Revenue-Maximizing Rankings for Online Platforms with Quality-Sensitive Consumers

Pierre L'Ecuyer

Département d'Informatique et de Recherche Opérationnelle, Université de Montréal, C.P. 6128, Succ. Centre-Ville, Montréal, H3C 3J7, Canada; and Inria Rennes – Bretagne Atlantique, lecuyer@iro.umontreal.ca

Patrick Maillé

Telecom Bretagne, 2 rue de la Châtaigneraie CS 17607, 35576 Cesson Sévigné Cedex, France, patrick.maille@telecom-bretagne.eu

Nicolás E. Stier-Moses

Universidad Torcuato Di Tella, Saenz Valiente 1010, Buenos Aires, Argentina; and CONICET Argentina, nstier@utdt.edu

Bruno Tuffin

Inria Rennes – Bretagne Atlantique, Campus Universitaire de Beaulieu, 35042 Rennes Cedex, France, bruno.tuffin@inria.fr

When a keyword-based search query is received by a search engine, a classified ads website, or an online retailer site, the platform has exponentially many choices in how to sort the search results. Two extreme rules are (a) to use a ranking based on estimated relevance only, which improves customer experience in the long run because of perceived quality, and (b) to use a ranking based only on the expected revenue to be generated immediately, which maximizes short-term revenue. Typically, these two objectives, and the corresponding rankings, differ. A key question then is what middle ground between them should be chosen. We introduce stochastic models that yield elegant solutions for this situation, and we propose effective solution methods to compute a ranking strategy that optimizes long-term revenues. This strategy has a very simple form and is easy to implement if the necessary data is available. It consists in ordering the output items by decreasing order of a score attributed to each, similarly to value models used in practice by e-commerce platforms. This score results from evaluating a simple function of the estimated relevance, the expected revenue of the link, and a real-valued parameter. We find the latter via simulation-based optimization, and its optimal value is related to the endogeneity of user activity in the platform as a function of the relevance offered to them.

Key words: E-commerce, Linear ordering, Revenue management, Dynamic programming.

1. Introduction

Electronic commerce via the Internet has increased and evolved tremendously in recent years. Marketplaces in which participants can conveniently buy, sell, or rent a huge variety of objects and services are now common. The Internet has evolved into a complex ecosystem of companies for which various business models have proved profitable. Among them, we find search engines (SEs) such as Google, that allow

users to find content of their interest on the web, and use these interactions to create opportunities to sell ads; online retailers such as Amazon.com that act as intermediaries between producers and consumers; and classified ad websites such as eBay that allow sellers or service-providers, and buyers or service-consumers, respectively, to meet and conduct transactions. To be profitable, those marketplaces typically rely on one or more of the following revenue streams. Some charge a commission equal to a percentage of the agreed price-tag (e.g., eBay or Airbnb). Some marketplaces provide a basic service for free but charge sellers to display their items in premium locations or for keeping them on for additional time (e.g., leboncoin.fr in France, or Mercado Libre in Latin America). Some increase their revenue by offering additional services such as insurance or delivery for a fee. Finally, some also rely on third-party advertisers that display text, banners, images, videos, etc., within the pages of the marketplace in exchange for payment.

A common feature in all those platforms is that when a user connects to them and enters a query, the site provides a list of relevant items (e.g., links, products, services, classified ads) that may match what the user wants. To provide the best value to users, the platform would ideally present the relevant items by decreasing order of (estimated) relevance, so the user is more likely to find the most appropriate ones. By doing this, the site can increase its reputation and attract more user visits. Measures of relevance can be based on various criteria, which are sometimes selected by the user. For example eBay provides relevance-based rankings that can account for time until the end of the auction, distance, price, etc. How to assign a relevance value to each item returned by a query depends on the intrinsic details of the platform. For example, eBay may use the string distance between the query and the item descriptions as well as the rating of the seller, Amazon may use the number of conversions for a product and its quality, and Google may use the Pagerank algorithm (Google 2011). Methods to define and compute relevance indices have been the subject of several studies, especially for SEs. Examples include Avrachenkov and Litvak (2004), Austin (2006), Auction Insights (2008), Williams (2010). In this paper, we are not concerned with how to define and compute these measures of relevance (this is outside our scope); we assume that they are given as part of the input data.

In addition, each matching item returned by a query has an expected revenue that could be obtained directly or indirectly by the platform owner when the user visits the item. The platform may have interest in taking this expected revenue into account when ranking the items, by placing highly-profitable ones in prominent positions. However, a myopic approach that ranks the items only in terms of expected revenue and not on relevance would decrease the reputation in the long run, and eventually decrease the number of user visits and future revenues. A good compromise should account for both relevance and expected revenue. Our aim in this paper is to offer a model of this situation and an algorithm to compute optimal rankings that can balance immediate profit with long-term user activity.

A request is abstracted out in our model as a random vector that contains a relevance index and an expected revenue for each matching item. For the purpose of this study, the distribution of this random

vector is assumed to be fixed and known (i.e., it is time-stationary and does not depend on any other element of the model). Estimating (or learning) this distribution from actual data is of course important for practical implementations, but is outside the scope of this paper. In real applications, this distribution is likely to change with time, at a slower time scale than the arrivals of requests, and the ranking strategy would be updated accordingly whenever deemed appropriate. This aspect is also beyond our scope.

In addition to the regular output that includes *organic results*, most platforms also display paid ads (also referred to as *sponsored results*). Our study focuses on the ordering of the organic results only. We assume that the average arrival rate of search requests is an increasing function of the average relevance of organic results, and is not affected by the choice and ordering of the sponsored results. This makes sense because the latter ordering is not likely to have much impact on the future arrival rate of requests. On the other hand, the total expected revenue from sponsored search depends on the arrival rate of requests. Our model accounts for this with a coefficient that represents the expected ad revenue per request, which we multiply by the arrival rate. There is an extensive literature on pricing and ranking sponsored results. For details, we refer the reader to Varian (2007), Edelman et al. (2007), Lahaie et al. (2007), Athey and Ellison (2011), Maillé et al. (2012), and the references therein. However, the impact of using alternative rankings to classify organic results has not yet received a similar level of attention.

The object of our work is to characterize optimal ordering strategies to rank the items returned by a query, to maximize the long-term expected revenue per unit of time, and to find an effective way of computing such a strategy. We want a model whose solution has a simple and elegant form, and that can inform the design of ranking policies, as opposed to a detailed and complicated model whose solution has no simple form. We propose a ranking policy that relies on a single real-valued parameter. This value can be optimized efficiently using simulation-based methods. Our model and algorithms also permit one to compare the optimal policy to other possible rankings—such as those based on relevance only or those based on short-term revenue only—in terms of expected revenue for the platform, expected revenue for the various content providers, and consumer welfare which is captured by the resulting relevance.

The expected relevance and expected income per request depend on the *ranking policy* used to select a permutation (ranking) of the items returned by a query (request). The ranking can be based on the estimated relevance and the expected revenue of each matching item. A *deterministic ranking policy* assigns to each possible request a single permutation, which is a deterministic function of the request. However, we will give examples in which no deterministic policy can be optimal. This motivates us to consider a richer class of *randomized ranking policies* which to each request assign a probability distribution on the set of permutations of all matching items. Whenever a request arrives, the platform selects a ranking using the probabilities specified by the policy for this request. Of course, computing and implementing such general policies, deterministic or randomized, appears impractical, because the number of possible requests is typically huge, so there would be way too many permutations or probabilities to compute and store. For

4

L'Ecuyer et al.: *Revenue-Maximizing Rankings*
Article submitted to ; manuscript no. (Please, provide the manuscript number!)

this reason, we are interested in developing a model for which we can prove that an optimal policy has a much simpler form, and is easier to implement.

Our main contribution is the characterization of such an optimal ranking policy. We show that for our model, an optimal policy must always rank the relevant items by decreasing order of their *score*, where the score of each item is a real number defined as a linear combination of a function of the estimated relevance and a function of the expected gain, in which the first coefficient can be taken as 1 and the second coefficient (the same for all items and requests) can be optimized. If the scores of matching items are all distinct with probability 1 (i.e., for almost all requests), finding the optimal coefficient specifies an optimal deterministic policy which has a very simple form, so we have reached our goal. This generally happens if the requests have a continuous probability density. But one may argue that in reality, the requests have a discrete distribution, in which case equalities between two or more scores occur with positive probability. The bad news is that in that case, only randomized policies can be optimal in general. Any optimal policy would still sort the matching items by order of score, but it must randomize the order of those having the same score, with specific optimal probabilities. In practice, if the probability of equality is small, to avoid computing the optimal probabilities for randomization, one may opt to forget the randomization to break ties and just use an arbitrary ordering in case of equality, as an approximation. We propose a more robust strategy: add a small random perturbation (uniform over a small interval centered at 0) to the expected revenue of each item, so scores are all distinct with probability 1. The impact of this perturbation on the expected long-term revenue can be made arbitrarily small by taking the size of the interval small enough. The modified model admits a deterministic optimal policy and one can just use this policy. This can also be viewed as a different (simpler) way of randomizing the policy.

Balancing immediate revenue and long-term impact on future arrivals when choosing a policy has been discussed by Mendelson and Whang (1990), Maglaras and Zeevi (2003), Besbes and Maglaras (2009). In those articles, one selects the price of a service (or the prices for different classes) to maximize the long-term revenue given that each arriving customer has a random price threshold under which she takes the service. The systems have capacity constraints and there can be congestion, which degrades the service quality. The strategy corresponds to the selected prices, which can be time-dependent. The derivations in those papers differ from what we do here in many aspects. The authors use approximations, e.g, by heavy-traffic limits, to select the prices. Aflaki and Popescu (2014) also compute a continuous policy (for the service level of a single customer) in a dynamic context, using dynamic programming (DP). Their solutions are algorithmic.

The model considered here obviously simplifies reality, as do virtually all models whose solution has a simple form. While there are many other "simple" heuristics that platforms may use to factor in profitability in their algorithms, the one we obtain here is not only clear and simple, but is also proved to be optimal for a reasonable model. We think this is a nice insight that can inform platforms on how to better position their results to tradeoff relevance with profits.

An important motivation of our work is the *search neutrality debate*: Some argue that SEs should be seen as a public service and therefore should be regulated to have their organic search results based only on objective measures of relevance, while others think they should be free to rank items in the way they think is best and to compete against each other freely. One key issue in this debate is whether platforms use rankings that depend on revenue-making ingredients (Crowcroft 2007). For example, since Google owns YouTube, it might favor YouTube pages for the extra revenue they generate. This type of search bias has been amply documented in experiments (Edelman and Lockwood 2011, Wright 2012, Maillé and Tuffin 2014). In this context, the framework we introduce can be of interest both to platform owners, to improve their ranking rules, and to regulators, to study the impact of various types of regulations on users and on overall social welfare, if they have access to the appropriate data to build the model.

The rest of the article is organized as follows. In Section 2, we present our modeling framework and state the optimization problem in terms of randomized policies. In Section 3, we derive a general characterization of the optimal policies and we obtain optimality conditions for the two situations where the requests have a discrete and a continuous distribution. For the continuous case, in which the requests have a density, we show that the optimal policy is completely specified by a unique scalar. This number is used to combine relevance and revenues into scores, which are used to rank the items in decreasing order. This works because all the matching items for each request have different scores with probability 1. This policy is very easy to implement and one does not need to consider the exponentially-many possible orderings. We provide algorithms to appropriately compute or estimate this scalar number. In Section 4, we provide numerical examples to illustrate the algorithms and what could be done with the model. Finally, we offer concluding remarks in Section 5.

2. Model Formulation

We define our model in the context of a SE that receives keyword-based queries and generates a list of organic results using links to relevant and/or profitable web pages. By changing the interpretation, the model applies to other marketplaces such as electronic retailers and classified-ad websites, as described in the introduction.

For each arriving request (i.e., a query sent to the SE by a user), different content providers (CPs) host pages that can be relevant to that request. Let $M \geq 0$ be the (random) number of pages that *match* the arriving request, i.e., deemed worthy of consideration for this particular request, out of the universe of all pages available online. We assume that M has a global deterministic upper bound $m_0 < \infty$, independent of the request. When $M > 0$, each page $i = 1, \ldots, M$ has a relevance value $R_i \in [0, 1]$, and an expected revenue per click for the SE of $G_i \in [0, \Gamma]$, where Γ is a positive constant. The variable R_i represents the SE's assessment of how page i would please the author of the request. The variable G_i contains the total expected revenue that the SE might receive directly or indirectly, from third-party advertisement displayed

6

L'Ecuyer et al.: *Revenue-Maximizing Rankings*
Article submitted to ; manuscript no. (Please, provide the manuscript number!)

on page i if the user clicks on that link. In particular, if the CP of page i receives an expected revenue per click for page i, and a fixed fraction of this revenue is transferred to the SE, then G_i represents the expected revenue that the SE receives. If the CP of a page is the SE, then the fraction is 1 and the SE receives all the revenue generated by it.

Thus, a request can be encoded as a random vector $Y = (M, R_1, G_1, \ldots, R_M, G_M)$ whose components satisfy the conditions just given. We denote a realization of Y by $y = (m, r_1, g_1, \ldots, r_m, g_m)$. We have:

ASSUMPTION A. *The vector Y has a fixed probability distribution (discrete or continuous) over a compact (closed and bounded) subspace $\Omega \subseteq \cup_{m=0}^{m_0}(\{m\} \times ([0, 1] \times [0, \Gamma])^m)$.*

After receiving a request $y \in \Omega$, the SE selects a permutation $\pi = (\pi(1), \ldots, \pi(m)) \in \Pi_m$ of the m matching pages, where Π_m is the set of permutations of $\{1, \ldots, m\}$, and displays the links in the corresponding order. The link to page i is presented in position $j = \pi(i)$.

The *click-through-rate* (CTR) of a link that points to a page is defined as the probability that the user clicks on that link (Hanson and Kalyanam 2007, Chapter 8). This probability generally depends on the relevance of the link, its position in the ranking, and perhaps on the relevance and positions of other links as well. For a given request y, we denote the CTR of page i placed in position j by $c_{i,j}(y)$. Our results will be derived under the following standard assumption:

ASSUMPTION B. The CTR function has the separable form $c_{i,j}(y) = \theta_j \, \psi_i(y)$, where $1 \geq \theta_1 \geq \theta_2 \geq \cdots \geq \theta_{m_0} > 0$ is a non-increasing sequence of fixed positive constants that describe the importance of each position in the ranking, and each $\psi_i : \Omega \to [0, 1]$ is a measurable function that gives a position-independent click "probability" for page i for any request y.

This separability assumption of the CTR is pervasive in the e-Commerce literature (Varian 2007, Edelman et al. 2007, Maillé et al. 2012). We will rely on it to derive optimality conditions in Section 3. In fact, the usual assumption in the cited works has $\psi(r_i)$ instead of $\psi_i(y)$; that is, this function is the same for all i and it is a function of r_i only. Our assumption is much less restrictive: it allows each ψ_i to be a function of the entire vector y, and therefore of the relevance of all other items that are offered. However, it is not allowed to depend on the rankings of these other items. For any given y, we define a quality-adjusted relevance $\tilde{r}_i := \psi_i(y) r_i$ and a quality-adjusted revenue $\tilde{g}_i := \psi_i(y) g_i$, so that we can write $c_{i,j}(y) r_i = \theta_j \tilde{r}_i$ and $c_{i,j}(y) g_i = \theta_j \tilde{g}_i$. Similarly, we define the corresponding quality-adjusted random variables $\tilde{R}_i := \psi_i(Y) R_i$ and $\tilde{G}_i := \psi_i(Y) G_i$.

If we select permutation π for request y, the corresponding expected relevance (the *local relevance*) is defined by

$$r(\pi, y) := \sum_{i=1}^{m} c_{i,\pi(i)}(y) r_i = \sum_{i=1}^{m} \theta_{\pi(i)} \tilde{r}_i. \tag{1}$$

It captures the attractiveness of this ordering π for this particular y, from the consumer's perspective. The expected revenue to the SE from the organic links for this request is

$$g(\pi, y) := \sum_{i=1}^{m} c_{i,\pi(i)}(y)g_i = \sum_{i=1}^{m} \theta_{\pi(i)}\tilde{g}_i. \tag{2}$$

A *(deterministic) stationary ranking policy* μ is a function that assigns a permutation $\pi = \mu(y)$ to each possible realization $y \in \Omega$. (We skip the technical issues of measurability of policies in this paper; this can be handled as in Bertsekas and Shreve (1978), for example.) By taking the expectation with respect to the distribution of input requests Y, we obtain the long-term value induced by a stationary ranking policy μ. The *expected relevance per request* for policy μ (which we use as proxy of the *reputation* of the SE) is

$$r := r(\mu) = \mathbb{E}_Y[r(\mu(Y), Y)] = \mathbb{E}_Y\left[\sum_{i=1}^{M} \theta_{\mu(Y)(i)}\tilde{R}_i\right] \tag{3}$$

and the *expected revenue per request* from the organic links for the SE is

$$g := g(\mu) = \mathbb{E}_Y[g(\mu(Y), Y)] = \mathbb{E}_Y\left[\sum_{i=1}^{M} \theta_{\mu(Y)(i)}\tilde{G}_i\right] \tag{4}$$

where $\mathbb{E}_Y$ represents the expectation taken with respect to the distribution of Y. Note that $0 \leq r \leq \sup_{y \in \Omega, \, \pi \in \Pi_m} \sum_{i=1}^{m} c_{i,\pi(i)}(y) \leq m_0$, where m corresponds to request y, and similarly $0 \leq g \leq m_0\Gamma$.

We also consider randomized policies, motivated by the fact that in some situations they can do better than the best deterministic policy (we will give an example of that). A *randomized stationary ranking policy* is a function $\check{\mu}$ that assigns a probability distribution over the set of permutations Π_m to each $y \in \Omega$. One has $\check{\mu}(y) = \{q(\pi, y) : \pi \in \Pi_m\}$, where $q(\pi, y) \geq 0$ is the probability of selecting π, and $\sum_{\pi \in \Pi_m} q(\pi, y) = 1$. The expressions for r and g for a policy $\check{\mu}$ are then

$$r := r(\check{\mu}) = \mathbb{E}_Y\left[\sum_{\pi \in \Pi_M} q(\pi, Y)\sum_{i=1}^{M} c_{i,\pi(i)}(Y)R_i\right] = \mathbb{E}_Y\left[\sum_{\pi \in \Pi_M} q(\pi, Y)\sum_{i=1}^{M} \theta_{\pi(i)}\tilde{R}_i\right] \tag{5}$$

and

$$g := g(\check{\mu}) = \mathbb{E}_Y\left[\sum_{\pi \in \Pi_M} q(\pi, Y)\sum_{i=1}^{M} c_{i,\pi(i)}(Y)G_i\right] = \mathbb{E}_Y\left[\sum_{\pi \in \Pi_M} q(\pi, Y)\sum_{i=1}^{M} \theta_{\pi(i)}\tilde{G}_i\right]. \tag{6}$$

Notice that although r and g depend on μ or $\check{\mu}$, as defined in (3) and (4) or in (5) and (6), if understood from the context, we omit the dependency to simplify notation.

ASSUMPTION C. *The objective for the SE is to maximize a long-term utility function of the form $\varphi(r, g)$ where φ is a strictly increasing function of r and g, with bounded second derivatives over $[0, m_0] \times [0, m_0\Gamma]$.*

An *optimal policy* from the perspective of the SE is a stationary ranking policy μ in the deterministic case, or $\check{\mu}$ in the randomized case, that maximizes $\varphi(r, g)$. That is, in the randomized case, we have an optimization problem whose decision variables are all the probabilities $q(\pi, y)$ that define the distributions

$\check{\mu}(y)$, for all $y \in \Omega$. Always ranking the pages by decreasing order of $\tilde{R}_i$ would maximize r, whereas ranking them by decreasing order of $\tilde{G}_i$ would maximize g. An optimal policy must usually make a compromise in between these two extremes.

Although our results in the paper are derived for a general φ that satisfies Assumption C, in our examples and much of the discussion, we will use the following specific form of φ, for the sake of concreteness and because this is the form of objective that we have in mind for applications. One may suppose that requests arrive according to a point process whose average arrival rate (per unit of time) is $\lambda(r)$, where $\lambda : [0, m_0] \to [0, \infty)$ is an increasing, positive, continuously differentiable, and bounded function. The expected gain per request (on average) from the ads and sponsored links placed on the page that provides the organic links, is assumed to be a constant β that does not depend on the ordering of the organic links. This expected gain adds up to g, so the long-term expected revenue per unit of time is

$$\varphi(r,g) = \lambda(r)(\beta + g). \tag{7}$$

Again, the properties we will derive stand for a more general φ and are not a consequence of the separability in terms of r and g as in (7). The definition of $\lambda(r)$ implies that it does not depend on the ordering of sponsored links. That is, sponsored links (paid ads) do not drive users to the website in the long term.

REMARK 1. According to Assumption A, the distribution of Y does not depend on the ranking policy and does not change with time. This represents a good approximation over a short time scale, although in real life, the distribution of Y may change over a longer time scale. We also do not distinguish requests at a finer granularity than the definition of the global distribution of Y (e.g., if Y represents the aggregate user population, we do not distinguish individual users). The measures r and g are averages across all queries. Another relevant issue related to the previous point is whether one can assume, as we do, that this distribution remains (approximately) the same when we change the ordering policy. In real life, the choice of policy can have an impact on the distribution of Y, e.g., by attracting more queries of certain types only. To address this situation, one can segment the space of queries (partition by user type, demographics, topic, etc.) and apply the model to each segment. Then each segment can have its own distribution Y, pair (r, g), and policy. This can be useful if the optimal policies differ markedly across segments. Developing effective ways of making this segmentation is a direction of further research. In principle, one could have a very large number of small segments, even a single IP address or user may be a segment, but in practice one must also have enough data to estimate the distribution of Y for each segment. So there would be a tradeoff between the accuracy of the model (more segments) and the ability to estimate the parameters (more data per segment).

For the model defined so far, implementing a general deterministic policy μ or randomized policy $\check{\mu}$ in a realistic setting may appear hopeless, because it involves too many permutations and probabilities. Our goal

L'Ecuyer et al.: *Revenue-Maximizing Rankings*
Article submitted to ; manuscript no. (Please, provide the manuscript number!)

9

in the next section is to characterize optimal policies and show that they have a nice and simple structure, under our assumptions. In particular, we show that an optimal policy ranks the relevant pages of a request y by decreasing value of a *score* defined as $\tilde{r}_i + \rho^* \tilde{g}_i$, for an appropriate constant $\rho^* \geq 0$ common to all items and requests, found by optimization. We will show this property for the general randomized policies. The property will imply that for the optimal ρ^*, randomization should be used only to order the pages having the same score. If the probability of such an equality is zero, then we can have a deterministic optimal policy. We will also show how the situations in which equality occurs with positive probability can be handled by adding a small perturbation to the G_i's, as mentioned in the introduction.

For the rest of the paper, we suppose that Assumptions A to C hold.

3. Optimality Conditions for Ranking Policies

To characterize an optimal ranking, we reformulate randomized policies so we can rely on convex analysis techniques.

3.1. Reformulation of Randomized Policies

Recall that a randomized policy assigns a probability $q(\pi, y)$ to each permutation $\pi \in \Pi_m$, for each request y. These $m!$ probabilities $q(\pi, y)$ determine in turn the probabilities

$$z_{i,j}(y) = \sum_{\pi \in \Pi_m} q(\pi, y) \mathbb{I}[\pi(i) = j] \geq 0 \tag{8}$$

that page i ends up in position j, for each (i,j). Here, $\mathbb{I}$ denotes the indicator function. Let $\mathbf{z}(y)$ denote the $m \times m$ doubly stochastic matrix whose (i,j) entry is $z_{i,j}(y)$, for $1 \leq i, j \leq m$ (in this matrix, each row sums to 1 and each column sums to 1). Using (8), we have

$$\sum_{\pi \in \Pi_M} q(\pi, Y) \sum_{i=1}^{M} c_{i,\pi(i)}(Y) R_i = \sum_{\pi \in \Pi_M} q(\pi, Y) \sum_{i=1}^{M} \sum_{j=1}^{M} c_{i,j}(Y) \mathbb{I}[\pi(i) = j] R_i = \sum_{i=1}^{M} \sum_{j=1}^{M} z_{i,j}(y) c_{i,j}(Y) R_i,$$

from which we can rewrite

$$r = r(\check{\mu}) = \mathbb{E}_Y \left[\sum_{i=1}^{M} \sum_{j=1}^{M} z_{i,j}(Y) c_{i,j}(Y) R_i \right] = \mathbb{E}_Y \left[\sum_{i=1}^{M} \sum_{j=1}^{M} z_{i,j}(Y) \theta_j \tilde{R}_i \right]. \tag{9}$$

Similarly, we can rewrite

$$g = g(\check{\mu}) = \mathbb{E}_Y \left[\sum_{i=1}^{M} \sum_{j=1}^{M} z_{i,j}(Y) c_{i,j}(Y) G_i \right] = \mathbb{E}_Y \left[\sum_{i=1}^{M} \sum_{j=1}^{M} z_{i,j}(Y) \theta_j \tilde{G}_i \right]. \tag{10}$$

Given that the objective function depends only on r and g and that those can be written directly in terms of $z_{i,j}(y)$ instead of $q(\pi, y)$, we can reformulate the optimization problem equivalently with $z_{i,j}(y)$ as decision variables. The objective is again to maximize $\varphi(r, g)$, where r and g are defined in (9) and (10). This reduces significantly the number of independent decision variables, since for any given y, we have $m! - 1$ degrees

10

L'Ecuyer et al.: *Revenue-Maximizing Rankings*
Article submitted to ; manuscript no. (Please, provide the manuscript number!)

of freedom for choosing the set $\{q(\pi, y), \pi \in \Pi_m\}$ and only $(m-1)^2$ degrees of freedom for choosing the matrix $\mathbf{z}(y)$.

In view of this equivalence, we can define a randomized policy as a rule $\tilde{\mu}$ that assigns, for each $y \in \Omega$, a doubly stochastic probability matrix $\tilde{\mu}(y) = \mathbf{z}(y) \in \mathbb{R}^{m \times m}$. We adopt this definition for the rest of this paper. Let $\tilde{\mathcal{U}}$ be the class of such randomized policies. A deterministic policy μ is just a special case of this for which the entries of each matrix are all 0 or 1, with a single 1 in each row and each column (such a matrix defines a permutation). To implement a randomized policy $\tilde{\mu}$, for any given request y, one can easily generate a ranking (a random permutation) that satisfies the probabilities specified in $\mathbf{z}(y) = \tilde{\mu}(z)$ by first generating the page at position 1 using the probabilities in the first column of $\mathbf{z}(y)$, then for each position $j = 2, \ldots, m$ in succession, select the page at position j using the conditional probabilities given the selections made at positions before j. To update the conditional probabilities, whenever a page i is selected for position j, we remove the corresponding row i in the matrix and rescale the remaining entries in row $j + 1$ so their sum equals 1.

The optimization problem for the SE can be summarized as

$$(\text{OP}) \qquad \max \varphi(r, g)$$

$$\text{subject to (9), (10), and } \mathbf{z}(y) \text{ doubly stochastic, for each } y \in \Omega.$$

The decisions variables in OP are the $z_{i,j}(y)$'s, whose values define a policy. This large-scale nonlinear optimization problem is not easy to solve directly in this general form, but we can characterize its optimal solutions via (standard) convex analysis arguments, as follows. Each randomized policy $\tilde{\mu}$ can be mapped to a corresponding pair $(r, g) = (r(\tilde{\mu}), g(\tilde{\mu}))$. Let $\mathcal{C}$ be the set of all points (r, g) that correspond to some $\tilde{\mu} \in \tilde{\mathcal{U}}$.

LEMMA 1. The set $\mathcal{C}$ is a convex and compact set.

PROOF. If two pairs (r^1, g^1) and (r^2, g^2) are in $\mathcal{C}$, they must correspond to two randomized policies $\tilde{\mu}^1$ and $\tilde{\mu}^2$ in $\tilde{\mathcal{U}}$. Suppose $\tilde{\mu}^1(y) = \mathbf{z}^1(y) = \{z_{i,j}^1(y) : 1 \le i, j \le m\}$ and $\tilde{\mu}^2(y) = \mathbf{z}^2(y) = \{z_{i,j}^2(y) : 1 \le i, j \le m\}$ for each $y \in \Omega$. For any given $\alpha \in (0, 1)$, let $\tilde{\mu} = \alpha \tilde{\mu}^1 + (1 - \alpha)\tilde{\mu}^2$ be the policy defined via $\tilde{\mu}(y) = \mathbf{z}(y) = \{z_{i,j}(y) : 1 \le i, j \le m\}$ where $z_{i,j}(y) = \alpha z_{i,j}^1(y) + (1 - \alpha)z_{i,j}^2(y)$, for all i, j and $y \in \Omega$. This policy provides a feasible solution that corresponds to the pair $(r, g) = \alpha(r^1, g^1) + (1 - \alpha)(r^2, g^2)$, so this pair must belong to $\mathcal{C}$. This proves the convexity of the set.

To prove that $\mathcal{C}$ is compact, we can use the fact that a continuous mapping of a compact set is compact (Rudin 1974). Consider the mapping which to each policy $\tilde{\mu}$ defined by $\{\tilde{\mu}(y) = \mathbf{z}(y), y \in \Omega\}$ assigns the corresponding pair (r, g). This mapping is continuous and maps the compact set of all doubly stochastic matrices $\mathbf{z}(y)$ for all possible requests $y \in \Omega$ to the set $\mathcal{C}$. □

Even though the tuple (r, g) does not contain decisions variables in our problem, it is convenient to define the two-dimensional auxiliary problem

$$\text{(OP2)} \qquad \max \varphi(r, g)$$
$$\text{subject to } (r, g) \in \mathcal{C},$$

whose optimal solutions correspond to optimal policies for OP. Suppose (r^*, g^*) is an optimal solution to OP2, which means that the pair belongs to $\mathcal{C}$, it can be obtained by an optimal policy for OP, and has optimal value $\varphi^* = \varphi(r^*, g^*)$. We know that such an optimal solution always exists, because $\mathcal{C}$ is compact and φ is continuous and bounded, from Assumption C. Let $\varphi_r(r, g)$ and $\varphi_g(r, g)$ denote the partial derivatives of $\varphi(r, g)$ with respect to r and g, respectively, and $\nabla \varphi(r, g)$ its gradient. Define $h(r, g) := \varphi_g(r, g)/\varphi_r(r, g)$ and $\rho^* := h(r^*, g^*)$, which is a constant. Note that with our specific objective function $\varphi(r, g) = \lambda(r)(\beta + g)$, we have $\rho^* = h(r^*, g^*) = \lambda(r^*)/((\beta + g^*)\lambda'(r^*))$. Let us define a modified version of OP2, named OP2-L, in which we replace the objective function $\varphi(r, g)$ by the linear function $r + \rho^* g$:

$$\text{(OP2-L)} \qquad \max r + \rho^* g$$
$$\text{subject to } (r, g) \in \mathcal{C}.$$

PROPOSITION 1. *Any optimal solution (r^*, g^*) to OP2 is also an optimal solution to OP2-L.*

PROOF. Given that $\mathcal{C}$ is convex and $\varphi(r, g)$ is increasing in both r and g, from standard convexity theory, one must have $(r - r^*, g - g^*) \cdot \nabla \varphi(r^*, g^*) \leq 0$ for all $(r, g) \in \mathcal{C}$, otherwise (r^*, g^*) would not be optimal. Therefore, since it gives 0 (the maximal value over $\mathcal{C}$) for this linear form, the solution $(r, g) = (r^*, g^*)$ for OP2 remains an optimal solution to the modified problem if we replace $\varphi(r, g)$ in OP2 by the linear function $(r - r^*, g - g^*) \cdot \nabla \varphi(r^*, g^*)$, or equivalently by $\varphi_r(r^*, g^*)r + \varphi_g(r^*, g^*)g$, which becomes $r + \rho^* g$ after dividing by the constant $\varphi_r(r^*, g^*)$. $\qquad \square$

The converse of Proposition 1 is not true: an optimal solution (r, g) to OP2-L is not necessarily optimal for OP2. However, if OP2-L has a unique optimal solution, then it must be optimal for OP2 as well. This happens if and only if the line $(r - r^*, g - g^*) \cdot \nabla \varphi(r^*, g^*) = 0$ intersects $\mathcal{C}$ only at one point (r^*, g^*).

Both an optimal solution (r^*, g^*) for OP2, and the corresponding ρ^*, are also not necessarily unique in general. For example, multiple optimal solutions would occur if $\varphi(r, g)$ is constant along one segment of the boundary of $\mathcal{C}$ (a finite curve) and the optimum is reached on that segment. Then, any point (r^*, g^*) on that segment, with the corresponding ρ^*, would satisfy the proposition. For a sufficient condition for uniqueness, consider the contour curve defined by $\varphi(r, g) = \varphi^*$ in the r-g plane. If this contour curve represents the graph of g as a strictly convex function of r, then (r^*, g^*) and ρ^* must be unique, because the set $\mathcal{C}$ is convex. For $\varphi(r, g) = \lambda(r)(\beta + g)$, if $\lambda(r) = r^\alpha$ with $\alpha > 0$, for example, then the contour curve obeys the equation $g = g(r) = \varphi^* r^{-\alpha} - \beta$. Differentiating twice, we find that $g''(r) = \varphi^* \alpha(\alpha + 1)r^{-\alpha-2} > 0$, so the contour curve represents a strictly convex function of r for $0 < r \leq 1$, and the solution is unique.

12

L'Ecuyer et al.: *Revenue-Maximizing Rankings*
Article submitted to ; manuscript no. (Please, provide the manuscript number!)

Proposition 1 does not yet tell us the form of an optimal policy. We will now build on it to provide a characterization of these optimal solutions to OP.

3.2. Necessary Optimality Conditions Under a Discrete Distribution for Y

Here we consider the situation in which Y has a discrete distribution over a finite set Ω, with $p(y) = \mathbb{P}[Y = y]$ for all $y = (m, r_1, g_1, \ldots, r_m, g_m) \in \Omega$. In this case, the optimization problem OP can be rewritten in terms of a finite number of decision variables, as follows:

$$\text{(OP-D)} \qquad \max \varphi(r, g)$$

$$\text{subject to} \quad r = \sum_{y \in \Omega} p(y) \sum_{i=1}^{M} \sum_{j=1}^{M} z_{i,j}(y) \theta_j \tilde{r}_i$$

$$g = \sum_{y \in \Omega} p(y) \sum_{i=1}^{M} \sum_{j=1}^{M} z_{i,j}(y) \theta_j \tilde{g}_i$$

$$\sum_{j=1}^{M} z_{i,j}(y) = \sum_{i=1}^{M} z_{i,j}(y) = 1 \qquad \text{for all } y \in \Omega, \ 1 \le i, j \le M$$

$$0 \le z_{i,j}(y) \le 1 \qquad \text{for all } y \in \Omega, \ 1 \le i, j \le M$$

in which the $z_{i,j}(y)$ are the decision variables. Since Ω is typically very large, this is in general a hard-to-solve large-scale nonlinear optimization problem.

Suppose that the current solution $\tilde{\mu}$ is optimal for the linear objective $r + \rho g$ for a given $\rho > 0$. Then we should not be able to increase $r + \rho g$ by changing the probabilities $z_{i,j}(y)$ in this optimal solution, for any given request y with $p(y) > 0$. In particular, if $\delta := \min(z_{i,j}(y), z_{i',j'}(y)) > 0$, decreasing those two probabilities by δ and increasing the two probabilities $z_{i,j'}(y)$ and $z_{i',j}(y)$ by δ gives another feasible solution (or policy) $\tilde{\mu}'$. In view of the expressions for r and g in problem OP-D, we find that this probability swap would change r and g by $\Delta r = \delta p(y)(\theta_{j'} - \theta_j)(\tilde{r}_i - \tilde{r}_{i'})$ and $\Delta g = \delta p(y)(\theta_{j'} - \theta_j)(\tilde{g}_i - \tilde{g}_{i'})$, respectively. Since the current solution is optimal for the objective $r + \rho g$, it cannot increase this objective, so we must have $\Delta r + \rho \Delta g \le 0$. This translates into the conditions

$$(\theta_{j'} - \theta_j) \left[(\tilde{r}_i - \tilde{r}_{i'}) + \rho(\tilde{g}_i - \tilde{g}_{i'}) \right] \le 0 \tag{11}$$

whenever $\min(z_{i,j}(y), z_{i',j'}(y)) > 0$, for all i, j, i', j', y with $p(y) > 0$. Without loss of generality, suppose that $j' > j$, so we know that $\theta_{j'} \le \theta_j$. If $\theta_{j'} = \theta_j$, the condition is always trivially satisfied (these two positions are equivalent, so the two pages at these positions can be permuted in any way and this changes nothing to r and g). If $\theta_{j'} < \theta_j$, one must have

$$\tilde{r}_i + \rho \tilde{g}_i \ge \tilde{r}_{i'} + \rho \tilde{g}_{i'} \,. \tag{12}$$

That is, if there is a positive probability that page i is ranked at a strictly better position j than the position j' of page i', then Condition (12) must hold. We call a ranking policy that satisfies this property for a given

L'Ecuyer et al.: *Revenue-Maximizing Rankings*
Article submitted to ; manuscript no. (Please, provide the manuscript number!)

13

$\rho \geq 0$ a *linear ordering (LO) policy with ratio* ρ (or LO-ρ policy, for short). Thus, in this discrete distribution setting, an LO-ρ policy always ranks the pages by decreasing order of their score $\tilde{r}_i + \rho\tilde{g}_i$, for some constant $\rho \geq 0$ common to all the requests y for which $p(y) > 0$, with the exception that if $\theta_{j'} = \theta_j$ for two positions j and j', the order between the pages at these positions does not matter (it can be arbitrary), because these two positions are equivalent. When $\rho = 0$, the ordering is based only on $\tilde{r}_i$, whereas in the limit as $\rho \to \infty$, the ordering is based only on $\tilde{g}_i$.

PROPOSITION 2. In the setting where Y has a discrete distribution, any optimal randomized policy with corresponding tuple (r^*, g^*) must be an LO-ρ policy with $\rho = \rho^* = h(r^*, g^*)$.

PROOF. We know from Proposition 1 that any optimal randomized policy must also be optimal for OP2-L, i.e., for the linear objective $r + \rho^* g$. From the argument given in the previous paragraph, this implies that it must be an LO-ρ^* policy. $\square$
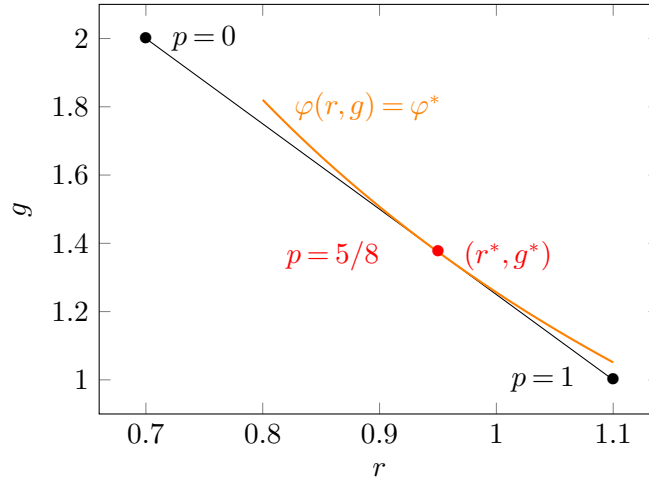
We emphasize that finding an optimal ρ might not be enough to specify an optimal policy in the case where, with positive probability, two or more pages have the same score $\tilde{R}_i + \rho\tilde{G}_i$. If the way we order those pages when this happens would not matter, then we could obtain an optimal deterministic ranking policy simply by choosing an arbitrary deterministic rule to order the pages having the same score. Given ρ, this would be very easy to implement, just by sorting the m matching pages by decreasing order of their score, for each y. However, the ordering in case of equality does matter, as illustrated by the following small example.

EXAMPLE 1. We consider an instance with a unique request type and two matching pages, $Y = y = (m, r_1, g_1, r_2, g_2) = (2, 1, 0, 1/5, 2)$ with probability 1. We take $\psi_i(y) = 1$ for all i and y, $\lambda(r) = r$, $(\theta_1, \theta_2) = (1, 1/2)$, and $\varphi(r, g) = r(1 + g)$. At each request we select a ranking, either $(1, 2)$ or $(2, 1)$. With a randomized policy, we select $(1, 2)$ with probability $z_{1,1}(y) = p$ and $(2, 1)$ with probability $1 - p$. In this simple case, one can write r, g, and $\varphi(r, g)$ as functions of p, and optimize. We have

$$r = p(\theta_1 r_1 + \theta_2 r_2) + (1 - p)(\theta_1 r_2 + \theta_2 r_1) = (7 + 4p)/10,$$
$$g = p(\theta_1 g_1 + \theta_2 g_2) + (1 - p)(\theta_1 g_2 + \theta_2 g_1) = 2 - p,$$
$$\varphi(r, g) = r(1 + g) = (7 + 4p)(3 - p)/10 = (21 + 5p - 4p^2)/10.$$

This objective function is quadratic and has its maximum at $p^* = 5/8$. Evaluating, we obtain $r^* = 19/20$, $g^* = 11/8$, and $\varphi(r^*, g^*) = 361/160$. Note that by taking $p = 0$ we get $(r, g) = (7/10, 2)$ and $\varphi(r, g) = 21/10 = 336/160$, whereas by taking $p = 1$ we get $(r, g) = (11/10, 1)$ and $\varphi(r, g) = 22/10 = 352/160$. Thus, the optimal randomized policy does strictly better than any deterministic one. The feasible set $\mathcal{C}$ here is the line segment that goes from $(7/10, 2)$ to $(11/10, 1)$. Figure 1 illustrates this line segment.

With the optimal $p^* = 5/8$, we also obtain $\rho^* = r^*/(1 + g^*) = 2/5$, and it turns out that $\tilde{r}_1 + \rho^*\tilde{g}_1 = \tilde{r}_2 + \rho^*\tilde{g}_2 = 1$. Thus, any ordering (and any randomized policy) satisfies the LO-ρ^* rule! That is, in this

14

L'Ecuyer et al.: *Revenue-Maximizing Rankings*
Article submitted to ; manuscript no. (Please, provide the manuscript number!)

Figure 1 The reachable pairs (r, g) **for Example 1 (black line), a level curve of the objective (in red) and the optimal pair** (r^*, g^*)**. Deterministic policies can only reach the two black dots labeled** $p = 0$ **and** $p = 1$**.**

simple example, the LO-ρ^* rule (and knowing ρ^*) tells us nothing about the optimal policy. Note that the entire segment $\mathcal{C}$ belongs to the line $r + \rho^* g = 1$, so maximizing the linear objective is not sufficient to obtain an optimal solution. □

3.3. A Model With a Continuous Distribution for Y

In this section, we extend the discussion to a model in which the requests Y have a continuous distribution over $\Omega \subseteq \cup_{m=1}^{m_0} (\{m\} \times ([0,1] \times [0, \Gamma])^m)$. The goal is to have a model for which ρ^* is unique and not difficult to compute or estimate, and for which knowing ρ^* is sufficient to specify an optimal policy. This type of continuous model can be used as an approximation to a model with discrete distribution. We will show that the approximation error can be made arbitrarily small.

We suppose that Ω is a measurable set and that Y has a probability measure ν over the class $\mathcal{B}$ of Borel subsets of Ω, so that $\nu(\{y\}) = 0$ for all $y \in \Omega$ and $\nu(B) = \int_B d\nu(y)$ for each $B \in \mathcal{B}$. Lemma 1 and Proposition 1 still apply in this case. However, the argument that led to Proposition 2 in the discrete case no longer applies, because each y has probability 0. We modify this argument to prove a similar result for the continuous case. We first adapt the notion of LO-ρ policy to this setting, and we establish that any optimal randomized policy must be of that form. In the continuous case, a randomized policy $\tilde{\mu}$ is called an *LO-ρ policy* if for almost all Y (with respect to the measure ν), $\tilde{\mu}$ sorts the pages by decreasing order of $\tilde{R}_i + \rho \tilde{G}_i$, except perhaps at positions j and j' where $\theta_j = \theta_{j'}$, at which the order does not matter (it can be arbitrary). That is, if page i is at position j, page i' is at position j', and $\theta_j > \theta_{j'}$, then one must have $\tilde{R}_i + \rho \tilde{G}_i \geq \tilde{R}_j + \rho \tilde{G}_j$.

PROPOSITION 3. *In the continuous model, if the tuple* (r^*, g^*) *corresponds to an optimal policy, then this policy must be an LO-ρ policy with* $\rho = \rho^* = h(r^*, g^*)$.

PROOF. Let $\tilde{\mu}$ be an optimal randomized policy, with its corresponding values of r^* and g^*, and let $B \in \mathcal{B}$ be a set with $\nu(B) > 0$ such that $\delta := \inf_{y \in B} \min(z_{i,j}(y), z_{i',j'}(y)) > 0$ for some arbitrary pages $i \neq i'$ and positions $j \neq j'$, for this policy $\tilde{\mu}$. Suppose we change $\tilde{\mu}$ into $\tilde{\mu}'$ by decreasing the probabilities $z_{i,j}(y)$ and $z_{i',j'}(y)$ by δ, and increasing $z_{i,j'}(y)$ and $z_{i',j}(y)$ by δ, for all $y \in B$. This gives another admissible randomized policy. The changes on r and g coming from this probability switch are

$$\Delta r = \delta \int_B (\theta_{j'} - \theta_j)(\tilde{r}_i - \tilde{r}_{i'}) d\nu(y),$$

$$\Delta g = \delta \int_B (\theta_{j'} - \theta_j)(\tilde{g}_i - \tilde{g}_{i'}) d\nu(y),$$

and we must have

$$0 \geq \Delta r + \rho \Delta g = (\theta_{j'} - \theta_j)\delta \int_B \left[(\tilde{r}_i - \tilde{r}_{i'}) + \rho(\tilde{g}_i - \tilde{g}_{i'}) \right] d\nu(y). \tag{13}$$

Now take $j' > j$ such that $\theta_j - \theta_{j'} > 0$. Suppose that there exists $\epsilon > 0$, $\delta > 0$, and $B \in \mathcal{B}$ such that $\nu(B) > 0$, for which for all $y \in B$,

$$\tilde{r}_i + \rho^* \tilde{g}_i \leq \tilde{r}_{i'} + \rho^* \tilde{g}_{i'} - \epsilon \tag{14}$$

and $\min(z_{i,j}(y), z_{i',j'}(y)) \geq \delta$ under policy $\tilde{\mu}$. That is, there is a set of positive probability on which the two pages i and i' are not placed by decreasing order of $\tilde{R}_i + \rho \tilde{G}_i$. We can modify policy $\tilde{\mu}$ to a policy $\tilde{\mu}'$ that first orders the pages according to policy $\tilde{\mu}$, and then if $y \in B$, page i is at position j, and page i' is at position j', the pages i and i' are swapped positions. This swapping occurs with probability at least $\nu(B)\delta^2$, and when it occurs it improves

$$(\theta_{j'} - \theta_j) \left[(\tilde{r}_i - \tilde{r}_{i'}) + \rho^*(\tilde{g}_i - \tilde{g}_{i'}) \right] \tag{15}$$

by at least $(\theta_{j'} - \theta_j)\epsilon$. The modification therefore improves the linear objective $r + \rho^* g$ by at least $\nu(B)\delta^2(\theta_{j'} - \theta_j)\epsilon > 0$, which contradicts the assumption that $\tilde{\mu}$ is optimal. $\qquad \square$

Proposition 3 tells us that any optimal policy must satisfy the LO-ρ conditions for $\rho = \rho^*$. But we need an additional assumption to make sure that knowing ρ^* and using an LO-ρ^* rule is sufficient to completely specify an optimal policy. For the remainder of this section, we make the following assumption for the continuous model.

ASSUMPTION D. *For any $\rho \geq 0$, and any $j > i > 0$, $\mathbb{P}[M \geq j \text{ and } \tilde{R}_i + \rho \tilde{G}_i = \tilde{R}_j + \rho \tilde{G}_j] = 0$.* $\qquad \square$

Under Assumption D, for any fixed $\rho \geq 0$, a deterministic LO-ρ policy $\mu = \mu(\rho)$ provides a unique ranking of the pages with probability 1, i.e., for almost any $Y \in \Omega$. If $\rho = \rho^* = h(r^*, g^*)$ for an optimal solution (r^*, g^*), such a deterministic LO-ρ^* policy is optimal. We also know that there always exists such an optimal (r^*, g^*), because $\mathcal{C}$ is compact. An LO-ρ policy $\mu = \mu(\rho)$ has corresponding values of $(r, g) = (r(\mu), g(\mu))$ and of $h(r, g)$ that are uniquely defined. To refer to $h(r, g)$ as a function of ρ, we write $\tilde{h}(\rho) := h(r(\mu(\rho)), g(\mu(\rho)))$.

16

L'Ecuyer et al.: *Revenue-Maximizing Rankings*
Article submitted to ; manuscript no. (Please, provide the manuscript number!)

PROPOSITION 4. In the continuous setting, under Assumption D and with deterministic LO-ρ policies, we have: (i) The fixed-point equation

$$\tilde{h}(\rho) = \rho \tag{16}$$

always has ρ^* as a solution.

(ii) If $\tilde{h}'(\rho) < 1$ for all $\rho > 0$, then ρ^* must be the unique fixed point.

(iii) For the particular objective function $\varphi(r,g) = \lambda(r)(\beta + g)$ in (7), if $\lambda(r)/\lambda'(r)$ is non-decreasing in r, then $\tilde{h}'(\rho) \leq 0$ for all $\rho > 0$ and ρ^* is the unique fixed point.

PROOF. We know that an optimal policy exists, and from Proposition 3, any optimal policy μ must be an LO-ρ policy with $\rho = \rho^* = h(r(\mu(\rho^*)), g(\mu(\rho^*))) = \tilde{h}(\rho^*)$. Thus, the fixed-point equation has at least one solution $\rho^* \in [0, \infty)$ and this proves (i). If $\tilde{h}'(\rho) < 1$ for all $\rho > 0$, then the function $\tilde{h}(\rho)$ cannot cross the graph of the identity function (the line of slope 1 that crosses the origin) more than once, so the fixed point must be unique, which proves (ii). Under the conditions of (iii), we have

$$\tilde{h}(\rho) = \lambda(r(\mu(\rho)))/[\lambda'(r(\mu(\rho)))(\beta + g(\mu(\rho)))].$$

If $\lambda(r)/\lambda'(r)$ is non-decreasing in $r = r(\mu(\rho))$, then it is non-increasing in ρ because $r(\mu(\rho))$ is non-increasing in ρ. Additionally, since we know that $g(\mu(\rho))$ is non-decreasing in ρ, it follows that $\tilde{h}(\rho)$ is non-increasing in ρ, so $\tilde{h}'(\rho) \leq 0$ and the fixed point must be unique. $\square$

This condition that $\lambda(r)/\lambda'(r)$ is non-decreasing is clearly stronger than what we need to guarantee uniqueness of the root, but it is quite reasonable. To illustrate when this condition is satisfied, take $\lambda(r) = a_0 + b_0 \ln(c_0 + r)$ for some constants $a_0 \geq 0$, $b_0 > 0$, and $c_0 \geq 1$. Then, $\lambda'(r) = b_0/(c_0 + r)$, and therefore $\lambda(r)/\lambda'(r) = [a_0 + b_0 \ln(c_0 + r)](c_0 + r)/b_0$, which is bounded and increasing in $r \in [0, m_0]$. Other simple cases where this condition holds are the monomial forms $\lambda(r) = a_0 r^{b_0}$ for any positive values a_0 and b_0; which includes the case $\lambda(r) = r$ considered in some examples in this paper.

When $G_i = 0$ for all i, or $\tilde{G}_i$ is equal to the same constant for all i, the choice of ρ does not matter, and it is always optimal to sort the pages by decreasing order of $\tilde{R}_i$. If we also have that $r_i \geq r_{i'}$ implies $\psi_i(y) \geq \psi_{i'}(y)$, which seems natural, then it is easy to see that it is always optimal to sort the pages by order of relevance R_i; i.e., the SE has the incentive to conform to search neutrality. Under this ordering, r takes its maximal possible value $r_0 \leq m_0$. We always have $r \in [0, r_0]$.

We now provide examples showing why we need Assumption D. In particular, having a density for Y is not sufficient for the optimal policy to be deterministic and uniquely defined by ρ^*.

EXAMPLE 2. Starting from Example 1, we add a third page with relevance R_3 uniformly distributed over $[0, \epsilon]$ for some small $\epsilon > 0$, and revenue $G_3 = 0$. We assume that $\theta_3 = 1/4$. Since R_3 has a density, $p(y) = 0$ for all $y \in \Omega$. For any $\rho > 0$, if ϵ is small enough, this third page will always be ranked last, and its impact on $h(r,g)$ is very small. Then the problem of ranking the first two pages becomes the same as Example 1, so the optimal policy must be randomized. $\square$

L'Ecuyer et al.: *Revenue-Maximizing Rankings*
Article submitted to ; manuscript no. (Please, provide the manuscript number!)

17

One might think that if no $\tilde{R}_i$ or $\tilde{G}_i$ has a probability mass at a given point, then Assumption D must hold, but this is also not sufficient, because (unless we assume independence) one may construct an example in which with positive probability, one has $\tilde{R}_i = \tilde{R}_j$ and $\tilde{G}_i = \tilde{G}_j$ and then $\tilde{R}_i + \rho\tilde{G}_i = \tilde{R}_j + \rho\tilde{G}_j$.

3.4. Continuous Approximation to the Discrete Case

The continuous model under Assumption D has an important advantage over a discrete model, in terms of the simplicity of an optimal policy. This motivates the idea that when Y has a discrete distribution, instead of solving OP-D, we can approximate the distribution by a continuous one by adding a random perturbation to each G_i for each Y. The perturbations are all independent and uniform over the interval $(-\epsilon, \epsilon)$ for a very small ϵ. The perturbed model then satisfies Assumption D, and it suffices to compute $\rho^*(\epsilon)$, an optimal ρ for this perturbed problem, to obtain an optimal LO-ρ ranking policy for it. Let $r^*(\epsilon) = r(\rho^*(\epsilon))$ and $g^*(\epsilon) = g(\rho^*(\epsilon))$ be the average relevance and gain when using the optimal policy on the perturbed problem. The next proposition compares the optimal values of the original and perturbed problems, and shows that by taking ϵ small enough, an optimal policy for the perturbed problem will be ϵ'-optimal for the original problem for an arbitrarily small ϵ'. Let φ^* and $\varphi^*(\epsilon)$ be the optimal values of the original and perturbed problem, and let $\varphi^{**}(\epsilon) = \varphi(r^*(\epsilon), g^*(\epsilon))$ be the value of using the LO-$\rho^*(\epsilon)$ policy on the original problem. What we lose (per unit of time) by using this policy instead of an optimal one for the original problem is $\varphi^* - \varphi^{**}(\epsilon)$, which is shown to be bounded in the next proposition.

PROPOSITION 5. For the specific form of φ in (7), we have

$$0 \leq \varphi^* - \varphi^{**}(\epsilon) \leq \varphi^*(\epsilon) - \varphi^{**}(\epsilon) \leq \epsilon' := \lambda(r^*(\epsilon))(\theta_1 + \cdots + \theta_{m_0})\epsilon. \tag{17}$$

PROOF. For the original problem, applying the LO-$\rho^*(\epsilon)$ policy cannot do better than the optimal policy. This gives $\varphi^{**}(\epsilon) \leq \varphi^*$. For the perturbed problem, imagine that G_i is observed both before and after the perturbation, so we can apply an optimal randomized policy $\tilde{\mu}$ for the original problem to the perturbed problem by ignoring the realized perturbation before making the decision. This would give the same pair $(r(\tilde{\mu}), g(\tilde{\mu}))$ and the same value φ^* as for the original problem, because the expected value of the perturbation is zero, so adding it does not change g. But this policy ignores some relevant information that is available in the perturbed problem, namely the perturbed G_i's, which are the relevant values for the perturbed problem. Therefore it cannot beat an optimal policy for the perturbed problem, i.e., we must have $\varphi^* \leq \varphi^*(\epsilon)$. To show the last inequality in (17), note that, following the same method, if we take the rankings (randomized or not) obtained by the LO-$\rho^*(\epsilon)$ policy run on the original requests, and apply them to both the original and the perturbed problems, the values of r and $\lambda(r)$ are the same in both cases, and the difference in g cannot exceed $(\theta_1 + \cdots + \theta_{m_0})\epsilon$. Therefore the difference $\varphi^*(\epsilon) - \varphi^{**}(\epsilon)$ in $\varphi(r, g) = \lambda(r)(\beta + g)$ cannot exceed $\lambda(r)(\theta_1 + \cdots + \theta_{m_0})\epsilon$. $\square$

Since λ is a bounded function, (17) shows that by using an LO-$\rho^*(\epsilon)$ policy instead of an optimal policy for the original problem, the loss in value is $\mathcal{O}(\epsilon)$, and can be made negligible by taking ϵ sufficiently small.

Table 1   Optimal values for Example 3 as a function of ϵ.

ϵ	$p^*(\epsilon)$	$\rho^*(\epsilon)$	$r^*(\epsilon)$	$g^*(\epsilon)$	$\varphi^*(\epsilon)$	$\varphi^{**}(\epsilon)$
0.0	0.625	0.4	0.95	1.375	2.25625	2.25625
0.001	0.62491	0.39995	0.94996	1.37521	2.25636	2.256155
0.01	0.62411	0.39950	0.94964	1.37006	2.25736	2.255395
0.1	0.61705	0.39537	0.94682	1.39476	2.26741	2.248697
0.5	0.59771	0.38137	0.93908	1.46240	2.31240	2.230315

In practice, one can estimate $\rho^*(\epsilon)$ for a very small ϵ, such as 10^{-10} for example. When implementing the policy, for each request Y, one would add the random perturbations and use the perturbed values to rank the pages. In fact, it would also suffice to generate the perturbations only for the pages for which there is an equality.

As an illustration, we apply this perturbation method to Example 1. In fact, we will add a perturbation only to G_2, since it is sufficient for eliminating equalities.

EXAMPLE 3. In Example 1, suppose now that G_2 has the uniform distribution over the interval $(2 - \epsilon, 2 + \epsilon)$. The expectation of G_2 is unchanged, but now we obtain slightly more accurate information on the revenue G_2 before making the ranking decision. This modified (perturbed) model satisfies Assumption D, so finding $\rho^*(\epsilon)$ is sufficient to completely specify an optimal policy for the perturbed model. Since the perturbed G_2 is observed before making the ranking decision and can be used for making the decision, we expect that $\varphi^*(\epsilon) - \varphi^* > 0$ and increases with ϵ. It also converges to 0 when $\epsilon \to 0$. Figure 2 confirms this.

For the perturbed model, we can write $G_2 = 2 + (2V - 1)\epsilon$ where $V \sim U(0,1)$. An LO-ρ policy then selects the order $(1, 2)$ if and only if $R_1 + \rho G_1 > R_2 + \rho G_2$, if and only if $V < p = p(\epsilon) := 2/(5\rho\epsilon) - 1/\epsilon + 1/2$, which occurs with probability p. We have $r = (7 + 4p)/10$ as before and

$$g = 2 - p + \epsilon \int_0^p (v - 1/2)\mathrm{d}v + \epsilon \int_p^1 (2v - 1)\mathrm{d}v \ = \ 2 - p + \epsilon p(1 - p)/2,$$

$$\varphi(r, g) = r(1 + g) = \frac{7 + 4p}{10}(3 - p + \epsilon p(1 - p)/2),$$

$$\frac{\partial}{\partial p}\varphi(r, g) = [5 - 8p + \epsilon(7/2 - 3p(1 + 2p))]/10.$$

We can find the optimal p, say $p^* = p^*(\epsilon)$, as a root of this equation and we have

$$\rho = \rho(p) = \frac{2}{5(1 + \epsilon(p - 1/2))}$$

from the definition of p. We see that when $\epsilon \to 0$, $p^* = p^*(\epsilon) \to 5/8 = 0.625$ and $\rho^*(\epsilon) \to 2/5 = 0.4$. Table 1 gives the optimal values as a function of ϵ. It shows how $\rho^*(\epsilon)$ and $\varphi^*(\epsilon)$ converge to ρ^* and φ^* when $\epsilon \to 0$.

Figure 2 pictures the reachable values of (r, g) for this example, for some large values of ϵ. The upper curve (in blue) represents the pairs (r, g) for LO-ρ policies for all values of $\rho \in [0, \infty]$ (or $0 \le p \le 1$), for $\epsilon = 0.5$. The optimal point (r^*, g^*) on that curve is marked by a blue dot. The corresponding curve and point for $\epsilon = 0.1$ are in green (the middle curve). The lower line (a straight line) represents the pairs (r, g) for the policies that select the ranking $(1, 2)$ with probability p independently of G_2. Those are the randomized policies of Example 1. For any given ϵ, the region delimited by the lower line and the curve is the set $\mathcal{C}$.
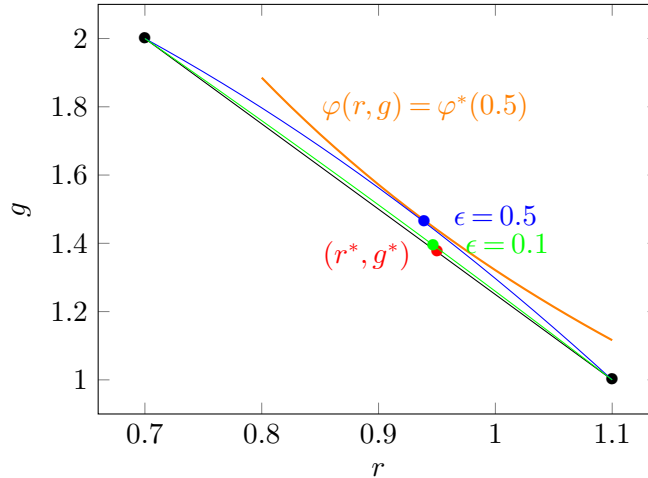
Figure 2   The upper boundary of $\mathcal{C}$ and the optimal solution with an LO-ρ policy for $\epsilon = 0.5$ **(blue line) and**
$\epsilon = 0.1$ **(green line) for the perturbed problem in Example 3**

3.5.   Computing or Estimating ρ^*

We now discuss how to compute ρ^*, an optimal ρ for an LO-ρ policy, under Assumption D. We saw that
the optimal ρ can be written as a function of r and g, which in turn depend on the selected policy μ and are
unknown a priori. In the previous small examples, we were able to derive explicit analytical expressions for
$r(\rho)$ and $g(\rho)$, and use them to find the optimal ρ. Unfortunately, instances of real size do not admit such
closed-form derivations and they would usually have to be estimated through simulation. Then the search
for ρ^* is a *stochastic root-finding* problem: estimate a root of $\tilde{h}(\rho) - \rho = 0$ when only noisy estimates of $\tilde{h}$
can be obtained, via simulation. Several algorithms have been designed and studied for this type of problem.
Two prominent classes of approaches are *sample average optimization* and *stochastic approximation*; see,
e.g., Pasupathy and Kim (2011) and the references therein. These methods generally assume that a root
exists and is unique, which is typically the case (we gave conditions for that earlier). If $\tilde{h}$ is a contraction
mapping, we can apply a different method which we now describe. It is very simple, easy to implement, and
has worked very nicely in all the examples we tried (including all the numerical examples in this paper). It
requires an estimator of $\tilde{h}(\rho)$.

We can define and compute an estimator $\hat{h}_n(\rho)$ of $\tilde{h}(\rho)$ at any given value of ρ as follows. We generate
n independent realizations $Y_1, \ldots, Y_n$ of Y, with $Y_k = (M_k, R_{k,1}, G_{k,1}, \ldots, R_{k,M_k}, G_{k,M_k})$. For each k,
we order the M_k pairs $(R_{k,i}, G_{k,i})$ by decreasing order of $\tilde{R}_{k,i} + \rho \tilde{G}_{k,i}$, and let π_k be the corresponding
permutation. We then compute the following unbiased estimators of $r(\rho)$ and $g(\rho)$:

$$\hat{r}_n(\rho) = \frac{1}{n} \sum_{k=1}^{n} \sum_{i=1}^{M_k} \theta_{\pi_k(i)} \tilde{R}_{k,i} \text{ and } \hat{g}_n(\rho) = \frac{1}{n} \sum_{k=1}^{n} \sum_{i=1}^{M_k} \theta_{\pi_k(i)} \tilde{G}_{k,i}.$$

They lead to the following estimator of $\tilde{h}(\rho)$:

$$\hat{h}_n(\rho) = \varphi(\hat{r}_n(\rho), \hat{g}_n(\rho)) = \lambda(\hat{r}_n(\rho))(\beta + \hat{g}_n(\rho)), \tag{18}$$

20

L'Ecuyer et al.: *Revenue-Maximizing Rankings*
Article submitted to ; manuscript no. (Please, provide the manuscript number!)

which is biased because φ is nonlinear, but is consistent. The bias decreases as $\mathcal{O}(1/n)$ and a confidence interval for $\tilde{h}(\rho)$ can be computed via the Delta method; see Asmussen and Glynn (2007).

Recall that $\rho \to \tilde{h}(\rho)$ is a *contraction mapping* if there is a constant $\gamma \in [0,1)$ such that $|\tilde{h}(\rho) - \tilde{h}(\rho')| \leq \gamma|\rho - \rho'|$ for all $\rho, \rho' \geq 0$. A sufficient condition for this is that $|\tilde{h}'(\rho)| \leq \gamma$ for all ρ (in the region of interest). It is very common in our setting that $\rho \to \tilde{h}(\rho)$ is a contraction mapping. In particular, it is true in all the examples considered in this paper. When this holds, we can start from some $\rho_0 > 0$ and iterate: $\rho_j = \tilde{h}(\rho_{j-1})$, for $j = 1, 2, \ldots$. Then, the fixed-point theorem for contraction mappings (Bertsekas and Shreve 1978) guarantees that $\rho_j \to \rho^*$ at a geometric rate: $|\rho_j - \rho^*| \leq \gamma^j|\rho_0 - \rho^*|$, which provides very fast convergence when $\gamma \ll 1$. In practice, we can replace $\tilde{h}(\rho_j)$ by $\hat{h}_{n_j}(\rho_j)$, and convergence to ρ^* will occur if $n_j \to \infty$ when $j \to \infty$. Note that if n_j does not increase with j, ρ_j will generally not converge to ρ^*. If n_j is fixed to some large constant n and we use independent random numbers (IRN) across the different steps j for the simulations, ρ_j will never converge but will wander around in a small neighborhood of ρ^*. If we use common random numbers (CRN) for the simulations (i.e., exactly the same n realizations Y_k at all steps j), it will converge to a value close to ρ^*, but generally different.

3.6. LO-ρ Policies for More General Situations

In situations where the assumptions made in this paper are not satisfied, and for which computing an optimal policy can be much too difficult, one can still use an LO-ρ policy as a heuristic. The parameter ρ can be optimized via simulation-based optimization. This can be done even if $\tilde{h}(\rho) = \rho$ has multiple roots, by using random search to approximate the optimal ρ. This type of policy has the advantage of being easy to implement for any observed request Y once ρ is selected. It constitutes a very attractive heuristic for general usage.

4. Some Illustrative Examples

This section introduces some examples that, although very simple and stylized, capture some real world concerns and illustrate our method. To simplify the exposition, our examples assume that $\psi_i(y) = 1$ for all i, which means that the CRT is $c_{i,j}(y) = \theta_j$ and depends only on the position of the page. We take $\varphi(r, g) = \lambda(r)(\beta + g)$ with $\lambda(r) = r$, so $\lambda(r)/\lambda'(r) = r$ is non-decreasing and $\tilde{h}(\rho) = \rho$ always has a unique fixed point ρ^*. These simple choices are by no means realistic, but they simplify the exposition. Assumption D is also satisfied in all our examples, so ρ^* always defines the order uniquely with probability 1. Example 4 illustrates how to apply the methodology in a simple case, while Example 5 illustrates the fact that revenue-maximizing ranking strategies can have significant economic consequences on other actors, e.g., in terms of fairness among content providers (the search neutrality debate). Such bias in ranking has been observed in real life, as mentioned in the introduction. Whether regulators or governments should take action and what they should do is a political issue which is beyond the scope of this paper.

4.1. Computing the optimal ranking: an illustrative example

This example shows the results of applying the framework put forward in this paper to an instance inspired by a platform that has the choice of showing its own or third-party results for which its own content generates profits while the external content does not. The example illustrates the tradeoffs in terms of total revenue, relevance and short-term gains when one transitions from a relevance-based ranking into a profit-based ranking. The final outcome is the coefficient ρ^* that resolves that tradeoff, which can be used to rank the results optimally.

EXAMPLE 4. This example depicts a situation in which the SE has a positive expected revenue G_i for a page for which it is also the CP, and $G_i = 0$ otherwise. An alternative interpretation is that all the content is served by other CPs, and some of those CPs agree to pay the SE a fixed price, normalized to 1, for each click to their pages served from the SE's output. This price does not depend on the ranking of the link; it just gives an incentive for the SE to favor links with $G_i = 1$ in its ranking.

Suppose there are two matching pages ($M = 2$). For $i = 1, 2$, R_i has a uniform distribution over $[0, 1]$, the revenue G_i is a Bernoulli random variable with parameter $p = \mathbb{P}[G_i = 1] = 1 - \mathbb{P}[G_i = 0]$, and these four random variables are independent. The density of Y is thus well defined and Assumption D holds. Let $(\theta_1, \theta_2) = (1, 0)$, which amounts to assuming that the SE displays only one page. Focusing on LO-ρ policies, we derive explicit formulas for $r = r(\rho)$, $g = g(\rho)$, and $\varphi(r(\rho), g(\rho))$. The fixed point ρ^* can then be computed from these formulas.

To derive the formula for $r = r(\rho)$, we distinguish two cases for (G_1, G_2): If $G_1 = G_2$, only the most relevant link is displayed, resulting in conditional expected relevance $\mathbb{E}[\max(R_1, R_2) \mid G_1 = G_2] = 2/3$. If $G_1 \neq G_2$, we can assume (possibly by swapping the pages) that $G_1 = 1$ and $G_2 = 0$. If $R_1 + \rho \geq R_2$, link 1 is displayed and the relevance is R_1; otherwise, the relevance is R_2. If $\rho > 1$, link 1 is always shown, leading to an expected observed relevance of $1/2$. If $\rho \leq 1$, the expected relevance conditional on $(G_1, G_2) = (1, 0)$ is

$$\int_{r_1=0}^{1} \int_{r_2=0}^{1} r_1 \mathbb{I}[r_1 + \rho > r_2] + r_2 \mathbb{I}[r_1 + \rho \leq r_2] \mathrm{d}r_2 \mathrm{d}r_1 = \frac{2}{3} - \frac{\rho^2}{2} + \frac{\rho^3}{3}.$$

By combining all cases, we obtain $r = r(\rho) = 2/3 + p(1-p)\bar{\rho}^2 (2\bar{\rho}/3 - 1)$ where $\bar{\rho} := \min(1, \rho)$. Similarly, to compute the expected revenue $g = g(\rho)$ per request, we consider the same two cases: If $G_1 = G_2$, the expected revenue is 0 if $G_1 = 0$, and 1 otherwise. If $G_1 \neq G_2$, we can assume again that $G_1 = 1$ and $G_2 = 0$. If $\rho > 1$, link 1 is displayed and the revenue is 1. If $\rho \leq 1$, the expected revenue conditional on $(G_1, G_2) = (1, 0)$ is

$$\int_{r_1=0}^{1} \int_{r_2=0}^{1} \mathbb{I}[r_1 + \rho > r_2] \mathrm{d}r_2 \mathrm{d}r_1 = 1 - \frac{(1-\rho)^2}{2}.$$

By regrouping all cases, we obtain $g = g(\rho) = p^2 + p(1-p)\left(1 - (2 - \bar{\rho})^2\right)$.

Both $r(\rho)$ and $g(\rho)$ are constant for $\rho \geq 1$, so we can restrict the search for ρ^* to the interval $[0, 1]$, and we have $\bar{\rho} = \rho$ in that interval. With $\lambda(r) = r$, the expected revenue per unit of time is

$$\varphi(r(\rho), g(\rho)) = r(\rho)(\beta + g(\rho)) = \left(2/3 + p(1-p)\rho^2 (2\rho/3 - 1)\right)\left(\beta + \left(p^2 + p(1-p)\left(2 - (1-\rho)^2\right)\right)\right).$$

Table 2 **Values of ρ_j at the first six iterations of the contraction mapping for Example 4, with IRN and CRN**

Method	ρ_1	ρ_2	ρ_3	ρ_4	ρ_5	ρ_6
IRN	0.4444471	0.377720	0.387115	0.3857771	0.3860725	0.3859246
CRN	0.4444471	0.377670	0.387079	0.3857318	0.3859223	0.3858940

Figure 3 depicts the expected revenue as a function of ρ, along with $r(\rho)$ and $g(\rho)$, for $\beta = 1$ and $p = 1/2$. While $g(\rho)$ increases and $r(\rho)$ decreases with ρ, the maximal revenue is obtained by taking ρ around 0.4. This optimal ρ uniquely determines the optimal policy (with probability 1).
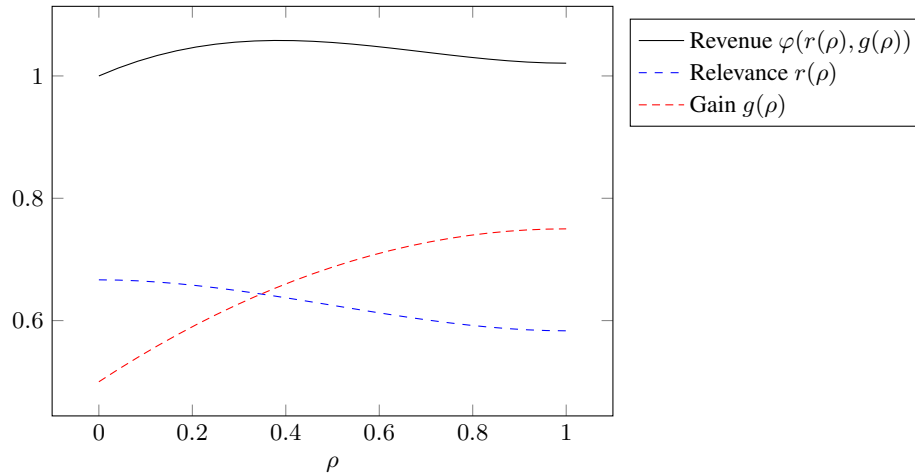
Figure 3 **Expected SE revenue per unit time for $\beta = 1$ and $p = 1/2$**

For this example, with the expressions previously derived for r and g, we get

$$\tilde{h}(\rho) = \frac{2/3 + p(1-p)\bar{\rho}^2(2\bar{\rho}/3 - 1)}{\beta + p^2 + 2p(1-p)(1 - (1-\bar{\rho})^2/2)}$$

and

$$\tilde{h}'(\rho) = -\frac{2}{3}p(1-p)(1-\bar{\rho})\frac{(3\bar{\rho}\beta + 3p\bar{\rho} + 3p\bar{\rho}^2 - p\bar{\rho}^3 - 3p^2\bar{\rho}^2 + p^2\bar{\rho}^3 + 2)}{(\beta + p + 2p\bar{\rho} - p\bar{\rho}^2 - 2p^2\bar{\rho} + p^2\bar{\rho}^2)^2}.$$

For $\beta = 1$ and $p = 1/2$, one can verify numerically that for $0 \leq \rho \leq 1$, $\tilde{h}'(\rho)$ is negative and achieves a maximum absolute value of approximately $0.15 < 1$ (although the derivative is not monotone). Hence, $\rho_j \to \tilde{h}(\rho_j)$ is a contraction mapping with $\gamma \approx 0.15$ in that area. We applied this contraction mapping for six iterations, starting with $\rho_0 = 0$, with a fixed sample size of $n_j = 10^7$ for all j. We did this with both IRN and CRN. The results are in Table 2. In both cases, ρ_j provides a good approximation to ρ^* very quickly. We find that $\rho^* \approx 0.3859$. Thus, the optimal strategy is to display the page for which the score $R_i + 0.3859G_i$ is largest. For the pages that provide no direct revenue to the SE ($G_i = 0$), the score is equal to the relevance, while for the pages that provide revenue ($G_i = 1$), the score is determined by adding the constant 0.3859 to the relevance, which gives a competitive advantage that helps uprank those pages.

L'Ecuyer et al.: *Revenue-Maximizing Rankings*
Article submitted to ; manuscript no. (Please, provide the manuscript number!)

23

4.2. Applying the methodology to study the impact of non-neutral search engines

The next example illustrates how our method could help answer specific questions about whether regulations are beneficial or not. The framework we introduced can indeed be of high interest to regulators who study the impact of search neutrality on users and on overall social welfare. It can help them determine if intervention is warranted, study the consequences of doing so, and provide arguments for or against non-neutral SEs by computing welfare measures of all parties involved in the interaction with the platform.

These questions were discussed by the Federal Trade Commission in the US (Brill 2013) and in a Senate hearing (Rushe 2012). Search biases have been amply documented in experiments (Edelman and Lockwood 2011, Wright 2012, Maillé and Tuffin 2014). Data indicate that a search for a video in Google is likely to generate more organic links to YouTube pages, which contain ads that directly benefit Google, than in another SE. Since videos in competitors' platforms do not generate additional revenue, Google has a financial interest for the user to click on YouTube content. Similarly, Google's expected revenue may increase if a link to a Google map is included in the output instead of a link to MapQuest, Yahoo Maps, Bing Maps, etc. The debate about whether SEs should be regulated or not has ignited public interest (Crowcroft 2007, Inria 2012). A neutral SE should only use relevance to construct its rankings, and ignore revenues. This would allow new entrants that perform well (i.e., that are commonly clicked) to be listed near the top of the list of organic search results. The risk of a non-neutral ranking is that it may slow down innovation by favoring the incumbents that are known to generate profits, thereby preventing new applications/content from being shown, and hence from becoming known and successful.

EXAMPLE 5. *A Vertically Integrated SE with a CP.* Here we focus on a specific type of request which can be served by either third-party CPs or by the SE itself. Assume that a limited number of CPs compete with the SE, and that the parameters r, g, and $\lambda(r)$ for the instance correspond to just this type of request. We assume that each request has $M = 10$ matching pages, that one of them (say Page 1) is served directly by the SE, while the other nine are served by third-party CPs. In addition to the revenue coming from Page 1, the SE receives an expected revenue of $\beta = 1$ per request from sponsored links in the results page. We assume that $R_1, \ldots, R_{10}, G_1$ are independent random variables, all uniformly distributed over $(0,1)$, whereas $G_2 = \cdots = G_{10} = 0$. The CTRs θ_j, in Table 3, are proportional to the observed relative numbers of clicks as a function of position j, given in the first table of Dejarnette (2012). The multiplicative proportionality constant has no impact on our derivations, so we take it equal to 1. Each third-party CP also receives an expected revenue C_i, for $i = 2, \ldots, 10$, where the C_i are independent and uniform over $(0,1)$. Those C_i are not considered by the optimization performed by the SE, but we will use them to estimate the impact of various strategies on the CP revenues.

The $M = 10$ pages are ranked by the SE by decreasing value of $\tilde{R}_i + \rho \tilde{G}_i$, for some $\rho \geq 0$. Figure 4 shows the SE revenue, the relevance $r(\rho)$, the revenue and the visit rate for CP 1 and for each other (third-party) CP, all as a function of ρ. The revenues are per unit of time. When ρ increases, the SE favors CP 1

24

L'Ecuyer et al.: *Revenue-Maximizing Rankings*
Article submitted to ; manuscript no. (Please, provide the manuscript number!)

Table 3   CTR values θ_j used in Example 5

θ_1	θ_2	θ_3	θ_4	θ_5	θ_6	θ_7	θ_8	θ_9	θ_{10}
0.364	0.125	0.095	0.079	0.061	0.041	0.038	0.035	0.03	0.022

more, which decreases the overall relevance and increases the visit rate to CP 1. The optimal tradeoff for the SE is attained with $\rho^* \approx 0.55$. The ranking bias from choosing $\rho > 0$ only affects Page 1. The relative positions of the other pages remain the same as in the neutral ranking. Consequently, the relevance $r(\rho)$ is only marginally affected by ρ in this case. If R_1 was stochastically much smaller than the other R_i's (e.g., uniform over $[0, \epsilon]$ for a small ϵ), then the impact of ρ on r would be larger. When $\rho \to \infty$, Page 1 is always ranked first, and the relevance $r(\rho)$ becomes

$$r(\infty) = \left(\frac{\theta_1}{2} + \sum_{i=1}^{9} \theta_{i+1} \mathbb{E}[U_{(10-i)}] \right) = \frac{\theta_1}{2} + \sum_{i=1}^{9} \theta_{i+1} \frac{(10-i)}{10} \approx 0.517,$$

where $U_{(1)}, \ldots, U_{(9)}$ are independent random variables uniformly distributed over $[0, 1]$ sorted by increasing order (the order statistics), and the visit rate to Page 1 is $\theta_1 r(\infty) \approx 0.188$.
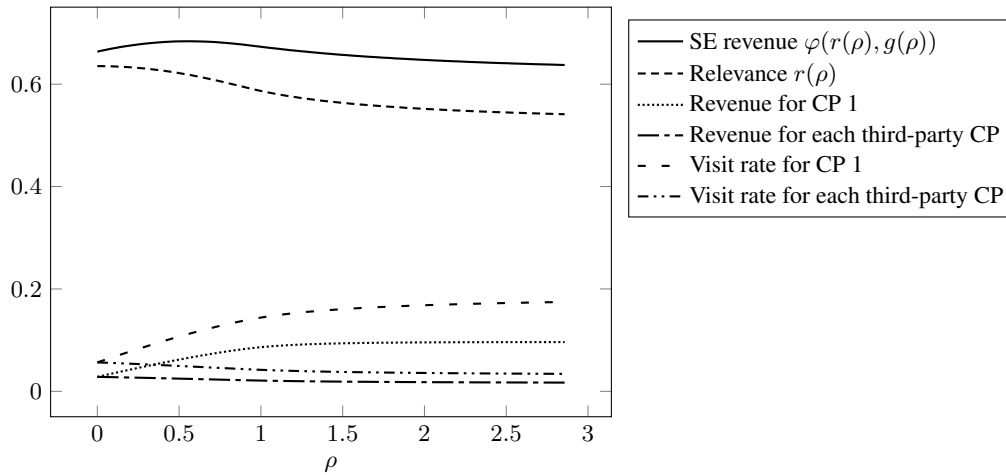
Figure 4   Performance measures as a function of ρ (simulation results)

To assess the sensitivity of the SE strategy to advertising, we now examine how the results depend on β. This shows the tradeoff that the SE faces for different types of requests. For requests related to, e.g., airline tickets, hotel reservations, or retailer products, the SE could make more profit by showing its own content among organic links rather than through sponsored search, because requests of this kind may produce conversions, whereas for requests that are appealing in the sponsored search market, the SE may try to offer the most relevant links, to boost that revenue stream. Figure 5(a) plots ρ^* as a function of β, while Figure 5(b) plots the ensuing revenue for CP 1 and for each third-party CP. Those functions were estimated by simulation, using the iterative fixed-point method to find ρ^*, with a fixed sample size of $n = 10^7$ at each

(a) Optimal ρ (b) CP revenues
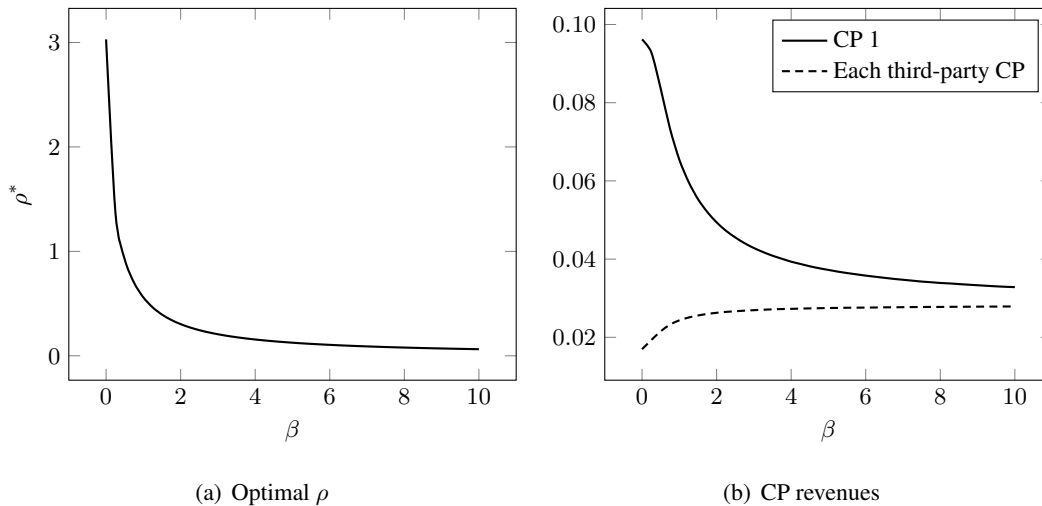
Figure 5   **Optimal ρ for the ranking, and CP revenue per unit of time, as functions of β (from simulation)**

step. When β grows, ρ^* tends to zero, because the revenue from sponsored links dominates, making it more rewarding for the SE to improve its reputation to attract more users. The impact of non-neutrality is small in this example because biasing the ranking only attracts limited additional revenue. When β is small, in contrast, sponsored links do not pay much and it becomes worthwhile for the SE to sacrifice relevance to some extent for immediate profits. In the extreme case when $\beta = 0$, we have $\rho^* = \infty$, so Page 1 is always placed at the top, and the other pages are sorted by decreasing order of relevance. This gives an average revenue of 0.09619 for CP 1 and of 0.01695 for each other CP. Although not shown in the figure, the expected SE revenue $\varphi(r^*, g^*)$ grows almost linearly with β, which means that the increasing revenues of sponsored search dominate the additional revenue to the SE coming from Page 1.

To illustrate the impact of non-neutrality, Table 4 reports the variations of the most relevant performance metrics when using $\rho = \rho^*$ instead of $\rho = 0$ (neutral ranking), for different values of β. For our parameters, we see that while the variation of the perceived quality (relevance) remains small (around 10%), the impact on the visibility and the revenues of the SE-owned CP is substantial: by being non-neutral, the SE can multiply the revenues of its CP by a factor of 2.8 and its visit rate by a factor larger than 3. On the other hand, the other CPs see their revenues and visit rates reduced by 14% to 32%, a significant decrease that may impact their long-term profitability. Of course, the results may be different with real data, but the framework and method can be applied similarly and can be used to study the impact of different types of regulations and optimization strategies on the various actors and on public welfare.

5.   Conclusion

We have introduced a new modeling framework that allows online platforms to rank items in a way that maximizes long-term revenues. The long-term impact is captured by the arrival rate of requests, which is an increasing function of the average relevance of the displayed pages clicked by the users. We proved that

Table 4   **The impact of non-neutral rankings in Example 5**

	Relevance	CP 1 revenue	other CP revenue	CP 1 visit rate	other CP visit rate
Neutral, $\rho = 0$ (reference case, optimal for $\beta = \infty$)	0.635	0.028	0.0283	0.057	0.057
Non-neutral, $\rho = 0.559$ (optimal for $\beta = 1$)	0.618 (-3%)	0.066 ($+136\%$)	0.0243 (-14%)	0.112 ($+96\%$)	0.049 (-14%)
Non-neutral, $\rho = 0.924$ (optimal for $\beta = .5$)	0.592 (-7%)	0.084 ($+200\%$)	0.0215 (-24%)	0.140 ($+146\%$)	0.043 (-25%)
Non-neutral, $\rho = 1.374$ (optimal for $\beta = .25$)	0.568 (-11%)	0.093 ($+232\%$)	0.0193 (-32%)	0.158 ($+177\%$)	0.039 (-32%)

although we have to choose an ordering among a large number of possibilities for each request and the objective function is nonlinear, an optimal ranking must satisfy some simple conditions: the items must be sorted by order of a score defined as a simple function of the (expected) relevance and short-term profits. This function depends on a single real-valued parameter. Under the additional assumption that the scores are all different with probability 1, this provides a unique ordering for each request. Then, the whole problem reduces to optimizing the value of this parameter. We have provided an algorithm for that that relies on simulating the process to obtain estimates.

Our model and results might prove useful to platform owners (SEs, classified ads websites, online retailers) to navigate the tradeoff between short-term and long-term effects when defining their ranking strategies. They can also be of interest to regulators, who can run experiments with the model (based on real data if available) to better understand the behavior of revenue-oriented platforms and to anticipate the impact of regulatory interventions, which is of particular importance with regard to the current *search neutrality* debate.

Acknowledgments

References

Aflaki, S., I. Popescu. 2014. Managing retention in service relationships. *Management Science* **60**(2) 415–433.

Asmussen, S., P. W. Glynn. 2007. *Stochastic Simulation*. Springer-Verlag, New York.

Athey, S., G. Ellison. 2011. Position auctions with consumer search. *Quarterly Journal of Economics* **126**(3) 1213–1270.

Auction Insights. 2008. Decoding eBay's best match. `https://web.archive.org/web/*/http://www.auctioninsights.info/decoding-ebays-best-match.html`. Last accessed Oct 2014.

L'Ecuyer et al.: *Revenue-Maximizing Rankings*
Article submitted to ; manuscript no. (Please, provide the manuscript number!)

27

Austin, D. 2006. How google finds your needle in the web's haystack. *American Mathematical Society Feature Column* **10**(12). http://www.ams.org/samplings/feature-column/fcarc-pagerank.

Avrachenkov, K., N. Litvak. 2004. Decomposition of the Google PageRank and Optimal Linking Strategy. Rapport de recherche RR-5101, INRIA. URL http://hal.inria.fr/inria-00071482.

Bertsekas, D. P., S. E. Shreve. 1978. *Stochastic Optimal Control: The Discrete Time Case*. Academic Press, New York.

Besbes, O., C. Maglaras. 2009. Revenue optimization for a make-to-order queue in an uncertain market environment. *Operations Research* **57**(6) 1438–1450.

Brill, J. 2013. Statement of the Commission regarding Google's search practices. http://www.ftc.gov/public-statements/2013/01/statement-commission-regarding-googles-search-practices. Last accessed Oct 2014.

Crowcroft, J. 2007. Net neutrality: The technical side of the debate: A white paper. *ACM SIGCOMM Computer Communication Review* **7**(1).

Dejarnette, R. 2012. Click-through rate of top 10 search results in Google. URL http://www.internetmarketingninjas.com/blog/search-engine-optimization/click-through-rate.

Edelman, B., B. Lockwood. 2011. Measuring bias in "organic" web search. http://www.benedelman.org/searchbias.

Edelman, B., M. Ostrovsky, M. Schwarz. 2007. Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. *The American Economic Review* **97**(1) 242–259.

Google. 2011. Facts about Google and competition. http://web.archive.org/web/20111104131332/http://www.google.com/competition/howgooglesearchworks.html. Last accessed Oct 2014.

Hanson, W.A., K. Kalyanam. 2007. *Internet Marketing and E-Commerce*. International student edition, Thomson/South-Western, Mason, Ohio.

Inria. 2012. Inria's response to ARCEP consultation about network neutrality.

Lahaie, S., D. Pennock, A. Saberi, R. Vohra. 2007. Sponsored search auctions. N. Nisan, T. Roughgarden, E. Tardos, V. Vazirani, eds., *Algorithmic Game Theory*, chap. 28. Cambridge University Press, 699–716.

Maglaras, C., A. Zeevi. 2003. Pricing and capacity sizing for systems with shared resources: Scaling relations and approximate solutions. *Management Science* **49**(8) 1018–1038.

Maillé, P., E. Markakis, M. Naldi, G. Stamoulis, B. Tuffin. 2012. Sponsored search auctions: an overview of research with emphasis on game theoretic aspects. *Electronic Commerce Research* **12** 265–300.

Maillé, P., B. Tuffin. 2014. *Telecommunication Network Economics: From Theory to Applications*. Cambridge University Press.

Mendelson, H., S. Whang. 1990. Optimal incentive-compatible priority pricing for the M/M/1 queue. *Operations Research* **38**(5) 870–883.

Pasupathy, R., S. Kim. 2011. The stochastic root finding problem: Overview, solutions, and open questions. *ACM Transactions on Modeling and Computer Simulation* **21**(3) Article 19.

Rudin, W. 1974. *Real and Complex Analysis*. 2nd ed. McGraw-Hill, New York.

Rushe, D. 2012. Eric Schmidt Google senate hearing – as it happened. `http://www.guardian.co.uk/technology/blog/2011/sep/21/eric-schmidt-google-senate-hearing` . Last accessed Oct 2014.

Varian, H.R. 2007. Position auctions. *International Journal of Industrial Organization* **25**(6) 1163–1178.

Williams, H. 2010. Measuring search relevance. `http://www.ebaytechblog.com/2010/11/10/measuring-search-relevance`.

Wright, J. D. 2012. Defining and measuring search bias: Some preliminary evidence. George Mason Law & Economics Research Paper 12-14, George Mason University School of Law.