

Modeling the Internet Search Market: a Step in Bringing the User into the Picture

Natalia Kudryashova*

Cambridge Centre for Analysis, University of Cambridge, Wilberforce Road, Cambridge, CB3 0WA, United Kingdom.

Received: 4 Jul. 2014, Revised: 5 Oct. 2014, Accepted: 6 Oct. 2014

Published online: 1 Apr. 2015

Abstract: Success in the internet search market is bound to the popularity of a search platform with the search users. In the absence of switching costs, the users can easily switch between the search platforms, which for them are literally just a click away. Users' switching behaviour reflects changes in their perception of the relative values the competing offerings. The user value may be influenced by strategic decisions of the market players and, in turn, can significantly modify intended effects of the platform's strategies. Users' preferences are not necessarily rational, and their dynamics is affected by various factors, such as habitual attraction. We propose to place strategic optimisation in the internet search market into a framework accounting for a switching feedback of learning users. We demonstrate importance of accounting for habitual attraction for understanding the market dynamics and discuss applications of the proposed approach in economics and law.

Keywords: Internet Search Market, Modeling, User Behaviour

1 Introduction

The market success of both publishers and advertisers in the internet search market is ultimately determined by their popularity with the users. The users' "feedback" to the strategic decisions of the market players has the potential to significantly modify the market landscape. However, despite the clear importance of users' dynamics, it has not received due attention in the context of strategic optimisation and decision making in the market.

The internet search market also became a subject of interest in a range of other disciplines. There are various studies focusing on strategies to optimize revenue of a publisher (i.e. a search platform) or optimal bidding strategies of advertisers competing in online auctions of sponsored links ([15],[7],[4], [9], [8], [1] and [3]).

Considerable attention has been devoted to the discussion of mechanisms used by search platforms to allocate sponsored links, which are most often modelled using game theoretical considerations([15],[7],[4] and [5]). The discussion has also been placed in the framework of dynamic stochastic games([9] [8] and [1]).

Recently, the internet search market came into the focus of public attention and scholarly work in law and

economics in conjunction with the legal proceedings against Google, the most popular general purpose search engine in the EU and the USA¹. There is little agreement regarding the merits of the case, and orthogonal views have been expressed on whether Google abuses its market position to the detriment of competition, innovations and users.

To enhance legal argumentation, there have been some attempts to place the legal discussion in the context of economics, and several modelling attempts were undertaken. These attempts are based on the game theoretical approach, in which market players simultaneously optimise their control variables to maximize their utility functions. There is however no agreement regarding the assumptions adopted in different papers, which also do not seem to reflect the business model and mechanisms operating in the market ([2],[6],[12]). Unsurprisingly, these models fail to capture realities of the market and its dynamics and appear not to be applicable to the legal points they are supposed to illustrate. Also, competition between the search platforms and switching dynamics of the end users between the

¹ Market share reports: <http://marketshare.hitslink.com/>, comScore, Inc. <http://www.comscore.com>, StatCounter, <http://gs.statcounter.com/>

* Corresponding author e-mail: nk375@cam.ac.uk

platforms do not seem to have received due attention, despite the apparent need to capture the complex and rapidly changing market environment and users' response [14] and despite the fact that interaction between the users and the search platform plays a crucial role in its competitiveness.

In this paper we describe a user centred modelling approach, which relies on minimal assumptions, takes into account specifics of the market and allows capturing the overall picture. Unlike many of the previous modelling attempts, our approach is dynamic and is based on explicit treatment of users' switching in the environment set by the platform's strategic decisions. The proposed approach reflects the role of the users in the internet search market and the strong asymmetry between the users and advertisers. While the users are not paying customers and the main source of the platforms revenue comes from the advertisers, it is collective choices of the users that ultimately determine the platform's success. Having a conflicting motivation, the platform therefore has to balance optimization of the user value and the platform's revenue, thereby affecting user value and, hence, users' preferences and dynamics between the platforms.

The model focuses on the net effects of the platform's decisions, though leaving space for more detailisation, and aims to help establishing a "user-centred" picture, which is required in the context of anti-trust proceedings. It can also aid in elucidating effects of the natural to the market factors and to distinguish them from the consequences of the market players' strategic decisions that can be deemed anti-competitive, and can be employed to illustrate legal argumentation used, for instance, in the context of Google's cases, thereby assisting in probing its validity.

2 The Model In A Nutshell

We consider a large ensemble of users (indexed by $i = 1, \dots, I$), searching the web according to a Poisson process $\{N_i(t), t \geq 0\}$ of a constant intensity λ_i . The random variables $\{\lambda_i\}_{i=1, \dots, I}$ are assumed to be identically distributed with some distribution function f_λ .

At Poisson times $t_{ik}, k = 1, \dots, N_i(T)$ user i chooses a search platform (indexed by $m = 1, \dots, M$), performs a search, and evaluates the received customer value on the scale 0 – 1 with 0 corresponding to full dissatisfaction and 1 corresponding to full satisfaction.

These evaluations together with expected valuations of the other platforms form a set α_{ik}^m which will be further referred to as user's satisfaction rates.

The variables $\{\alpha_{ik}^m\}_{i=1, \dots, I}$ for each feasible m are assumed to be distributed with distribution functions denoted as f_k^m independently from the frequency of the search λ_i . The distribution functions f_k^m are specific to the platforms and evolve in time in response to the platforms decisions that affect the user value. User's preferences for the search

platforms are described by a set of probabilities $\{p_{ik}^m\}_{m=1, \dots, M}$ of choosing the platforms at times t_{ik} . The probability that user i chooses platform m at times t_{ik+1} is a sum of the probability that he or she stays with the same platform as at times t_{ik} and the probability that he or she switches to it:

$$p_{ik+1}^m = Pr(stay)p_{ik}^m + Pr(switch)(1 - p_{ik}^m). \quad (1)$$

If the users are fully informed and rational they choose with probability 1 the platform with the highest perceived satisfaction rate, otherwise the probabilities to stay and to switch in eq. (1) depend on various factors that govern users choices, including (perceived) quality of the offerings, evaluation of the (recent) experience, and habitual attraction to a particular platform.

If decision making is assumed to be a Markov process, i.e. based only on the recent valuation of the offerings, the market is free from users' bias. In the absence of switching costs the user stays with the platform when he or she is satisfied with the previous experience and switches to another platform when dissatisfied, i.e. the competition in the biasfree market is literally a click away.

However, in a mature market the user's decision is influenced by the previous search history: the user is not unbiased and tends to be attracted to a particular platform that he or she uses habitually. Then, even when dissatisfied, the user may stay with the platform with some probability that characterizes the strength of habitual attraction.

Then, eq. (1) can be written in the following form:

$$p_{ik+1}^m = (1 - \delta_{ik} \beta_{ik}^m \gamma_{ik}^m) p_{ik}^m + \sum_{l=1}^M p_{ik}^l \beta_{ik}^l T_{ik}^{ml} \gamma_{ik}^m (1 - \delta_{lm}), \quad (2)$$

where $\sum_{m=1}^M p_{ik}^m = 1, m = 1, \dots, M, i = 1, \dots, I, k = 1, \dots, N_i(t)$, β_{ik}^m denotes probability that at time t_{ik} user i is dissatisfied with the search with platform m to the extent that warrants switching to another platform; and γ_{ik}^m is the probability that the user still stays with platform m in case he or she is dissatisfied. Being a measure of habitual attraction on the scale 0-1, the parameters γ_{ik}^m are assumed to correspond to the average probability that the platform was used previously and not to depend on α_{ik}^m and λ_i . The set $\{T_{ik}^{ml}\}_{m, l=1, \dots, M, m \neq l}$ denote matrix elements of the transition matrix that governs users preferences upon switching. To ensure the detailed balance:

$$\sum_{l=1}^M T_{ik}^{ml} = 1, \quad (3)$$

for any feasible i, k and m . The parameters β_{ik}^m may be viewed as a measure of lack of quality or relevance of the search results, i.e. of the user's dissatisfaction with the platform, and therefore are equal to $1 - \alpha_{ik}^m$. The elements of the transition matrix $\{T_{ik}^{ml}\}_{m, l=1, \dots, M, m \neq l}$ depend on the

dissatisfaction parameters and may also depend on the parameters describing habitual attraction and other factors governing users decision making. For instance, a rational user always goes for the most satisfactory offering, while a biasfree user, when dissatisfied, chooses the least dissatisfactory alternative. In the limit of a very large ensemble of users ($I \rightarrow \infty$) no individual choice, time or frequency of search or valuation matters and therefore affects the average preferences. The time evolution of the averaged user preferences $p^{(m)}(t)$ is then a continuous process described by a Kolmogorov equation with the expected values for the corresponding parameters in eq. 2. Then, knowledge of the distribution functions is no longer required, while the expected values can be obtained, for instance, from a dedicated empirical study. From the evolution of averaged users' preferences, the dynamics of the market distribution between the platforms can be straightforwardly obtained.

Since the expected values of the dissatisfaction rates have to refer to the quality (generally perceived quality) of the offerings, they are susceptible to strategic decision of the platforms that affect the users value. Then, equations describing dynamics of the users can be combined with specific optimization problems of the market players, to provide a dynamic feedback through the change in user preferences and the market share.

To give an example how the modelling settings can be further detailed, we can distinguish between dynamics of user preferences for organic and sponsored search (i.e. $p^{(m)} \rightarrow p_{org}^{(m)}, p_{spons}^{(m)}$). It is easy to see that in this case the dissatisfaction rates $\beta^{(m)}(t)$ are the product of dissatisfaction with the organic and the sponsored search: $\beta^{(m)}(t) \rightarrow \beta_{org}^{(m)}(t), \beta_{spons}^{(m)}(t)$. We can introduce a learning environment (Figure 1), for instance, by assuming a dynamic relation between the probability of using the sponsored search when searching the web and the (perceived) quality of the sponsored search (further referred to as a learning equation):

$$\frac{dp_{spons}^{(m)}(t)}{dt} = -\frac{\beta_{spons}^{(m)}(t)}{dt} \tag{4}$$

Furthermore, the modelling settings can be tuned to various circumstances, for instance to those which are of interest in the legal context. For instance, dynamics of innovative start-ups can be inferred from dynamics of user preferences/ market share for the new entrant with considerably smaller dissatisfaction rates, while expansion in the content segment can be described by coupling dynamics of the two markets.

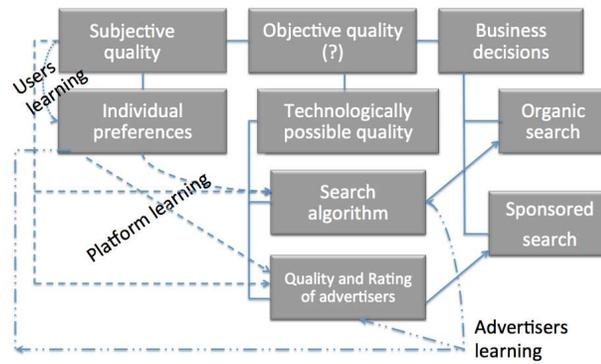


Fig. 1: Learning environment for the market participants

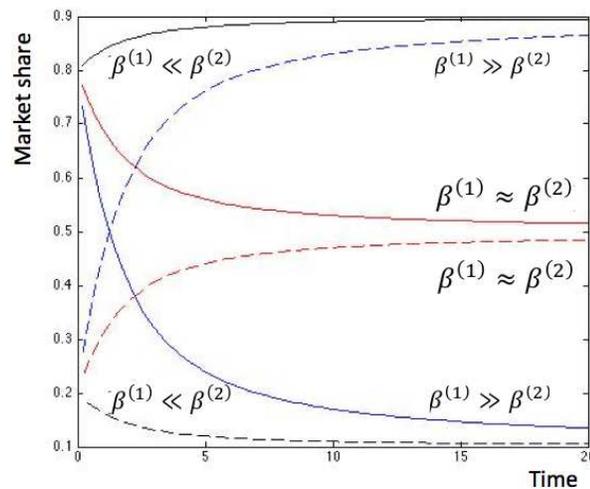


Fig. 2: Dynamics in the biasfree market

3 Results and Discussion

3.1 Verifying the model

In the first step we show that the model of the biasfree market is not applicable to the market of keywords search in the EU and the USA in which Google competes with Bing and Yahoo! ², while the model of market with habitual attraction reflects well the distribution of the real market and its dynamics.

Consider the market consisting of two search platforms. The stationary distribution of the market share ($n^{(1)}$ and $n^{(2)}$, $n^{(1)} + n^{(2)} = 1$) in the biasfreemodel

² who are part of the strategic partnership

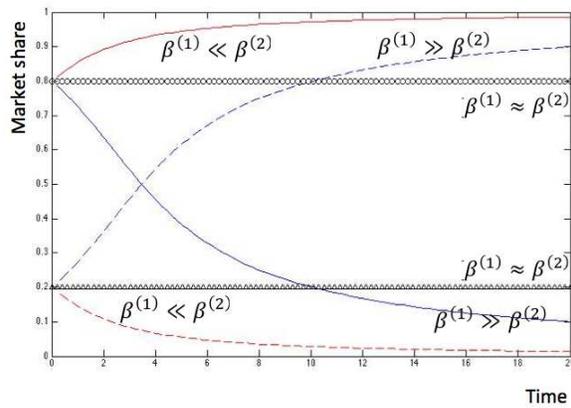


Fig. 3: Dynamics in the market with habitual attraction

(Figure 3) is given by the inverse ratio of the dissatisfaction parameters of the competing offerings:

$$\frac{n^{(1)}}{n^{(2)}} = \frac{\beta^{(2)}}{\beta^{(1)}}. \quad (5)$$

It does not depend on the initial distribution of the market and is fair in a sense of classical understanding of fair competition based on the users values of the offerings.

Then, the market share for the offerings with similar users value should be also similar, whereas the market dominance should reflect substantial superiority of one of the offerings. The real market, however, is clearly dominated by the Google's product despite it is not radically better than that of the competitor. Hence, the model of the biasfree market (in which the competition is literally a click away) is not applicable. This may not be surprising, but does not mean that Google behaves anti-competitively, since in a mature market, which the current internet search market of keywords is, the notion of fair competition is considerably modified by habitual attraction, as we will see below.

In the presence of habitual attraction, the stationary solutions to the Kolmogorov equation correspond to either the monopoly of one of the players or indefinite propagation of the initial preference/ market distribution in case the offerings are of the same user value (ure 3); the latter seems to reflect well the current market situation. Initial distribution can result, for instance, from the effect of the "first move" upon introduction of the innovation to the market, as it seems to be the case for Google.

The model of fully rational users results in the monopoly of the platform offering the best product and implies fast reaction by the users, which does not correspond to the market situation and will not be discussed further.

Comparison with the biasfree model shows that habitual attraction considerably slows down the market dynamics. Furthermore, in the biasfree market the condition for gaining the market share (and market dominance), say by platform 1, corresponds to minimization of the perceived user dissatisfaction rates, $\beta^{(1)} \ll \beta^{(2)}$, thereby providing incentives for radical innovations. For comparison, in the model with habitual attraction the condition for market dominance corresponds to a marginally smaller dissatisfaction with the product than with the alternative offering: $\beta^{(1)} < \beta^{(2)}$. Notably, Microsoft is investing considerable efforts in improving (perceived) user value of its product Bing through an extensive marketing and advertising campaign, which however largely targets habitual attraction to Google's search instead. Modification of the model by temporarily reducing the parameter of habitual attraction to the dominant offering suggests that due to the strong dominance of Google and slow market dynamics, the effect of the campaign, which has to be sustained over prolonged time, may not be significant.

3.2 Selected applications in the context of market competition

It is clear from the above that habitual attraction modifies the notion of fair competition (as well as its manifestations in the dynamics of the market), so that strong dominance of the product of a similar quality over competing offerings may be a natural consequence of habitual attraction and initial distribution of the market, and therefore is not necessarily unfair. Also, explicit consideration of the user dynamics in response to introduction of an innovative product (with lower dissatisfaction rate and no habitual attraction to it) shows that innovative start-ups are facing a much tougher competition to gain market share than the incumbent players and have to offer a radically better product (Figure 4). Moreover, slow market dynamics increases the probability that innovative start-ups may have to leave the market before having realized full potential of their innovations. Furthermore, the calculations show that the incumbent players appeared to have a chance of getting full credit for the radical innovation (i.e. increase their market share) in case their responses are sufficiently prompt. The above is in line with real market practices when innovative start-ups found it optimal not to develop their innovation to its full market potential, but to sell it to an incumbent player.

Similar considerations are applicable to the expansion of a search platform's operations into the content market, for instance, of internet street maps or vertical search.

In this case the Kolmogorov equations for the search market and the content market are coupled. During the introductory period of the platform's offering users of the content market may be attracted to it as a brand and, in

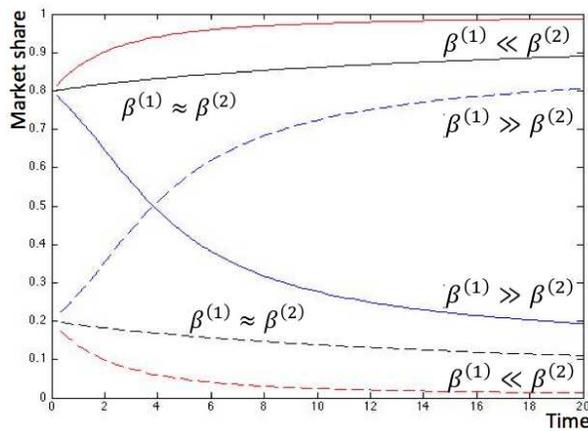


Fig. 4: New entrant in the market with habitual attraction

addition, temporarily overvalue it. Then, due to habitual attraction to the brand, the market share of the platform’s offering in the content market is expected to be substantially higher than in the biasfree market (which corresponds to the perceived fair value) or in the market with habitual attraction for a new entrant. However, if the platform’s product is inferior to that of the incumbent players its market success in the content market is expected to be short lived.

To give an example of the application in the legal context, the legal arguments in the case Streetmap EU Ltd. v Google Inc. ³ are based on the fact that the Google’s share in the market of internet street maps is higher than the perceived fair value in the context of classical fair competition. However, our modelling leads to the conclusion that it may not be disproportionately high, because of Google’s dominance in the internet search market and resulting strong habitual attraction to Google as a brand. Consequently, legal arguments built upon comparison of the user value of the competing products and the corresponding market share cannot be necessarily considered as supporting allegation of anti-competitive behaviour.

In the context of the legal proceedings ⁴, we can also investigate effect of the proposed public regulation of the dominant search platform [11], which in effect would constrain the user value for the platform’s offering. Accounting for the depreciation of the product and recalling that the other player in the market with habitual attraction has incentive to offer only a marginally better product, it is straightforward to deduce that the proposed

public regulation could in the long run harm market dynamics, innovations and the users.

3.3 Selected applications in economics

Explicit treatment of user dynamics helps to remove a number of controversies associated with admittedly counter-intuitive conclusions resulting from a static approach, as for instance in [13]

The first conclusion of [13] is that a search platform has incentive to deliberately deteriorate the quality of the organic search results.

By adopting a dynamic approach it is, however, straightforward to show that the above strategy may not be compatible in the long-run with sustainable growth and profitability.

This can be seen from the condition for the platform’s marginal revenue: $\pi^{(m)} = p^{(m)}(\langle u \rangle p_{spons}^{(m)} - k)$ (where $\langle u \rangle$ denotes the average revenue from an auction of sponsored links, $\langle u \rangle$ and k are fixed costs) not to decrease with time ($\frac{d\pi^{(m)}}{dt} \geq 0$), the Kolmogorov equation for the $p^{(m)}$ and the learning equation for $p_{spons}^{(m)}$.

The second conclusion of [13], stating that large switching costs for the users would stimulate improvement of organic search quality, can be disproved by observing that large switching costs entrench habitual attraction and recalling that the condition for the market dominance (for the platform 1) is $\beta^{(1)} < \beta^{(2)}$

Since $\beta^{(1)} = \beta_{org}^{(1)} \beta_{spons}^{(1)}$ and $\beta_{spons}^{(1)}$ is to be decreased by the learning equation, there is some room to increase $\beta_{org}^{(1)}$, i.e. to decrease quality of the organic search rather to improve it.

4 Conclusions

We propose a user centered dynamic approach to modelling the internet search market. We show, using a family of simplified models, how the approach can be used to illustrate and critically assess some of the arguments arising in conjunction with legal proceedings against Google. We demonstrate the importance of accounting for habitual attraction, which modifies the notion of fair competition from the general perception of fairness in the context of the biasfree market, and discuss some consequences for the market dynamics and propagation of innovations. Furthermore, we demonstrate that the proposed approach is helpful in removing controversies arising as a result of static modelling.

References

[1] K. M. K. P. Amin, K. and A. Schwaighofer. Budget optimization for sponsored search:censored learning in mdps. *Uncertainty in Artificial Intelligence (UAI)*, 2012.

³ Streetmap EU Ltd. v. Google Inc., case no 13-1013, High Court of Justice,Chancery Division

⁴ AT39740 Google, http://europa.eu/rapid/press-release-SPEECH-12-372_en.htm

- [2] C. Argenton and J. Prufer. Search engine competition with network externalities. *Journal of Competition Law & Economics*, 8(1):73–105, 2012.
- [3] S. Athey and G. Ellison. Position auctions with consumer search. *Quarterly Journal of Economics*, 126(3):1213–1270, 2011.
- [4] S. Athey and D. Nekipelov. A structural model of sponsored search advertising auctions. *In Sixth Ad Auctions Workshop*, 2010.
- [5] Y. Chen and C. He. Paid placement: Advertising and search on the internet. *Economic Journal*, 121(556):F309–F328, 2011.
- [6] K. L. Devine. Preserving competition in multi-sided innovative markets: How do you solve a problem like google? 2008.
- [7] M. O. Edelman, B. and M. Schwarz. Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. *American Economic Review*, 97(1):242–259, 2007.
- [8] K. P. Gummadi, R. and A. Proutiere. Optimal bidding strategies and equilibria in dynamic auctions with budget constraints. 2012.
- [9] J. R. Iyer, K. and M. Sundararajan. Mean field equilibria of dynamic auctions with learning. *In proc. of ACM Conference on Electronic Commerce*, 2011.
- [10] N. Kudryashova. The market of internet sponsored links in the context of competition law: Can modeling help? *Proceedings of the Companion Publication of the 23rd International Conference on World Wide Web Companion*, pages 331–332, 2014.
- [11] I. Lianos and E. Motchenkova. Market dominance and search quality in the search engine market. *Journal of Competition Law and Economics*, 9(2):419–455, 2011.
- [12] G. Luchetta. Is the google platform a two-sided market? *Journal of Competition Law & Economics*, doi: 10.1093/joclec/nht026, 2013.
- [13] G. Taylor. Search quality and revenue cannibalization by competing search engines. *Journal of Economics & Management Strategy*, (22):445–467, 2013.
- [14] E. van Couvering. The history of the internet search engine: Navigational media and the traffic commodity. *Web Search, Information Science and Knowledge Management*, 14:177–206, 2008.
- [15] H. R. Varian. Position auctions. *International Journal of Industrial Organization*, 25(6):1163–1178, 2007.



Natalia Kudryashova has a multidisciplinary background combining degrees in Mathematics, Business Administration and International Business Law. She is particularly interested in expanding application of mathematical modelling in the context of e-commerce, patent and ant-trust law, combining knowledge from seemingly different subject areas to build a "bigger" picture.