

CONDITIONAL SOJOURN TIMES AND THE VOLATILITY OF PAYMENT SCHEMES FOR BANDWIDTH SHARING IN PACKET NETWORKS

FABRICE M. GUILLEMIN,* *Orange Labs*
RAVIR. MAZUMDAR,** *University of Waterloo*

Abstract

Predictability of revenue and costs to both operators and users is critical for payment schemes. We study the issue of the design of payment schemes in networks with bandwidth sharing. The model we consider is a processor sharing system that is accessed by various classes of users with different processing requirements or file sizes. The users are charged according to a Vickrey–Clarke–Groves mechanism because of its efficiency and fairness when logarithmic utility functions are involved. Subject to a given mean revenue for the operator, we study whether it is preferable for a user to pay upon arrival, depending on the congestion level, or whether the user should opt to pay at the end. This leads to a study of the volatility of payment schemes and we show that opting for prepayment is preferable from a user point of view. The analysis yields new results on the asymptotic behavior of conditional response times for processor sharing systems and connections to associated orthogonal polynomials.

Keywords: Pricing; M/M/1 processor sharing; orthogonal polynomials

2010 Mathematics Subject Classification: Primary 60K25
Secondary 90B22; 80M35; 91B26

1. Introduction

Pricing is a critical issue for any facility provider offering resources to customers competing for their use. This is notably a key issue for internet service providers (ISPs) offering transmission capacities (wired or wireless) to residential or business customers sharing the bandwidth of the network, especially bandwidth of peering links and of radio cells. In this paper we investigate pricing policies by considering both fairness and efficiency. As in [4], we pay special attention to the Vickrey–Clarke–Groves (VCG) auction model [15], [17] defined as follows.

Assume that a single auctioneer (an ISP in the present context) offers to N bidders (customers) an infinitely divisible commodity of capacity c (network capacity). The bidder n has a valuation function $W_n(x)$ of the fraction x of the resource ($W_n(x)$ is often referred to as the utility function). The first step of the auctioneer is to maximize the social welfare. The optimal allocation is $\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_N)$ such that

$$\tilde{x} = \arg \max_x \sum_{j=1}^N W_j(x_j) \quad (1)$$

Received 21 February 2014; revision received 29 August 2014.

* Postal address: Orange Labs, 2 Avenue Pierre Marzin, F-22300 Lannion, France.

** Postal address: Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, N2L 3G1, Canada. Email address: mazum@uwaterloo.ca

subject to the conditions $x_j \geq 0$ for all j and $\sum_{j=1}^N x_j \leq c$. The solution to (1) is said to be a social optimal solution (that lies on the Pareto boundary) and when $W_j(x) = \log x$ then the corresponding solution is said to be proportionally fair or a Nash bargaining point. Solving (1) when $W_j(x) = \log x$ yields $\tilde{x}_i = c/N$, which corresponds to the processor discipline equally sharing amongst all users. This service discipline as a means of sharing bandwidth is natural in packet networks where data is transferred using transmission control protocol, aiming to achieve an equal share of the bandwidth of bottlenecks in networks [13]. This assumption can be verified by using measurements from real networks when bottlenecks are clearly identified (mostly at peering links). This is often referred to as max-min bandwidth sharing.

After having identified the optimal allocation, the price charged to bidder n_0 to use the resource per unit of time denoted by π_{n_0} , is evaluated as

$$\pi_{n_0} = \max_{x \mid x_{n_0}=0} \left(\sum_{j \neq n_0} W_j(x_j) \right) - \sum_{j \neq n_0} W_j(\tilde{x}_j) \tag{2}$$

and measures the maximum decrease in social welfare suffered by customers upon arrival of the marked customer n_0 .

Pricing of network resources has been studied in a number of works. Most analyses deal with static scenarios when the number of users is fixed; see, for example [8] and [16]. A detailed survey of pricing and incentives can be found in [10].

The VCG pricing mechanism is often referred to as an *efficiency* auction mechanism as opposed to a revenue maximization mechanism where the price charged is the Lagrange multiplier of the optimization problem (1). Bulow and Klemperer [5] have shown that the total revenue from an efficiency auction with $(n + 1)$ players is at least as high as the total revenue obtained from revenue maximization with n players. In this paper we refer to the revenue maximization price as the congestion or Lagrangian price.

Once prices per unit of time have been fixed, the next issue is to determine how to fix the total price of a transaction (say, the transmission of a certain amount of data, knowing that bandwidth is shared among active flows). More precisely, let us denote by $p(t)$ the price to pay per unit of time at time t when the number of users is $n(t)$. Either payment can be fixed at the time of arrival or once the customer has completed its service. In the latter case, the total price for using the service facility is given by the path integral

$$\Pi_{\text{post}} = \int_0^S p(t) dt, \tag{3}$$

where S is the time spent by the customer in the system. We refer to this as a *postpayment* scheme.

If the price is fixed upon arrival, the total cost depends on the system state upon arrival multiplied by the expected sojourn time. This is what we call *prepayment* and the total price is the random variable

$$\Pi_{\text{pre}} = p_o \mathbb{E}_0(S), \tag{4}$$

where p_o is the price upon the arrival of customer and $\mathbb{E}_0(s)$ is the mean sojourn time of the customer, depending on the state of the system at arrival time. In the case of the processor sharing disciplines with general service times, $\mathbb{E}_0(s)$ depends on the number of customers and their residual service times.

Remark 1. We could also consider a variation of (4) by considering

$$\Pi_{\text{modpre}} = p_o S,$$

where p_o is the price and S the conditional sojourn time, depending on the number present at arrival time. Such a scheme gives the same mean payment but has higher volatility as is shown.

Let L be the number in the system at the time of arrival.

Then

$$\mathbb{E}[p_o(L)S] = \mathbb{E}[p_o(L)\mathbb{E}[S | L]] = \mathbb{E}[p_o(L)\mathbb{E}_0(S)] = \mathbb{E}[\Pi_{\text{pre}}].$$

The second moment is given by

$$\mathbb{E}[(p_o(L)S)^2] = \mathbb{E}[p_o(L)^2\mathbb{E}[S^2 | L]] \geq \mathbb{E}[p_o(L)^2(\mathbb{E}_0(S))^2] = \mathbb{E}[\Pi_{\text{pre}}^2]$$

and, therefore, the payment scheme will have higher volatility than the prepayment scheme (4). We do not consider it in this paper.

The objective of this paper is to compare these two payment schemes under various pricing policies, the VCG outlined above as well as other congestion based prices and max-min bandwidth sharing. Recently, Birmiwal *et al.* [4] presented an analysis of the VCG pricing and Lagrangian pricing for the processor sharing model where the behavior of prepayment and postpayment strategies was analyzed through simulation. In this paper we present a detailed analytical analysis of such schemes that leads to very interesting connections between conditional sojourn times in processor sharing systems and some orthogonal polynomials; the analysis leads us to establish new results for the joint distribution of the service and the sojourn times in an M/M/1 processor sharing queue. In Section 2 we describe the network model under consideration and various pricing policies. The computations of the prices is performed in Section 3 under the assumption that the flow size is exponentially distributed. Numerical results are presented in Section 4 where we study the volatility of prices conditioned to have the same mean value. Concluding remarks are presented in Section 5.

2. Network model

2.1. Price under processor sharing

We consider a servicing facility with capacity c . Customers arrive according to a Poisson process with rate λ and require independent and identically distributed service times with probability distribution function $B(x)$ and mean $\mathbb{E}(B)$.

We assume that the service capacity is shared among those customers in the system, according to the processor sharing discipline (max-min fairness). In other words, if there are n customers in the system at time t , each of them is served with a service rate equal to c/n at time t .

It is well known in the literature that the processor sharing queue is insensitive to the service time distribution, and in the stationary regime the number of customers in the queue is geometric with mean $\rho/(1 - \rho)$, where $\rho = \lambda\mathbb{E}(B)/c$. In the following, let $L(t)$ denote the number of customers in the processor sharing queue at time t .

Let us assume that customers are charged according to the VCG policy. From (2), it is easily checked that when there are $L(t)$ customers in the system, the price at time t is

$$p(t) = \pi_{n_0} = L(t) \log \left(1 + \frac{1}{L(t)} \right).$$

It is worth noting that the function $x \rightarrow x \log(1 + 1/x)$ is concave and monotonic increasing and tends to 1 when $x \rightarrow \infty$. Thus, a customer will have to pay more when arriving in a congested system.

The major issue with VCG pricing is that a customer arriving in an empty system will pay nothing since it causes no perturbation of the social welfare. But, from the servicing facility it is essential to charge the customer even if the system is empty. To remedy this shortcoming we shall assume that there is always a (fictitious) customer in the system. This leads us to consider the modified VCG price.

Definition 1. (*Modified VCG pricing.*) When there is a permanent fictitious customer in the system, the decrease in social welfare caused by an arriving customer is evaluated by considering the decrease in social welfare of the fictitious customer. In such a case, the price to pay per unit of time is computed as

$$p(t) = (L(t) + 1) \log \left(1 + \frac{1}{L(t) + 1} \right).$$

Note that we consider normalized prices. In theory, we should multiply $p(t)$ by a constant expressed in monetary units per second.

Remark 2. As mentioned in the introduction, the modified VCG pricing will yield a total revenue at least as large as the revenue obtained from the revenue maximization problem with L users. Of course for large L the two VCG solutions will coincide.

We easily observe that $(1 - 1/2(L(t) + 1)) \leq p(t) \leq 1$. The upper and lower bounds in this equation are of the following form.

Definition 2. (*Homographic pricing.*) The price to pay per unit of time for homographic pricing is defined by

$$\tilde{p}(t) = a + \frac{b}{L(t) + 1} \tag{5}$$

for some real constants a and b such that $a + b > 0$.

Beyond approximating the VCG policy, homographic pricing has the advantage of encompassing other pricing policies.

- *Fixed price.* The price to pay per unit of time is set equal to constant ($b = 0$ in (5)).
- *Usage pricing.* The price to pay per unit of time is proportional to the bandwidth received by the customer ($a = 0$ in (5)).

Finally, let us consider congestion pricing which consists of more charging customers when the system is loaded.

Definition 3. (*Congestion pricing.*) The price to pay per unit of time in the case of congestion pricing is defined as $\hat{p}(t) = L(t) + 1$.

All the pricing policies considered in this paper depend on the state of the system (namely on the number of customers). Generally speaking, when there are n customers in the system, the price to pay per unit of time is p_n . For modified VCG pricing,

$$p_n = (n + 1) \log \left(1 + \frac{1}{n + 1} \right)$$

and for congestion pricing $p_n = n + 1$. For homographic pricing, $p_n = a + b/(n + 1)$.

2.2. Postpayment versus prepayment

In the case of prepayment, the customer pays upon its arrival into the system. In that case, the price a customer has to pay under the VCG pricing policy is defined by (4). In a stationary M/G/1 processor sharing queue, the mean sojourn time $\mathbb{E}_0(S)$ is a linear function of the number L of customers in the system upon arrival [2], namely

$$\mathbb{E}_0(S) = A_0 + A_1 L \tag{6}$$

for some known constants A_0 and A_1 . This can be used to compute the distribution of the random variable Π_{pre} once the price p_o is known.

In the case of postpayment, the price to pay is given by (3). A general method of computing the above integral is given in the next section for exponential service times. For the specific case of homographic pricing defined by (5), postpayment leads us to consider the integral

$$\tilde{\Pi}_{\text{post}} = \int_0^S \tilde{p}(t) dt = aS + b\sigma,$$

where σ is the service requirement of the customer and where we have used the fact that for the processor sharing policy $\int_0^S dt / (L(t) + 1) = \sigma$. Thus, we have to determine the random variable $aS + b\sigma$, which requires the study of the joint distribution of the random variables S and σ .

3. Price computations

In this section we assume that service times (or, equivalently, flow sizes) are exponentially distributed. To simplify the notation, we take the mean service time of a flow $\mathbb{E}(B)/c = 1$ so that the flow arrival rate $\lambda = \rho$, which is the load of the system.

3.1. Prepayment

In the case of exponential service times, the constants A_0 and A_1 introduced in (6) are given in [14],

$$A_0 = \frac{2}{2 - \rho}, \quad A_1 = \frac{1}{2 - \rho}. \tag{7}$$

3.1.1. *Modified VCG pricing.* The prepayment price under the VCG pricing policy is given by

$$\Pi_{\text{pre}} = (L + 1) \log \left(1 + \frac{1}{L + 1} \right) (A_0 + A_1 L).$$

Let us define the function

$$f(x) = (x + 1)(A_0 + A_1 x) \log \left(1 + \frac{1}{x + 1} \right).$$

The function $f(x)$ is monotonic increasing on $[0, \infty)$, $f(x) \sim A_1 x$, and $f^{-1}(x) \sim x/A_1$ for large x .

The first and second moment of the random variable Π_{pre} are clearly given by

$$\mathbb{E}(\Pi_{\text{pre}}) = (1 - \rho) \sum_{n=0}^{\infty} f(n)\rho^n, \quad \mathbb{E}(\Pi_{\text{pre}}^2) = (1 - \rho) \sum_{n=0}^{\infty} f(n)^2 \rho^n,$$

respectively. Finally, we have

$$\mathbb{P}(\Pi_{\text{pre}} > x) = \mathbb{P}(L > f^{-1}(x)) = \rho^{\lfloor f^{-1}(x) \rfloor} \sim \exp\left(-x \frac{\log(1/\rho)}{A_1}\right)$$

for large x . This shows that the tail of the distribution of the prepayment price has an exponential decay with rate $(2 - \rho) \log(1/\rho)$.

3.1.2. *Homographic pricing.* When considering the pricing policy defined by (5), we have

$$\tilde{\Pi}_{\text{pre}} = \left(a + \frac{b}{L + 1}\right)(A_0 + A_1 L).$$

The moments of the random variable $\tilde{\Pi}_{\text{pre}}$ can be easily computed. The mean value is given by

$$\mathbb{E}(\tilde{\Pi}_{\text{pre}}) = aA_0 + (b - a)A_1 + \frac{aA_1}{1 - \rho} - \frac{b(A_0 - A_1)(1 - \rho) \log(1 - \rho)}{\rho}$$

and the second moment by

$$\begin{aligned} \mathbb{E}(\tilde{\Pi}_{\text{pre}}^2) &= -(aA_0 + (b - a)A_1)^2 + 2abA_1(A_0 - A_1) + 2(aA_0 + (b - a)A_1)\mathbb{E}(\tilde{\Pi}_{\text{pre}}) \\ &\quad + \frac{a^2A_1^2(1 + \rho)}{(1 - \rho)^2} + \frac{b^2(A_0 - A_1)^2}{\rho} \text{Li}_2(1 - \rho), \end{aligned}$$

where $\text{Li}_2(x)$ is the dilogarithm (or Spence's function) [1].

Taking $a = 1$ and $b = -\frac{1}{2}$ and using (7) for the expression of constants A_0 and A_1 , we obtain an approximation of the first and second moments of the random variable $\tilde{\Pi}_{\text{pre}}$ in the exponential case:

$$\mathbb{E}(\tilde{\Pi}_{\text{pre}}) = \frac{1}{2(2 - \rho)} \left(\frac{3 - \rho}{1 - \rho} + \frac{1}{\rho}(1 - \rho) \log(1 - \rho) \right)$$

and

$$\mathbb{E}(\tilde{\Pi}_{\text{pre}}^2) = \frac{5 + 6\rho - 3\rho^2}{4(2 - \rho)^2(1 - \rho)^2} + \frac{(1 - \rho) \log(1 - \rho)}{2(2 - \rho)^2\rho} + \frac{(1 - \rho)\text{Li}_2(1 - \rho)}{4\rho(2 - \rho)^2}.$$

3.1.3. *Congestion pricing.* For congestion pricing, the mean value of prepayment can be expressed as

$$\mathbb{E}(\hat{\Pi}_{\text{pre}}) = \frac{A_0}{1 - \rho} + \frac{2\rho A_1}{(1 - \rho)^2} = \frac{2}{(2 - \rho)(1 - \rho)^2} \tag{8}$$

and the second moment

$$\mathbb{E}(\hat{\Pi}_{\text{pre}}^2) = \frac{(1 + \rho)A_0^2}{(1 - \rho)^2} + \frac{4\rho A_0 A_1(2 + \rho)}{(1 - \rho)^3} + \frac{4\rho(1 + 4\rho + \rho^2)A_1^2}{(1 - \rho)^4} = \frac{4(1 + 4\rho + \rho^2)}{(2 - \rho)^2(1 - \rho)^4}. \tag{9}$$

3.1.4. *Fixed pricing.* In the case of fixed pricing and exponential service times, we have

$$\mathbb{E}(\hat{\Pi}_{\text{pre}}) = \mathbb{E}(\hat{\Pi}_{\text{post}}) = \frac{a}{1 - \rho}$$

and

$$\mathbb{E}(\hat{\Pi}_{\text{pre}}^2) = \frac{(4 - 3\rho + \rho^2)a^2}{(1 - \rho)^2(2 - \rho)^2} \leq \mathbb{E}(\hat{\Pi}_{\text{post}}^2) = \frac{4a^2}{(1 - \rho)^2(2 - \rho)}.$$

The squared coefficient of variation for prepayment is ρ and that for postpayment is $2(1 + \rho)/(2 - \rho)$ showing that postpayment is more volatile than prepayment since the mean values are equal.

3.2. Postpayment

To compute the price with postpayment, we have to compute the path integral (3). We first describe a general method of evaluating this quantity for a general price structure, applicable to modified VCG pricing. We then consider the case of homographic pricing, notably when $a = 1$ and $b = -\frac{1}{2}$, thus approximating the modified VCG pricing policy.

3.2.1. *General price structure.* To compute the path integral (3), let us condition on the number of customers in the system upon arrival of the outstanding customer for which we want to compute the postpayment price. This leads us to consider the quantity defined by

$$\mathcal{P}_n = \int_0^S p(t) dt \mid L(0) = n.$$

By using the same technique as in [7], it is easily verified that the complementary distribution functions $P_n(y) = \mathbb{P}(\mathcal{P}_n > y)$, $n \geq 0$, verify the recursion

$$p_n \frac{\partial P_n}{\partial y} = \frac{n}{n+1} P_{n-1} - (1 + \rho) P_n + \rho P_{n+1}. \tag{10}$$

The above equation can be written in matrix form as

$$D \frac{\partial P}{\partial y} = M P, \tag{11}$$

where P is the infinite column vector with n th entry is equal to $P_n(y)$, D is the infinite diagonal matrix whose n th diagonal element is p_n , and M is the infinite tridiagonal matrix whose nonnull elements are given for $n \geq 0$ by

$$m_{n,n-1} = \frac{n}{n+1}, \quad m_{n,n} = -(1 + \rho), \quad m_{n,n+1} = \rho. \tag{12}$$

Set $\pi_n = (n + 1)\rho^n$ and define the Hilbert space

$$H = \left\{ (f_n) \in \mathbb{R}^{\mathbb{N}} : \sum_{n=0}^{\infty} f_n^2 p_n \pi_n < \infty \right\},$$

equipped with scalar product

$$(f, g) = \sum_{n=0}^{\infty} f_n g_n p_n \pi_n \quad \text{for } f, g \in H.$$

Let us make the following assumption.

Assumption 1. *The prices p_n are lower bounded and such that*

$$K_p = \sup_{n \geq 0} \left(\sqrt{\frac{p_{n+1}}{p_n}} \right) + \sup_{n \geq 0} \left(\sqrt{\frac{(n+1)p_n}{(n+2)p_{n+1}}} \right) < \infty. \tag{13}$$

Note that the above assumption is satisfied when p_n is increasing and the ratio p_{n+1}/p_n tends to a limit when $n \rightarrow \infty$, which is the case for modified VCG, homographic, and congestion pricing. In the case of usage pricing, $K_p = 2$.

Proposition 1. *Under the assumption*

$$1 + \rho - K_p \sqrt{\rho} > 0, \tag{14}$$

the operator $\mathcal{M} = D^{-1}M$ is symmetric and bounded and, hence, selfadjoint. The operator $-\mathcal{M}$ is, moreover, monotonic, i.e. for all $f \in H$, $(-\mathcal{M}f, f) \geq 0$.

Proof. Symmetry is readily verified by using the fact that for $n \geq 0$, $\rho\pi_n = ((n + 1)/(n + 2))\pi_{n+1}$. For $f \in H$, we have

$$(-\mathcal{M}f, f) = \sum_{n=0}^{\infty} \left((1 + \rho)f_n^2 - \rho f_{n+1}f_n - \frac{n}{n + 1}f_{n-1}f_n \right) \pi_n.$$

By using the Schwarz inequality, it is easily shown that

$$(-\mathcal{M}f, f) \geq (1 + \rho - \sqrt{\rho}K_p)(D^{-1}f, f),$$

where the constant K_p is defined by (13). Under assumption (14), the operator $-\mathcal{M}$ is monotonic. By using the same arguments, we have

$$|(\mathcal{M}f, f)| \leq (1 + \rho + \sqrt{\rho}K_p)(D^{-1}f, f).$$

By using the fact that prices are lower bounded, we deduce that the operator \mathcal{M} is bounded.

Assumption (14) is satisfied for modified VCG, homographic pricing with $a \leq 1$, congestion, and usage pricing since $K_p \leq 2$ and $1 + \rho - K_p \sqrt{\rho} > (1 - \sqrt{\rho})^2 > 0$. Monotonicity implies that the spectrum of the operator \mathcal{M} is included in $(-\infty, 0]$.

To compute the spectrum of the selfadjoint operator \mathcal{M} , we consider the following problem. Find $f(x) \in H$ and $x \in \mathbb{R}$ such that $\mathcal{M}f(x) = xf(x)$. This leads us to consider the following recursion: $f_0(x)$ is a constant (say, equal to 1), $f_{-1}(x) = 0$, and, for $n \geq 0$,

$$f_{n+1}(x) = \frac{1}{\rho}(p_n x + 1 + \rho)f_n(x) - \frac{n}{\rho(n + 1)}f_{n-1}(x).$$

The above recursion is a three-term recurrence relation for which Favard’s condition is obviously satisfied [3]. Let us introduce the family of polynomials $Q_n(x)$ defined as follows: $Q_0(x) = 1$, $Q_{-1}(x) = 0$, and, for $n \geq 0$,

$$\rho Q_{n+1}(x) - (p_n x + 1 + \rho)Q_n(x) + \frac{n}{n + 1}Q_{n-1}(x) = 0.$$

Since Favard’s condition is satisfied, the polynomials $(Q_n(x))$ are orthogonal with some weight measure $\psi(dx)$ with support in $(-\infty, 0]$ so that

$$\int_{-\infty}^0 Q_j(x)Q_k(x)\psi(dx) = \frac{p_0}{p_j\pi_j}\delta_{j,k},$$

where $\delta_{j,k}$ is the Kronecker symbol. The roots of the polynomial $Q_n(x)$ for $n > 0$ are real and simple. From the monotonicity property, these roots are nonpositive.

The solution to (11) then reads, for $n \geq 0$,

$$P_n(y) = \sum_{m=0}^{\infty} \frac{\pi_m p_m}{p_0} \int_{-\infty}^0 Q_m(x)Q_n(x)e^{xy}\psi(dx),$$

which comes from the fact that $P = e^{y\mathcal{M}}e$, where e denotes the vector with all entries being 1; see [7] for details.

Computing the exponential $e^{y\mathcal{M}}$ can be very difficult in practice. Nevertheless, by using the fact that the operator \mathcal{M} is selfadjoint, we can obtain approximations for the vector P by truncating the matrices M and D to the subspace spanned by the first N components for $N > 0$ chosen appropriately to keep the error small.

The above results can be applied to VCG pricing but the results are not explicit. By approximating this scheme by the homographic pricing policy (5) with $a = 1$ and $b = -\frac{1}{2}$, we can obtain a closed form expression by means of some orthogonal polynomials and their associated orthogonality measure.

3.2.2. Homographic pricing. When the price to pay per unit of time is \tilde{p}_n given by (5), the postpayment price is $\tilde{\Pi}_{\text{post}} = aS + b\sigma$. When service times are exponentially distributed, it is possible to compute the Laplace transform of the random variable $\tilde{\Pi}_{\text{post}}$ by using the transform of the sojourn time $S(x)$ conditioned on the service time (i.e. the sojourn time of a customer when the service is x); see [12] for details.

The mean value of the random variable $S(x)$ for general service times is given by [9],

$$\mathbb{E}(S(x)) = \frac{x}{1 - \rho}$$

and then

$$\mathbb{E}(\tilde{\Pi}_{\text{post}}) = \frac{a\mathbb{E}(\sigma)}{1 - \rho} + b\mathbb{E}(\sigma), \quad \mathbb{E}(\tilde{\Pi}_{\text{post}}^2) = a^2\mathbb{E}(S^2) + \frac{2ab\mathbb{E}(\sigma^2)}{1 - \rho} + b^2\mathbb{E}(\sigma^2).$$

It is worth noting that the second moment of postpayment price depends on the second moment of the service. This is not the case for prepayment. Hence, postpayment may be subject to more fluctuations than prepayment. This will be illustrated in the next section, even for exponential service times.

In the case of exponential service times with unit mean, since

$$\mathbb{E}(S) = \frac{1}{1 - \rho}, \quad \mathbb{E}(S^2) = \frac{4}{(1 - \rho)^2(2 - \rho)}, \quad \mathbb{E}(\sigma S) = \mathbb{E}\left(\frac{\sigma^2}{1 - \rho}\right) = \frac{2}{1 - \rho},$$

we obtain

$$\mathbb{E}(\tilde{\Pi}_{\text{post}}^2) = \frac{4a^2}{(1 - \rho)^2(2 - \rho)} + \frac{4ab}{1 - \rho} + 2b^2.$$

Note that when $a = 1$ and $b = -\frac{1}{2}$, we have

$$\mathbb{E}(\tilde{\Pi}_{\text{post}}) = -\frac{1}{2} + \frac{1}{1 - \rho}, \quad \mathbb{E}(\tilde{\Pi}_{\text{post}}^2) = \frac{4}{(1 - \rho)^2(2 - \rho)} - \frac{2}{1 - \rho} + \frac{1}{2}.$$

To estimate the tail of the distribution of the random variable $\tilde{\Pi}_{\text{post}}$, the use of the Laplace transform of the conditional service time $S(x)$ given in [12] seems to be very difficult. Instead, we give below an alternative expression for the Laplace transform $\tilde{\Pi}_{\text{post}}^*(s)$ by using sieved Pollaczek polynomials [6] defined by the following recursion: $P_1^\lambda(x, \alpha, \beta) \equiv 0$, $P_0^\lambda(x, \alpha, \beta) \equiv 1$, and, for $n \geq 0$,

$$(n + 1)P_{n+1}^\lambda(x; \alpha, \beta) - 2[(n + \lambda + \alpha)x + \beta]P_n^\lambda(x; \alpha, \beta) + nP_{n-1}^\lambda(x; \alpha, \beta) = 0. \quad (15)$$

When $\lambda = \frac{1}{2}$, we simply set $P_{n+1}^\lambda(x; \alpha, \beta) = P_n(x; \alpha, \beta)$.

For $x \in [-1, 1]$ and $\theta \in [0, \pi]$ such that $x = \cos \theta$, we set

$$\tau = \frac{\alpha \cos \theta + \beta}{\sin \theta}.$$

For $x = \cos \theta \in [-1, 1]$, Pollaczek polynomials $P_n(x; \alpha, \beta)$ have the following generating function:

$$\sum_{n=0}^{\infty} P_n(x; \alpha, \beta) z^n = (1 - ze^{i\theta})^{-1/2+i\tau} (1 - ze^{-i\theta})^{-1/2-i\tau}. \tag{16}$$

These polynomials are, furthermore, orthogonal with respect to a measure which has a component absolutely continuous with respect to the Lebesgue measure and that can have an atomic component. The density of the continuous component is given by

$$w(x; \alpha, \beta) = \frac{1}{2 \cosh \pi \tau} e^{(2\theta-\pi)\tau}$$

so that

$$\int_{-1}^1 P_n(x; \alpha, \beta) P_m(x; \alpha, \beta) w(x) dx = \frac{1}{n + 1/2 + \alpha} \delta_{m,n},$$

where $\delta_{m,n}$ is the Kronecker symbol. The discrete component has atoms at points x_n and y_n defined by (22) and (23), respectively.

We have the following result for the Laplace transform $\tilde{\Pi}_{\text{post}}^*(s)$ of the random variable $\tilde{\Pi}_{\text{post}}$.

Proposition 2. *The Laplace transform $\tilde{\Pi}_{\text{post}}^*(s)$ is given by*

$$\tilde{\Pi}_{\text{post}}^*(s) = (1 - \rho) \int_{-\infty}^{\infty} \frac{\tilde{\mathcal{Q}}(x; \rho)^2}{s - x} d\tilde{\psi}(x), \tag{17}$$

where $d\tilde{\psi}(x)$ is the spectral measure of the polynomials $\tilde{Q}_n(x)$ defined by

$$\tilde{Q}_n(x) = \frac{1}{(\sqrt{\rho})^n} P_n\left(\frac{ax + 1 + \rho}{2\sqrt{\rho}}; \frac{b}{a} + \frac{1}{2}, -\frac{b}{a}\left(\frac{1 + \rho}{2\sqrt{\rho}}\right)\right). \tag{18}$$

The function $\tilde{\mathcal{Q}}(x; \rho)$ is given by

$$\tilde{\mathcal{Q}}(x; \rho) = \frac{1}{\sqrt{1 + \rho - 2\sqrt{\rho} \cos \tilde{\theta}}} e^{2\tilde{\phi}\tilde{\tau}}$$

with $\tilde{\phi}$ being defined by

$$\tilde{\phi} = \arctan\left(\frac{\sqrt{\rho} \sin \tilde{\theta}}{1 - \sqrt{\rho} \cos \tilde{\theta}}\right) \tag{19}$$

and

$$\tilde{\tau} = \frac{(2(b/a) + 1) \cos \tilde{\theta} - b(1 + \rho)/a\sqrt{\rho}}{2 \sin \tilde{\theta}}. \tag{20}$$

Proof. We apply the method described in Section 3.2.1. When the instantaneous prices are given by (5) ($\tilde{p}_n = a + b/(n + 1)$), we have to deal with the infinite tridiagonal matrix $\tilde{\mathcal{M}}$ whose nonnull entries are given by

$$\tilde{m}_{n,n-1} = \frac{n}{a(n + 1) + b}, \quad \tilde{m}_{n,n-1} = \frac{-(1 + \rho)(n + 1)}{a(n + 1) + b}, \quad \tilde{m}_{n,n+1} = \frac{\rho(n + 1)}{a(n + 1) + b}.$$

A vector $(f_n(x))$ is an eigenvector for the matrix \mathcal{M} with eigenvalue x (i.e. $\mathcal{M}f(x) = xf(x)$) if, for $n \geq 0$,

$$\left(a + \frac{b}{n + 1}\right)xf_n(x) = \frac{n}{n + 1}f_{n-1}(x) - (1 + \rho)f_n(x) + \rho f_{n+1}(x),$$

which is equivalent to

$$\rho(n + 1)f_{n+1}(x) - ((a(n + 1) + b)x + (n + 1)(1 + \rho))f_n(x) + nf_{n-1}(x) = 0.$$

The above recurrence relation can be written as

$$(n + 1)(\sqrt{\rho})^{n+1}f_{n+1}(x) - 2((n + \lambda + \alpha)X + \beta)(\sqrt{\rho})^n f_n(x) + n(\sqrt{\rho})^{n-1}f_{n-1}(x) = 0$$

with $\lambda = \frac{1}{2}$,

$$X = \frac{ax + 1 + \rho}{2\sqrt{\rho}}, \quad \alpha = \frac{b}{a} + \frac{1}{2}, \quad \beta = -\frac{b}{a}\left(\frac{1 + \rho}{2\sqrt{\rho}}\right).$$

By using the recurrence relation (15), we deduce that

$$f_n(x) = \frac{1}{(\sqrt{\rho})^n} P_n\left(\frac{ax + 1 + \rho}{2\sqrt{\rho}}; \frac{b}{a} + \frac{1}{2}, -\frac{b}{a}\left(\frac{1 + \rho}{2\sqrt{\rho}}\right)\right).$$

The polynomials $\tilde{Q}_n(x)$ defined by (18) are orthogonal with respect to the measure $d\tilde{\psi}(x)$, which may take various forms, depending on the parameters a and b . The measure $d\tilde{\psi}(x)$ has a component absolutely continuous with respect to Lebesgue measure and can have a discrete component. The support of the absolutely continuous component is the interval $[-(1 + \sqrt{\rho})^2/a, -(1 - \sqrt{\rho})^2/a]$ and for

$$x = -\frac{1 + \rho - 2\sqrt{\rho} \cos \tilde{\theta}}{a} \tag{21}$$

with $\tilde{\theta} \in [0, \pi]$, the absolutely continuous component is $\tilde{\psi}_c(dx)$ given by

$$\frac{d\tilde{\psi}_c(x)}{dx} = \frac{1}{2\sqrt{\rho} \cosh(\pi \cot \tilde{\theta}/2)} \exp\left(\cot \tilde{\theta} \left(-\frac{\pi}{2} + \tilde{\theta}\right)\right),$$

where $\tilde{\tau}$ is defined by (20).

When $b/a + \frac{1}{2} < -(b/a)(1 + \rho)/2\sqrt{\rho}$ (which is the case when $a = 1$ and $b = -\frac{1}{2}$), the discrete component has atoms at the points $\xi_n = -(1 + \rho - 2\sqrt{\rho}x_n)/a$ and $\chi_n = -(1 + \rho - 2\sqrt{\rho}y_n)/a$, where

$$x_n = \frac{-\alpha\beta - (n + \lambda)\sqrt{(n + \lambda)^2 + \beta^2 - \alpha^2}}{\alpha^2 - (n + \lambda)^2}, \tag{22}$$

$$y_n = \frac{-\alpha\beta + (n + \lambda)\sqrt{(n + \lambda)^2 + \beta^2 - \alpha^2}}{\alpha^2 - (n + \lambda)^2}, \tag{23}$$

where α and β are given above.

The polynomials $\tilde{Q}_n(x)$ are such that

$$\int_{-\infty}^{\infty} \tilde{Q}_n(x)\tilde{Q}_m(x) d\tilde{\psi}(x) = \frac{1}{\tilde{p}_n\pi_n} \delta_{n,m}.$$

We apply the general method of Section 3.2.1. From the spectral theorem, any $f \in H$ can be decomposed as $f = \int_{-\infty}^{\infty} f(x)\tilde{Q}(x) d\tilde{\psi}(x)$, where $\tilde{Q}(x)$ is the vector whose n th entry is $\tilde{Q}_n(x)$; the function $f(x)$ appearing in this decomposition is given by

$$f(x) = \sum_{n=0}^{\infty} \tilde{p}_n\pi_n f_n \tilde{Q}_n(x). \tag{24}$$

The complementary distribution functions of the prices \tilde{P}_n defined by

$$\tilde{P}_n = \int_0^S \tilde{p}(t) dt \mid L(0) = n \quad \text{for } n \geq 0$$

satisfy (10) with $p_n = \tilde{p}_n = a + b/(n + 1)$. The vector $\tilde{\mathcal{P}}^*(s)$ whose n th component is the Laplace transform $\tilde{\mathcal{P}}_n^*(s) = \mathbb{E}(e^{-s\tilde{P}_n})$ then satisfies $(s\mathbb{I} - \tilde{\mathcal{M}})\tilde{\mathcal{P}}^*(s) = \mathbf{e}_{\tilde{p}}$, where $\mathbf{e}_{\tilde{p}}$ is the vector whose n th entry is $1/(n + 1)\tilde{p}_n$ and \mathbb{I} is the identity operator. In other words, $\tilde{\mathcal{P}}^*(s) = (s\mathbb{I} - \tilde{\mathcal{M}})^{-1}\mathbf{e}_{\tilde{p}}$.

The Laplace transform $\tilde{\Pi}_{\text{post}}^*(s)$ of the random variable $\tilde{\Pi}_{\text{post}}$ is, by definition, given by

$$\tilde{\Pi}_{\text{post}}^*(s) = (\mathbf{e}_{\rho}, \tilde{\mathcal{P}}^*(s)) = (\mathbf{e}_{\rho}, (s\mathbb{I} - \tilde{\mathcal{M}})^{-1}\mathbf{e}_{\tilde{p}}) = \int_{-\infty}^{\infty} \frac{(\mathbf{e}_{\rho}, \mathbf{e}_{\tilde{p}}(x))}{s - x} d\tilde{\psi}(x),$$

where we have used the spectral resolvent identity in the last step [11] and where $\mathbf{e}_{\tilde{p}}(x)$ is the projection of the vector $\mathbf{e}_{\tilde{p}}$ on the vector space generated by the vector $\tilde{Q}(x)$ and \mathbf{e}_{ρ} is the vector with components $(1 - \rho)\rho^n$ for $n \geq 0$.

By using (24), we have

$$\mathbf{e}_{\tilde{p}}(x) = \sum_{n=0}^{\infty} \tilde{Q}_n(x)\rho^n \tilde{Q}(x) = \tilde{\mathcal{Q}}(x; \rho)\tilde{Q}(x)$$

with $\tilde{\mathcal{Q}}(x; \rho)$ denoting the generating function of the polynomials $\tilde{Q}_n(x)$, given for $x \in [-(1 + \sqrt{\rho})^2/a, -(1 - \sqrt{\rho})^2/a]$ by

$$\tilde{\mathcal{Q}}(x; \rho) = \left(1 - \frac{t}{\sqrt{\rho}}e^{i\tilde{\theta}}\right)^{-1/2+i\tilde{\tau}} \left(1 - \frac{t}{\sqrt{\rho}}e^{-i\tilde{\theta}}\right)^{-1/2-i\tilde{\tau}},$$

where we have used relation (18) between the polynomials $\tilde{Q}_n(x)$ and Pollaczek polynomials, and the parameters $\tilde{\theta}$ and $\tilde{\tau}$ are as in (21) and (20), respectively.

The Laplace transform $\tilde{\Pi}_{\text{post}}^*(s)$ is eventually given by

$$\tilde{\Pi}_{\text{post}}^*(s) = \int_{-\infty}^{\infty} \frac{\tilde{\mathcal{Q}}(x; \rho)(\mathbf{e}_{\rho}, \tilde{Q}(x))}{s - x} d\tilde{\psi}(x) = (1 - \rho) \int_{-\infty}^{\infty} \frac{\tilde{\mathcal{Q}}(x; \rho)^2}{s - x} d\tilde{\psi}(x),$$

where the function $\tilde{\mathcal{Q}}(x; \rho)$ is given by

$$\tilde{\mathcal{Q}}(x; \rho) = (1 - \sqrt{\rho}e^{i\tilde{\theta}})^{-1/2+i\tilde{\tau}} (1 - \sqrt{\rho}e^{-i\tilde{\theta}})^{-1/2-i\tilde{\tau}}.$$

By using the fact that

$$1 - \sqrt{\rho}e^{i\tilde{\theta}} = \sqrt{1 + \rho - 2\sqrt{\rho} \cos \tilde{\theta}}e^{-i\tilde{\phi}},$$

where $\tilde{\phi}$ is defined by (19), (17) follows.

To conclude this section, let us discuss the tail asymptotics for the random variable $\tilde{\Pi}_{\text{post}}$. From (17), we have

$$\mathbb{P}(\tilde{\Pi}_{\text{post}} > y) = -(1 - \rho) \int_{-\infty}^{\infty} \frac{1}{x} \tilde{\mathcal{Q}}(x; \rho)^2 e^{xy} d\tilde{\psi}(x).$$

For arbitrary values of the parameters a and b , the above equation does not allow an explicit derivation of the tail asymptotics (see, for example, [7] for a discussion); for the sojourn time in the M/M/1 processor sharing queue, a singular perturbation analysis can be performed when $y \rightarrow 0$ and $\rho \rightarrow 1$. However, when $a = 1$ and $b = -\frac{1}{2}$, the spectrum of the measure $d\tilde{\psi}(x)$ has atoms. The atom with minimum absolute value is located at point $-(1 + \rho - 2\sqrt{\rho}x_0)/a$, where $x_0 \in (-1, 0)$ is defined by (22) (with $\alpha = 0$). Hence, the tail of the distribution of the random variable $\tilde{\Pi}_{\text{post}}$ approximating modified VCG pricing ($a = 1$ and $b = -\frac{1}{2}$) is exponential with decay rate $1 + \rho - 2\sqrt{\rho}x_0$. For the random variable Π_{pre} with exponential service times, the coefficient of the exponential decay is $(2 - \rho) \log(1/\rho)$, which is much larger than $1 + \rho - 2\sqrt{\rho}x_0 < (1 - \sqrt{\rho})^2$.

3.2.3. *Congestion pricing.* By taking $p_n = \hat{p}_n \triangleq n + 1$, we are led, by applying the method of Section 3.2.1, to consider the family of orthogonal polynomials $\hat{q}_n(x)$ defined by the recursion: $\hat{q}_{-1}(x) = 0$, $\hat{q}_1(x) = 1$, and, for $n \geq 0$,

$$\rho \hat{q}_{n+1}(x) - ((n + 1)x + 1 + \rho)\hat{q}_n(x) + \frac{n}{n + 1}\hat{q}_{n-1}(x) = 0.$$

To the best of the authors' knowledge, this family of orthogonal polynomials ($\hat{q}_n(x)$) seems to be unknown in the literature. In particular, their generating function defined by $\hat{q}(x; t) = \sum_{n=0}^{\infty} \hat{q}_n(x)t^n$ satisfies the following second-order differential equation:

$$xt^2 \frac{\partial^2 \hat{q}}{\partial t^2} - (t^2 - (3x + 1 + \rho)t + \rho) \frac{\partial \hat{q}}{\partial t} - (t - 1 - \rho - x)\hat{q} = 0.$$

The first two moments of the random variable $\hat{\Pi}_{\text{post}}$ can, however, be computed as follows. By using (10), we obtain

$$\begin{aligned} \rho \mathbb{E}(\hat{\Pi}_{\text{post}} \mid L(0) = n + 1) - (1 + \rho)\mathbb{E}(\hat{\Pi}_{\text{post}} \mid L(0) = n) + \frac{n}{n + 1}\mathbb{E}(\hat{\Pi}_{\text{post}} \mid L(0) = n - 1) \\ = -\hat{p}_n. \end{aligned}$$

The above equation can be written as $M\hat{m} = -\hat{p}$, where M is the infinite matrix defined by (12) and \hat{m} (respectively \hat{p}) is the column vector whose n th entry is equal to $\mathbb{E}(\hat{\Pi}_{\text{post}} \mid L(0) = n)$ (respectively \hat{p}_n). By using the results of [7], the eigenvectors of the matrix M are the vectors $\hat{Q}(x)$ whose n th component is

$$\hat{Q}_n(x) = \frac{1}{(\sqrt{\rho})^n} P_n\left(\frac{1 + \rho + x}{2\sqrt{\rho}}; 1, 0\right),$$

where the polynomials $P_n(x; \alpha, \beta)$ are the Pollaczek polynomials. The polynomials $\hat{Q}_n(x)$ are orthogonal with respect to the measure $\hat{\psi}(x)$ whose support is the interval $[-(1 + \sqrt{\rho})^2, -(1 - \sqrt{\rho})^2]$ and defined for $x = -1 - \rho - 2\sqrt{\rho} \cos \theta$ with $\theta \in [0, \pi]$ by

$$\frac{d\hat{\psi}}{dx} = \frac{\sin \theta}{\cosh((\pi \cot \theta)/2)} \exp\left(\cot \theta \left(-\frac{\pi}{2} + \theta\right)\right).$$

The polynomials $\hat{Q}_n(x)$ satisfy the orthogonality relation

$$\int_{-(1+\sqrt{\rho})^2}^{-(1-\sqrt{\rho})^2} \hat{Q}_n(x) \hat{Q}_m(x) d\hat{\psi}(x) = \frac{1}{\pi_n} \delta_{n,m}.$$

By using the spectral resolvent identity, and by setting e_n equal to the vector with all components equal to 0 except the n th one equal to 1, we obtain

$$\begin{aligned} \mathbb{E}(\hat{\Pi}_{\text{post}} \mid L(0) = n) &= \frac{1}{\pi_n} (\hat{m}, e_n) \\ &= \frac{-1}{\pi_n} (M^{-1} \hat{p}, e_n) \\ &= \int_{-(1+\sqrt{\rho})^2}^{-(1-\sqrt{\rho})^2} \frac{-1}{x} \sum_{m=0}^{\infty} \hat{p}_m \pi_m \hat{Q}_m(x) \hat{Q}_n(x) d\hat{\psi}(x). \end{aligned} \tag{25}$$

By deconditioning on the initial state, we eventually obtain

$$\mathbb{E}(\hat{\Pi}_{\text{post}}) = \int_{-(1+\sqrt{\rho})^2}^{-(1-\sqrt{\rho})^2} \frac{-(1 - \rho)}{x} \sum_{m=0}^{\infty} \hat{p}_m \pi_m \hat{Q}_m(x) \sum_{n=0}^{\infty} \rho^n \hat{Q}_n(x) d\hat{\psi}(x).$$

As shown in [7], the generating function $\hat{Q}(x; t) = \sum_{n=0}^{\infty} \hat{Q}_n(x) t^n$ satisfies the differential equation

$$(t^2 - (x + \rho + 1)t + \rho) \frac{\partial \hat{Q}}{\partial t} + (t - x - \rho - 1) \hat{Q} = 0. \tag{26}$$

We have

$$\sum_{m=0}^{\infty} \hat{p}_m \pi_m \hat{Q}_m(x) = \rho^2 \frac{\partial^2 \hat{Q}}{\partial t^2}(x; \rho) + 3\rho \frac{\partial \hat{Q}}{\partial t}(x; \rho) + \hat{Q} = \frac{x + 2 - \rho}{x^2} \hat{Q}(x; \rho), \tag{27}$$

where we have used (26) in the last step of (27). Hence,

$$\mathbb{E}(\hat{\Pi}_{\text{post}}) = \int_{-(1+\sqrt{\rho})^2}^{-(1-\sqrt{\rho})^2} \frac{-(1 - \rho)(x + 2 - \rho)}{x^3} \hat{Q}(x; \rho)^2 d\hat{\psi}(x). \tag{28}$$

From [7], we know that the Laplace transform $S^*(z)$ of the sojourn time is given by

$$S^*(z) = (1 - \rho) \int_{-(1+\sqrt{\rho})^2}^{-(1-\sqrt{\rho})^2} \hat{Q}(x; \rho)^2 \frac{d\hat{\psi}(x)}{z - x};$$

hence, we have

$$\mathbb{E}(\hat{\Pi}_{\text{post}}) = \frac{2 - \rho}{2} \mathbb{E}(S^2) - \mathbb{E}(S) = \frac{1 + \rho}{(1 - \rho)^2}.$$

Similarly, by using (10), we easily obtain the matrix equation $M\hat{v} = -\hat{p}'$, where \hat{v} (respectively \hat{p}') is the column vector whose n th entry is $\mathbb{E}(\hat{\Pi}_{\text{post}}^2 \mid L(0) = n)$ (respectively $2p_n\mathbb{E}(\hat{\Pi}_{\text{post}} \mid L(0) = n)$) and, consequently,

$$\mathbb{E}(\hat{\Pi}_{\text{post}}^2) = \int_{-(1+\sqrt{\rho})^2}^{-(1-\sqrt{\rho})^2} \frac{-(1-\rho)}{x} \sum_{m=0}^{\infty} \hat{p}'_m \pi_m \hat{Q}_m(x) \hat{Q}(x; \rho) d\hat{\psi}(x).$$

By using (25), we have

$$\begin{aligned} \mathbb{E}(\hat{\Pi}_{\text{post}} \mid L(0) = n) &= \int_{-(1+\sqrt{\rho})^2}^{-(1-\sqrt{\rho})^2} \frac{-1}{x} \sum_{m=0}^{\infty} \hat{p}'_m \pi_m \hat{Q}_m(x) \hat{Q}_n(x) d\hat{\psi}(x) \\ &= \int_{-(1+\sqrt{\rho})^2}^{-(1-\sqrt{\rho})^2} -\frac{x+2-\rho}{x^3} \hat{Q}(x; \rho) \hat{Q}_n(x) d\hat{\psi}(x), \end{aligned}$$

where we have used (27) in the last step. Now, from [7], we know that

$$\begin{aligned} \int_{-(1+\sqrt{\rho})^2}^{-(1-\sqrt{\rho})^2} \frac{1}{x^2} \hat{Q}(x; \rho) \hat{Q}_n(x) d\hat{\psi}(x) &= \frac{n+2}{2-\rho}, \\ \int_{-(1+\sqrt{\rho})^2}^{-(1-\sqrt{\rho})^2} \frac{-1}{x^3} \hat{Q}(x; \rho) \hat{Q}_n(x) d\hat{\psi}(x) &= \frac{n+2}{(2-\rho)(3-2\rho)} \left(n+1 + \frac{4}{2-\rho} \right), \end{aligned}$$

and then

$$\begin{aligned} \mathbb{E}(\hat{\Pi}_{\text{post}} \mid L(0) = n) &= \frac{n+2}{3-2\rho} \left(n+1 + \frac{1+2\rho}{2-\rho} \right), \\ \hat{p}'_n &= \frac{2(n+1)(n+2)}{3-2\rho} \left(n+1 + \frac{1+2\rho}{2-\rho} \right). \end{aligned}$$

Consequently, we have

$$\sum_{m=0}^{\infty} \hat{p}'_m \pi_m \hat{Q}_m(x) = \frac{2}{3-2\rho} \sum_{m=0}^{\infty} (m+1)^2(m+2) \left(m+1 + \frac{1+2\rho}{2-\rho} \right) \rho^m \hat{Q}_m(x).$$

On the one hand, we have

$$\begin{aligned} &\sum_{m=0}^{\infty} (m+1)^3(m+2)\rho^m \hat{Q}_m(x) \\ &= \rho^4 \frac{\partial^4 \hat{Q}}{\partial t^4}(x; \rho) + 11\rho^3 \frac{\partial^3 \hat{Q}}{\partial t^3}(x; \rho) + 31\rho^2 \frac{\partial^2 \hat{Q}}{\partial t^2}(x; \rho) + 22\rho \frac{\partial \hat{Q}}{\partial t}(x; \rho) + 2\hat{Q}(x; \rho) \\ &= \frac{(2-\rho)(3-2\rho)(4-3\rho) + (30-15\rho-2\rho^2)x + 8x^2}{x^4} \hat{Q}(x; \rho). \end{aligned} \tag{29}$$

On the other hand,

$$\begin{aligned} & \sum_{m=0}^{\infty} (m + 1)^2(m + 2)\rho^m \hat{Q}_m(x) \\ &= \rho^3 \frac{\partial^3 \hat{Q}}{\partial t^3}(x; \rho) + 7\rho^2 \frac{\partial^2 \hat{Q}}{\partial t^2}(x; \rho) + 10\rho \frac{\partial \hat{Q}}{\partial t}(x; \rho) + 2\hat{Q}(x; \rho) \\ &= -\frac{6 - 7\rho + 2\rho^2 + 4x}{x^3} \hat{Q}(x; \rho). \end{aligned} \tag{30}$$

By using the fact that [7], for all $m \geq 0$, can be expressed as

$$\int_{-(1+\sqrt{\rho})^2}^{-(1-\sqrt{\rho})^2} \frac{-1}{x} \hat{Q}(x; \rho) \hat{Q}_m(x) d\hat{\psi}(x) = 1, \quad \sum_{m=0}^{\infty} (m + 1)^2(m + 2)\rho^m = \frac{2(1 + 2\rho)}{(1 - \rho)^4},$$

we deduce that, from (30),

$$\int_{-(1+\sqrt{\rho})^2}^{-(1-\sqrt{\rho})^2} \frac{1}{x^4} \hat{Q}(x; \rho)^2 d\hat{\psi}(x) = \frac{2(2 + \rho)}{(2 - \rho)^2(1 - \rho)^4},$$

where we have used the identity from [7],

$$2(1 - \rho) \int_{-(1+\sqrt{\rho})^2}^{-(1-\sqrt{\rho})^2} \frac{-1}{x^3} \hat{Q}(x; \rho)^2 d\hat{\psi}(x) = \mathbb{E}(S^2) = \frac{4}{(1 - \rho)^2(2 - \rho)}.$$

Similarly, by using (29) and the fact that

$$\sum_{m=0}^{\infty} (m + 1)^3(m + 2)\rho^m = \frac{2(1 + 7\rho + 4\rho^2)}{(1 - \rho)^5},$$

we have

$$\int_{-(1+\sqrt{\rho})^2}^{-(1-\sqrt{\rho})^2} \frac{-1}{x^5} \hat{Q}(x; \rho)^2 d\hat{\psi}(x) = \frac{4(6 + 5\rho - 4\rho^2 - \rho^3)}{(2 - \rho)^3(1 - \rho)^5(3 - 2\rho)}.$$

Hence,

$$\begin{aligned} & \sum_{m=0}^{\infty} (m + 1)^2(m + 2) \left(m + 1 + \frac{1 + 2\rho}{2 - \rho}\right) \rho^m \hat{Q}_m(x) \\ &= \frac{(2 - \rho)^2(3 - 2\rho)(4 - 3\rho) + (2 - \rho)(27 - 19\rho + 2\rho^2)x + 4(3 - 4\rho)x^2}{(2 - \rho)x^4} \hat{Q}(x; \rho). \end{aligned}$$

It follows that

$$\begin{aligned} \mathbb{E}(\hat{\Pi}_{\text{post}}^2) &= \frac{-2(1 - \rho)}{(3 - 2\rho)(2 - \rho)} \\ &\times \left(\int_{-(1+\sqrt{\rho})^2}^{-(1-\sqrt{\rho})^2} \frac{(2 - \rho)^2(3 - 2\rho)(4 - 3\rho)}{x^5} \hat{Q}(x; \rho)^2 d\hat{\psi}(x) \right. \\ &\quad \left. + \int_{-(1+\sqrt{\rho})^2}^{-(1-\sqrt{\rho})^2} \frac{(2 - \rho)(27 - 19\rho + 2\rho^2) + 4(3 - 4\rho)x}{x^4} \hat{Q}(x; \rho)^2 d\hat{\psi}(x) \right) \end{aligned}$$

and, hence,

$$\mathbb{E}(\hat{\Pi}_{\text{post}}^2) = \frac{2}{(3 - 2\rho)(2 - \rho)} \left((2 - \rho)^2(3 - 2\rho)(4 - 3\rho) \frac{\mathbb{E}(S^4)}{24} - (2 - \rho)(27 - 19\rho + 2\rho^2) \frac{\mathbb{E}(S^3)}{6} + 4(3 - 4\rho) \frac{\mathbb{E}(S^2)}{2} \right).$$

By using the fact that

$$\mathbb{E}(S^3) = 6(1 - \rho) \int_{-(1+\sqrt{\rho})^2}^{-(1-\sqrt{\rho})^2} \frac{1}{x^4} \hat{\mathcal{Q}}(x; \rho)^2 d\hat{\psi}(x) = \frac{12(2 + \rho)}{(2 - \rho)^2(1 - \rho)^3},$$

$$\mathbb{E}(S^4) = 24(1 - \rho) \int_{-(1+\sqrt{\rho})^2}^{-(1-\sqrt{\rho})^2} \frac{-1}{x^5} \hat{\mathcal{Q}}(x; \rho)^2 d\hat{\psi}(x) = \frac{96(6 + 5\rho - 4\rho^2 - \rho^3)}{(2 - \rho)^3(1 - \rho)^4(3 - 2\rho)},$$

we eventually obtain

$$\mathbb{E}(\hat{\Pi}_{\text{post}}^2) = \frac{4(3 + 16\rho + \rho^2 - 8\rho^3)}{(3 - 2\rho)(2 - \rho)(1 - \rho)^4}. \tag{31}$$

4. Numerical results

In Figure 1 we have displayed the mean value of prepayment and postpayment pricing as a function of the load ρ for exponential service times with unit mean. The approximation of VCG pricing can satisfactorily be approximated by homographic pricing with $a = 1$ and $b = -\frac{1}{2}$. In addition, for this latter pricing policy the mean values of postpayment and prepayment prices are remarkably close to one another. This means that under these pricing policies, the revenue of the operator and the price to pay by customers do not significantly depend on the postpayment versus prepayment policy.

To study the volatility of prices, we compute the coefficients of variation (i.e. the ratio of the variance to the mean) of approximated VCG prepayment pricing and postpayment pricing.

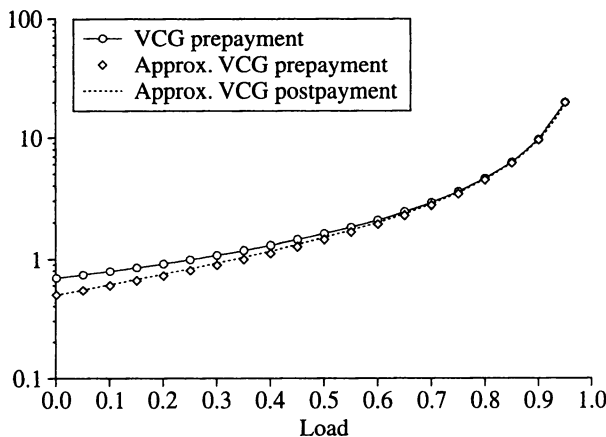


FIGURE 1: Mean values of VCG and approximated VCG prepayment and postpayment pricing payment as a function of the load ρ .

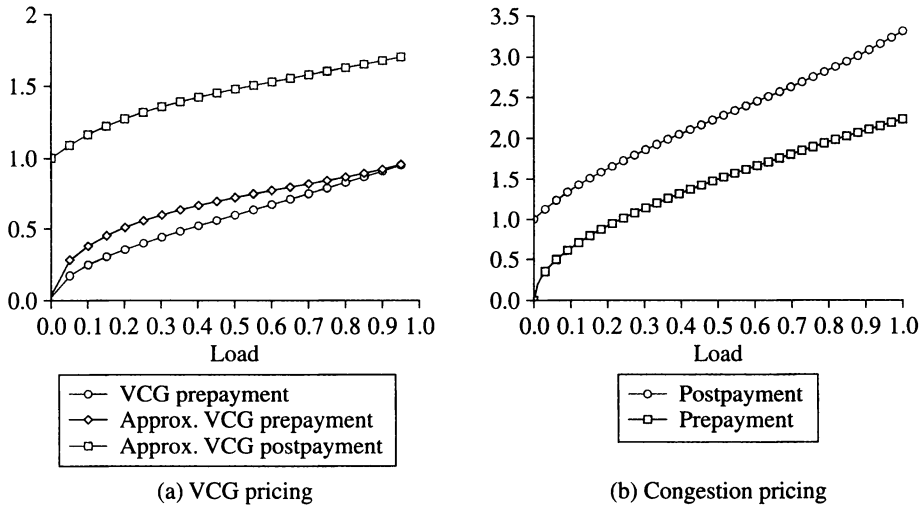


FIGURE 2: Coefficients of variation of prepayment and postpayment pricing as a function of the load ρ .

It is worth noting that if we rescale prices so as to equate the means, the comparison of the second moments amounts to comparing the coefficients of variation of the original prices. We have also reported the coefficient of variation of the pure VCG prepayment pricing, which can easily be numerically evaluated. The results are reported in Figure 2(a).

We deduce from Figure 2(a) that the approximating VCG scheme is sufficiently close to the pure VCG scheme. On the basis of this observation, we can observe that postpayment is much more volatile than prepayment. This can easily be understood by examining limiting cases. When $\rho \rightarrow 0$, $S \sim \sigma$ and, hence, $\tilde{\Pi}_{\text{post}} \sim \sigma/2$ and then the coefficient of variation of is equal $\sqrt{\mathbb{E}(\sigma^2)/\mathbb{E}(\sigma)^2 - 1} \sim 1$, while $\tilde{\Pi}_{\text{pre}} \sim A_0/2$ and has a very small coefficient of variation.

Similarly when $\rho \rightarrow 1$, $\tilde{\Pi}_{\text{post}} \sim S$ and the coefficient of variation is the coefficient of variation of the sojourn time, and is approximately equal to $\sqrt{3}$. For prepayment, we have $\tilde{\Pi}_{\text{pre}} \sim A_1 L$ and the coefficient of variation tends to 1.

From the above analysis it appears that for VCG pricing, prepayment is less volatile than postpayment. From a practical point of view, prepayment thus appears more suitable for customers because prices are less volatile and the scheme is much more simpler to implement.

Another key feature of VCG pricing is the domination of prepayment over postpayment for all load values. This property is also satisfied for congestion pricing. From (8) and (9), the coefficient of variation for prepayment in the case of congestion pricing is given by $\sqrt{4\rho + \rho^2}$ and for postpayment, from (28) and (31), the coefficient of variation is

$$\sqrt{\frac{6 + 59\rho + 10\rho^2 - 29\rho^3 - 2\rho^4}{(2 - \rho)(1 + \rho)^2(3 - 2\rho)}}.$$

We observe that coefficients of variation are greater under congestion pricing than those under VCG pricing. This latter policy is, hence, less volatile than congestion pricing.

5. Conclusion

In this paper we have developed a general method of studying pricing policies in packet networks where bandwidth is shared according to the equal share principle. The method has particularly been applied when flows have an exponential volume of data but some results hold for general volume distributions.

From the numerical results obtained in this paper, it appears that prepayment is both less volatile (less fluctuations in the price for the customer) and much easier to implement. The operator does not have to track the flow for all its lifetime. The price can be computed at flow arrival. Further studies are necessary to understand whether these results hold under other bandwidth sharing principles (for example, in the presence of head-of-line priority traffic or weighted fairness).

References

- [1] ABRAMOWITZ, M. AND STEGUN, I. A. (1964). *Handbook of Mathematical Functions, with Formulas, Graphs, and Mathematical Tables* (Nat. Bureau Standards Appl. Math. Ser. 55). U.S. Government Printing Office, Washington, DC.
- [2] ASARE, B. K. AND FOSTER, F. G. (1983). Conditional response times in the $M/G/1$ processor-sharing system. *J. Appl. Prob.* **20**, 910–915.
- [3] ASKEY, R. AND ISMAIL, M. (1984). Recurrence relations, continued fractions and orthogonal polynomials. *Mem. Amer. Math. Soc.* **49**, No. 300.
- [4] BIRMIWAL, S., MAZUMDAR, R. R. AND SUNDARAM, S. (2012). Processor sharing and pricing implications. In *Proc. ITC 24* (Krakow, Poland).
- [5] BULOW, J. AND KLEMPERER, P. (1996). Auctions versus negotiations. *Amer. Econom. Rev.* **86**, 180–194.
- [6] CHARRIS, J. A. AND ISMAIL, M. E. H. (1987). On sieved orthogonal polynomials. V. Sieved Pollaczek polynomials. *SIAM J. Math. Anal.* **18**, 1177–1218.
- [7] GUILLEMIN, F. AND BOYER, J. (2001). Analysis of the $M/M/1$ queue with processor sharing via spectral theory. *Queueing Systems* **39**, 377–397.
- [8] KELLY, F. (1997). Charging and rate control for elastic traffic. *Europ. Trans. Telecommun.* **8**, 33–37.
- [9] KLEINROCK, L. (1975). *Queueing Systems*, Vol. II. John Wiley, New York.
- [10] OZDAGLAR, A. AND SRIKANT, R. (2007). Incentives and pricing in communications networks. In *Algorithmic Game Theory*, Cambridge University Press, pp. 571–591.
- [11] REED, M. AND SIMON, B. (1980). *Methods of Modern Mathematical Physics, Vol. II, Fourier Analysis, Selfadjointness*. Academic Press, New York.
- [12] ROBERT, P. (2003). *Stochastic Networks and Queues* (Appl. Math. (New York) **52**). Springer, Berlin.
- [13] ROBERTS, J. W. AND MASSOULIÉ, L. (2000). Bandwidth sharing and admission control for elastic traffic. *Telecommun. Systems* **15**, 185–201.
- [14] SENGUPTA, B. AND JAGERMAN, D. L. (1985). A conditional response time of the $M/M/1$ processor-sharing queue. *AT&T Tech. J.* **64**, 409–421.
- [15] VICKREY, W. (1961). Counterspeculation, auctions, and competitive sealed tenders. *J. Finance* **16**, 8–37.
- [16] YAICHE, H., MAZUMDAR, R. R. AND ROSENBERG, C. (2000). A game theoretic framework for bandwidth allocation and pricing in broadband networks. *IEEE/ACM Trans. Networking* **8**, 667–678.
- [17] YANG, S. AND HAJEK, B. (2007). VCG-Kelly mechanisms for allocation of divisible goods: adapting VCG mechanisms to one-dimensional signals. *IEEE J. Selected Areas Commun.* **25**, 1237–1243.