Revenue Management for Online Advertising: Impatient Advertisers

Kristin Fridgeirsdottr
Management Science and Operations, London Business School, Regent’s Park, London, NW1 4SA, United Kingdom, kristin@london.edu

Sami Najafi Asadolahi
Management Science and Operations, London Business School, Regent’s Park, London, NW1 4SA, United Kingdom, snajafi@london.edu

The Internet is currently the fastest growing advertising medium. Online advertising brings new opportunities and has many different characteristics from advertising in traditional media that support efficient and mechanized decision making. We consider a web publisher that generates revenues from displaying advertisements on its website. The advertisers approach the web publisher, request their ad to be displayed to a certain number of visitors to the website, and are charged according to the so-called pay-per-impression pricing scheme. The advertisers are impatient and want their ads to be posted right away otherwise they approach another web publisher. We model the advertising operation of the web publisher as a queueing system with no waiting spaces (a loss system) where advertising slots correspond to servers. This system is different from known multi-server systems. We derive a closed-form solution for its steady state probabilities and analyze the system properties. We determine the optimal advertising price and provide managerial insights such that the optimal price is increasing in the number of impressions made to the viewers, which goes against the economy-of-scale intuition. The queueing model is compared to known queueing systems such as the bulk system and we provide additional results for those.

Key words: marketing: advertising and media, pricing; queues; inventory policies: marketing/pricing

1. Introduction

The Internet is currently the fastest growing advertising medium (TNS (2007)). It provides an access to an enormous consumer base and companies are increasingly allocating a larger portion of their marketing budget towards online advertising (IAB (2007)). Online advertising brings new opportunities and has different characteristics from advertising in traditional media such as television and newspapers. In online advertising it is possible to detect immediately how a visitor to a
website responds to advertisements (ads) and his actions can be kept track of. Furthermore, with the Internet’s capabilities to identify website’s visitors it is feasible to target ads on a level not possible before. These characteristics and others make online advertising suitable for more quantitative and automated decision making.

Online advertising can be divided into two domains: sponsored search advertising and display advertising. Sponsored search advertising involves advertisers paying a fee to appear next to search results for particular search words. Search engines such as Google and Yahoo generate a large portions of their revenues from this type of advertising and the pricing and allocation mechanism are usually based on bidding (for an overview and modeling see Feng et. al (2008)). Display advertising involves a web publisher, e.g., Yahoo and MSN, that provides services or content and instead of charging a subscription fee it generates revenues from posting ads on its site. The most popular display ad is the banner (other ad types are, e.g., pop-ups, pop-unders, and in-stream video).

In this paper we focus on display advertising and consider a web publisher that sells advertising space on its website. The web publisher has to manage the uncertain demand of the advertisers for the advertising space as well as the uncertain supply of the visitors to its website. In order to maximize its revenues it determines the optimal price to charge to attract a suitable amount of advertisers given the stream of visitors. Often in practise, the pricing is determined via estimations and negotiations. However, the nature of the Internet supports a more efficient and automated pricing approach. The objective of this paper is to provide a decision making tool as well as insights based on a stylized model of the operation of the web publisher to fill this gap.

Managing the advertising operation can be a challenging task. The web publisher usually manages multiple webpages within the website. The web publisher not only needs to decide on the price to offer for a certain number of impressions but is also faced by decisions such as how many advertising slots to have and how many advertisers to serve at the same time. The web publisher can in fact serve more advertisers than there are slots. One can notice this when refreshing a webpage that a new set of ads is often displayed. Also if a visitor stays for a long time on a website sometimes a new set of ads is displayed.
Web publishers use different pricing schemes to charge advertisers. The most common one is “pay-per-impression” where the advertiser is charged for each time a viewer visits the site where the ad is displayed. The fee is usually quoted as cost per thousand impressions or CPM (the M stands for the Roman M for milli). The cost varies with web publishers but one can commonly see a CPM of $4 - $60. Another frequently used pricing scheme is “pay-per-click”. In this type of a contract the web publisher charges a certain price (cost per click or CPC) for every time a visitor clicks on an ad. Usually the CPC is below a dollar. Other less frequently used schemes are pay-per-lead, pay-per-acquisition and pay-per-sale.

In this paper we develop a stylized model that captures many of the issues faced by a web publisher running an advertising operation. The model can also be used as a building block for other issues such as interactions between web publishers. We build on the novel modelling framework developed by Araman and Fridgeirsdottir (2007) for a web publisher that generates revenues from posting ads on its website and charges according to the most common pay-per-impression pricing scheme. As in Araman and Fridgeirsdottir (2007) we consider “arrivals” of advertisers interested in posting their ads and “arrivals” of viewers visiting the website. Given the dynamics of how the advertisers are served in a “synchronized” way by the viewers (illustrated in detail later) the resulting system is not a traditional multi-server system. In this paper, in contrast with Araman and Fridgeirsdottir (2007), the advertisers are assumed to be impatient and not willing to wait for their ads to be posted rather they would choose to go to another web publisher. The operation of the web publisher is modeled as a queueing system with no waiting spaces (often called a loss system). Usually in practice, there is a large number of web publishers that offer similar services and therefore, the advertisers do not have a reason to wait if their ad cannot be posted when they make their request. Hence, assuming an advertiser is lost if his demand is not met is a reasonable assumption in the competitive online advertising world. (For the same assumption in another setting see Savin et. al (2005).) From a technical point of view it improves the tractability of the problem enabling closed form solutions that can give further insights.
The first contribution of the paper is the development and analysis of a new queueing model for an operation of a web publisher that deals with impatient advertisers. A closed form solution is derived for the probability distribution of the number of advertisers in the system for a general number of advertising slots and impressions. This enables us to solve the revenue maximizing problem of the web publisher and we derive the optimal price to charge per impression. The model also serves as a building block for more other settings such as competition and advertising networks serving multiple advertisers and website (see a discussion on extensions in Section 7). The second contribution is the managerial insights derived for a web publisher on the optimal revenue and price. For example, we show that the optimal price decreases with the number of advertising slots on the page, which is quite intuitive. However, the optimal price increases with the number of impressions requested by the advertisers, which is counter intuitive compared to the usual economies-of-scale perception. Overall, we provide a pricing tool for a web publisher to determine the price to charge based on the demand for the advertising space and the web traffic, which can replace the current often ad-hoc approaches. Thirdly, we compare our queueing system to known systems. For the Erlang loss system we show that the average number of jobs is greater than for our system. For the bulk system we generate some additional results and insights. Finally, we provide additional insights through numerical examples.

The paper is organized as follows. In the following section the relevant literature is reviewed. Section 3 presents the model developed of the operation of a web publisher. The optimal price to charge for advertising is derived in Section 4 and numerical examples with further insights are provided in Section 5. In Section 6 we compare the queueing model developed to known queueing models. Finally, we conclude in Section 7 and give some insights for numerous future directions.

2. Literature Review

The literature on online advertising within the marketing area is quite extensive. Novak and Hoffman (2000) provide an overview of advertising pricing models for the internet. However, there is limited literature on analytical models for pricing and other decision making for a web publisher.
with an advertising operation. (For issues faced by advertisers such as predicting audience for advertising campaigns see, e.g., Danaher (2007) and papers referenced within.)

Savin et. al (2005) consider revenue management for rental businesses with two customer classes. As in our case traditional revenue management models are not adequate in their setting as those assume capacity to be sold by a finite horizon. Furthermore, in our case the capacity is random (the incoming viewers), which makes traditional revenue management approaches even less applicable. Even though the dynamics of their problem is different from ours resulting in different models there are some similarities such as the arrivals of the customers.

The paper by Araman and Popescu (2007) is conceptually related to our paper in the context of media. They study revenue management for traditional media, more specifically broadcasting. The issues considered there are of similar nature as in this paper. However, the setting is quite different. They consider the challenge faced by a media broadcasting company of allocating limited advertising space between up-front contracts and the scatter market in order to maximize revenues and meet contractual commitments. Even though the demand side of their problem is somewhat similar to ours the supply side of fixed commercial times with uncertain number of viewers is different from our arrival stream of viewers.

Many web publishers not only generate revenues from advertising but also from subscriptions. Several papers consider the trade-off between those two revenue streams. Baye and Morgan (2000) develop a simple economic model of online advertising and subscription fees. Prasad et. al (2003) model two offerings to viewers of a website: A lower fee with more ads and a higher fee with fewer ads. Kumar and Sethi (2006) study the problem of dynamically determining the subscription fee and the size of advertising space on a website. They use optimal control theory to solve the problem and obtain the optimal subscription fee and the optimal advertisement level over time. The optimal price to charge per impression of an ad is not determined.

In this paper we consider the pay-per-impression pricing scheme. Another pricing scheme is pay-per-click. Baye and Morgan (2003) develop an analytical approach to model consumer response to ad exposures at a website. Mangani (2003) compares the expected revenues from the pricing
strategy of pay-per-click with pay-per-impression using a simple deterministic model. Chickering and Heckerman (2003) develop a delivery system that maximizes click-through rate given inventory-management constraints in the form of advertisement quotas. None of these papers consider pricing decisions. In Fridgeirsdottir and Najafi (2008) we determine optimal prices for pay-per-click pricing schemes, which requires different models as the click-through probability depends on the number of ads displayed.

Scheduling of ads on a website is one of the most popular topics for using operations research tools in online advertising. Kumar et. al (2006) develop a model that determines how ads on a website should be scheduled in a planning horizon to maximize revenue. They consider geometry and display frequency as the two important factors specifying the ads. Their problem belongs to the class of NP-hard problems and they develop a heuristic to solve it. They also provide a good overview of other papers on scheduling.

The model developed in this paper uses the modeling framework developed by Araman and Fridgeirsdottir (2007). However, we consider the advertisers to be impatient, i.e., they are not willing to wait for an advertising slot to be available. (Savin et. al (2005) make the same assumption for customers of a rental business.) In addition to being a realistic assumption for the competitive advertising world and for most types of advertisers that require generic web publishers, it is analytically appealing. In this setting we can develop closed form solutions for the systems that does not seem possible when waiting is allowed.

3. Online Advertising Model

We consider a web publisher that generates revenues from posting ads on its website. We assume the web publisher is managing a single webpage but multiple webpages within the same website can be handled as proposed in Araman and Fridgeirsdottir (2007). The web publisher faces uncertain demand from advertisers wanting to post their ad and uncertain traffic from viewers visiting the website. There are slots for $n$ ads and the price for each is the same. (Different prices are briefly discussed in Section 7). The advertisers are assumed to approach the web publisher according to a
Poisson process with rate $\lambda_a$ and each advertiser requests their ad to be shown to $x$ viewers. The viewers are assumed to visit the website according to a Poisson process with rate $\lambda_v$. Both Poisson assumptions are required for tractability. The first is common in the literature (see e.g. Savin et. al (2005)). The second one can be criticized based on the fact that some research supports that web traffic shows self similarity, long range dependence and heavy tailed distribution (see Gong et. al (2005)), which are not properties of the Poisson process. However, other studies recognize that a Poisson distribution is a reasonable assumption (see Cao et. al (2002)) and we will make that assumption here.

When an advertiser approaches the web publisher two things typically happen. Either a slot is available for his ad or all slots are occupied by other advertisers. (Note a web publisher does not usually leave a slot empty rather it places its own ad in there. However, it would immediately free this slot up for a revenue generating advertiser.) If a slot is available the ad is posted and it stays on the website until $x$ viewers have visited. If no slot is available the advertiser approaches another web publisher and is lost. Given the fact that the advertisers are not willing to wait we refer to them as impatient. Here we are assuming that the number of advertisers that can be served cannot exceed the number of slots. However, in reality there are often more advertisers being served than there are slots. One can notice this when refreshing a webpage a new set of ads is often displayed. Araman and Fridgeirsdotir (2007) discuss how this can be incorporated into the modeling framework by introducing “versions” of the website. The same could be done for our model. However, we will be focusing here on the fundamental dynamics of a single version of a website.

The setting described above is quite simple and at first sight it seems equivalent to traditional queueing settings with slots corresponding to servers. However, what makes the system different is the service mechanism. When a viewer visits the website all the advertisers that have their ads posted are served with the arrival of the viewer. This means that the impressions left for all the ads go down by one at the same time. Hence, we can consider the system to have synchronized servers as in Araman and Fridgeirsdotir (2007).
It takes $x$ visitors with exponential interarrival time to service one advertiser, which means the service time of an advertiser has an $Erlang(\lambda_a, x)$ distribution. The fact that there are $n$ separate advertising slots gives this system the flavour of an $M/E_x/n/n$ queueing system (with $n$ independent servers and no waiting space). However, the fact that the slots operate in a synchronized manner and the service is initiated by one viewer at a time provides a distinctive feature. We compare our system to known queueing systems in Section 6.

The goal of the web publisher is to maximize the revenues from the ads. Each advertiser pays a price $p$ per impression made to a visitor. With $x$ impressions requested the total payment per advertiser is $px$. To capture the price sensitivity of the advertiser we assume a price demand relationship $p(\lambda_a)$ (more details to follow in Section 4). Then the revenue rate generated by the web publisher is:

$$R(\lambda_a) = \lambda_a (1 - P_n(\lambda_a)) p(\lambda_a) x$$

where $\lambda_a (1 - P_n(\lambda_a))$ is the actual arrival rate consisting of $\lambda_a$, the arrival rate of advertisers checking whether a slot is available, and $1 - P_n(\lambda_a)$ the probability of their request being accepted. $P_n(\lambda_a)$ is the probability that $n$ advertisers are being served, i.e., the system is full. In order to determine the optimal price for the web publisher to charge we need to derive $P_n(\lambda_a)$.

The next sections are dedicated to deriving $P_n$. The case of a single advertising slot is trivial as it is equivalent to the $M/E_x/1/1$ queueing system. With two slots and more, the synchronized service starts playing a role.

### Summary of Notations

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_a$</td>
<td>Arrival rate of advertisers</td>
</tr>
<tr>
<td>$\lambda_v$</td>
<td>Arrival rate of viewers</td>
</tr>
<tr>
<td>$r$</td>
<td>$\frac{\lambda_a}{\lambda_v}$</td>
</tr>
<tr>
<td>$n$</td>
<td>Number of advertising slots</td>
</tr>
<tr>
<td>$x$</td>
<td>Number of impressions</td>
</tr>
<tr>
<td>$P_i$</td>
<td>Probability of having $i$ advertisers in the system</td>
</tr>
<tr>
<td>$\pi_{i_1,i_2,...,i_n}$</td>
<td>Probability of having $i_k$ impressions left to satisfy in slot $k$</td>
</tr>
<tr>
<td>$\pi_{(i_1,i_2,...,i_n)}$</td>
<td>Probability of having $i_1, i_2, ..., i_n$ impressions left in any of the $n$ slots</td>
</tr>
<tr>
<td>$L$</td>
<td>Average number of advertisers in the system</td>
</tr>
<tr>
<td>$L_i$</td>
<td>Average number of advertisers in slot $i$ or occupation rate of slot $i$</td>
</tr>
<tr>
<td>$p$</td>
<td>The price charged per impression</td>
</tr>
</tbody>
</table>
3.1. The Case of a Single Advertising Slot

As mentioned earlier the case of a single slot is trivial. With advertisers arriving according to a Poisson process with rate $\lambda_a$ wanting $x$ impressions, and viewers visiting according to a Poisson process with rate $\lambda_v$, the resulting queueing system is a traditional $M/E_x/1/1$ queue with one server and no waiting space. Note that with only one slot for an ad there is no notion of synchronized or independent servers. For this case the probability of having the system full is $P_1 = \frac{rx}{1+rx}$ (see Gross and Harris (1998)) where $r = \frac{\lambda_a}{\lambda_v}$.

3.2. The Case of $n$ Advertising Slots

When there are two or more slots on the website the concept of synchronization starts playing a role and results for known queueing systems no longer apply. The system has a notion of a queueing system with signals (see Chao et. al (1999)) where the arrival of a viewer triggers a signal for the slots to work together. However, this literature includes mainly signals across nodes in networks and to our knowledge it does not provide us with solution approaches for our system.

Having Markovian arrival and service processes we model the system using Markov Chains. Even though we are ultimately interested in keeping track of the number of advertisers in the system, in order to form a Markov chain we need to keep track of the system at a more detailed level; i.e., of the number of impressions left for each slot. The slots are considered to be equivalent (in terms of price) and when an advertiser arrives to the system he is randomly assigned to any of the available slots with equal probability. (In Section 3.2.1 we consider another type of an ad-to-slot allocation.)
We define the state of the system as the number of impressions left to satisfy in each slot. The advertising slots are labeled from 1 to \( n \) and \( \pi_{k_1,k_2,...,k_n} \) is the probability that there are \( k_1, k_2, ..., k_n \) impression left in slots 1 to \( n \). The random ad-to-slot allocation means that the slots are symmetrical and we can keep track of the dynamics of the system without distinguishing between them. With this in mind we introduce the concept of permutation probabilities.

**Definition 1.** For a given vector \((k_1, k_2, ..., k_n)\) where each element represents the number of impressions left to satisfy in a slot we define \( \pi_{(k_1,k_2,...,k_n)} \), the permutation probability, as the sum over the probabilities of all possible permutations of \( k_1, k_2, ..., k_n \).

To illustrate let us look at an example of the permutation probabilities defined above for the case of three slots. If \( k_1, k_2 \) and \( k_3 \) are different numbers then the number of permutations are \( 3! = 6 \). Therefore \( \pi_{(k_1,k_2,k_3)} = \pi_{k_1,k_2,k_3} + \pi_{k_1,k_3,k_2} + \pi_{k_3,k_1,k_2} + \pi_{k_2,k_1,k_3} + \pi_{k_3,k_1,k_2} + \pi_{k_2,k_3,k_1} \). However, if \( k_1 = k_2 = a \) then the number of permutations among \( k_1, k_2 \) and \( k_3 \) are \( \frac{3!}{2!} = 3 \) and therefore \( \pi_{(k_1,k_2,k_3)} = \pi_{(a,a,k_3)} = \pi_{a,a,k_3} + \pi_{a,k_3,a} + \pi_{k_3,a,a} \).

**Remark 1.** The permutation probability \( \pi_{(k_1,k_2,...,k_n)} \) is the probability of finding \( k_1, k_2, ..., k_n \) impressions left in the \( n \) slots.

In the following proposition we state the closed form solution of the probabilistic properties of the system. Given the complexity of the system (see the flow balance equation in the proof of the proposition below) it is quite attractive and perhaps not expected to get closed form results.

**Theorem 1.** The probability of having \( k_1, k_2, ..., \) and \( k_i \) impressions left to satisfy in \( i \) slots and \( n - i \) empty slots is:

\[
\pi_{(k_1,k_2,...,k_i,0,...,0)} = \frac{r^i(1+r)^{n-i-1}}{\sum_{j=0}^{n} \binom{x+n-1}{j} r^j} \quad \text{for } i = 0, 1, 2, ..., n-1 \tag{2}
\]

\[
\pi_{(k_1,k_2,...,k_n)} = \frac{r^n}{\sum_{j=0}^{n} \binom{x+n-1}{j} r^j} \tag{3}
\]

Furthermore, the steady-state probability of having \( i \) advertisers in the system is:

\[
P_i = \frac{\binom{x+i}{i} r^i(1+r)^{n-i-1}}{\sum_{j=0}^{n} \binom{x+n-1}{j} r^j} \quad \text{for } i = 0, 1, 2, ..., n-1 \tag{4}
\]
\[ P_n = \frac{(\frac{x+n-1}{n})^n}{\sum_{j=0}^{n} \left( \frac{x+n-1}{j} \right)^j} \] (5)

All proofs are provided in the technical appendix. Note that even though \( \pi(k_1,k_2,\ldots,k_i,0,\ldots,0) \) does not depend on the actual number of impressions left in each slot only the number of filled slots, we need to keep track of the number of impressions left in each slot when deriving the formula. Therefore, we keep this notation. To give a simplified overview of the proof it involves listing the flow-balance equations and then verifying that Equations (2) and (3) satisfy them. Then the relevant terms from Equations (2) and (3) are added to prove Equations (4) and (5). (The proof of Proposition 1 gives further insights into the process of identifying equations such as Equations (2) and (3).)

With the proposition above we have fully characterized the probabilistic properties of the system with the closed form solution of the steady state probabilities.

Using Little’s law we can find the average number of advertisers in the system:

\[ L = \lambda_a (1 - P_n) \frac{x}{\lambda_v} \] (6)

Note that since the slots are equivalent, i.e., we do not distinguish between them, the average number of advertisers in each slot (the occupation rate) is simply \( L/n \).

The propositions below state some structural properties of the average number of advertisers in the system and the busy probability that will be useful when considering the pricing problem of the web publisher.

**Corollary 1.** The probability of all slots being occupied, \( P_n \), is decreasing in \( n \).

**Corollary 2.** The probability of all slots being occupied, \( P_n \), is increasing in \( \lambda_a \).

**Corollary 3.** The probability of all slots being occupied, \( P_n \), is increasing in \( x \).

**Corollary 4.** The average number of advertisers in the system, \( L \), is concave increasing in \( \lambda_a \).

**Corollary 5.** The average number of advertisers in the system, \( L \) is convex increasing in \( x \).

**Corollary 6.** The average number of advertisers in the system, \( L \) is increasing in \( n \).

Numerical examples illustrate that \( P_n \) is not necessarily concave in \( x \).
3.2.1. Ordered Ad-to-Slot Allocation

One can think of other realistic mechanism to allocate ads to slots in addition to the random allocation considered until now. For example, the advertisers might have preferences for some of the slots. We do not model the preferences explicitly in this paper as we are considering homogeneous advertisers. However, it is interesting to identify how the probabilistic properties of the system changes. We consider an ordered allocation to slots such that the slots are not considered to be equivalent. Rather if slots $i$ and $j$, $i > j$, are available then an ad is placed in slot $i$. This lack of symmetry makes the problem less tractable. We can no longer avoid keeping track of each slot specifically with an aggregated approach as we did for the random ad-to-slot allocation using the permutation probabilities.

We focus on the case of two slots and derive $\pi_{n,m}$, the steady-state probability of finding $n$ impressions left in slot 1 and $m$ impressions left in slot 2. A similar approach can be used for three slots but for a larger number of slots the problem seems nontractable. Note however, that no matter what the allocation of ads to slots is the probability of having $i$ ads on the website, $P_i$, is the same.

In the following proposition we state the closed form solution for the probability $\pi_{n,m}$.

**Proposition 1.** In a system with two advertising slots and an ordered ad-to-slot allocation, the steady-state probabilities for the number of impressions left to satisfy in each slot are:

$$
\begin{align*}
\pi_{0,0} &= \frac{1 + r}{1 + (x+1)r + \left(\frac{x+1}{2}\right)r^2} \\
\pi_{i,0} &= \frac{r(1 + ir)}{(1 + xr)(1 + (x+1)r + \left(\frac{x+1}{2}\right)r^2)} \\
\pi_{0,j} &= \frac{(x - j)^2}{(1 + xr)(1 + (x+1)r + \left(\frac{x+1}{2}\right)r^2)} \\
\pi_{i,j} &= \frac{(1 + (x + i - j)r)r^2}{(1 + xr)(1 + (x+1)r + \left(\frac{x+1}{2}\right)r^2)} \\
\pi_{i,i} &= \frac{(i - j)r^3}{1 + (x+1)r + \left(\frac{x+1}{2}\right)r^2}
\end{align*}
$$

where $i, j = 1, 2, ..., x$.

The proof of Proposition 1, provided in the technical appendix, gives through simple examples
thorough insights into how the equations above are derived in addition to verifying them using the flow-balance equations.

Having the probability distribution on hand we can derive further characteristics of the system. In the following corollary we obtain the occupation rate of each slot, which can also be considered as the average number of advertisers in a slot.

**Corollary 7.** Let $L_1$ and $L_2$ be the occupation rate of slots 1 and 2 respectively. Then $L_1 = \frac{rx}{1+x}$ and $L_2 = \frac{x^2r^2+x\left(\frac{x+1}{2}\right)r^3}{(1+rx)(1+rx)+x}.

We note that when an advertiser arrives and there are two empty slots the default is to allocate the ad to slot 1. That makes the slots asymmetric and we have $L_1 \neq L_2$. However, the sum of the occupation rates, $L_1 + L_2$, is the same as the average number of advertisers in the system for the random allocation mechanism, $L$.

Note that the occupation rate of the first slot is equal to the occupation rate in a one slot system. Hence, the first slot operates as the only slot in the single slot system and the second slot deals with some of the advertisers that would have been rejected in the single slot case. This observation can be generalized to the $n$ slot system with the ordered ad-to-slot allocation. In other words in an $n$ slot system the occupation rate of slot $i$ is the same as the occupation rate for slot $i$ in any $m$ slot system such that $i \leq m \land n$.

As mentioned before the probability of $i$ advertisers in the system, $P_i$, does not depend on how the ads are allocated to slots. Therefore, those probabilities for the ordered ad-to-slot allocation is the same as for the random allocation even though the probabilities of finding certain number of impressions in each slot and the occupation rates of the slots are not the same.

**3.2.2. Service Policy** We have made the natural assumption that each advertiser does not occupy more than one slot on a website. In addition to this service arrangement being more effective from the advertising point of view we can show that it is also more attractive from a service point of view. Note that based on the objective function in Equation (1) and Equation (6) higher average number of advertisers, $L$, gives more revenues.
Let us focus on a website with two slots. The web publisher could give both slots to one advertiser and instead of $x$ impressions per slot, $x/2$ impressions would be made for each of the two ads. This system is equivalent to a single slot system with $x/2$ impressions. We denote the average number of advertisers (the occupation rate) of this setting as $L_1(x/2)$. We compare this service policy to the one we have considered as the natural one of having only one advertiser in each slot with $x$ impressions each. We denote the average number of advertisers of this setting as $L_2(x)$. Using the fact that $L$ is increasing in $x$ and $n$ (see Corollaries 5 and 6) we note that $L_1(x/2) \leq L_1(x) \leq L_2(x)$. Furthermore, using Equations (1) and (6) we can show that the optimal revenue is higher when the same advertiser occupies only a single slot. Hence, there are not only marketing reasons for the service policy we have used so far but also operational reasons.

4. The Optimal Price

Having fully characterized the probabilistic properties of the web publisher’s operation we now focus on the decision making problem of the web publisher. The web publisher’s objective is to determine the price to charge per impression in order to maximize the revenue rate. (As most cost components are fixed we do not consider those. However, loss-of-goodwill cost for rejected advertisers can be incorporated into the model.) As stated in Section 3 the revenue rate generated is $R(\lambda_a) = \lambda_a(1 - P_n(\lambda_a))px$. According to Equation (6) it is equivalent to $R(\lambda_a) = L(\lambda_a)\lambda vp$.

Advertisers are usually sensitive to the price offered by the web publisher. Therefore, a web publisher offering lower prices can expect more interest from advertisers than a web publisher offering high prices. We capture this behavior by defining a continuous price-demand function $p(\lambda_a)$ that is assumed to be decreasing. For the moment we assume the price does not depend on the number of slots on the website or the number of impressions. In Section 5 we illustrate examples with these dependencies. As mentioned before the objective of the web publisher is to determine the price that maximizes the revenue rate. Since, we have a one-to-one relationship between the price and the arrival rate of the advertisers we can as well optimize with respect to $\lambda_a$ and then
determine the price from the price-demand function, \( p(\lambda_a) \). The optimization problem of the web publisher can be written as follows:

\[
\max_{\lambda_a} R(\lambda_a) = \lambda_a (1 - P_n) p(\lambda_a) x = L(\lambda_a) \lambda_a p(\lambda_a)
\]

\[
st \lambda_a \geq 0
\]

The following proposition ensures the uniqueness of the optimal solution and gives the optimal price.

**Proposition 2.** If \( p(\lambda_a) \) is concave then the revenue rate, \( R(\lambda_a) \), is concave in \( \lambda_a \) and the optimal price is \( p^* = p(\lambda_a^*) \) where \( \lambda_a^* \) satisfies the equation \( L'(\lambda_a^*) p(\lambda_a^*) + L(\lambda_a^*) p'(\lambda_a^*) = 0 \).

Note that in order to ensure concavity we need \( p(\lambda_a) \) to be concave. Even though this might seem a restrictive assumption it includes a linear price and numerical analysis indicates that many convex pricing function give a unimodal revenue function. (Other “weaker” conditions such assuming concave payment rate \( \lambda_a p(\lambda_a) \) or monotonicity of the price elasticity \( -\frac{d\lambda_a p}{dp} \) \( \lambda_a \) do not seem sufficient.)

The proposition below gives the intuitive results that the web publisher is better off with having more number of slots, offering higher number of impressions and having more traffic to its website. Note that here the advertisers are not considered to be sensitive to the number of slots and number of impressions, which we consider in Section 5.

**Proposition 3.** The optimal revenue rate, \( R(\lambda_a^*) \), is increasing in the number of slots, \( n \), the number of impressions, \( x \), and the arrival rate of the viewers, \( \lambda_v \).

The following proposition states the counter-intuitive result that the optimal price does not follow the economies-of-scale property with respect to \( x \).

**Proposition 4.** The optimal price, \( p^* \), is increasing in the number of impressions requested, \( x \) and the optimal revenues are increasing in \( x \).

The proposition above is quite interesting as one could expect the opposite, i.e., the price to be lower when more impressions are offered. In order to understand what drives these results we
notice that there are two competing forces. First, the higher the number of impressions the longer it takes to service each advertiser, which means that the web publisher does not need as many advertisers as before and can therefore charge a higher price. More impressions mean more demand on the capacity, i.e., the viewers, and therefore, less advertisers are needed. Second, in general a high demand helps to fill quickly any slot that becomes available. However, the first effect seems dominating, which results in higher price with lower demand when the number of impressions is high. Practically speaking, the web publisher should not offer quantity discounts from an operational point of view. However, there could be marketing reasons for offering quantity discount. We will explore those in the following section.

5. Numerical Examples

5.1. Advertising Slots

We first consider sensitivity with respect to the number of advertising slots. The viewers arrival rate is set $\lambda_v = 1000$ (which can be considered as number of viewers per hour). Each advertiser requests $x = 50000$ impressions. The price-demand function (per impression) for the advertisers is chosen to be $p(\lambda_a) = 0.02 - 0.2\lambda_a^c$, where $c = 0.8$, 1 or 1.2, i.e., the price function is convex, linear or concave. We explore the optimal price and the optimal revenues.

As illustrated in Proposition 3 and on Figure 2 there is a decreasing relationship between the
optimal revenues and the number of slots. The intuition behind this is that by increasing the number of slots we are in fact increasing the capacity of the system. In that case the web publisher can handle more advertisers at a time and to attract more of them he charges a lower price as indicated on Figure 3. However, this effect levels off as indicated for $c = 0.8$ in Figures 2 and 3.

Note that in our price demand function we have not taken into account that advertisers might not be willing to pay as much for their ad to be posted on a website with five ads compared to a website with only one ad. In general, advertisers are likely to be willing to pay less when websites have a high number of ads as the effective impact of each ad on the incoming viewers is expected to be less. Therefore, it is realistic to model the price per impression to be decreasing with the number of slots. The same applies to number of impressions (see Section 5.2). To capture this effect we set the price to depend not only on the arrival rate of advertisers $\lambda_a$ but also the number of slots $n$. We consider the following price function:

$$ p(\lambda_a) = 0.02 - 0.2\lambda_a^c - 0.001n $$

We continue to explore the sensitivity with respect to the number of advertising slots. We set the number of impressions as $x = 50000$. Figures 4 and 5 show the optimal revenue and the optimal price vs. number of advertising slots.
Comparing Figures 2 and 4 we can see that the optimal revenue does not continue to increase with the number of slots as before. Instead after a certain number of slots the impact of the price sensitivity with respect to number of slots starts playing a role and the revenue starts decreasing. Hence, here the optimal number of slots to choose would be three or four slots depending on the price-demand relationship. The optimal price decreases now faster with the number of slots; i.e., the web publisher has to lower the price to attract the customers lost due to the impact of the number of slots.
5.2. Impressions

Next we consider the sensitivity with respect to the number of impressions requested by the advertisers. We suppose that there are two slots on the website and the price-demand function for the advertisers is chosen as before to be \( p(\lambda_a) = 0.02 - 0.2\lambda_a^c \), where \( c = 0.8, 1 \) or 1.2. Figures 6 and 7 show how the number of impressions affects the optimal revenue and the optimal price.

As shown in Proposition 4 and illustrated in Figures 6 and 7 the optimal revenue and the optimal price increase with the number of impressions. Hence, the effect of economies-of-scale does not apply here. The intuition behind this insight is, as we mentioned before, that more impressions
means the web publisher needs less advertisers and thus can charge higher price.

Note that here we have not taken any quantity discounts into account in our pricing function that advertisers might expect and are often offered in reality, i.e., the web publisher could offer lower price per impression for higher total number of impressions. This phenomenon is incorporated in the price-demand function, \( p(\lambda_a) = 0.02 - 0.2\lambda_a^c - 10^{-7}x \). Using this function we explore how the optimal price and revenues change with the number of impressions with number of slots as \( n = 4 \).

By incorporating quantity discounts in the pricing, the optimal revenue does not continue to increase as before, instead it starts decreasing, indicating an optimal value for the number of

---

**Figure 8**  Optimal revenue vs. impressions with price depending on number of impressions

**Figure 9**  Optimal price vs. impressions with price depending on number of impressions
impressions to offer.

6. Comparison With Known Queuing Models

In this section we compare our model of the web publisher’s operation with some related models from the existing literature.

6.1. Erlang Loss System

We first compare the web publisher’s model with the $M/E_x/n/n$ queue, which is the so-called Erlang’s loss model. In the Erlang loss model the system does not provide any space for jobs to wait and the only jobs in the system are the ones being served by one of the $n$ servers. From this respect this system is very similar to our system. However, there is a substantial difference.

In the Erlang system the service channels operate independently. However, in our system the slots are synchronized, i.e., the advertisers receive service simultaneously. The Erlang’s loss formula that represents the probability distribution of the number of jobs in the system is the following:

$$P_i^E = \frac{(x r)^n}{n!} \sum_{j=0}^{i} \frac{(x r)^j}{j!}, \quad 0 \leq i \leq n$$

which we can compare to the distribution for the web publisher’s system:

$$P_i = \frac{(x + n - 1)^i}{\sum_{j=0}^{n} (x + n - 1)^j} (1 + r)^{n-i-1}, \quad 0 \leq i \leq n - 1$$

$$P_n = \frac{(x + n - 1)^n}{\sum_{j=0}^{n} (x + n - 1)^j} r^n$$

If $n = 1$ the two formulas yield the same results as we have explored before. This is because with only one slot in the system there is no notion of interdependence among slots and our model is reduced to the $M/E_x/1/1$ model. However, when there is more than one slot, $n > 1$, the two systems are different.

It is interesting to notice how the interaction between the empty slots and the occupied slots comes through in the online system. In the formula for $P_i$, $r^i$ seems to play the role of the $i$ occupied slots while $(1 + r)^{n-i-1}$ plays the role of the $n-i$ empty slots. The multiplication of those two terms
captures in some sense the effect of the interaction of \( i \) occupied slots with \( n - i \) empty slots. Since in \( P_n \) all the \( n \) slots are occupied there is no interaction between the empty and the occupied slots. Therefore, \( P_n \) does not have a term of the form \((1 + r)\). This is different from the \( M/E_x/n/n \) model where the formula for \( P_i \), \( 0 \leq i \leq n \), has the same format even though there are empty servers.

In the following proposition we specifically compare the probability of the system being full for the Erlang loss system and the web publisher’s system

**Proposition 5.** The probability of a fully occupied system is higher for the web publisher than for the Erlang loss system, i.e., \( P_n \geq P_n^E \). In addition, the average number of people in the web publisher’s system is less than the average number of jobs in the Erlang loss system, \( L \leq L_E \).

The proposition above indicates that the online system is in some sense less efficient than the Erlang loss system with independent servers. This means that synchronized servers are less efficient than independent servers, which is intuitive as the synchronization imposes a restriction.

### 6.2. Bulk Service

A bulk service system, often denoted \( M/M^{[n]}/1 \), has some similarities to the web publisher’s system. The set up of this system is the following. The arrivals are Poisson and the service time is exponential. There are \( n \) slots for service and infinite waiting space. When \( n \) or less jobs are in the system they are all served at the same time and if a job arrives during the service and a slot is empty that job is also served and finishes at the same time as the others (memoryless service property). If there are more than \( n \) jobs in the system only \( n \) are served (all at the same time) and the rest waits. This system with the additional assumption of no waiting space is the same as the online system with one impression. We can denote it by \( M/M^{[n]}/1/n \). Since having one impression is not realistic for the online setting there does not seem to be much to gain for us from the bulk service literature. However, let us find out whether we can use the results from the online system to learn more about the bulk service system. First, the solution of the system \( M/M^{[n]}/1 \) (Gross and Harris (1998)) is the following:

\[
P_0 = (1 - x_0)
\]
\[ P_i = (1 - x_0)x_i^i \quad i = 1, 2, 3, ... \]

where \( x_0 \) is the unique solution (between zero and one) of the characteristic equation: \( \mu x^{n+1} - (\lambda + \mu)x + \lambda = 0 \) with \( \lambda \) as the customers arrival rate and \( \mu \) as the service rate. The characteristic equation is of order \( (n + 1) \) and has at most \( (n + 1) \) roots. However, in most applications it has a unique solution. The drawback of this formula is when \( n \) is large or approaches infinity, i.e., the system has infinite capacity, the characteristic equation will be hard to solve. However, when \( n \to \infty \) the assumption of no waiting space does not play a role and the result from the online model can be used as seen in the following proposition.

**Proposition 6.** If the service capacity in the bulk service system is infinite then the probability distribution of the number of jobs is:

\[ P_i = \frac{r^i}{(1 + r)^{i+1}} \quad \text{for } i \geq 0 \]

In addition, the average number of jobs in the system is \( r \).

We can potentially take advantage of the proposition above to approximate bulk service systems with large \( n \). Let us explore this in a numerical example where we consider two systems \( M/M^{[n]}/1 \) and \( M/M^{[n]}/1/n \) (which is the same as the online system with \( n \) slots and one impression) with \( \lambda = 15 \) and \( \mu = 10 \) and different values for \( n \). When calculating accurately the average number of jobs in \( M/M^{[n]}/1 \) (using the characteristic equations above from Gross and Harris (1998)) and in \( M/M^{[n]}/1/n \) (using Equations (5) and (6) with \( x = 1 \)) we obtain a difference in \( L \) of less than 1.2% with \( n \geq 10 \). Figure 10 illustrates the difference.

The speed of the converges depends on the systems parameters. Note that \( L \) for the \( M/M^{[n]}/1 \) system is higher as could be expected since there can be jobs waiting in a queue ready to go into service while the \( M/M^{[n]}/1/n \) system needs to wait for the next arrival.

The full state probability is illustrated in Figure 11. The convergence of \( P_n \) seems to be a bit slower and there is less than 1.4% difference for \( n \geq 10 \).
Figure 10  Average number of jobs vs. the service capacity

Figure 11  The full state probability vs. the service capacity

Note that the bulk service system with Erlang service time is not the same as our online advertising system. In the online system the “jobs” can leave and enter the “bulk”; i.e., the jobs being served simultaneously are not necessarily in the same phase of the Erlang distribution. However, in the bulk service system all jobs belonging to the same bulk have the same service time.
7. Conclusion

In this paper we consider a web publisher that generates revenue by displaying ads on its website and charges according to the “pay-per-impression” pricing scheme. The web publisher’s operation is modeled as a queueing system where the arrival process corresponds to the advertisers interested in posting their ads and the service process corresponds to the viewers visiting the website. The advertisers are assumed to be impatient and not willing to wait for their ads to be posted rather they would choose to go to another web publisher. In queueing terms this corresponds to a system with no waiting spaces (a loss system). Given the dynamics of how the advertisers are served the service is “synchronized”.

Despite the complexity of the model we are able to derive a closed form solution of the probability distribution of the number of advertisers in the system for any number of advertising slots and any number of impressions made of each ad. We determine the optimal price to charge per impression and show, e.g., that it is increasing in the number of impressions, which goes against the economy-of-scale intuition.

The model developed is different from models existing in the literature and we compare this model to some known queueing models. It has similarities to bulk service models and we use our results to get further insights for special cases of the bulk service settings.

This paper provides a modeling framework for a web publisher’s operations that can be expanded beyond the results of the paper. We have provided closed form solutions and managerial insights for the case of homogeneous advertisers. Adding some heterogeneity to the characteristics of the advertisers, such as the number of impressions requested, would be interesting and would enable price differentiation. This might not be feasible for a general number of advertising slots but might be possible for two slots. Even though pay-per-impression is the most common pricing scheme others exist such as pay-per-click or even a mixture of these two. Pay-per-click contracts require different models and comparison of them to pay-per-impression contracts can be found in Fridgeirsdottir and Najafi (2008). As there can be many web publishers for an advertiser to choose from,
a competition setting would be realistic. In addition, given how often it can be easy to keep track of viewers’ behavior and profile, targeted advertising is very attractive to advertisers as well as to web publishers that can charge a higher price for more targeted audience. Furthermore, there exist advertising networks (such as AdEngage.com and Advertising.com) that give advertisers access to many different websites. The models developed in this paper can serve as building blocks for the operation of an advertising network.

We have analyzed the operation of the web publisher from a steady state point of view. Dynamic pricing would be interesting and possible to implement as advertisers often buy their advertising space online, which makes it feasible for the web publisher to change the prices dynamically.

In summary, the modeling framework developed in this paper with its closed-form solutions can provide a basis for multiple research directions that would explore analytically many relevant issues in online advertising.

8. Technical Appendix

Proof of Theorem 1 To prove Equations (2) and (3) we list the flow balance equations and show that the probabilities are of the form:

\[
\pi(k_1,k_2,...,k_i,0,...,0) = Ar^i(1+r)^{n-i-1} \quad i = 0, 1, 2, ..., n-1
\]

\[
\pi(k_1,k_2,...,k_n) = Ar^n \quad j = 1, ..., n
\]

where \(k_1, k_2, ..., k_n > 0\). Then we show by summing over the relevant \(\pi's\) that the probabilities of finding a certain number of advertisers in the system are of the form:

\[
P_i = \binom{x+i-1}{i} Ar^i(1+r)^{n-i-1} \quad i = 0, 1, 2, ..., n-1
\]

\[
P_n = \binom{x+n-1}{n} Ar^n \quad j = 1, ..., n
\]

Finally we use the fact that \(\sum_{i=0}^{n} P_i = 1\) and solve for \(A\).

We first consider a Markov chain where the state of the system is the vector \((k_1, k_2, ..., k_n)\) and \(k_i\) represents the number of impressions left to satisfy in a slot. (We do not distinguish between the
slots.) After identifying the possible transitions of the system we list the flow balance equations.

The equations are of five types:

1. \( r\pi_{(0,0,0,...,0)} = \pi_{(1,0,0,...,0)} + \pi_{(1,1,0,...,0)} + \pi_{(1,1,1,0,...,0)} + \cdots + \pi_{(1,1,1,...,1)} \)

2. \( (1 + r)\pi_{(k_1,k_2,...,k_i,0,0,...,0)} = \pi_{(k_1+1,k_2+1,...,k_i+1,0,0,...,0)} + \pi_{(k_1+1,k_2+1,...,k_i+1,1,0,...,0)} + \cdots \)

   \[ 1 \leq k_j \leq x - 1, \ j = 1,...i, \ 1 \leq i \leq n - 1 \]

3. \( (1 + r)\pi_{(x,x,...,x,k_1,k_2,...,k_i,0,0,...,0)} = r\pi_{(x,x,...,x,k_1,k_2,...,k_i,0,0,...,0)} \)

   \[ l \geq 1, \ 1 \leq k_j \leq x - 1, \ j = 1,...i, \ 1 \leq n - l - i \]

4. \( \pi_{(x,x,...,x,k_1,k_2,...,k_i)} = r\pi_{(x,x,...,x,k_1,k_2,...,k_i)} \)

   \[ l \geq 1, \ 1 \leq k_j \leq x - 1, \ j = 1,...i, \ l + i = n \]

5. \( \pi_{(k_1,k_2,...,k_i)} = \pi_{(k_1+1,k_2+1,...,k_i+1,0,...,0)} \)

   \[ 1 \leq k_j \leq x - 1, \ j = 1,...i \]

Next we verify that the functional form stated in Equations (7) and (8) satisfies the Flow Balance Equations i) - v):

i) By inserting Equations (7) and (8) into the flow balance equation we obtain a left hand side of \( Ar(1 + r)^{n-1} \) and a right hand side of \( A[r(1 + r)^{n-2} + r^2(1 + r)^{n-3} + \cdots + r^{n-1}(1 + r)^0 + r^n] \).

We use induction to show that both sides are equal, i.e., \( \sum_{j=1}^{n-1} r^j(1 + r)^{n-j-1} + r^n = r(1 + r)^{n-1} \).

We start with \( n = 1 \) and note that both sides are equal to \( r \). We now assume that the equality holds for \( n = k \), i.e., \( \sum_{j=1}^{k-1} r^j(1 + r)^{k-j-1} + r^k = r(1 + r)^{k-1} \). In order to show that the equality then holds for \( n = k + 1 \) we need to show that \( \sum_{j=1}^{k} r^j(1 + r)^{k-j} + r^{k+1} = r(1 + r)^{k} \). We have that \( 1 + (1 + r)^{k_j} + r^{k+1} = (1 + r) \sum_{j=1}^{k-1} r^j(1 + r)^{k-j-1} + r^k + r^{k+1} \). Using the induction assumption we obtain \( (1 + r)[r(1 + r)^{k-1} - r^k] + r^k + r^{k+1} = r(1 + r)^{k} - r^{k+1} + r^{k+1} = r(1 + r)^{k} \), which completes the induction proof.

ii) Using a similar approach as for Case i) we need to show that \( A(\sum_{j=0}^{n-i} r^{i+j}(1 + r)^{n-i-j-1} + r^n) = (1 + r)Ar^i(1 + r)^{n-i-1} \), i.e., \( \sum_{j=0}^{n-i-1} r^j(1 + r)^{n-i-j-1} + r^{n-i} = (1 + r)^{n-i} \). To simplify the notation we set \( m = n - i \). We then need to show that \( \sum_{j=0}^{m-1} r^j(1 + r)^{m-j-1} + r^m = (1 + r)^m \). We prove this equality by induction. If \( m = 1 \) both sides of the equality are \( 1 + r \). Let us now assume that the equality holds for \( m = k \), i.e., \( \sum_{j=0}^{k-1} r^j(1 + r)^{k-j-1} + r^k = (1 + r)^{k} \). We now need to show that the equality holds for \( m = k + 1 \), i.e., \( \sum_{j=0}^{k} r^j(1 + r)^{k-j} + r^{k+1} = (1 + r)^{k+1} \). We have
\[ \sum_{j=0}^{k} \pi^j (1 + r) (1 + r)^{k-j} + r^{k+1} = (1 + r) \sum_{j=0}^{k-1} \pi^j (1 + r)^{k-j-1} + r^k + r^{k+1}. \]

Using the induction assumption this equals to \((1 + r) ((1 + r)^k - r^k) + r^k + r^{k+1} = (1 + r)^{k+1}\), which completes the induction proof.

iii) We need to verify that \((1 + r) A r^{k+j} (1 + r)^{n-k-j-1} = r^j A r^{k+1} (1 + r)^{n-k-j}\), which always holds.

iv) We need to verify that \(A r^n = r A r^{n-1}\), which always holds.

v) This equation always holds.

When deriving \(A\) we first need to formulate \(P_i\), the probability that there are \(i\) advertisers in the system, and then use the fact that \(\sum_{i=0}^{n} P_i = 1\) to solve for \(A\). First, we know that \(P_0 = \pi_{(0,\ldots,0)} = A (1 + r)^{n-1}\). Let us then consider \(i, 0 < i < n\). The probability of having \(i\) advertisers in the system where each has \(k_j\) impressions left where \(0 < k_j \leq x, j = 1, \ldots, i\) is \(\pi_{(k_1, k_2, \ldots, k_i, 0, \ldots, 0)}\). \(P_i\) is the sum over all possible values of \(k_j\), \(j = 1, \ldots, i\). Note we have to be careful with the counting as \(\pi_{(k_1, k_2, \ldots, k_i, 0, \ldots, 0)}\) includes all the permutations of having \(k_j\) impressions left. We have that \(P_i = \sum_{k_1=1}^{x} \sum_{k_2=k_1}^{x} \ldots \sum_{k_i=k_{i-1}}^{x} \pi_{(k_1, k_2, \ldots, k_i, 0, \ldots, 0)} = B_i \pi_{(k_1, k_2, \ldots, k_i, 0, \ldots, 0)}\) where \(B_i\) represents the number of terms in the multiple sum. Determining the number of terms for \(i = 1\) is straightforward with \(P_2 = \sum_{k_1=1}^{x} \pi_{(k_1, 0, \ldots, 0)} = x \pi_{(1, 0, \ldots, 0)}\) and thus \(B_1 = x\). For \(i = 2\) we have \(P_2 = \sum_{k_1=1}^{x} \sum_{k_2=k_1}^{x} \pi_{(k_1, k_2, 0, \ldots, 0)}\).

To determine the number of terms, \(B_2\), we divide them into two parts: The terms with \(k_1 = k_2\) and the terms with \(k_1 \neq k_2\). We have \(x\) terms of the first type and \(\binom{x}{2}\) of the second. Therefore \(B_2 = x + \binom{x}{2} = \binom{x+1}{2}\).

To illustrate further the counting of the terms we give another example and consider

\[ P(5) = \sum_{k_1=1}^{x} \sum_{k_2=k_1}^{x} \ldots \sum_{k_5=k_4}^{x} \pi_{(k_1, k_2, \ldots, k_5, 0, \ldots, 0)} . \]

The case of \(k_1 = k_2 = \ldots = k_5\) gives \(x\) terms. The case of \(k_1 = k_2 = k_3 = k_4 \neq k_5\) gives \(\binom{x}{1} \binom{x-1}{1}\) terms.

The case of \(k_1 = k_2 = k_3 \neq k_4 = k_5\) gives \(\binom{x}{1} \binom{x-1}{1}\) terms. The case of \(k_1 = k_2 = k_3\) and \(k_4\) and \(k_5\) taking different values gives \(\binom{x}{1} \binom{x-1}{2}\) terms. The case of \(k_1 = k_2\) and \(k_3, k_4,\) and \(k_5\) taking different values gives \(\binom{x}{1} \binom{x-1}{3}\) terms. The case of \(k_1 = k_2, k_3 = k_4\) and the two values being different from \(k_5\) gives \(\binom{x}{2} \binom{x-2}{1}\) terms. Finally, the case of \(k_1 \neq k_2 \neq k_3 \neq k_4 \neq k_5\) gives \(\binom{x}{3}\) terms. Therefore, \(B_5\) will
be $B_5 = x + \binom{x}{1} \left( \frac{x-1}{1} \right) + \binom{x}{2} \left( \frac{x-1}{2} \right) + \binom{x}{3} \left( \frac{x-1}{3} \right) + \binom{x}{4} \left( \frac{x-2}{1} \right) + \binom{x}{5}$. After some algebra we can show that $B_5 = \frac{1}{120} x(x+1)(x+2)(x+3)(x+4) = \left( \frac{x+4}{5} \right)$.

According to Lemma 1, $B_i = \binom{x+i-1}{i}$. Therefore, we have $P_i = \binom{x+i-1}{i} Ar^i (1+r)^{n-i-1}$ for $i < n$ and $P_n = \binom{x+n-1}{n} Ar^n$. Since $\sum_{i=0}^{n} P_i = 1$ we have that $\sum_{i=0}^{n-1} \binom{x+i-1}{i} Ar^i (1+r)^{n-i-1} + \binom{x+n-1}{n} Ar^n = 1$, which gives $A = \frac{1}{\sum_{j=0}^{n} \binom{x+j-1}{j} r^j (1+r)^{n-j-1}(\binom{x+n-1}{n})r^n}$. Finally using Lemma 3 we have that $A = \frac{1}{\sum_{j=0}^{n} \binom{x+j-1}{j} r^j}$, which completes the proof. □

**Lemma 1.** If $B_i$ is the total number of the terms in the summation

$$P_i = \sum_{k_1=1}^{x} \sum_{k_2=k_1}^{x} ... \sum_{k_i=k_{i-1}}^{x} \pi(k_1,k_2,...,k_i,0,0,...,0) = B_i \pi(k_1,k_2,...,k_i,0,0,...,0)$$

then $B_i = \binom{x+i-1}{i}$, $i = 1, 2, ..., n$.

**Proof** We prove the lemma with induction. For the case $i = 1$, as discussed earlier, $B_1 = x = \binom{x+1-1}{1}$. Now let us assume that the formula for $B_i$ holds for $i = s$, i.e., $B_s = \binom{x+s-1}{s}$ for any $x$. We then need to show that it holds for $i = s + 1$, i.e., $B_{s+1} = \binom{x+s}{s+1}$. Let us condition our counting of terms on the value of $k_{s+1}$. We first assume $k_{s+1}$ takes the value 1. The number of the terms in this case will be exactly the same as for the problem with $s$ filled slots which is equal to $\binom{x+s-1}{s}$ according to the induction assumption. If $k_{s+1} = 2$ the other indices can vary from 2 to $x$. They can not take 1 anymore because all the states with 1 are already counted for in the case with $k_{s+1} = 1$.

The number of terms in this case will be similar as the first case except we only have $x-1$ values to choose from, i.e., $\binom{x+s-2}{s}$. With a similar reasoning for $k_{s+1} = 3$ we obtain $\binom{x+s-3}{s}$. Repeating the same reasoning we can see that $B_{s+1} = \binom{x+s-1}{s} + \binom{x+s-2}{s} + \binom{x+s-3}{s} + ... + \binom{a}{a}$. By using Lemma 2 we obtain that this summation is equal to $\binom{x+s}{s+1}$, which completes the proof. □

**Lemma 2.** For a fixed $k$, $\sum_{i=k}^{x+k-1} \binom{i}{k} = \binom{x+k}{k+1}$ for all $x$.

**Proof** We prove the lemma by induction. For $x = 1$ we have both sides equal to 1. Let us assume that for $x = s$ we have $\sum_{i=k}^{s+k-1} \binom{i}{k} = \binom{s+k}{k+1}$. We then need to show that for $x = s + 1$ we have $\sum_{i=k}^{s+k} \binom{i}{k} = \binom{s+k+1}{k+1}$. We can see that $\sum_{i=k}^{s+k} \binom{i}{k} = \sum_{i=k}^{s+k-1} \binom{i}{k} + \binom{s+k}{k}$ and by using the induction assumption we have $\sum_{i=k}^{s+k} \binom{i}{k} = \binom{s+k+1}{k+1} + \binom{s+k}{k}$. Using the Pascal’s rule, $\binom{a-1}{b} + \binom{a-1}{b-1} = \binom{a}{b}$, we obtain $\sum_{i=k}^{s+k} \binom{i}{k} = \binom{s+k+1}{k+1}$, which completes the proof. □
Lemma 3. Let \( x \) and \( n \) be two integer numbers and \( r \) be a real number then the following equality holds

\[
\sum_{i=0}^{n-1} \binom{x+i-1}{i} r^i (1+r)^{n-i-1} = \sum_{i=0}^{n-1} \binom{x+n-1}{i} r^i
\]

Proof We prove the lemma by induction. If \( n = 1 \) then both sides are equal to 1. Let us assume the equality holds for \( n = k \), i.e., \( C = \sum_{i=0}^{k-1} \binom{x+i-1}{i} r^i (1+r)^{k-i-1} - \sum_{i=0}^{k-1} \binom{x+k-1}{i} r^i = 0 \). Then we need to show it also holds for \( n = k + 1 \), i.e.,

\[
C = (1+r) \sum_{i=0}^{k} \binom{x+i-1}{i} r^i (1+r)^{k-i} - \sum_{i=0}^{k} \binom{x+k}{i} r^i = 0
\]

We use induction assumption to get

\[
C = (1+r) \sum_{i=0}^{k} \binom{x+k-1}{i} r^i + (x + k) r^k - \sum_{i=0}^{k} \binom{x+k}{i} r^i - (x+k) r^k = \sum_{i=0}^{k-1} \binom{x+k-1}{i} r^i - \sum_{i=0}^{k} \binom{x+k-1}{i} r^i + (x+k) r^k - (x+k) r^k = \sum_{i=0}^{k-1} \binom{x+k-1}{i} r^i + \sum_{j=0}^{k-1} \binom{x+k-1}{j} r^j + \sum_{i=0}^{k} \binom{x+k-1}{i} r^i = 0,
\]

which completes the proof.

Proof of Corollary 1 We prove \( P_{n+1} \geq P_n \) using contradiction. Suppose \( P_{n+1} > P_n \) then

\[
\frac{(x+n+1)}{r} \left( \frac{x+n}{n+1} \right)^r > \frac{(x+n)}{r} \left( \frac{x+n}{n} \right)^r.
\]

This gives \( (x+n+1) \sum_{i=0}^{n} \binom{x+n}{i} r^i > (x+n+1) \sum_{i=0}^{n} \binom{x+n}{i} r^i \). After some simplifications we have \( r \sum_{i=0}^{n} \binom{x+n}{i} r^i > \frac{r(n+1)!}{(n+1)!} r^n \sum_{i=0}^{n} \binom{x+n}{i} r^i \). Reindexing the sum on the right hand side by setting \( i = j+1 \) gives \( \sum_{j=0}^{n} \frac{r^j}{(x+n-j)!} > \frac{1}{r} \sum_{j=0}^{n} \frac{r^j}{(x+n-j)!} \). By comparing the sums term by term we see that each term on the left hand side is smaller than the corresponding one on the right hand side, which leads to a contradiction. Hence, we must have \( P_{n+1} \leq P_n \).

Proof of Corollary 2 By differentiating \( P_n \) with respect to \( r \) we obtain:

\[
\frac{dP_n}{dr} = \frac{n(x+n-1) \sum_{i=0}^{n} \binom{x+n-1}{i} r^{i-1} - (x+n-1) \sum_{i=0}^{n} \binom{x+n-1}{i} r^{i-1} + \sum_{i=0}^{n} \binom{x+n-1}{i} r^{i-1}}{\sum_{i=0}^{n} \binom{x+n-1}{i} r^{i-1}} \geq 0
\]

Hence, \( P_n \) is increasing in \( r \).

Proof of Corollary 3 After some calculations (see more details in the proof of Corollary 5) we have that \( P_n(x+1) - P_n(x) = \frac{\sum_{i=0}^{n} r^i (x+n-1) \binom{x+n}{i} \frac{r^i}{(x+n-i)!}}{\sum_{i=0}^{n} \binom{x+n-1}{i} r^{i-1}}. \) Therefore, \( P_n(x+1) - P_n(x) \geq 0 \).

Proof of Corollary 4 As \( r = \lambda_a / \lambda_v \) it is equivalent to show that \( L \) is concave increasing in \( r \). We know that \( L = rx(1 - P_n) = rx - rxP_n \). Hence, \( \frac{dL}{dr} = -x \frac{dP_n}{dr} \) and \( \frac{d^2L}{dr^2} = -x \frac{d^2P_n}{dr^2} \). We first show that \( \frac{dL}{dr} \geq 0 \). We have that:
\[
\frac{d(rP_n)}{dr} = \frac{\left(\sum_{i=0}^{n} r^i\right)^2}{\sum_{i=0}^{n} r^i(n+1-i)}.
\]

Hence, in order to ensure that \( \frac{dL}{dr} \geq 0 \) we need:

\[
\sum_{i=0}^{n} \left(\sum_{i=0}^{n} r^i(n+1-i)\right)\left[\sum_{i=0}^{n} \left(\sum_{i=0}^{n} r^i\right)^2\right] = \sum_{i=0}^{n} \sum_{j=0}^{n} \left(x^{n-1}\right) \left(x^{n-1}\right)^2 - \left[\sum_{i=0}^{n} \left(x^{n-1}\right)^2(n+1-i)\right]
\]

\[
= \sum_{i=0}^{n} \sum_{j=0}^{n} \left(x^{n-1}\right)^2(n+1-i)\left(x^{n-1}\right)^2 - \sum_{i=0}^{n} \left(x^{n-1}\right)^2(n+1-i) \geq 0,
\]

which is true according to Lemma 4. Hence, we have proved that \( L \) is increasing in \( r \).

Now let us show that \( \frac{d^2L}{dr^2} \leq 0 \), which is equivalent to showing \( \frac{d^2(rP_n)}{dr^2} \geq 0 \). From above we have that

\[
\frac{d^2(rP_n)}{dr^2} = \frac{\left(\sum_{i=0}^{n} r^i\right)^2}{\sum_{i=0}^{n} r^i(n+1-i)}.
\]

Therefore,

\[
-2\left[\sum_{i=0}^{n} \left(x^{n-1}\right)^2(n+1-i)\left(x^{n-1}\right)^2 - \sum_{i=0}^{n} \left(x^{n-1}\right)^2(n+1-i)\right] / \left[\sum_{i=0}^{n} \left(x^{n-1}\right)^2\right]^3 \quad \text{and after some algebra we have that}
\]

\[
\frac{d^2(rP_n)}{dr^2} = \frac{\left(\sum_{i=0}^{n} r^i\right)^2}{\sum_{i=0}^{n} r^i(n+1-i)}.
\]

\[
-2\left[\sum_{i=0}^{n} \left(x^{n-1}\right)^2(n+1-i)\left(x^{n-1}\right)^2 - \sum_{i=0}^{n} \left(x^{n-1}\right)^2(n+1-i)\right] / \left[\sum_{i=0}^{n} \left(x^{n-1}\right)^2\right]^3 \quad \text{and after some algebra we have that}
\]

\[
\frac{d^2(rP_n)}{dr^2} = \frac{\left(\sum_{i=0}^{n} r^i\right)^2}{\sum_{i=0}^{n} r^i(n+1-i)}.
\]

From Lemma 6 we know that \( \sum_{i=0}^{n} \sum_{j=0}^{n} \left(x^{n-1}\right)^2(n+1-i)(n+i-2j) \geq 0 \), which ensures \( \frac{d^2(rP_n)}{dr^2} \geq 0 \). Hence, \( L \) is concave increasing in \( r \). \( \square \)

**Lemma 4.** For \( x \geq 1, n \geq 1 \), we have

\[
\sum_{i=0}^{n} \sum_{j=0}^{n} \left(x^{n-1}\right)(x+n-1)(x+n-1) \geq \sum_{i=0}^{n} \left(x^{n-1}\right)i(n+1-i).
\]

**Proof** We prove this lemma by selecting a few “convenient” terms from the double sum on the left hand side and then show that their sum is greater than the sum on the right hand side.

We focus on the double sum on the left hand side of the inequality and notice since all its terms are positive this double sum is greater than a sum over a few of its terms. We first list the terms where \( i + j = 2n \), then the term with \( i + j = 2n - 1 \), etc:

\[
\sum_{i=0}^{n} \sum_{j=0}^{n} \left(x^{n-1}\right)(x+n-1) r^i j \geq \sum_{i=0}^{n} \left(x^{n-1}\right)i(n+1-i) \left(x^{n-1}\right)^n + \left(\frac{x^{n-1}}{n-1}\right)^n + \left(\frac{x^{n-1}}{n-2}\right)^n + \left(\frac{x^{n-1}}{n-3}\right)^n.
\]
For any \( i \leq j \leq n \) and \( x \geq 1 \), we have

\[
\sum_{i=0}^{n} r^{n+i} \left[ \sum_{j=0}^{n} (x+n-1) \frac{(x+n-1)}{i} \right] r^{i+j} (n+1-j) = 0,
\]
which completes the proof. \( \square \)

**Lemma 5.** For \( 0 \leq j \leq i \leq n \) and \( x \geq 1 \) we have

\[
(x+n-1)\frac{(x+n-1)}{i} \geq (x+n-1)\frac{(x+n-1)}{j}.
\]

**Proof** We prove the lemma by contradiction and assume \((x+n-1)\frac{(x+n-1)}{i} < (x+n-1)\frac{(x+n-1)}{j}\). After some algebra we have \(n!(x-1)!j!(x+n-1-j)! < i!(x+n-1-i)!j!(x+n-1-j+i)!\).

With further simplifications we get \(\Pi_{k=j+1}^{i} (n-i+k) \cdot \Pi_{k=j+1}^{i} (x+n-k) < \Pi_{k=j+1}^{i} (x+i-k)\), which is a contradiction as \(n \geq i\). Hence, we conclude that \((x+n-1)\frac{(x+n-1)}{i} \geq (x+n-1)\frac{(x+n-1)}{j}\). \( \square \)

**Lemma 6.** For any \( n \geq 0 \), \( \sum_{i=0}^{n} \sum_{j=0}^{n} (x+n-1)\frac{(x+n-1)}{i} r^{i+j} (n+1-i)(n+i-2j) \geq 0 \).

**Proof** The lemma can be proved using a similar approach as in the proof of Lemma 4. \( \square \)

**Proof of Corollary 5** We set \( L(x) = rx(1-P_n(x)) \). We first need to show that \( \Delta L = L(x+1) - L(x) \geq 0 \) and then we show that \( \Delta L \) is decreasing in \( x \).

We have \( \Delta L = rx(1-P_n(x+1)) + r(1-P_n(x+1)) - rx(1-P_n(x)) = rx(P_n(x) - P_n(x+1)) + r(1-P_n(x+1)) \). Focusing on the first part we have

\[
x(P_n(x) - P_n(x+1)) = x\left(\frac{(x+n-1)x}{\sum_{i=0}^{n} (x+n-1)i} - \frac{(x+n)x}{\sum_{i=0}^{n} (x+n)i}\right)
\]
\[ r^n \sum_{i=0}^{n} \left( \binom{x+n}{i} \right) r_i \left( \frac{x+n}{n} \right)^{n-i} \sum_{i=0}^{n} \left( \frac{x+n}{n} \right)^i \]

\[ = r^n \sum_{i=0}^{n} \left( \frac{x+n}{n} \right)^i \left( \frac{x+n}{n} \right)^{n-i} \sum_{i=0}^{n} \left( \frac{x+n}{n} \right)^i \]

\[ = \frac{r^n \sum_{i=0}^{n} \left( \frac{x+n}{n} \right)^i \left( x+n \right)^i \left( x+n \right)^{n-i}}{r^n \sum_{i=0}^{n} \left( \frac{x+n}{n} \right)^i \left( x+n \right)^{n-i}} \]

\[ = \frac{\sum_{i=0}^{n} \left( \frac{x+n}{n} \right)^i \left( x+n \right)^i \left( x+n \right)^{n-i}}{\sum_{i=0}^{n} \left( \frac{x+n}{n} \right)^i \left( x+n \right)^{n-i}} \]

\[ = \frac{\sum_{i=0}^{n} \left( \frac{x+n}{n} \right)^i \left( x+n \right)^i \left( x+n \right)^{n-i}}{\sum_{i=0}^{n} \left( \frac{x+n}{n} \right)^i \left( x+n \right)^{n-i}} \]

\[ = \frac{\sum_{i=0}^{n} \left( \frac{x+n}{n} \right)^i \left( x+n \right)^i \left( x+n \right)^{n-i}}{\sum_{i=0}^{n} \left( \frac{x+n}{n} \right)^i \left( x+n \right)^{n-i}} \]
For any $z$, $0 \leq z < n$, the coefficient for $r^z$ in $A$ is

$$
\sum_{i=0}^{z} \sum_{j=0}^{n-z} B(i, j)
$$

where we set $k = z - i - j$ and

$$
B(i, j) = \frac{(x+n-1)}{x(x+n-z+i+j)} \frac{(x+n-1)}{z(x+n-z+i+j)} \frac{(x+n-1)}{z(x+n-z+i+j+1)} \frac{(x+n-1)}{(x+n+1-i)(x+n-j+1)} + 1
$$

Since $z = i + j + k$ and $0 \leq z < n$ we have

$$
\frac{1}{x(x+1)(x+n-z+i+j)} - \frac{1}{x(x+n-i)(x+n-j+1)} = \frac{(-i-jn+n^2)+zn+2iz-2jz+(z-i-j)}{x(x+1)(x+n-z+i+j)(x+n-i)(x+n-j+1)} \geq 0
$$

as all three terms in the numerator are positive. Similarly

$$
\frac{(n-z+i+j)}{x(x+n-z+i+j)(x+n-z+i+j+1)} - \frac{(n-z+i+j)}{z(x+n-i)(x+n+1-j)(x+n-z+i+j)} = \frac{(n-z+i+j)(n-i)+(n^2-in-jn)+(nx+zx-2iz-2jz)+ij}{x(x+n-z+i+j)(x+n-z+i+j+1)(x+n-i)(x+n+1-j)} \geq 0.
$$

Therefore, the coefficient of $r^z$ is positive, which completes the proof for $0 \leq z < n$.

Proof of Corollary 6  The proof follows from Corollary 1 and Equation (6).

Proof of Proposition 1  We prove this proposition by first stating the flow balance equations. Then through examples we determine the functional structure of the probabilities $\pi_{n,m}$, and finally we verify the proposed equation for $\pi_{n,m}$ by ensuring the flow balance equations are satisfied.

We define the state of the system as the number of impressions left to satisfy in each slot. We consider all the possible states of the system and list their stationary flow-balance equations with the steady-state probability of finding $n$ impressions left in slot 1 and $m$ impressions left in slot 2, denoted by $\pi_{n,m}$:

<table>
<thead>
<tr>
<th>State</th>
<th>Stationary flow balance equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(0,0)$</td>
<td>$\lambda_a \pi_{0,0} = \lambda_v (\pi_{1,0} + \pi_{0,1} + \pi_{1,1})$ (9)</td>
</tr>
<tr>
<td>$(n,0)$</td>
<td>$0 &lt; n &lt; x$</td>
</tr>
<tr>
<td>$(0,m)$</td>
<td>$0 &lt; m &lt; x$</td>
</tr>
<tr>
<td>$(n,m)$</td>
<td>$0 &lt; n, m &lt; x$</td>
</tr>
</tbody>
</table>
\[(x, 0) \quad (\lambda_a + \lambda_v)\pi_{x,0} = \lambda_a\pi_{0,0} \quad (13)\]
\[(0, x) \quad (\lambda_a + \lambda_v)\pi_{0,x} = 0 \Rightarrow \pi_{0,x} = 0 \quad (14)\]
\[(x, x) \quad \lambda_a\pi_{x,x} = \lambda_a(\pi_{x,0} + \pi_{0,x}) \quad (15)\]
\[(x, m) \quad 0 < m < x \quad \lambda_a\pi_{x,m} = \lambda_a\pi_{0,m} \quad (16)\]
\[(n, x) \quad 0 < n < x \quad \lambda_a\pi_{n,x} = \lambda_a\pi_{n,0} \quad (17)\]

In addition we also know that the sum over all possible states should yield one:
\[
\sum_{i=0}^{x} \sum_{j=0}^{x} \pi_{i,j} = 1 \quad (18)
\]

Equations (9) - (17) are obtained by equating the flow into a state to the flow out of a state. To illustrate the derivation of the probability distribution let us consider the case with 3 impressions, \(x = 3\). Using Equations (9) to (17) we can write the probability matrix \(\Pi = (\pi_{ij})\) as:
\[
\Pi = \begin{bmatrix}
(1+r)(1+3r) & 2 & 1 & 0 \\
\frac{1+4r}{2} & 1+3r & 1+2r & 1+r \\
\frac{1+2r}{r} & r & 1+3r & 1+2r \\
\frac{1+3r}{2} & r & 2r & 1+3r
\end{bmatrix} \cdot \pi_{0,2} \quad (19)
\]

Using Equation (18) we can solve for \(\pi_{0,2}\) and then the other probabilities:
\[
\pi_{0,0} = \frac{1+r}{1+4r+6r^2}
\]
\[
\pi_{i,0} = \frac{r(1+i)r}{(1+3r)(1+4r+6r^2)}
\]
\[
\pi_{0,j} = \frac{(3-j)r^2}{(1+3r)(1+4r+6r^2)}
\]
\[
\pi_{i,j} = \frac{(1+(4+i-j)r)r^2}{(1+3r)(1+4r+6r^2)} \quad 0 < i < j \leq x
\]
\[
\pi_{i,j} = \frac{(i-j)r^3}{(1+3r)(1+4r+6r^2)} \quad 0 < j < i \leq x
\]
\[
\pi_{i,i} = \frac{1+4r+6r^2}{r^2}
\]

with \(i,j = 1,2,3\). Even though the general case of \(x\) impressions generates a very large problem we can take advantage of the pattern that emerged in Equation (19). Based on that pattern we propose a general version of the matrix in (19) for the case of \(x\) impressions. By ensuring that the probabilities satisfy the Flow Balance Equations (9) - (18) we can conclude that we have
determined the unique values for the probabilities. Based on the pattern that emerged for the case of 3 impressions (see Equation (19)) we claim that for \( x \) impressions \( \Pi = (\pi_{ij}) = A \cdot \pi_{0,x-1} \) with:

\[
A = \begin{bmatrix}
\frac{1+r}{r} & x-1 & x-2 & \cdots & 2 & 1 & 0 \\
\frac{1+rx}{r} & 1+rx & 1+(x-1)r & \cdots & 1+3r & 1+2r & 1+r \\
\frac{1+2r}{r} & r & 1+rx & \cdots & 1+4r & 1+3r & 1+2r \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
\frac{1+(x-2)r}{r} & (x-3)r & (x-4)r & \cdots & 1+xr & 1+(x-1)r & 1+(x-2)r \\
\frac{1+(x-1)r}{r} & (x-2)r & (x-3)r & \cdots & r & 1+xr & 1+(x-1)r \\
\frac{1+xr}{r} & (x-1)r & (x-2)r & \cdots & 2r & r & 1+xr
\end{bmatrix}
\]

To proof our claim we check whether all the \( \pi_{ij} \) satisfy Equations (9) to (17):

\[
\text{State} \quad \textrm{Stationary flow balance equation}
\]

\begin{align*}
(0,0) & \quad \pi_{1,0} + \pi_{0,1} + \pi_{1,1} = \left[ 1+\frac{x}{r} + x - 1 + 1 + xr \right] \pi_{0,x-1} = r \pi_{0,0} \\
(n,0) & \quad 0 < n < x \quad \pi_{n+1,0} + \pi_{n,1} = \left[ 1+\frac{(n+1)x}{r} + nr \right] \pi_{0,x-1} = \left( 1 + \frac{r}{n} \right) \pi_{n,0} \\
(0,m) & \quad 0 < n < x \quad \pi_{0,m+1} + \pi_{1,m+1} = \left[ x - (m+1) + 1 + (x-m)r \right] \pi_{0,x-1} = \left( 1 + \frac{r}{m} \right) \pi_{0,m} \\
(n,m) & \quad 0 < n, m < x \quad \pi_{n,m} = \pi_{n+1,m+1} \\
(x,0) & \quad \pi_{x,0} = \frac{1+rx}{r} \pi_{0,x-1} = \left( 1 + \frac{r}{x} \right) \pi_{0,0} \\
(0,x) & \quad \pi_{0,x} = 0 \\
(x,x) & \quad r \pi_{x,0} + \pi_{0,x} = (1 + xr) \pi_{0,x-1} = \pi_{x,x} \\
(x,m) & \quad 0 < m < x \quad r \pi_{0,m} = r (x-m) \pi_{0,x-1} = \pi_{x,m} \\
(n,x) & \quad 0 < n < x \quad r \pi_{n,0} = (1 + nr) \pi_{0,x-1} = \pi_{n,x}
\end{align*}

In order to determine \( \pi_{0,x-1} \) we then use the fact that \( \sum_{i=0}^{x} \sum_{j=0}^{x} \pi_{i,j} = 1 \). Notice that the summation of the probabilities in first the column of \( \Pi \) where \( j = 0 \) is equal to \( \left( \frac{x+x(x+1)}{r^2} \right) + \left( \frac{1+r}{r} \right) \pi_{0,x-1} \). Furthermore, the summation of each of the other columns is the same namely \( (x + \frac{x(x+1)}{r^2}) \pi_{0,x-1} \) and since we have \( x \) of those we have that the sum over all the probabilities is \( \left( \frac{x+x(x+1)}{r^2} \right) + \left( \frac{1+r}{r} \right) \pi_{0,x-1} \) and \( x(x + \frac{x(x+1)}{2}) \pi_{0,x-1} = 1 \). After some algebra we can solve for \( \pi_{0,x-1} \):

\[
\pi_{0,x-1} = \frac{r^2}{\left( 1 + \frac{x}{1} \right) \left( 1 + \frac{x+1}{1} \right) r + \left( \frac{x+1}{2} \right) r^2}
\]

In order to determine the rest of the probabilities we multiply the formula for \( \pi_{0,x-1} \) to the matrix \( A \) above, which gives the required result:

\[
\begin{align*}
\pi_{0,0} & = \frac{(1+r)(1+xr)}{r^2} \pi_{0,x-1} = \frac{1+r}{\left( 1 + \frac{x}{1} \right) \left( 1 + \frac{x+1}{1} \right) r + \left( \frac{x+1}{2} \right) r^2} \\
\pi_{i,0} & = \frac{1+ir}{r} \pi_{0,x-1} = \frac{r(1+ir)}{\left( 1 + \frac{x}{1} \right) \left( 1 + \frac{x+1}{1} \right) r + \left( \frac{x+1}{2} \right) r^2} \\
\pi_{0,j} & = (x-j) \pi_{0,x-1} = \frac{(x-j)r^2}{[r+xr][1+(x+1)r + \left( \frac{x+1}{2} \right) r^2]}
\end{align*}
\]
Revenue Management for Online Advertising: Impatient Advertisers

According to Corollary 2, the expression for the optimal price follows from the FONC. We get that

\[ p = \left(1 + \frac{(x+1)(x+2)r^2}{(x+1)r^2 + (x+2)r^2}\right) \]

For the third part of the proof, we note that the busy probability \( P_n \) depends only on \( r = \lambda_a/\lambda_v \), not on \( \lambda_a \) and \( \lambda_v \) separately. Adapting our notation we denote the optimal revenues with \( \lambda_v \) as the arrival rate of the viewer as \( R^*(\lambda_v) = R(\lambda_v^*(\lambda_v), \lambda_v) = \lambda_v^*(\lambda_v)(1 - P_n(\lambda_v^*(\lambda_v)/\lambda_v))p(\lambda^*(\lambda_v)). \)

According to Corollary 2, \( P_n \) is increasing in \( \lambda_a \) (and \( r \)) and thus decreasing in \( \lambda_v \). Using that fact and optimality we have for \( \lambda_a^1 \geq \lambda_v^2 \) that \( R^*(\lambda_v^1) \geq \lambda_v^2(1 - P_n(\lambda_v^2/\lambda_v))p(\lambda^*(\lambda_v^2)) \geq \lambda_v^2(1 - P_n(\lambda_v^2/\lambda_v))p(\lambda^*(\lambda_v^1)) = R^*(\lambda_v^1) \), which completes the proof. □
Proof of Proposition 4  The proof involves using the FONC and the Implicit Function Theorem as well as comparing terms in multiple sums.

We need to show that:

\[
\frac{d\lambda^*_a}{dx} = - \frac{dF}{d\lambda^*_a} = - \frac{\frac{d}{dx}(L'(\lambda^*_a))p(\lambda^*_a) + \frac{d}{dx}(L(\lambda^*_a))p'(\lambda^*_a)}{L''(\lambda^*_a)p(\lambda^*_a) + L'(\lambda^*_a)p'(\lambda^*_a) + L(\lambda^*_a)p''(\lambda^*_a)} \leq 0.
\]

Note that since \( x \) is discrete we slightly abusing the Implicit Function Theorem. However, we treat \( x \) for the remainder as discrete and, e.g., \( \frac{d}{dx}(L'(\lambda^*_a)) \) corresponds to \( \Delta(L'(\lambda))\big|_{\lambda=\lambda^*_a} \). Since, \( p(\lambda^*_a) > 0 \), \( p'(\lambda^*_a) < 0 \), \( p''(\lambda^*_a) < 0 \) and \( L(\lambda^*_a) > 0 \), \( L'(\lambda^*_a) > 0 \), \( L''(\lambda^*_a) < 0 \) the denominator is negative. Hence, we need to show that \( \frac{d}{dx}(L'(\lambda^*_a))p(\lambda^*_a) + \frac{d}{dx}(L(\lambda^*_a))p'(\lambda^*_a) \leq 0 \). Using the FONC, \( L'(\lambda^*_a)p(\lambda^*_a) + L(\lambda^*_a)p'(\lambda^*_a) = 0 \), we are left with showing that

\[
g(\lambda^*_a) = \frac{d}{dx}(L(\lambda^*_a))L'(\lambda^*_a) - \frac{d}{dx}(L'(\lambda^*_a))L(\lambda^*_a) \geq 0
\]

Without loss of generality we set \( \lambda = 1 \) and thus \( \lambda^*_a = r \). Now we have \( L = rx(1-P_x) \) and then \( \frac{dx}{dx} = r(x+1)(1-P_{x+1}) - rx(1-P_x) \). (We denote \( P_n \) with \( P_x \) to emphasize the dependence on \( x \).) Also from the proof of Corollary 4 we have that \( L' = x(1-f_x) \) where

\[
f_x = \frac{\sum_{i=0}^{n} \sum_{j=0}^{n} \sum_{k=0}^{n} \sum_{l=0}^{n} \binom{x+n+1}{i} \binom{x+n-1}{j} \binom{x+n}{k} \binom{x+n}{l} r^{i+j+n+k(n-k+1)}}{\sum_{i=0}^{n} \sum_{j=0}^{n} \sum_{k=0}^{n} \sum_{l=0}^{n} \binom{x+n-1}{i} \binom{x+n-1}{j} \binom{x+n}{k} \binom{x+n}{l} r^{i+j+n+k(n-k+1)}}
\]

and then \( \frac{df}{dx} = 1 - x(f_{x+1} - f_x) - f_{x+1} \). Hence, after some algebra we have that \( g(\lambda^*_a) = (1 - P_x)(f_{x+1} - f_x) - (1 - f_x)(P_{x+1} - P_x) \). Next we will calculate each term in \( g(\lambda^*_a) \) by inserting the relevant functions. We have

\[
f_{x+1} - f_x = \frac{\sum_{i=0}^{n} \sum_{j=0}^{n} \sum_{k=0}^{n} \sum_{l=0}^{n} \binom{x+n-1}{i} \binom{x+n-1}{j} \binom{x+n}{k} \binom{x+n}{l} r^{i+j+n+k(n-k+1)}}{\sum_{i=0}^{n} \sum_{j=0}^{n} \sum_{k=0}^{n} \sum_{l=0}^{n} \binom{x+n-1}{i} \binom{x+n-1}{j} \binom{x+n}{k} \binom{x+n}{l} r^{i+j+n+k(n-k+1)}}
\]

Therefore,

\[
(1 - P_x)(f_{x+1} - f_x) = \frac{\sum_{i=0}^{n} \sum_{j=0}^{n} \sum_{k=0}^{n} \sum_{l=0}^{n} \binom{x+n-1}{i} \binom{x+n-1}{j} \binom{x+n}{k} \binom{x+n}{l} r^{i+j+n+k(n-k+1)}}{\sum_{i=0}^{n} \sum_{j=0}^{n} \sum_{k=0}^{n} \sum_{l=0}^{n} \binom{x+n-1}{i} \binom{x+n-1}{j} \binom{x+n}{k} \binom{x+n}{l} r^{i+j+n+k(n-k+1)}}
\]

\[
\frac{\sum_{i=0}^{n} \sum_{j=0}^{n} \sum_{k=0}^{n} \sum_{l=0}^{n} \binom{x+n-1}{i} \binom{x+n-1}{j} \binom{x+n}{k} \binom{x+n}{l} r^{i+j+n+k(n-k+1)}}{\sum_{i=0}^{n} \sum_{j=0}^{n} \sum_{k=0}^{n} \sum_{l=0}^{n} \binom{x+n-1}{i} \binom{x+n-1}{j} \binom{x+n}{k} \binom{x+n}{l} r^{i+j+n+k(n-k+1)}}}
\]

Knowing that

\[
P_{x+1} - P_x = \frac{\sum_{i=0}^{n} \sum_{j=0}^{n} \binom{x+n-1}{i} \binom{x+n-1}{j} r^{i+j}(n-i)}{\sum_{i=0}^{n} \sum_{j=0}^{n} \binom{x+n-1}{i} \binom{x+n-1}{j} r^{i+j}}
\]
and

\[ 1 - f_z = \frac{\sum_{i=0}^{n} \sum_{j=0}^{n} \binom{x+n-1}{i} \binom{x+n-1}{j} r^{i+j} - \sum_{i=0}^{n} \binom{x+n-1}{i} \binom{x+n}{n} r^n i (n-i+1)}{\sum_{i=0}^{n} \sum_{j=0}^{n} \binom{x+n-1}{i} \binom{x+n}{j} r^{i+j}} \]

we have that

\[
(1 - f(x)) (P_{x+1} - P_x) = \left\{ \begin{array}{l}
\sum_{i=0}^{n} \sum_{j=0}^{n} \sum_{k=0}^{n} \binom{x+n-1}{i} \binom{x+n-1}{j} \binom{x+n}{k} \frac{(n-k)}{x} r^{i+j+k+1} (n-k+1) \\
\sum_{i=0}^{n} \sum_{j=0}^{n} \sum_{k=0}^{n} \binom{x+n-1}{i} \binom{x+n-1}{j} \binom{x+n}{k} \frac{(n-k)}{x} r^{i+j+k+1} (n-k+1) \\
\sum_{i=0}^{n} \sum_{j=0}^{n} \sum_{k=0}^{n} \binom{x+n-1}{i} \binom{x+n-1}{j} \binom{x+n}{k} \frac{(n-k)}{x} r^{i+j+k+1} (n-k+1) \\
\sum_{i=0}^{n} \sum_{j=0}^{n} \sum_{k=0}^{n} \binom{x+n-1}{i} \binom{x+n-1}{j} \binom{x+n}{k} \frac{(n-k)}{x} r^{i+j+k+1} (n-k+1) \end{array} \right\}
\]

Adding equations (21) and (22) we have that

\[ g(\lambda^*_g) = \left\{ \begin{array}{l}
x \sum_{i=0}^{n} \sum_{j=0}^{n} \sum_{k=0}^{n} \binom{x+n-1}{i} \binom{x+n-1}{j} \binom{x+n}{k} \frac{(n-k)}{x} r^{i+j+k+1} (n-k+1) \\
x \sum_{i=0}^{n} \sum_{j=0}^{n} \sum_{k=0}^{n} \binom{x+n-1}{i} \binom{x+n-1}{j} \binom{x+n}{k} \frac{(n-k)}{x} r^{i+j+k+1} (n-k+1) \\
x \sum_{i=0}^{n} \sum_{j=0}^{n} \sum_{k=0}^{n} \binom{x+n-1}{i} \binom{x+n-1}{j} \binom{x+n}{k} \frac{(n-k)}{x} r^{i+j+k+1} (n-k+1) \\
x \sum_{i=0}^{n} \sum_{j=0}^{n} \sum_{k=0}^{n} \binom{x+n-1}{i} \binom{x+n-1}{j} \binom{x+n}{k} \frac{(n-k)}{x} r^{i+j+k+1} (n-k+1) \end{array} \right\}
\]

To show that \( g(\lambda^*_g) \geq 0 \) we need to systematically group the terms in the four sums together and show that the sum of the terms in each group are positive. We do the grouping according to the power of \( r \). Let us assume that the power of \( r \) is \( z \) where \( 0 \leq z \leq 4n \). If \( g(\lambda^*_g) \geq 0 \) then the coefficient of \( r^z \) for each and every \( z \), \( 0 \leq z \leq 4n \), needs be positive. We divide the range into four parts: \( 0 \leq z < n \), \( n \leq z < 2n \), \( 2n \leq z < 3n \), \( 3n \leq z \leq 4n \). Here we will illustrate the proof for \( 0 \leq z < n \). The other ranges can be proved using a similar approach.

Let \( B(x, n, z) \) be the coefficient of \( r^z \) for any given \( z \). If \( 0 \leq z < n \) and \( l = z - i - j - k \) then after some algebra we have that \( B(x, n, z) = B(x, n, z) = \sum_{i=0}^{z} \sum_{j=0}^{z} \sum_{k=0}^{z} H_{i,j,k}(x, n) C_{i,j,k}(x, n) \) where

\[ H_{i,j,k}(x, n) = \left( \frac{x+n-1}{i} \right) \left( \frac{x+n}{j} \right) \left( \frac{x+n}{k} \right) \left( \frac{x+n}{z} \right)^2 \geq 0 \] and \( C_{i,j,k}(x, n) = \frac{(n-k+1)}{(x+n-k)} - \frac{1}{x+n-k} \left( \frac{1}{z-t-j-k} \right) \).

After some algebra we have that

\[ C_{i,j,k}(x, n) \geq \frac{ij(n-l-1)+n^2(n-i-j-l-1)+kz(n-l)+n(2n-2i-2j-l-2)+x^2(n-i-j-1)}{(n+x-l)(x+n-j)(n+x-k)(n+x-k)} \geq 0. \]

Keeping in mind that \( z = i + j + k + l \) and \( 0 \leq z < n \) we notice that each term in the numerator (and the denominator) is positive. Hence, \( B(x, n, z) \geq 0 \). Given that other ranges for \( z \) hold we have that \( g(\lambda^*_g) \geq 0 \), which ensures \( \frac{d\lambda^*_g}{dx} \leq 0 \). As the price is decreasing in \( \lambda_a \), we have proved that the optimal price is increasing in \( x \).
Proof of Proposition 5 In the Erlang loss system with \( n \) servers the probability of a full system is \( P^E_n = \frac{(xr)^n}{\sum_{i=0}^{n} \binom{n}{i} (x+i-1)^i} = \frac{r^n}{\sum_{i=0}^{\infty} \binom{n}{i} (x+i-1)^i}. \) For the web publisher’s system it is \( P_n = \frac{(x+n-1)^n_i}{\sum_{i=0}^{n} \binom{n}{i} (x+i-1)^i} = \frac{r^n}{\sum_{i=0}^{\infty} \binom{n}{i} (x+i-1)^i} \geq \frac{r^n}{\sum_{i=0}^{n} \binom{n}{i} (x+i-1)^i} = P^n_E. \)

From Equation (6) we have that the average number of advertisers in the web publisher’s system is \( L = rx(1 - P_n). \) According to Harel (1985) the same formula holds for the average number of jobs in the Erlang loss system, i.e., \( L^E = rx(1 - P^E_n). \) As \( P_n \geq P^n_E \) we have that \( L \leq L^E. \)

Proof of Proposition 6 For the web publisher’s system we showed in Theorem 1 that the probability distribution of the number of the advertisers in the system is:

\[
P_i = \frac{(x+i-1)^i(1+r)^{n-i-1}}{\sum_{i=0}^{n} \binom{n}{i} (x+i-1)^i} \quad i = 0, 1, 2, ..., n - 1
\]

\[
P_n = \frac{r^n}{\sum_{i=0}^{\infty} \binom{n}{i} (x+i-1)^i}
\]

With \( x = 1 \) and \( n \to \infty \) we will get the distribution of the bulk service system with infinite capacity. The distribution with \( x = 1 \) and finite \( n \) is:

\[
P_i = \frac{r^i(1+r)^{n-i-1}}{\sum_{i=0}^{n} \binom{n}{i} (1+r)^i} = \frac{r^i(1+r)^{n-i-1}}{(1+r)^n} \quad \text{for} \ i = 0, 1, 2, ..., n - 1
\]

\[
P_n = \frac{r^n}{\sum_{k=0}^{\infty} \binom{n}{k} (1+r)^k} = \frac{r^n}{(1+r)^n}
\]

as \( \sum_{k=0}^{\infty} \binom{n}{k} (1+r)^k = (1+r)^n. \) Therefore, we have:

\[
P_i = \frac{r^i(1+r)^{n-i-1}}{(1+r)^n} = \frac{r^i}{(1+r)^{i+1}} \quad i = 0, 1, 2, ..., n - 1
\]

\[
P_n = \frac{r^n}{\sum_{k=0}^{n} \binom{n}{k} (1+r)^k} = \left(\frac{r}{1+r}\right)^n
\]

With \( n \to \infty \) then \( P(n) \to 0, \) i.e., the probability of a full system is zero. However, \( P_i = \frac{r^i}{(1+r)^{i+1}} \forall i \in \mathbb{N}. \)

For the second part, as \( n \to \infty \) then \( P_n \to 0 \) and therefore \( L = r(1 - P_n) \to r. \) \( \square \)

References


TNS Media Intelligence. 2007.