Les Cahiers du GERAD ISSN: 0711–2440

Revenue-maximizing rankings for online platforms with quality-sensitive consumers

P. L'Ecuyer, P. Maillé,
N. Stier-Moses, B. Tuffin

G–2015–73

August 2015

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

Revenue-maximizing rankings for online platforms with quality-sensitive consumers

Pierre L'Ecuyer [a]

Patrick Maillé [b]

Nicolás Stier-Moses [c]

Bruno Tuffin [d]

[a] *GERAD & Département d'informatique et de recherche opérationnelle, Université de Montréal, Montréal (Québec) Canada, H3C 3J7 & Inria Rennes – Bretagne Atlantique, France*

[b] *Telecom Bretagne, 35576 Cesson Sévigné Cedex, France*

[c] *Universidad Torcuato Di Tella Business School, Saenz Valiente 1010, Buenos Aires, Argentina & CONICET Argentina*

[d] *Inria Rennes – Bretagne Atlantique, Campus Universitaire de Beaulieu, 35042 Rennes Cedex, France*

lecuyer@iro.umontreal.ca
patrick.maille@telecom-bretagne.eu
nstier@utdt.edu
bruno.tuffin@inria.fr

August 2015

Les Cahiers du GERAD

G–2015–73

Abstract: When a keyword-based search query is received by a search engine (SE), a classified ads website, or an online retailer site, the platform has exponentially many choices in how to sort the output to the query. Two extreme rules are (a) to return a ranking based on relevance only, which attracts more requests (customers) in the long run because of perceived quality, and (b) to return a ranking based only on the expected revenue to be generated by immediate conversions, 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, which provides valuable insight. A key feature of our model is that customers are quality-sensitive and are attracted to the platform or driven away depending on the average relevance of the output. The proposed methods are of crucial importance in e-business and encompass: (i) SEs that can reorder their organic output and place their own content in more prominent positions than that provided by third-parties, to attract more traffic to their content and increase their expected earnings as a result; (ii) classified ad websites which can favor paid ads by ranking them higher; and (iii) online retailers which can rank products they sell according to buyers' interests and also the margins these products have. This goes in detriment of just offering rankings based on relevance only and is directly linked to the current *search neutrality debate*.

Acknowledgments: This work has been supported by Inria's associated team MOCQUASIN, as well as an Inria International Chair, a Canada Research Chair, an NSERC Discovery Grant to P. L'Ecuyer and by CONICET Argentina Grant PIP 112-201201-00450CO and FONCYT Argentina Grant PICT 2012-1324.

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 transactions as a chance to sell advertisement; 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. In some cases, they 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 classified ads in premium locations or for additional time (e.g., leboncoin.fr in France, or Mercado Libre in Latin America). They may also offer additional services such as insurance or delivery for a fee. Finally, revenue also comes from third-party advertisers that display text or banners 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 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 (expected) 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 users. 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 the time until the end of the auction, the distance, the price, etc. The details on how to assign a relevance value to each item for a query depend on the intrinsic details of the platform. For example, eBay may use the distance between the query string and the item description 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 PageRanks as input (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.

The matching items (or links) for a query also have different values of expected revenue that could be obtained directly or indirectly by the platform owner when the user visits the item, and the owner may have interest in taking this expected revenue into account when ranking the items, by placing highly-profitable ones in prominent positions, to increase its revenue. However, a myopic approach that ranks the items only in terms of expected revenue and not on relevance would decrease its reputation in the long run, and eventually decrease the arrival rate of requests and the future revenues. A compromise must be made that accounts for both relevance and expected revenue.

A request is abstracted out in our model as a random vector that contains a relevance index and an expected revenue for each matching page. For the purpose of this study, the distribution of this random vector is assumed to be known and time-stationnary. Estimating (or learning) this distribution from actual data is of course important for practical implementations, but is outside the scope of this paper.

In addition to the regular output (the list of matching items, usually referred to as *organic links*), most platforms also display paid ads (also named *sponsored links*). Our study here focuses on the ordering of the organic links only. We assume that the average arrival rate of search requests is influenced by the average relevance of organic links, and not by the choice and ordering of the sponsored links. This makes sense because the latter ordering is much less likely to impact the future arrival rate of requests than the ranking of organic links. On the other hand, the total expected revenue from sponsored search depends on the arrival rate of requests. Our model accounts for this, e.g., via 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 links (ads). 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 links has not yet received a similar level of attention.

The purpose of our work is to study the best compromise that can be made by the platform to account for both relevance and expected revenue when ranking the items for a query, to maximize the long-term expected revenue per unit of time. Our aim is to find an optimal ordering strategy that takes both effects into account. We want a model whose solution has a simple and elegant form, which can provide insight, as opposed to a detailed and complicated model whose solution has no simple form. Our results provide optimal ranking conditions and policies for long-term revenue maximization. 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 (captured by the expected quality).

To capture the impact of long-term reputation on the rate of visits, we assume that the (average) arrival rate of requests is an increasing function of the average relevance of displayed links, and that the objective (or utility) function of the platform is a smooth increasing function of the average relevance and of the average income per request. It can be the product of the arrival rate by the average revenue per request, for example, but a more general utility function is also allowed.

These expected relevance and income per request depend on the *ranking policy* used to select a permutation of matching pages for each request. The ranking can be based on both the relevance and the expected revenue of each matching page. We consider a rich class of *randomized ranking policies* which to each request assign a probability distribution on the set of permutations of all matching pages. Whenever a request arrives, the platform selects a ranking using the probabilities specified by the policy for this request. Of course, such randomized policies can be very complicated to optimize and implement in general. For this reason, we are interested in situations in which an optimal policy (provably) has a simpler form, and is easier to implement. A *deterministic ranking policy* assigns to each request a single permutation. This already gives a much smaller decision space, but such a policy can still be very complicated to specify and implement in general, given that the number of possible requests is typically huge. Our aim is to study and provide insight on how an optimal policy looks like, and to develop a model for which (under certain conditions) the optimal policy has a simple form. For this reason, we focus on time-stationary models and policies.

Our main contribution is the characterization of such an optimal ranking policy. Even in our stylized and simplified setting, finding an optimal policy is not obvious a priori, in part because the objective function is non-additive and classical dynamic programming tools cannot be used directly. We propose an algorithm that exploits our characterization of the optimal policy and computes a parameter that allows the platform to assign a figure of merit (a scalar number) to each matching item. When these figures of merit are all distinct for any request with probability 1, the optimal ranking is found simply by sorting them. This provides a very simple deterministic optimal ranking policy. When equality occurs with positive probability, only randomized policies can be optimal in general. An optimal policy would still sort the matching items by order of figure of merit, but it must randomize the order of those having the same value, with specific optimal probabilities. In practice, when the probability of an equality is small, to avoid computing the optimal probabilities for randomization, one may opt to forget the randomization and 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 page, so the figures of merit 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 just a different way of randomizing the policy.

Our model can be seen as an infinite-horizon time-stationary sequential decision process that fits the general framework of stochastic dynamic programming (DP), so one could think of using DP methodology to characterize and compute an optimal ranking policy. In this paper, we restrict ourselves to time-stationary policies. In a dynamic setting we could also consider a time-dependent model, with an arrival rate and request distribution that varies randomly with time. This would be more complicated and is beyond our scope. Our model is not an ordinary discrete-time or discrete-event DP model because the arrival rate depends on the policy that is used: the objective function is not additive and classical DP tools do not readily apply. Despite this, our characterization of optimal policies was initially inspired by the derivation of optimal ranking conditions via an interchange argument in Bertsekas (2005, Section 4.5), in a DP setting. However, our solution is more involved, because of the impact of the policy on future arrivals.

Balancing between immediate revenue and long-term impact on future arrivals, when choosing a policy, in not a novel idea; see, e.g., 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 he 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 DP methods. Their solutions are algorithmic.

The model considered here obviously simplifies reality, as do virtually all models whose solution has a simple form, such as the Erlang models in queueing theory, the Black-Scholes model in finance, inventory models with (S, s) policies, etc. Despite not being fully realistic, these models are extremely helpful and insightful, and are widely used everyday. The model proposed here is in that spirit. In our case, reputation is built over a large time horizon and one can argue that a stationary model is a reasonable way to capture that market dynamic. 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.

A major motivation of our work is related to the *search neutrality debate*. In recent years, some SEs have been under scrutiny by individuals and organizations that oversee the Internet as well as by regulators in various countries because some believe that the organic search ranking is not based only on objective measures of relevance, but also account for some revenue-making ingredients (Crowcroft 2007). For example, it has been said that Google may favor YouTube and other of its own content because of the extra revenue it generates. This was discussed by the Federal Trade Commission in the US (Brill 2013) and in a Senate hearing (Rushe 2012). Search bias 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, etc. There are many other similar situations, including weather reports, movies and showtimes, product search, news articles, and so forth. Heterogenous characteristics, including different ownership, can be captured by explicitly associating each link within the organic output to the expected revenue attained when somebody clicks on it, and using those expected revenues as input to find the optimal ordering. The debate about whether SEs should or should not be allowed to enter into these considerations has ignited public interest (Crowcroft 2007, Inria 2012). It relates to other policy debates regarding whether and how to regulate the Internet; the most prominent example being *network neutrality*; see, e.g., Odlyzko (2009) for a discussion of both issues. A neutral SE should only use relevance to construct its rankings, and ignore the revenue parameter mentioned earlier. 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 to become known and successful. Motivated by the fact that many of these platforms belong to public companies that strive to maximize returns to stakeholders, and considering the search neutrality debate, it is important to study the impact of ranking policies on platforms' revenue, as well as on social welfare. The framework we introduce can also be of high interest to regulators who study the impact of search neutrality on users and on overall social welfare. It can help regulators determine if intervention is warranted, study the consequences of doing so, and provide arguments for or against non-neutral SEs.

The rest of the article is organized as follows. In Section 2, we present our modeling framework, state the optimization problem in terms of randomized policies, and derive a general characterization of the optimal solution. In Section 3, 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 and all the matching pages for each request have different figures of merit with probability 1, we show that the optimal policy is completely specified by a unique scalar number, and consists in sorting the items in terms of a linear combination of relevance and revenue, with slope equal to this scalar. This policy is very easy to implement: one does not need to consider the exponentially-many possible orderings. We provide numerical illustrations. In Section 4, we present some algorithms to compute or estimate this appropriate scalar number. Section 5 gives further numerical examples, to illustrate what could be done with the model and tools proposed here. Finally, Section 6 gives a conclusion.

2 Model formulation

We now define the model considered in this paper. We describe it 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, this model also applies to other marketplaces such as electronic retailers and classified-ad websites, for example.

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 and finite) 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 $G_i \in [0, K]$, where K is a positive constant. 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. If $M = 0$, then $Y = (0)$. We assume that Y has a probability distribution over a subspace Ω (discrete or continuous) of the space of vectors Y that satisfy these conditions. The variable R_i is a measure of how the SE thinks finding 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 on page i, for this request, if the user clicks on the link to page i. 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 transmitted to the SE, then G_i contains this expected revenue transferred to the SE. If the SE is also the CP for some pages, then the fraction is 1 for those pages. We will denote a fixed realization of the random vector Y by $y = (m, r_1, g_1, \ldots, r_m, g_m)$.

After receiving a request Y, 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 and its position of display. For a given request $y = (m, r_1, g_1, \ldots, r_m, g_m) \in \Omega$, we denote the CTR of page i placed in position j by $c_{i,j}(y)$. This $c_{i,j}(y)$ is assumed to be non-decreasing in r_i and non-increasing in j: A higher relevance or better position cannot decrease the CTR. If we select permutation π for this request, the corresponding expected relevance (the *local relevance*) is defined by

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

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

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

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 $\mathbb{E}$ 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 (which we also call the *reputation* of the SE) is

$$r := r(\mu) = \mathbb{E}[r(\mu(Y), Y)] \tag{3}$$

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

$$g := g(\mu) = \mathbb{E}[g(\mu(Y), Y)]. \tag{4}$$

We also consider randomized policies, motivated by the fact that in some situations they can do better than the best deterministic policy. A *randomized stationary ranking policy* is a function $\tilde{\mu}$ that assigns, to each $y = (m, r_1, g_1, \ldots, r_m, g_m) \in \Omega$, a probability distribution over the set of permutations Π_m. One has $\tilde{\mu}(y) = \{q(\pi, y) : \pi \in \Pi_m\}$, where $q(\pi, y)$ is the probability of selecting π. The $m!$ probabilities $q(\pi, y)$ determine in turn the probability $z_{i,j}(y)$ that page i ends up in position j, for each (i, j). This gives an $m \times m$ matrix of probabilities $z_{i,j}(y) \geq 0$, $1 \leq i, j \leq m$, for which each row sums to 1 and each column sums to 1 (a *doubly stochastic* matrix). Note that the correspondence between those two sets of probabilities is not one to one, because one has $m! - 1$ degrees of freedom for choosing $\tilde{\mu}(y)$ (we subtract 1 because the probabilities must sum to 1), and only $(m-1)^2$ degrees of freedom for choosing the matrix. That is, whenever $m > 2$, an infinite number of distinct distributions over Π_m give the same matrix of probabilities $z_{i,j}(y)$.

However, the expected relevance and revenue for a request y in our model depend on π only via $r(\pi, y)$ and $g(\pi, y)$, which are sums over i in which each term depends only on the position $\pi(i)$. Therefore, the expected relevance r per request (the reputation) and the expected SE revenue g per request from the organic links for a given randomized policy $\tilde{\mu}$ can be written equivalently by taking (for each y) the expectation either with respect to the probabilities $q(\pi, y)$ or with respect to the probabilities $z_{i,j}(y)$, in addition to taking the expectation over Y. That is, we have

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

and

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

In view of this equivalence, with a slight abuse of notation, a randomized policy $\tilde{\mu}$ can be defined equivalently as a rule that assigns, for each $y \in \Omega$, an $m \times m$ doubly stochastic matrix $\tilde{\mu}(y) = \mathbf{z}(y) = \{z_{i,j}(y) \geq 0 : 1 \leq i, j \leq m\}$. 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). For a given request y and a matrix of probabilities $\mathbf{z}(y)$, one can generate a random permutation that satisfies these probabilities by first generating the page at position 1, 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 $< j$.

Each randomized policy $\tilde{\mu}$ has 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 set.*

Proof. If the 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 $\tilde{\mu}^1(y) = \mathbf{z}^1(y) = \{z^1_{i,j}(y) : 1 \leq i, j \leq m\}$ and $\tilde{\mu}^2(y) = \mathbf{z}^2(y) = \{z^2_{i,j}(y) : 1 \leq i, j \leq m\}$ 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 $z_{i,j}(y) = \alpha z^1_{i,j}(y) + (1-\alpha) z^2_{i,j}(y)$, 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 $\mathcal{C}$. □

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 $[0,1] \times [0,K]$. An *optimal policy* from the perspective of the SE is a stationary ranking policy μ in the deterministic case, or $\tilde{\mu}$ in the randomized case, that maximizes $\varphi(r,g)$. Always ranking the pages by decreasing order of R_i would maximize r, whereas ranking them by decreasing order of G_i would maximize g. An optimal policy must usually make a compromise in between these two extremes. Typically, $\varphi(r,g)$ would represent the expected average revenue per unit of time in the long run, which can be written as the (time-average) arrival rate of requests multiplied by an expected revenue per request. For example, we can assume that requests arrive according to a point process whose average arrival rate (per unit of time) is $\lambda(r)$, where $\lambda : [0,1] \to [0,\infty)$ is an increasing, positive, smooth (continuously differentiable), and bounded function, and r is the average relevance corresponding to the policy in use, as defined in (3) or (5). In addition to the expected gains G_i from organic links, the SE can also place adds and other sponsored links on the page that provides the organic links. We can assume that the expected gain from those adds and sponsored links is β per request, on average. We then have

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

A key assumption here (for this special case) is that the average arrival rate is a function of the reputation, which depends only on the relevance and ordering of the proposed organic links, and not on the ordering of other (sponsored) links. That is, paid search is not going to drive users significantly to the website in the long term. Equation (7) also assumes that the expected gain β does not depend on the ordering of the organic links. Note that the arrival rate does not have to be constant; it can be periodic for example, with a time average of $\lambda(r)$.

Other important simplifying assumptions in our model are that the distribution of Y is stationary and does not depend on the ranking policy, and that the average arrival rate depends only on the single global relevance measure r. Those assumptions permit us to obtain simple-form results, which provide a first-order approximation to an optimal policy. In real life, the distribution of Y changes with time, but our model may be a good approximation over a period in which it does not change too much. One could also think of a model in which the distribution of Y would change with time, and ρ^* would be a function of time, but this is beyond the scope of this paper. We also do not model or distinguish individual "customers", but only consider a global distribution for Y. The measures r and g are averages across all queries. One relevant issue is whether it is reasonable to 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. To address this situation, one can segment (partition by user type, topic, etc.) the space of queries and apply the model to each segment. Then each segment can have its own distribution of Y and triple (r,g,ρ). This can be useful if ρ^* differs markedly across segments. Developing effective ways of making this segmentation can be a topic for further research. In principle, one could have a very large number of small segments, even a single IP address or user for a segment, but in practice one must also have enough data to estimate the distribution of Y in 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).

With all these ingredients, the optimization problem for the SE can be formulated as

$$\text{OP:} \qquad \max_{\tilde{\mu} \in \tilde{\mathcal{U}}} \quad \varphi(r,g)$$

$$\text{subject to}$$

$$r = \mathbb{E}\left[\sum_{i=1}^{M}\sum_{j=1}^{M} z_{i,j}(Y) c_{i,j}(Y) R_i\right]$$

$$\begin{aligned} g &= \mathbb{E}\left[\sum_{i=1}^{M}\sum_{j=1}^{M} z_{i,j}(Y)c_{i,j}(Y)G_i\right] \\ \tilde{\mu}(y) &= \mathbf{z}(y) = \{z_{i,j}(y) : 1 \le i,j \le m\} \quad \text{for all } y \in \Omega. \end{aligned}$$

This nonlinear optimization problem is not easy to solve in this general form.

Given that $\mathcal{C}$ is convex and $\varphi(r,g)$ is increasing in both r and g, if (r^*,g^*) is an optimal solution, with optimal value $\varphi^* = \varphi(r^*,g^*)$, there must be a line of negative slope passing through the point (r^*,g^*) and for which $\mathcal{C}$ is completely on the lower left side of the line. Let φ_r and φ_g denote the partial derivatives of φ with respect to r and g, so $\nabla\varphi(r,g) = (\varphi_r(r,g),\, \varphi_g(r,g))$ is the gradient of φ at (r,g). We define

$$h(r,g) := \varphi_g(r,g)/\varphi_r(r,g) \tag{8}$$

for all $(r,g) \in \mathcal{C}$. Consider now the line with equation

$$\varphi_r(r^*,g^*)(r-r^*) + \varphi_g(r^*,g^*)(g-g^*) = 0, \tag{9}$$

which passes through (r^*,g^*) and is orthogonal to the gradient of φ at that point, and define

$$\rho^* = h(r^*,g^*). \tag{10}$$

Proposition 1 *No point $(r,c) \in \mathcal{C}$ can be strictly above this line defined by (9). Moreover, if we replace the objective $\varphi(r,g)$ by the linear function $r + \rho^* g$ in the optimization problem OP, the point (r^*,g^*) remains an optimal solution to the modified problem. It is the unique solution to the modified problem if and only if the line (9) touches $\mathcal{C}$ at a single point.*

Proof. If the point $(r,c) \in \mathcal{C}$ is strictly above the line (9), one must have

$$(r-r^*, g-g^*) \cdot \nabla\varphi(r^*,g^*) > 0,$$

and the point $(r^*,g^*) + \epsilon(r-r^*, g-g^*)$ also belongs to $\mathcal{C}$ for any $\epsilon > 0$. By taking a Taylor expansion and ϵ small enough, this implies that the value of φ at this point must be strictly larger that $\varphi(r^*,g^*)$, which is a contradiction. This proves the first part. From this, it follows that if we replace the objective $\varphi(r,g)$ by the linear function $\varphi_r(r^*,g^*)r + \varphi_g(r^*,g^*)g$ in the optimization problem, the set of optimal solutions to the modified problem must contains the point (r^*,g^*). Because $\varphi_r(r^*,g^*) > 0$, we can divide this objective by $\varphi_r(r^*,g^*)$ to obtain the equivalent linear objective function $r + \rho^* g$. □

One nice feature of this proposition is that the condition involves only a single positive real number $\rho^* > 0$. It is true that when solving the optimization problem, this ρ^* is unknown, but we can replace the objective $\varphi(r,g)$ by the linear form $r + \rho g$ in which $\rho \ge 0$ is treated as another real-valued decision variable in the problem. Given the distribution of Y, one may think of estimating the value of the objective function for any fixed value of ρ by simulation, for example. This could be done at several values of ρ and a search could be done to estimate the best ρ. We will return to this in Section 4. One problem, however, is that knowing ρ^* is not enough in general to known the optimal strategy, as we will see later.

Note that (r^*,g^*) and ρ^* are not necessarily unique in general. For example, given the convex set $\mathcal{C}$, one can define the function φ in a way that $\varphi(r,g)$ is constant along the upper right boundary of $\mathcal{C}$ (i.e., so that this boundary is a contour line for φ). Then, (r^*,g^*) can be taken as any point on that upper right boundary, and the corresponding ρ^* would satisfy the proposition. Although having multiple optimal solutions with distinct values of ρ^* would not be a problem, it is unlikely to happen in large real-life applications. The next proposition gives a sufficient condition for uniqueness.

Proposition 2 *Consider the contour line defined by $\varphi(r,g) = \varphi^*$. If this line represents the graph of g as a strictly convex function of r, then (r^*,g^*) and ρ^* are unique.*

Proof. We know that no point of $\mathcal{C}$ can be strictly above this contour line, by definition of φ^*, and the point (r^*, g^*) is on it. Since the contour line is strictly convex and the upper right boundary of $\mathcal{C}$ defines a concave function, there can be no other point of intersection. □

As an example, if $\varphi(r, g) = \lambda(r)(\beta + g)$ where $\lambda(r) = r^\alpha$ with $\alpha > 0$, then the contour line 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 line is strictly convex for $0 < r \le 1$.

3 Optimality conditions for ranking policies

In this section, we characterize optimal ranking rules for slightly more specific models, in which the CTR has a separable form. We show that when Y has a discrete distribution, the optimal policy must be randomized in general. But in a model where Y has a continuous distribution, we show that there is an optimal deterministic policy that has a nice and simple form. Later, we propose algorithms to compute or approximate it.

3.1 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:

$$
\begin{aligned}
\text{OP-D:} \qquad & \max \quad \varphi(r, g) \\
& \text{subject to} \\
& r = \sum_{y\in\Omega} p(y) \sum_{i=1}^{M}\sum_{j=1}^{M} z_{i,j}(y) c_{i,j}(y) r_i \\
& g = \sum_{y\in\Omega} p(y) \sum_{i=1}^{M}\sum_{j=1}^{M} z_{i,j}(y) c_{i,j}(y) g_i \\
& \sum_{j=1}^{M} z_{i,j}(y) = 1 \qquad \text{for all } y \in \Omega, 1 \le i \le M \\
& \sum_{i=1}^{M} z_{i,j}(y) = 1 \qquad \text{for all } y \in \Omega, 1 \le j \le M \\
& 0 \le z_{i,j}(y) \le 1 \qquad \text{for all } y \in \Omega, 1 \le i, j \le M
\end{aligned}
$$

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, this probability swap would change r and g by

$$\Delta r = \delta p(y) \left[(c_{i,j'}(y) - c_{i,j}(y)) r_i + (c_{i',j}(y) - c_{i',j'}(y)) r_{i'}\right]$$

and

$$\Delta g = \delta p(y) \left[(c_{i,j'}(y) - c_{i,j}(y)) g_i + (c_{i',j}(y) - c_{i',j'}(y)) g_{i'}\right],$$

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

$$\begin{aligned} & c_{i,j}(y)r_i + c_{i',j'}(y)r_{i'} + \rho[c_{i,j}(y)g_i + c_{i',j'}(y)g_{i'}] \\ \geq \quad & c_{i,j'}(y)r_i + c_{i',j}(y)r_{i'} + \rho[c_{i,j'}(y)g_i + c_{i',j}(y)g_{i'}] \end{aligned}$$

whenever $\min(z_{i,j}(y),\, z_{i',j'}(y)) > 0$, for all i, j, i', j', y. Computing an optimal solution by using these general conditions is not easy because: (i) it must be done with $\rho = \rho^*$, which is unknown; (ii) even if we can approximate ρ^* (e.g., via a line search), finding a permutation that satisfies these conditions, for each y, is time-consuming; (iii) as we shall see later, there are often an infinite number of solutions that satisfy those conditions, and only one (or a few) of them are optimal. That is, these conditions are necessary for optimality, but they are not sufficient in general, so even if we knew ρ^* exactly this may not be sufficient. Next, we find simpler forms of these conditions, for a more specific model of CTR function, then we introduce a model in which those simplified conditions are necessary and sufficient.

3.2 Optimality conditions for a separable CTR function

In the rest of this article we assume that the CTR function has the separable form

$$c_{i,j}(y) = \theta_j\, \psi(r_i), \tag{11}$$

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 $\psi : [0,1] \to [0,1]$ is a non-decreasing function that maps the relevance r_i to a position-independent (base) click probability for the page. This separability assumption of the CTR is pervasive in the e-Commerce literature (Varian 2007, Maillé et al. 2012). We will rely on it to derive simple and convenient optimality conditions. We define $\tilde{R}_i := \psi(R_i)R_i$ and $\tilde{G}_i := \psi(R_i)G_i$, so that we can write $c_{i,j}(Y)R_i = \theta_j\tilde{R}_i$ and $c_{i,j}(Y)G_i = \theta_j\tilde{G}_i$, and similarly for the realization y, r_i, and g_i.

In the case where Y has a discrete distribution, with the separable form of the CTR, when $p(y) > 0$ the expressions for Δr and Δg given earlier become

$$\begin{aligned} \Delta r &= \delta p(y)(\theta_{j'} - \theta_j)(\tilde{r}_i - \tilde{r}_{i'}), \\ \Delta g &= \delta p(y)(\theta_{j'} - \theta_j)(\tilde{g}_i - \tilde{g}_{i'}), \end{aligned}$$

and the conditions (11) simplify to

$$(\theta_{j'} - \theta_j)\,[(\tilde{r}_i - \tilde{r}_{i'}) + \rho(\tilde{g}_i - \tilde{g}_{i'})] \leq 0 \tag{12}$$

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 $\theta_{j'} \leq \theta_j$. If $\theta_{j'} = \theta_j$, the condition is always trivially satisfied. If $\theta_{j'} < \theta_j$, one must have

$$\tilde{r}_i + \rho\tilde{g}_i \geq \tilde{r}_{i'} + \rho\tilde{g}_{i'}. \tag{13}$$

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 (13) must hold. This means that we have the following:

Proposition 3 *Any optimal randomized policy must rank the pages by decreasing order of their value of* $\tilde{r}_i + \rho\tilde{g}_i$ *whenever* $p(y) > 0$*, with the exception that the order at positions* $j' > j$ *with* $\theta_{j'} = \theta_j$ *does not matter.*

We call a ranking policy that satisfies this condition for a given $\rho > 0$ a *linear ordering (LO) policy with ratio* ρ (or LO-ρ policy, for short). 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 $\tilde{g}_i$.

We emphasize that specifying ρ is not enough to uniquely characterize an optimal policy in the case where, with positive probability, two or more pages have the same value of $\tilde{R}_i + \rho\tilde{G}_i$. 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 them. Given ρ, this would be very easy to implement, just by sorting the M matching pages by decreasing order of $\tilde{R}_i + \rho\tilde{G}_i$, for each y. However, the ordering in case of equality does matter, as illustrated by the following tiny 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 $\psi(r_i) = 1$ for all r_i, $\lambda(r) = r$, $(\theta_1, \theta_2) = (1, 1/2)$, 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

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

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 $(11/10, 1)$. See Figure 1.

With the optimal $p^* = 5/8$, we also obtain $\rho^* = r^*/(1 + g^*) = 2/5$, and it turn 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 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, defined by a probability measure ν over the class $\mathcal{B}$ of Borel subsets of the space Ω of vectors $y = (m, r_1, g_1, \ldots, r_m, g_m)$ where $m \in \{1, \ldots, m_0\}$ and $(r_i, g_i) \in [0, 1] \times [0, K)$ for each i. We assume that Y has a (finite) density function f over Ω, so that for each $B \in \mathcal{B}$, $\nu(B) = \int_B f(y)dy$.

Lemma 1 and Proposition 1 still apply in this case. However, the argument that led to (13) no longer applies, because $p(y) = 0$ for all $y \in \Omega$. We now prove a variant of this result by adapting the argument.

Let $\tilde{\mu}$ be an optimal randomized policy, with corresponding r^* and g^*, and let $B \in \mathcal{B}$ be a set with $\nu(B) > 0$ and such that

$$\delta := \sup_{y \in B} \min(z_{i,j}(y),\ z_{i',j'}(y)) > 0$$

for some 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 the probabilities $z_{i,j'}(y)$ 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

$$\begin{aligned} \Delta r &= \delta \int_B (\theta_{j'} - \theta_j)(\tilde{r}_i - \tilde{r}_{i'}) f(y)dy, \\ \Delta g &= \delta \int_B (\theta_{j'} - \theta_j)(\tilde{g}_i - \tilde{g}_{i'}) f(y)dy, \end{aligned}$$

and we must have

$$0 \geq \Delta r + \rho \Delta g = (\theta_{j'} - \theta_j)\delta \int_B \left[(\tilde{r}_i - \tilde{r}_{i'}) + \rho(\tilde{g}_i - \tilde{g}_{i'})\right] f(y)dy. \tag{14}$$

We now define a notion of LO-ρ policy in the setting of a density, and we establish that any optimal randomized policy must be of that form, for $\rho = \rho^*$.

Definition 1 *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 can be arbitrary.*

Proposition 4 *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. The proof is by contradiction. 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 $y \in B$,

$$\tilde{r}_i + \rho^* \tilde{g}_i \le \tilde{r}_{i'} + \rho^* \tilde{g}_{i'} - \epsilon \tag{15}$$

and $\min(z_{i,j}(y),\, z_{i',j'}(y)) \ge \delta$ 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

$$(\theta_{j'} - \theta_j)\left[(\tilde{r}_i - \tilde{r}_{i'}) + \rho(\tilde{g}_i - \tilde{g}_{i'})\right] \tag{16}$$

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. □

Proposition 4 tells us that any optimal policy must satisfy the LO-ρ conditions for ρ^* with probability 1. But we need further assumptions to make sure that this specifies an optimal policy. In the rest of this section, we make the following assumption.

Assumption 1 *For any $\rho \ge 0$, and any $j > i > 0$, $\mathbb{P}[M \ge j$ and $\tilde{R}_i + \rho\tilde{G}_i = \tilde{R}_j + \rho\tilde{G}_j] = 0$.* □

Proposition 5 *Under Assumption 1, for any fixed $\rho \ge 0$, there is a deterministic LO-ρ policy $\mu = \mu(\rho)$ that sorts almost any $Y \in \Omega$ (i.e., there is a unique order with probability 1). If $\rho = \rho^* = h(r^*, g^*)$ for an optimal solution (r^*, g^*), this deterministic policy is optimal.*

The LO-ρ ranking policy in Proposition 5 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 $\tilde{h}(\rho)$. From Proposition 4, if μ is an optimal policy, then one must have $\rho = \rho^* = h(r(\mu(\rho^*)), g(\mu(\rho^*))) = \tilde{h}(\rho^*)$. The next proposition says that under certain conditions such a fixed point exists and is unique.

Proposition 6 *(i) If $h(r, g)$ is bounded over $[0, 1] \times [0, K]$, then the fixed-point equation*

$$\tilde{h}(\rho) = \rho \tag{17}$$

has at least one solution in $[0, \infty)$.
(ii) If the derivative $\tilde{h}'(\rho) < 1$ for all $\rho > 0$, then the solution is unique.

Proof. (i) If $\tilde{h}(0) = 0$, then $\rho = 0$ is already a solution. Otherwise, we have $\tilde{h}(0) > 0$ and $\tilde{h}(\rho) \le K'$ for all $\rho \ge 0$ for some constant K'. In particular, $\tilde{h}(K') \le K'$. Since $\tilde{h}$ is a continuous function, there must be at least one point $\rho \in [0, K']$ at which $\tilde{h}(\rho) = \rho$.

(ii) The slope of the function $\tilde{h}(\rho)$ is always less than 1, so it cannot cross the line $f(\rho) = \rho$ more than once. □

Going back to the special case of our running examples where $\varphi(r, g) = \lambda(r)(\beta + g)$, Proposition 6 becomes:

Proposition 7 *Suppose $\varphi(r, g) = \lambda(r)(\beta + g)$.*
(i) If $\lambda(r)/\lambda'(r)$ is bounded for $r \in [0, 1]$ and $g(\rho(0)) > 0$, then (17) has at least one solution in $[0, \infty)$.
(ii) If $\lambda(r)/\lambda'(r)$ is also non-decreasing in r, then the solution is unique.

Proof. (i) In this case, we have

$$\tilde{h}(\rho) = \frac{\lambda(r(\mu(\rho)))}{\lambda'(r(\mu(\rho)))g(\mu(\rho))}.$$

Note that $g(\mu(\rho)) \geq g(\mu(0))$ for all $\rho \geq 0$. Therefore, the conditions in (i) imply that $\tilde{h}(\rho)$ is bounded and we can apply Proposition 6 (i).

(ii) If $\lambda(r)/\lambda'(r)$ is non-decreasing in r, then it is non-increasing in ρ since $r(\mu(\rho))$ 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 we can apply Proposition 6 (ii). □

The condition in (ii) that $\lambda(r)/\lambda'(r)$ is non-decreasing is a bit stronger than what we need to satisfy the condition of Proposition 6 (ii). 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 $r \in [0, 1]$. Other simple cases where the 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 several examples in this paper. We see that the uniqueness condition in Proposition 7 (ii) is quite reasonable.

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 for the ordering, and the pages are always sorted by decreasing order of $\tilde{R}_i$. If we assume that $\psi(r_i)$ is a non-decreasing function of r_i, which is natural, then the pages should always be sorted by order of relevance R_i. Then, the SE has the incentive to conform to search neutrality. The value of r obtained under this ordering, say r_0, is the maximal possible value, so we always have $r \in [0, r_0]$.

The next example establishes that 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. □

One might thing that if no $\tilde{R}_i$ or $\tilde{G}_i$ has a probability mass at a given point, then Assumption 1 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 $\tilde{R}_i + \rho\tilde{G}_i = \tilde{R}_j + \rho\tilde{G}_j$.

Example 3 Starting again from Example 1, suppose that G_2 now has the uniform distribution over the interval $(2 - \epsilon, 2 + \epsilon)$. The expectation of G_2 is the same as before, but now we obtain slightly more accurate information on the revenue G_2 before making the ranking decision. This modified (perturbed) model satisfies Assumption 1, so finding $\rho^* = \rho^*(\epsilon)$ is sufficient to completely specify an optimal policy for the modified model. Since the perturbed G_2 is observed before making the ranking decision and can be used for making the decision, we anticipate that the perturbed model will provide a larger φ^* than the original one with its optimal randomized policy, and that φ^* increases with ϵ. Figure 1 will confirm this. However, this gain on φ^* is negligible when ϵ is small enough.

We can write $G_2 = 2 + (2V - 1)\epsilon$ where $V \sim U(0, 1)$. The 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 $V < p = p(\epsilon) \stackrel{\text{def}}{=} 2/(5\rho\epsilon) - 1/\epsilon + 1/2$, which occurs with probability p. We have $r = (7 + 4p)/10$ as before and

$$
\begin{aligned}
g &= 2 - p + \epsilon \int_0^p (v - 1/2)\mathrm{d}v + \epsilon \int_p^1 (2v - 1)\mathrm{d}v \;=\; 2 - p + \epsilon p(1 - p)/2, \\
\varphi(r, g) &= r(1 + g) = \frac{7 + 4p}{10}(3 - p + \epsilon p(1 - p)/2), \\
\frac{\partial}{\partial p}\varphi(r, g) &= [5 - 8p + \epsilon(7/2 - 3p(1 + 2p))]/10.
\end{aligned}
$$

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

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

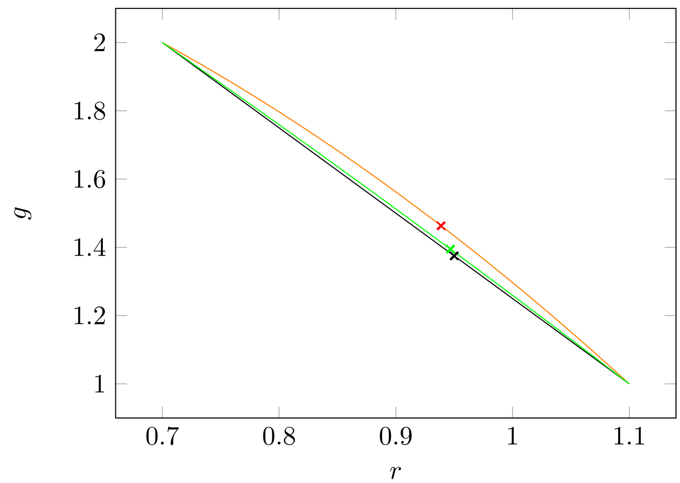


Figure 1: The reachable pairs (r, g) for Example 1 (black line), and for Example 3 with an LO-ρ policy for $\epsilon = 0.5$ (red line) and $\epsilon = 0.1$ (green line)

from the definition of p. We see that when $\epsilon \to 0$, $p^* = p^*(\epsilon) \to 5/8 = 0.625$ and $\rho^* \to 2/5 = 0.4$. Table 1 gives the optimal values as a function of ϵ. When ϵ is small enough, ρ^* and φ^* for the perturbed problem are almost identical to the optimal values for the original problem with a randomized policy.

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

ϵ	p^*	ρ^*	r^*	g^*	φ^*
0.0	0.625	0.4	0.95	1.375	2.25625
0.001	0.62491	0.39995	0.94996	1.37521	2.25636
0.01	0.62411	0.39950	0.94964	1.37006	2.25736
0.1	0.61705	0.39537	0.94682	1.39476	2.26741
0.5	0.59771	0.38137	0.93908	1.46240	2.31240

For a fixed $\epsilon > 0$, an alternative way of finding $p^*(\epsilon)$ numerically (instead of solving the quadratic equation) is to start say with $p = 5/8$, which is the optimal probability when $\epsilon = 0$, compute $\rho = \rho(p)$, and iterate the contraction mapping $\rho \leftarrow \tilde{h}(\rho) = r(\rho)/g(\rho)$ until ρ has converged to $\rho^*(\epsilon)$ with sufficient accuracy (see Section 4). This iterative method will also work for large models.

Figure 1 pictures the reachable values of (r, g) for this example, for some large values of ϵ. The upper curve (in red) 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^*) is shown by an x in red. 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 $\mathcal{C}$. For smaller values of ϵ, this region becomes very thin and the optimal LO-ρ policy gives almost the same value as the optimal randomized policy of Example 1. That is, the optimal randomized policy for the discrete case can be approximated arbitrarily closely by a deterministic LO-ρ policy after adding a small random perturbation to Y so its distribution satisfies Assumption 1.

3.4 Continuous approximation to the discrete case

Motivated by Example 3, when Y has a discrete distribution, instead of solving OP-D, we can approximate the problem 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. We know that the optimal ranking policy for the perturbed problem is an LO-ρ policy for some $\rho = \rho^* = \rho^*(\epsilon)$ and that ρ^* completely determines the policy (with probability 1). Note that the optimal value $\varphi^*(\epsilon)$ for the perturbed problem is always at least as large as the optimal value φ^* of the original problem, because its decisions are based on more information. In fact, any given (fixed) randomized policy $\tilde{\mu}$ for the original problem, say with corresponding pair $(r(\tilde{\mu}), g(\tilde{\mu}))$, can be applied to the perturbed problem by ignoring the realized perturbation before making the decision. This cannot beat the optimal policy for the perturbed problem. On the other hand, the optimal policy for the perturbed problem, which is an LO-ρ policy for $\rho = \rho^*(\epsilon)$, can be applied to the original problem as follows: whenever a request arrives, generate artificial random perturbations on the G_i's and use the rule with $\rho^*(\epsilon)$ to determine the ranking. This policy cannot beat the optimal randomized policy for the original problem, but at the same time its corresponding value of g for the original and perturbed problems cannot differ by more than $\epsilon(\theta_1 + \cdots + \theta_{m_0})$ and its corresponding value of r is the same for both problems. Therefore, the optimal values φ^* of the original and perturbed problems cannot differ by more than $\epsilon(\theta_1 + \cdots + \theta_{m_0}) \sup_{r,g} \varphi_g(r, g)$ and, since $\varphi_g(r, g)$ is bounded, this difference converges to 0 when $\epsilon \to 0$.

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

3.5 Two illustrative examples

This section provides two examples that, although very simple and stylized, capture some features of the search market in the real world. The first example can model an SE that has content that competes with third-party CPs. Imagine that Google receives a video-search request and there are two pages that match the search; the first is from YouTube, owned by Google, while the second belongs to a competitor such as Dailymotion. Google's revenue generated by the YouTube page is positively correlated with the page relevance because of higher advertisement revenue and higher YouTube perception. Instead, Google's revenue generated by the Dailymotion page is negatively correlated with the page relevance because the more relevant, the more it diverts traffic from YouTube.

To simplify the exposition, we assume in all our illustrative examples here and in the rest of the paper that $\lambda(r) = r$, $\varphi(r, g) = \lambda(r)(\beta + g)$ for some constant $\beta \geq 0$, and $\psi(R_i) = 1$. The latter means that $c_{i,j}(y) = \theta_j$, so the CTR depends only on the position of the page. These assumptions are by no means necessary or realistic, but they simplify the examples.

Example 4 Consider an instance with two pages where R_1 and R_2 are independent and uniformly distributed over $[0, 1]$, $G_1 = \alpha R_1$ and $G_2 = \alpha(1 - R_2)$ where $0 \leq \alpha \leq 1$, $\theta_1 = 1$, and $\theta_2 = 0$. Here, Y has a continuous distribution and Assumption A is satisfied.

For any given ρ, with probability 1, an LO-ρ policy places Page 1 before Page 2 if and only if $R_1 + \rho G_1 > R_2 + \rho G_2$; i.e., when $(1 + \alpha\rho)R_1 > \rho - (1 - \alpha\rho)R_2$; i.e., on the domain

$$D = \left\{ (R_1, R_2) : R_1 > \frac{\tilde{\rho}}{1 + \tilde{\rho}} + R_2 \frac{1 - \tilde{\rho}}{1 + \tilde{\rho}} \right\},$$

where $\tilde{\rho} := \alpha\rho$. Let $\bar{D} := [0, 1]^2 \setminus D$. By straightforward computations, we find that the LO-ρ policy gives

$$r = \int_D r_1 dr_1 dr_2 + \int_{\bar{D}} r_2 dr_1 dr_2 = \frac{2}{3} - \frac{\tilde{\rho}^2}{6(1 + \tilde{\rho})^2}$$

and

$$\frac{g}{\alpha} = \int_D r_1 dr_1 dr_2 + \int_{\bar{D}} (1 - r_2) dr_1 dr_2 = \frac{2}{3} - \frac{1}{6(1 + \tilde{\rho})^2}.$$

The SE revenue is thus

$$\begin{aligned} \varphi(r, g) &= r(\beta + g) = \left(\frac{2}{3} - \frac{\tilde{\rho}^2}{6(1 + \tilde{\rho})^2} \right) \left(\beta + \frac{2\alpha}{3} - \frac{\alpha}{6(1 + \tilde{\rho})^2} \right) \\ &= \frac{\alpha\tilde{\rho}^2/36 - (3\beta + 2\alpha)\tilde{\rho}^2(1 + \tilde{\rho})^2/18 - \alpha(1 + \tilde{\rho})^2/9}{(1 + \tilde{\rho})^4}, \end{aligned}$$

which we want to maximize with respect to $\tilde{\rho} \geq 0$. Taking the derivative with respect to $\tilde{\rho}$ and setting it to 0, we get the following equation, to which $\alpha\rho^*$ is a positive root (unless $\rho^* = 0$):

$$(3\beta + 2\alpha)\tilde{\rho}^3 + (6\beta + 5\alpha/2)\tilde{\rho}^2 + (3\beta - 5\alpha/2)\tilde{\rho} - 2\alpha = 0.$$

It follows from Proposition 7 (ii) that ρ^* is unique. Moreover, Assumption 1 is satisfied, so ρ^* defines the order uniquely with probability 1.

For a complete numerical illustration, let us take $\alpha = \beta = 1$. Then, ρ^* is the unique positive root of $5\rho^3 + 17\rho^2/2 + \rho/2 - 2$, which is $\rho^* \approx 0.412149553$. Figure 2 shows the expected SE revenue as a function of ρ. One can also verify that for $\rho = \rho^*$, we have $r = 0.6524696521$, $g = 0.583089554$, and then $h(r, g) = r/(1 + g) = 0.412149553 = \rho^*$, as expected. □

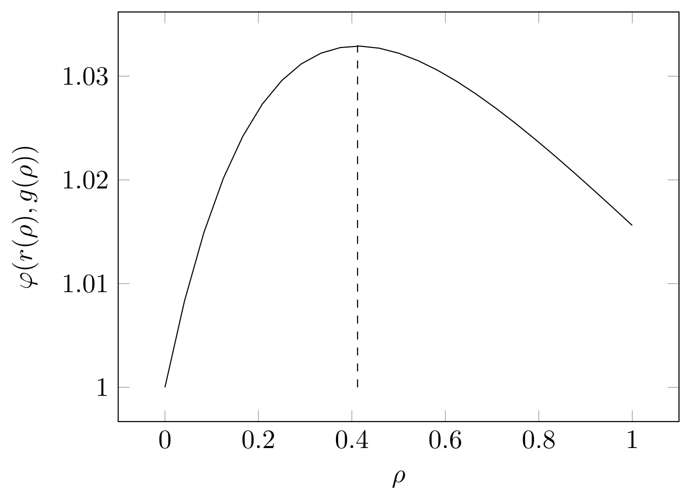


Figure 2: Expected SE Revenue in terms of ρ when $\alpha = \beta = 1$

In the next example, we consider a situation where the SE has a positive expected revenue G_i for a page when it is also the CP for that page, 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.

Example 5 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 a mixture of two uniform densities, and Assumption 1 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.

We first derive the formula for $r = r(\rho)$. We distinguish two cases for (G_1, G_2):

1. If $G_1 = G_2$, only the most relevant link is displayed, resulting in conditional expected relevance $\mathbb{E}[\max(R_1, R_2) \mid G_1 = G_2] = 2/3$.
2. 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

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

Combining the four possibilities for (G_1, G_2), the overall expected relevance for the LO-ρ policy is

$$r = r(\rho) = \frac{2}{3} + p(1-p)\bar{\rho}^2\left(\frac{2\bar{\rho}}{3} - 1\right), \tag{18}$$

where $\bar{\rho} := \min(1, \rho)$.

Similarly, to compute the expected revenue $g = g(\rho)$ per request, we consider the same two cases:

1. If $G_1 = G_2$, the expected revenue is 0 if $G_1 = 0$, and 1 otherwise.
2. If $G_1 \neq G_2$, we can assume again that $G_1 = 1$ and $G_2 = 0$. If $\rho > 1$, link 1 is always shown and the revenue is 1. If $\rho \leq 1$, the expected revenue conditional on $(G_1, G_2) = (1, 0)$ is

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

Regrouping all cases, we obtain

$$g = g(\rho) = p^2 + p(1-p)\left(1 - (2-\bar{\rho})^2\right). \tag{19}$$

Both $r(\rho)$ and $g(\rho)$ are constant for $\rho \geq 1$, so we can reduce the search for an optimal 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

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

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). □

This illustrates that to optimize the tradeoff between short-term revenue coming from immediate payments and long-term revenue coming from more exposure due to higher relevance, one must place the appropriate relative weights on relevance and on short-term revenues (1 and ρ, respectively).

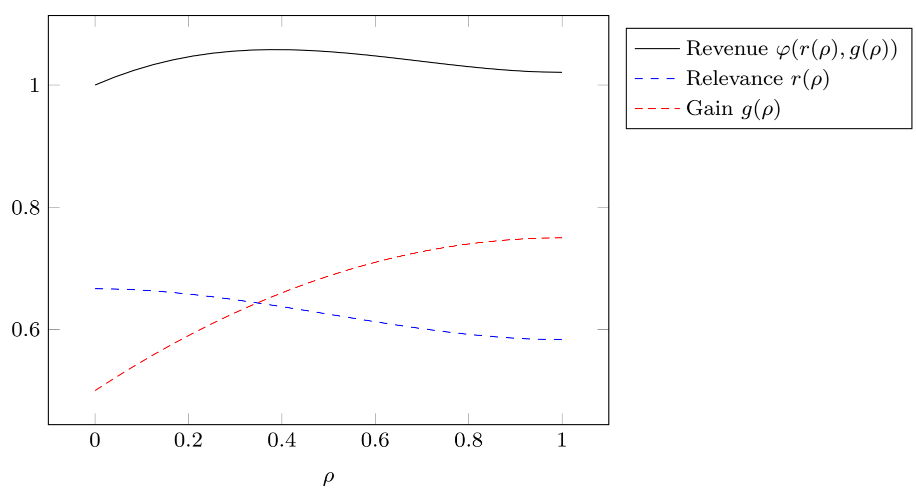


Figure 3: Expected SE revenue per unit time for $\beta = 1$ and $p = 1/2$

4 Finding optimal rankings by computing ρ^*

In this section, we discuss how to find ρ^*, the optimal ρ for the LO-ρ policy. Proposition 4 tells us that to maximize his revenue, the SE must follow an LO-ρ policy for a properly chosen $\rho = h(r, g)$. But this $h(r, g)$

depends on r and g, which in turn depend on the selected policy μ and are unknown a-priori. Moreover, typically, this dependence is not expressed in a closed-form formula. In the examples of Section 3.5, 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. Here we formulate the search for ρ^* as 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. The methods discussed here assume that a root exists and is unique. Proposition 7 gives sufficient conditions for this to hold. Several algorithms have been designed and studied for this type of problem; see, e.g., Pasupathy and Kim (2011) and the references therein. We explain how to adapt and use this technology in our setting and we end up recommending a simplified approach that turns out to be a contraction mapping in most cases.

An estimator $\hat{h}_n(\rho)$ of $\tilde{h}(\rho)$ at any given value of ρ can be defined and computed 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 $g(\rho)$:

$$\begin{aligned}\hat{r}_n(\rho) &= \frac{1}{n}\sum_{k=1}^{n}\sum_{i=1}^{M_k}\theta_{\pi(i)}\tilde{R}_{k,i},\\ \hat{g}_n(\rho) &= \frac{1}{n}\sum_{k=1}^{n}\sum_{i=1}^{M_k}\theta_{\pi(i)}\tilde{G}_{k,i}.\end{aligned}$$

They lead to the estimator

$$\hat{h}_n(\rho) = \varphi(\hat{r}_n(\rho), \hat{g}_n(\rho)) \tag{20}$$

for $h(\rho)$, which is generally biased when φ is nonlinear, but is consistent, and the bias typically decreases as $\mathcal{O}(1/n)$ (Asmussen and Glynn 2007). A confidence interval for $\tilde{h}(\rho)$ can be computed via the Delta method (Asmussen and Glynn 2007), under the assumption that $\hat{r}_n(\rho)$ and $\hat{g}_n(\rho)$ have (approximately) a normal distribution and φ is smooth enough. For the special case where $\varphi(r, g) = \lambda(r)(\beta + g)$, the estimator is $\hat{h}_n(\rho) = \lambda(\hat{r}_n(\rho))(\beta + \hat{g}_n(\rho))$.

When searching for a root of $\tilde{h}(\rho) - \rho$, or if we want to estimate the function $\tilde{h}$ over some interval, we need to compute $\hat{h}_n(\rho)$ at many values of ρ. This can be done using *common random numbers* (CRN), which means that we use exactly the same n realizations $Y_1, \ldots, Y_n$ at all values of ρ at which we perform a function evaluation, or using *independent random numbers* (IRN), in which case we draw a fresh independent sample $Y_1, \ldots, Y_n$ at each ρ where we estimate $\tilde{h}(\rho)$. In the CRN case, $\hat{h}_n(\rho)$ becomes a deterministic function of ρ and this function typically varies much less than in the IRN case. The *sample average optimization* method consists in optimizing this sample function $\hat{h}_n(\rho)$ defined with CRNs. However, for any fixed n, this sample function is piecewise-constant in ρ, because it depends on ρ only via the selected permutation for each i, and therefore only takes a finite number of values as a function of ρ. As a result, its derivative is zero almost everywhere and (in general) $\hat{h}_n(\rho) - \rho$ has no exact root. Therefore, the best we can do for fixed n is to compute an *approximate* root $\hat{\rho}_n^*$ of $\hat{h}_n(\rho) - \rho$, and for this, any method that relies on the derivative of $\hat{h}_n(\rho)$ must be ruled out. We can approximate the root either by a method that does not rely on derivatives (such as binary search), or by a derivative-based method (e.g., a Newton-type method) by approximating the derivative with finite differences. Thus, we can compute $\hat{\rho}_n^*$ such that $\epsilon_n = |\hat{h}_n(\hat{\rho}_n^*) - \hat{\rho}_n^*|$ is small, and do this for an increasing sequence of values of n, in a way that $\epsilon_n \to 0$ when $n \to \infty$. This is possible under the assumption that $\hat{h}_n \to \tilde{h}$ uniformly when $n \to \infty$, which usually occurs with CRNs (under mild conditions). For each considered sample size n, we would use the approximate root $\hat{\rho}_n^*$ as a starting point when finding the approximate root for the next (larger) value of n.

Another approach is a Robbins-Monro-type *stochastic approximation* (SA) iterative method; see Pasupathy and Kim (2011) for an overview and convergence results. For the situation where $\tilde{h}(\rho) - \rho$ is decreasing in ρ, SA starts from some ρ_0 and generates iterates of the form

$$\rho_{j+1} = \rho_j + a_j(\hat{h}_{n_j}(\rho_j) - \rho_j), \tag{21}$$

where $\hat{h}_{n_j}(\rho_j)$ is an estimate of $\tilde{h}(\rho_j)$ based on sample size n_j. These estimates are independent across values of j, and $\{a_j,\ j \geq 0\}$ is a slowly-decreasing sequence such that $\sum_{j=0}^{\infty} a_j = \infty$ and $\sum_{j=0}^{\infty} a_j^2 < \infty$. Note that there is no need to have $n_j \to \infty$; one can take n_j as a small constant independent of j. If we replace a_j by the inverse derivative $1/(\tilde{h}'(\rho_j) - 1)$ and the estimate of $\tilde{h}(\rho_j)$ by its exact value, we obtain the Newton method, which usually converges much faster, but requires knowledge of the function and of its derivative (or accurate estimators and $n_j \to \infty$), in contrast to SA. On the other hand, without a good choice of the a_j's, SA might converge extremely slowly.

If we replace a_j by 1 in (21), we obtain

$$\rho_{j+1} = \rho_j + (\hat{h}_{n_j}(\rho_j) - \rho_j) = \hat{h}_{n_j}(\rho_j). \tag{22}$$

If $n_j \to \infty$, this iteration becomes equivalent in the limit to the mapping $\rho \to \tilde{h}(\rho)$. 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 to hold is that $|\tilde{h}'(\rho)| \leq \gamma$ for all ρ (in the region of interest). When this holds, we can start from some $\rho_0 > 0$ and iterate: $\rho_{j+1} = \tilde{h}(\rho_j)$, 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 $j \to \infty$. On the other hand, if n_j does not increase with j, ρ_j will generally not converge to θ^*. If n_j is fixed to some large constant n and we use IRN, ρ_j will never converge but wander around in a small neighborhood of θ^*. If we use CRNs, it will converge to a value close to θ^*, but generally different.

It is very common in our model that $\rho \to \tilde{h}(\rho)$ is a contraction mapping. In particular, this holds in all the examples considered in this paper. Generating iterates of (22), we verified that in all cases it converged very quickly to a very good approximation of ρ^*.

To illustrate our discussion of stochastic root finding methods to estimate ρ^* in our setting, we use them to solve numerically the two examples given in Section 3.5.

Example 6 We revisit Example 4 and estimate $r(\rho)$, $g(\rho)$, and $\varphi(r(\rho), g(\rho))$ over a range of values of ρ. We take a sample of size $n = 10^7$, and ρ from 0 to 1 with a step size of 0.001. The plot on the left of Figure 4 shows the estimate of $\varphi(r(\rho), g(\rho))$ as a function of ρ, when using IRN. This functional estimator suffers from high-frequency noise. The true maximum is at $\rho^* = 0.41214955$ whereas the sample function (the estimate) has its maximum at $\rho = 0.437$. We see that, even with this large sample size, the noise is significant compared with the variation of $\varphi(r(\rho), g(\rho))$ around ρ^*. This illustrates the fact that sample-average optimization with IRN is a poor method to estimate the optimizer, because of the large noise.

We also applied the mapping (22) iteratively, starting at $\rho_0 = 0$, with a fixed sample size of $n_j = 10^7$ for all j, and IRN across iterations. This gave $\rho_1 = 0.44446$ at the first iteration and $\rho_4 = 0.4121$, already accurate to four digits after only four iterations. Thus, this method gets close to the optimum very quickly. For comparison, we ran the same method with CRN. The values of ρ_j with both methods are shown in Table 2, for $j = 1, \ldots, 6$. With both methods, ρ_j provides a good approximation of ρ^* very quickly. With CRN, it converges to 0.4121425 for $j \geq 7$, which is not the exact value but is accurate to five digits.

Table 2: Values of ρ_j at the first six iterations of (22) for Example 6, with IRN and CRN

Method	ρ_1	ρ_2	ρ_3	ρ_4	ρ_5	ρ_6
IRN	0.4444478	0.4102066	0.4122955	0.4121269	0.4121527	0.4121252
CRN	0.4444478	0.4101850	0.4122638	0.4121351	0.4121428	0.4121424

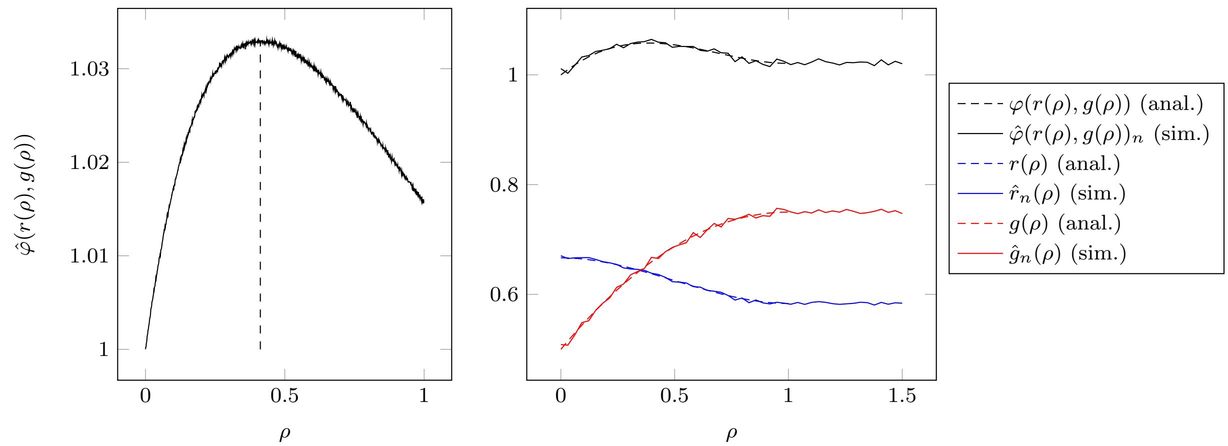


Figure 4: Estimate of expected SE revenue per unit time in terms of ρ for $\alpha = \beta = 1$ in Example 4 (left), and for $\beta = 1$ and $p = 1/2$ in Example 5 (right)

To prove that we indeed have a contraction mapping for this example, recall that $\lambda(r) = r$, $r = 2/3 - \rho^2/(6(1+\rho)^2)$, and $g = 2/3 - 1/(6(1+\rho)^2)$. This gives

$$\tilde{h}(\rho) = \frac{2/3 - \frac{\rho^2}{6(1+\rho)^2}}{\beta + 2/3 - \frac{1}{6(1+\rho)^2}} \qquad \text{and} \qquad \tilde{h}'(\rho) = -\frac{2(4\rho^2 + 6\beta\rho^2 + 6\beta\rho + 7\rho + 4)}{(6\beta + 12\beta\rho + 6\beta\rho^2 + 3 + 8\rho + 4\rho^2)^2}.$$

For $\beta = 1$, one can verify that $\tilde{h}'(\rho)$ is negative and increasing, with $|\tilde{h}'(\rho)| \leq |\tilde{h}'(0)| = 8/81 < 1$. Therefore the mapping $\rho \to \tilde{h}(\rho)$ is contracting with $\gamma = 8/81$. □

Example 7 We also revisit Example 5 and solve the problem numerically. The plot on the right of Figure 4 shows the estimates of $r(\rho)$, $g(\rho)$, and $\varphi(r(\rho), g(\rho))$, computed with IRN with $n = 10^5$. We also superimpose the corresponding exact curves. Again, we applied (22) for six iterations, starting with $\rho_0 = 0$, and a fixed sample size of $n_j = 10^7$ for all j. The results are in Table 3. We find that $\rho^* \approx 0.3859$.

Table 3: Values of ρ_j at the first six iterations of (22) for Example 7, 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

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, we have a contraction mapping with $\gamma \approx 0.15$ in that area. □

As illustrated by these examples, when (22) is a contraction mapping (which is typical), applying this mapping for just a few iterations provides a quick and efficient way of approximating ρ^*. The mapping is easy to apply when Y has a density and Assumption A holds. It converged very quickly in all the examples we tried. This is therefore the method we recommend to try first.

5 Comparing neutral and non-neutral ranking policies

In this section, we give further numerical examples to illustrate how the theory developed earlier could be used to study the impact of different ranking policies on various performance indicators such as consumer welfare (captured by expected relevance), SE, and CP revenue. In particular, we compare neutral ranking policies, where $\rho = 0$, with non-neutral ones, in which the SE chooses the optimal ρ^*. As in the previous examples, we take $\lambda(r) = r$, $\varphi(r, g) = \lambda(r)(\beta + g)$, and $\psi(R_i) = 1$.

5.1 A vertically integrated SE with a CP

Example 8 Here we focus on a specific type of request which can be served by either third-party CPs or by the SE itself. This is typical for many search categories where the SE also provides content (e.g., video, weather, finance, news, maps, flight information, and so on). In this case, a limited number of CPs compete with the SE, and 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. 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 4, 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 have no impact on the optimization.

Table 4: CTR values θ_j used in Example 8

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

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 5 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 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 $\theta_1 r(\infty) \approx 0.188$.

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 6(a) plots ρ^* as a function of β, while Figure 6(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 step. When β grows, ρ^* tends to zero, because the revenue from sponsored links dominates, making it more

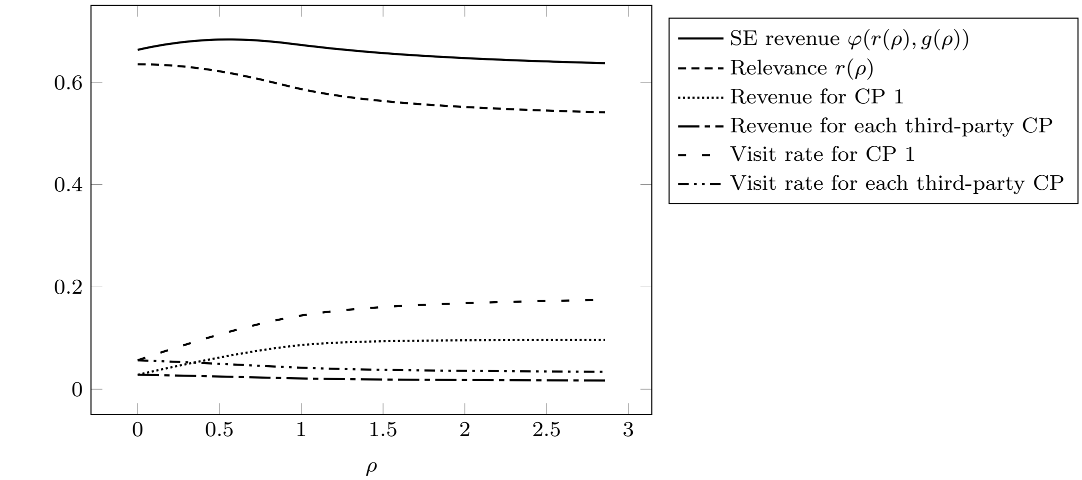
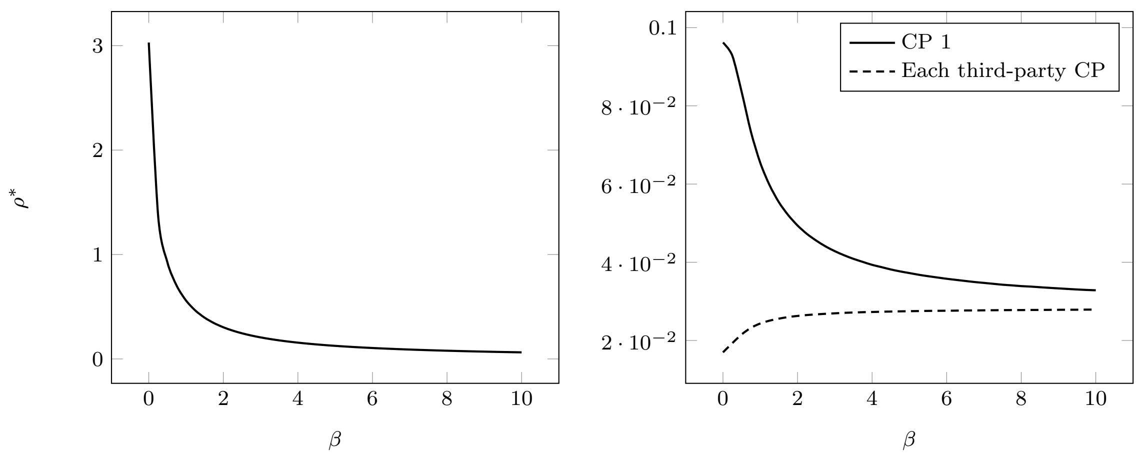


Figure 5: Performance measures as a function of ρ (simulation results)



(a) Optimal ρ

(b) CP revenues

Figure 6: Optimal ρ for the ranking, and corresponding CP revenue per unit of time, as functions of β (simulation results)

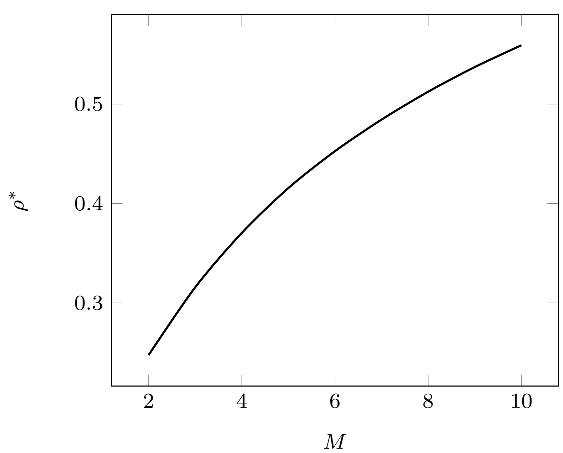
rewarding for the SE to improve its reputation to attract more users. The impact of non-neutrality is small in this example because biasing the ranking only attracts limited additional revenue. When β is small, in contrast, sponsored links do not pay much and it becomes worthwhile for the SE to sacrifice 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 $\varphi(r^*, g^*)$ grow 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 5 reports the variations of the most relevant performance metrics when using $\rho = \rho^*$ instead of $\rho = 0$ (neutral ranking), for different values of β. 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.

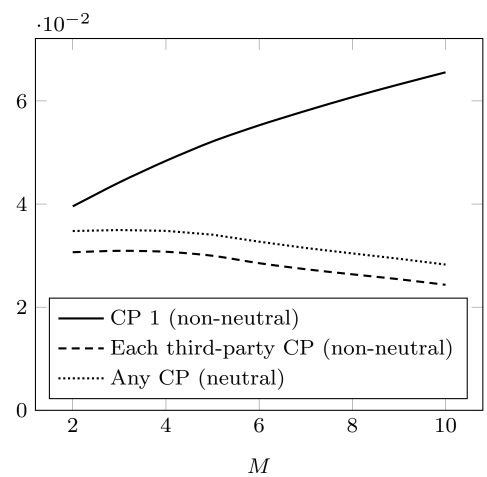
Table 5: Impacts of a non-neutral ranking for the scenario of Section 5.1

	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%)

Finally, we explore the sensitivity of outcomes to the number M of matching pages. Figure 7 plots ρ^* and the CP revenues as functions of M, for both the neutral ($\rho = 0$) and non-neutral ($\rho = \rho^*$) regimes. As before, ρ^* was estimated via the fixed-point algorithm with $n = 10^7$ at each step. As M increases, ρ^* increases: The SE can give more weight to CP 1 and increase its revenue while making less damage to the relevance, because placing CP 1 higher has less impact on the overall relevance when M is larger. As a result, the revenue of CP 1 when $\rho = \rho^*$ increases with M, and so does the advantage of CP 1 over the other CPs. The loss of revenue for the other CPs becomes almost constant as a function of M when M is large.



(a) Optimal ρ



(b) CP revenues

Figure 7: Optimal ρ for the ranking, and corresponding CP revenues per unit of time, as functions of M (simulation results)

5.2 Vertical integration and investment

Example 9 Continuing with the previous example of vertical integration, we now assume that one of the nine third-party CPs, say CP 2, invests in quality and manages to improve its relevance distribution. More specifically, we assume that when it invests $z > 0$, the relevance of CP 2 becomes uniformly distributed over $[0, 1 + 20z]$ (instead of over $[0, 1]$). The other parameters and distributions, including the distribution of G_2, are unchanged. Figures 8(a) and 8(b) show simulation results when the SE ranks CPs according to $\tilde{R}_i + \rho\tilde{G}_i$, for varying values of ρ, and when $z = 2$. For a neutral ranking ($\rho = 0$), CP 2 logically makes more revenue than the other CPs, since it regularly gets higher ranking. However, when ρ increases and exceeds about 0.8, CP 1 becomes the one with highest revenue, despite its (stochastically) lower relevance.

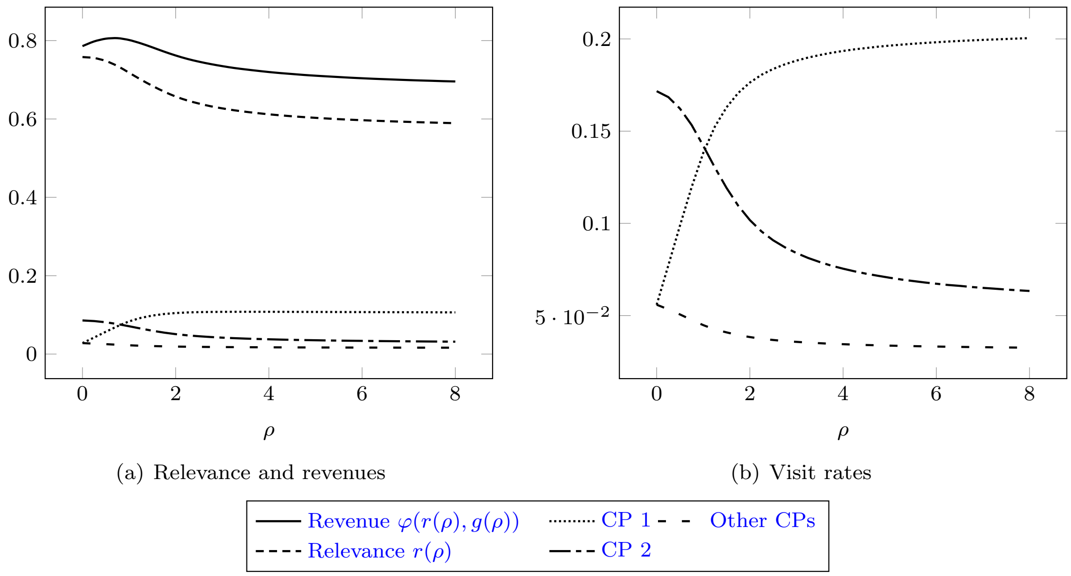


Figure 8: Average relevance, revenues and number of visits per unit of time for the case of vertical integration with investment

We now take the perspective of CP 2, and compute its optimal decision. CP 2 invests z in quality to modify its relevance distribution to $[0, 1 + 20z]$, anticipating that the SE is going to rank requests according to ρ^*. To simplify the example, we assume that an estimate of the distribution of Y is always immediately available to the SE (in real life there will be a delay to update the estimate when the distribution changes, which is fine if the distribution changes only slowly). The profit of CP 2 is the revenue from the search market, minus z. To optimize z, we simulated the outcomes for $z \in [0, 0.45]$. Figures 9(a) and 9(b) plot the resulting curves. In both figures, we find that the difference between neutral and non-neutral revenues is small, except for CP 1. This is particularly true for CP 2. Thus, at least in this example, non-neutrality does not deter innovation. Actually, the optimal investment level z under both regimes coincide and is equal to

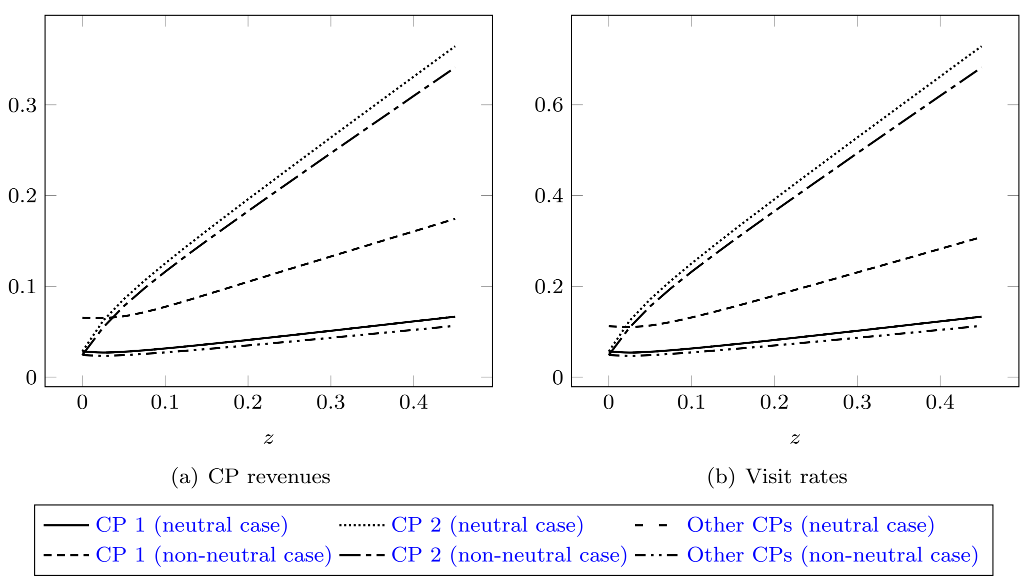


Figure 9: Revenues and visit rates to various CPs as a function of the investment from CP 2

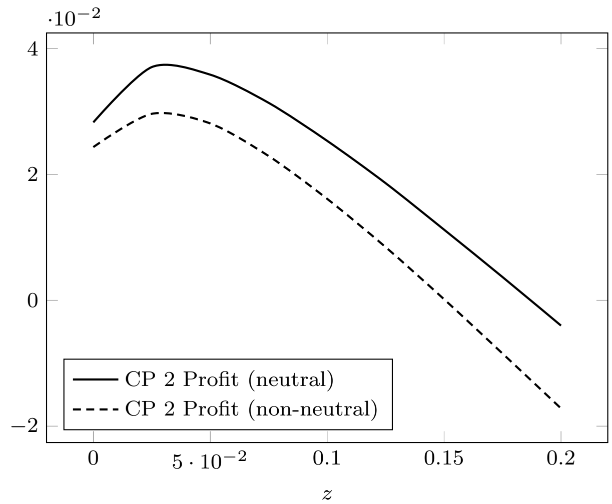


Figure 10: Profit per unit of time as a function of CP 2 investment

$z^* = 0.025$. Optimal profits, though, vary. They are 0.037 for the neutral case and 0.0296 for the non-neutral one; see Figure 10, which displays the profits of CP 2 as a function of z.

6 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 clicked by the users. We proved that 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 figure of merit that is a linear combination between relevance and short-term profits. This sorting depends only on the slope of the linear combination, a single real number. Under further assumptions, with probability 1, this provides a unique ordering for each request. Then, the whole problem reduces to finding the appropriate constant used in the linear combination. We provided algorithms to find this constant.

Our model and results might prove useful to platform owners (search engines, 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, seeking to 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.

Topics that would deserve further investigation include: (i) perform case studies with real data; (ii) based on real data, study the implications of profit-maximizing platforms on the online economy to shed light on the search-neutrality debate.

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.

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. 2005. Dynamic Programming and Optimal Control, Volume I. 3rd ed. Athena Scientific, Belmont, Mass.

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. 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. To appear.

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

Odlyzko, A. 2009. Network neutrality, search neutrality, and the never-ending conflict between efficiency and fairness in markets. Review of Network Economics 8(1) 40–60.

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.

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.