf{(10)}$$

For network parameters C, λ, μ, and θ, the solution and inversion of (8)–(10) will give the expected lost revenue in $[0,t]$ conditional on there being n connections at time 0.

To solve these equations, note that (8) and (9) can be written in the form

$$P_{n+1}(\xi) = (\xi - (dn+f))P_n(\xi) - n(gn+h)P_{n-1}(\xi), \quad n \geq 0, \quad \textbf{(11)}$$

where $d = -\mu/\lambda$, $f = -1$, $g = 0$, $h = \mu/\lambda$, and $\xi = s/\lambda$. The recurrence relation (11) describes the class of Meixner polynomials [2, p. 164]. If, as is the case here, the recurrence relation can further be written as

$$P_{n+1}(\xi) = (\xi - d(n + hd^{-2}))P_n(\xi) - hnP_{n-1}(\xi),$$

then the solution is known to be

$$P_n(\xi) = d^n C_n^{(a)}\left(\frac{\xi}{d}\right), \quad \textbf{(12)}$$

where $a = hd^{-2}$ and $C_n^{(a)}(\cdot)$ is a Charlier polynomial. Charlier polynomials can be generally expressed in terms of Laguerre polynomials via the relation

$$C_n^{(a)}(\xi) = n!\,L_n^{(\xi-n)}(a).$$

Hence,

$$P_n(\xi) = d^n n! L_n^{(\xi/d-n)}(a). \tag{13}$$

It follows that the solution to (8) and (9) is

$$\tilde{R}_n(s) = A(s)P_n\left(\frac{s}{\lambda}\right), \tag{14}$$

where $A(s)$ is a function of s and P_n is given by (13). To find $A(s)$, we use (10), which gives us

$$\tilde{R}_C(s) = A(s)P_C\left(\frac{s}{\lambda}\right) = \frac{1}{s+C\mu}\left(C\mu A(s)P_{C-1}\left(\frac{s}{\lambda}\right) + \frac{\theta\lambda}{s}\right)$$

and, therefore,

$$A(s) = \left(\frac{1}{s}\right)\left(\frac{\theta\lambda}{(s+C\mu)P_C(s/\lambda) - C\mu P_{C-1}(s/\lambda)}\right).$$

Thus, the general solution for $\tilde{R}_n(s)$ can be written as

$$\tilde{R}_n(s) = \left(\frac{1}{s}\right)\left(\frac{\theta\lambda}{(s+C\mu)P_C(s/\lambda) - C\mu P_{C-1}(s/\lambda)}\right)P_n\left(\frac{s}{\lambda}\right). \tag{15}$$

In order to derive $R_n(t)$, we need to invert $\tilde{R}_n(s)$. In our numerical investigations presented below, we use the Euler method outlined in Abate and Whitt [1]. However, there are a number of methods available in the literature. For a description of these alternate methods, the reader is referred to Abate and Whitt [1] and Garbow, Giunta, and Lyness [4].

We now present an example of the computation of $R_n(t)$ for a small system with $C = 6$.

*Example 1: M/M/*6/6 Loss System. Let $C = 6$, $\theta = 1$, $\lambda = 3$, and $\mu^{-1} = 0.5$. The values of $R_n(t)$ for $n = \{0,\dots,6\}$ and $t \in [0,10)$ are given in Figure 1. The function $R_0(t)$ is the lowest curve and the function $R_6(t)$ is the highest curve.

As we would expect, we observe that, for all t, $R_n(t)$ increases with n. We also observe that, with increasing t, $R_n(t)$ is well approximated by a linear function with a slope that is independent of n. This slope is, in fact, equal to $\theta\lambda E(1.5,6)$, where Erlang's function

$$E(\rho,C) \equiv \frac{\rho^C/c!}{\sum_{i=0}^{C} \rho^i/i!} \tag{16}$$

gives the equilibrium probability that the link is full when the traffic is $\rho = \lambda/\mu$ and the capacity is C. This observation makes sense, because $\theta\lambda E(1.5,6)$ is the equilibrium rate of losing revenue. The difference in the height of the linear part of the

https://doi.org/10.1017/S0269964802164084 Published online by Cambridge University Press

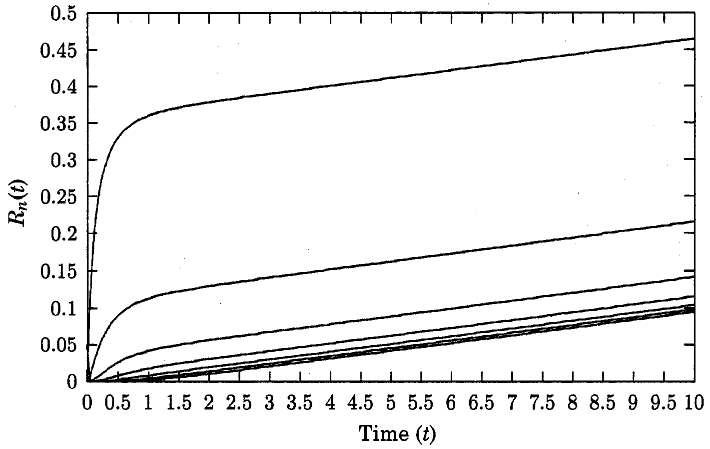

FIGURE 1. Lost revenue functions for $n = 0, \ldots, 6$ when $C = 6$, $\lambda = 3$, and $\mu^{-1} = 0.5$.

functions $R_{n+1}(t)$ and $R_n(t)$ reflects the difference in expected lost revenue incurred before equilibrium is reached when the system starts with $n + 1$ customers rather than n customers.

In Figure 1, we presented the lost revenue functions for a system with low blocking ($E(1.5,6) = 0.00353$). Figure 2 gives the lost revenue functions for a system with high blocking. This has been achieved by increasing the mean holding time μ^{-1} to 2. The blocking probability $E(6,6)$ is equal to 0.26492.

The traffic, and hence the equilibrium slope of the curves, is much greater in Figure 2 than in Figure 1. However, the latter is still given by $\theta\lambda E(\rho, C)$. The dif-

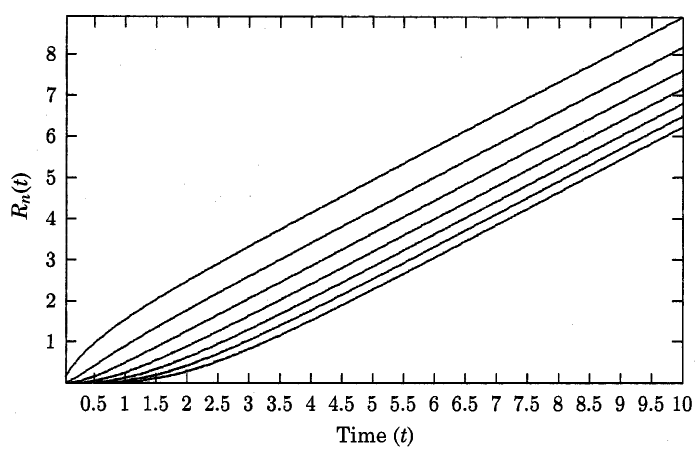

FIGURE 2. Lost revenue functions for $n = 0, \ldots, 6$ when $C = 6$, $\lambda = 3$, and $\mu^{-1} = 2$.

https://doi.org/10.1017/S0269964802164084 Published online by Cambridge University Press

ference in the equilibrium height of the functions $R_{n+1}(t)$ and $R_n(t)$ does not vary as much between $n = 0$ and $n = 5$ as for the low blocking system. This reflects the fact that, in the low blocking system, states with high occupancy are unlikely to be visited in the short term if the link does not start with a high occupancy. Thus, the penalty associated with starting in states with a high occupancy is high compared to the penalty associated with starting in states with low occupancy. In the high blocking system, the probability of moving to states with high occupancy in the short term is relatively higher even if the starting state has a low occupancy.

Example 2: $M/M/100/100$ Loss System. In this example, we consider a larger system with parameters $C = 100$, $\theta = 1$, $\lambda = 85$, and $\mu^{-1} = 1$. The values of $R_n(t)$ for $n \in \{0, 25, 50, 75, 90, 100\}$ and for $t \in [0, 10)$ are shown in Figure 3.

As with the smaller loss system, we immediately observe that after an initial period in which the starting state has an effect, the $R_n(t)$ increase linearly at the same rate. They are also increasing in n, with more pronounced increases as n becomes large.

3. SETTING THE PRICE OF BANDWIDTH

Having determined the expected lost revenue in the time interval $[0, t]$, we are left with the problem of converting this into prices at which one unit of extra capacity should be "bought" or "sold." Let us assume that the network manager is making buying and selling decisions for a planning horizon of T time units. The selection of T is a decision for the network manager. If the link is such that opportunities for capacity trading occur every u time units, one possibility would be to make $T = u$. Such a choice would be myopic in the sense that it does not take into account the

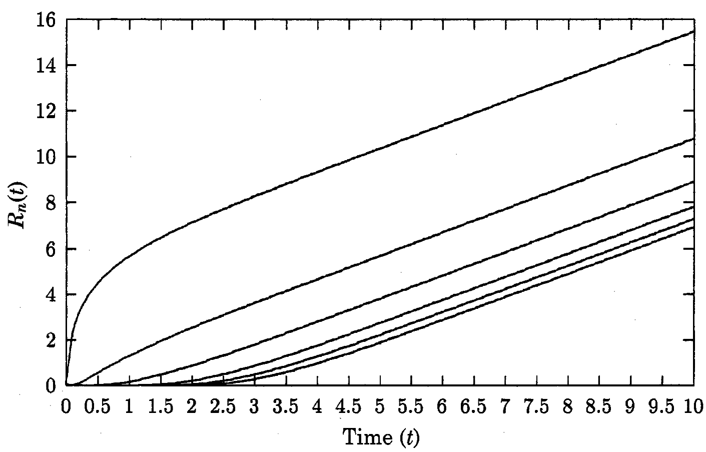

FIGURE 3. Lost revenue functions for $n = 0$, 25, 50, 75, 90, and 100 when $C = 100$, $\lambda = 85$, and $\mu^{-1} = 1$.

https://doi.org/10.1017/S0269964802164084 Published online by Cambridge University Press

number of connections on the link at the end of the time period nor, indeed, the residual value of having extra capacity at time u.

Once the manager has chosen T, we can regard the value of an extra unit of capacity as the difference in the total expected lost revenue over time $[0,T]$ if the system were to increase its capacity by one unit at time zero. Conversely, we can calculate the selling price of a unit of capacity as the difference in the total expected lost revenue over time $[0,T]$ if the system were to decrease its capacity by one unit.

The buying and selling prices, $B_n(T)$ and $S_n(T)$, respectively, of bandwidth when the initial state is n and the planning horizon is T can thus be written as

$$B_n(T) = R_{n,C}(T) - R_{n,C+1}(T), \tag{17}$$

$$S_n(T) = R_{n,C-1}(T) - R_{n,C}(T) \tag{18}$$

where the extra subscript in $R_{n,C}(T)$ indicates the initial capacity. We expect that, for all n and T, $S_n(T) > B_n(T)$. We give some examples of the computation of $B_n(T)$ and $S_n(T)$.

Example 3: $M/M/6/6$ Loss System. For the low blocking system with $C = 6$, $\theta = 1$, $\lambda = 3$, and $\mu^{-1} = 0.5$, the buying and selling prices of bandwidth, $B_n(T)$ (dotted lines) and $S_n(T)$ (continuous lines) for $n = 4, 5, 6$ are displayed in Figure 4. The same data for the high blocking system with $C = 6$, $\theta = 1$, $\lambda = 3$, and $\mu^{-1} = 2$ is shown in Figure 5.

From Figures 4 and 5, we immediately observe that $S_n(T)$ is greater than $B_n(T)$ for all n and T in both systems. We see that as n approaches link capacity C, the system places an increasingly higher value on the available capacity, both from buying and selling points of view. Also, both $S_n(T)$ and $B_n(T)$ are steeper for the

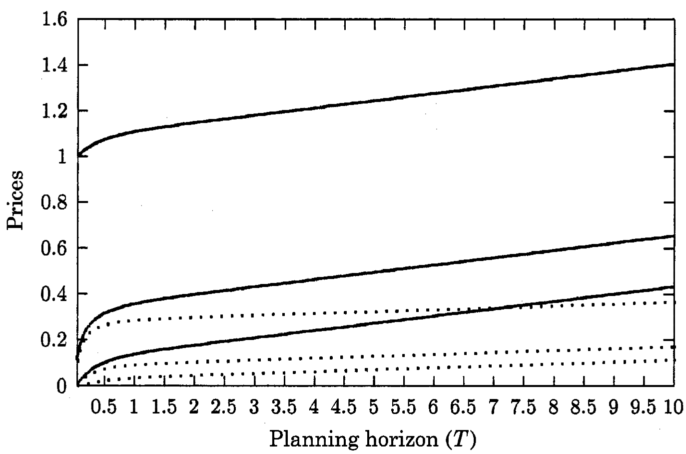

FIGURE 4. Buying and selling price functions for $n = 4, 5, 6$ when $C = 6$, $\lambda = 3$, and $\mu^{-1} = 0.5$.

https://doi.org/10.1017/S0269964802164084 Published online by Cambridge University Press

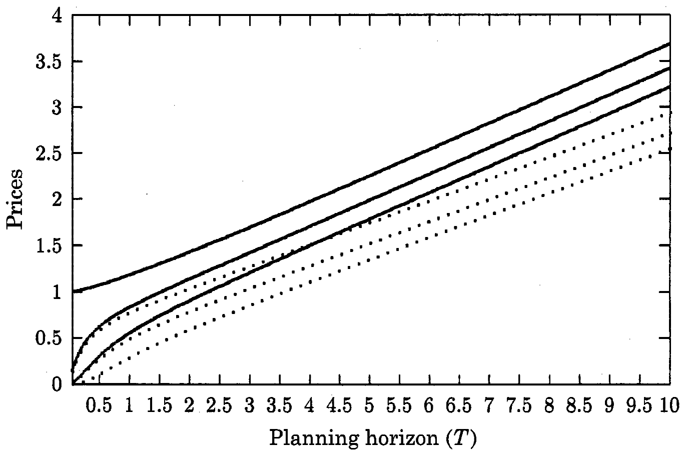

FIGURE 5. Buying and selling price functions for $n = 4, 5, 6$ when $C = 6$, $\lambda = 3$, and $\mu^{-1} = 2$.

link with greater load, but the distances $S_{n+1}(T) - S_n(T)$ and $B_{n+1}(T) - B_n(T)$ show less variation as n varies.

As with the lost revenue functions, it is useful to note that for large T, $S_n(T)$ and $B_n(T)$ are well approximated by linear functions. More precisely $S_n(T)$ can be approximated by a function of the form $\hat{S}_n(T) = s_n + \theta\lambda(E(\rho,7) - E(\rho,6))T$ and $B_n(T)$ can be approximated by a function of the form $\hat{B}_n(T) = b_n + \theta\lambda(E(\rho,6) - E(\rho,5))T$. The values of s_n and b_n reflect the total contributions to the buying and selling prices of capacity that are accumulated when the system is in its transient stage.

*Example 4: M/M/*100/100 Loss System. The values of $B_n(T)$ and $S_n(T)$ for $n \in \{50,75,90,100\}$ are given in Figure 6. Similar observations can be made as for the smaller system. The buying prices are always lower than the selling prices, and both increase markedly as the link nears full occupancy.

4. CONCLUSIONS

In this article, we have presented a model which, given the current state of a C-server loss system, computes the expected loss in revenue due to blocking. Assuming a given planning horizon, we then translated these expected losses into buying and selling prices of one unit of bandwidth. The main motivation for considering the model presented in this article is that it could be implemented in future work as an integral part of a network management scheme in which logical paths trade capacity between themselves.

We formulated a model on the basis of a system of renewal equations and then derived a system of recurrence relations satisfied by the Laplace transform of $R_n(t)$. It was shown that this system of recurrence relations was of Meixner type, for which

https://doi.org/10.1017/S0269964802164084 Published online by Cambridge University Press

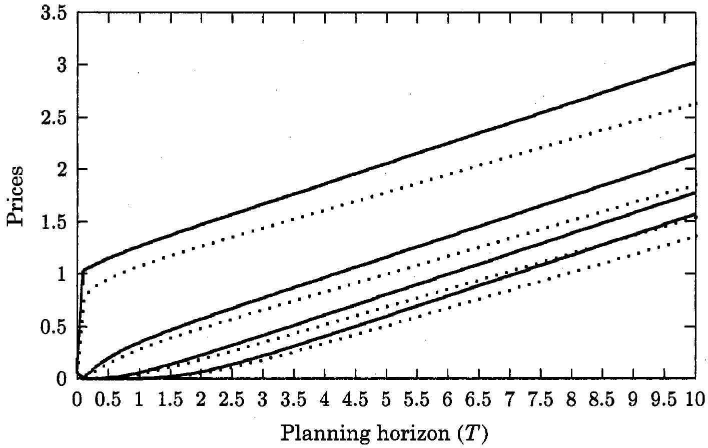

FIGURE 6. Buying and selling price functions for $n = 50, 75, 90, 100$ when $C = 100$, $\lambda = 85$, and $\mu^{-1} = 1$.

a solution could be determined in terms of Laguerre polynomials. We inverted these Laplace transforms numerically using the Euler method.

We demonstrated the computation of these prices for both a small system with $C = 6$ and a more realistically sized system with $C = 100$.

Acknowledgment

The authors would like to thank the Australian Research Council for supporting this work through grant A10033153.

References

1. Abate, J. & Whitt, W. (1995). Numerical inversion of Laplace transforms of probability distributions. *ORSA Journal on Computing* 7: 36–43.
2. Chihara, T. (1978). *An introduction to orthogonal polynomials*. New York: Gordon and Breach.
3. Fulp, E. & Reeves, D. (2000). Qos rewards and risks: A multi-market approach to resource allocation. In *Proceedings of the IFIP-TC6 Networking 2000 Conference*, pp. 945–956.
4. Garbow, B., Giunta, G., & Lyness, J. (1988). Software for an implementation of weeks' method for the inverse Laplace transform problem. *ACM Transactions on Mathematical Software* 14(2): 163–170.
5. Kleinrock, L. (1975). *Queuing systems*. Vol. 1: *Theory*. New York: Wiley.
6. Lanning, S., Massey, W., Rider, B., & Wang, Q. (1999). Optimal pricing in queueing systems with quality of service constraints. In *Proceedings of the 16th International Teletraffic Congress (ITC 16)*, pp. 747–756.
7. Wellman, M. (1993). A market-oriented programming environment and its application to distributed multicommodity flow problems. *Journal of Artificial Intelligence Research* 1: 1–23.
8. Yamaki, H., Wellman, M., & Ishida, T. (1996). A market-based approach to allocating qos for multimedia applications. In *Second International Conference of Multi-Agent Systems (ICMAS-96)*, pp. 385–392.
9. Yamaki, H., Yamaauchi, Y., & Ishida, T. (1998). Implementation issues on market-based qos control. In *International Conference on Multi-Agent Systems (ICMAS-98)*, pp. 357–364.