In Google we trust?

Roberto Burguet*, Ramon Caminal*, Matthew Ellman*

Preliminary Draft
(not for distribution)

June 2013

Abstract

This paper develops a unified and microfounded model of search and display advertising where consumers search for offline products as well as online content published on platforms that display ads. We identify novel incentives for a monopoly search engine to distort both its organic and its sponsored search results. First, distorting organic results towards publisher websites with less effective display advertising raises merchants' willingness to pay for sponsored search. Second, underweighting consumer relevance in sponsored search auctions cashes in on higher margin merchants. The interplay between these two incentives and the need to attract search participants determines search bias and welfare. We also characterize how the welfare consequences of search engine integration into ad intermediation and publishing depend on publisher asymmetry, the degrees of downstream monopolization and display ad targeting.

Keywords: Search engine bias, internet economics, vertical integration, two-sided markets, antitrust. *JEL Classifications*: L13, L41, L44, L86.

*Institut d'Anàlisi Econòmica CSIC, and Barcelona GSE.
Emails: roberto.burguet@iae.csic.es, ramon.caminal@iae.csic.es, matthew.ellman@iae.csic.es

1 Introduction

With market shares exceeding 90% in most European countries and a global average above 80%, Google dominates online search in most of the world (State-of-search, 2012). This may reflect widespread trust in Google's motives and ability to deliver reliable search results, but a growing chorus of voices argues against trusting in Google's "do no evil" promise to deliver unbiased results. These critics point to evidence of specific search biases that raise the ranking of Google's own content and services (see e.g., Edelman and Lockwood, 2011, Edelman and Lai, 2013, and www.FairSearch.org). Search biases, especially between websites not owned by Google, are hard to detect, but analyzing Google's incentives shows where to look for bias and a unified model can evaluate the welfare implications of Google's expansion into display advertising as ad intermediary (AdWords-AdSense) and publisher (e.g., Google Finance, YouTube and Zagat).

In this paper, we develop a micro-founded economic model that integrates a number of issues so far analyzed only individually. Our analysis features consumers seeking content and products, merchants selling offline products, publishers offering online content (or services), advertising intermediaries helping merchants advertise on publisher websites, and a monopoly search engine which we call Google that directs consumers to merchants and publishers via its search rankings.[1] We model both organic search (where rankings are not paid for) and sponsored search (where a position auction determines the ranking of links) and we characterize Google's incentives to distort each type of search result. We derive two types of bias - one affecting organic search and one affecting sponsored search. These biases operate independently but also interact. In addition, we investigate the bias and welfare consequences of Google's expansions into publishing or ad intermediation, which we represent as integrations with both vertical and horizontal features: vertical in that display advertising involves Google's organic search sending visitors to publishers that display ads; horizontal in that Google's sponsored search advertising is a substitute for display advertising.

The quality of search affects welfare by (a) matching consumers with products they may wish to buy and (b) matching consumers with content they may wish to consume online, which (c) determines the effectiveness of display advertising and (d) influences surplus appropriation and the investment incentives of the five different groups of actors. We distinguish between these consumer goods because display advertising is readily tied to online content but not to offline products: publishers can display third-party ads to visitors while they

[1] Alternative search engines may be "just one click away" as Google has argued, but indirect network externalities, combined with small switching costs, habit effects and delays in identifying search quality reductions, interfere with competitive forces. For example, Argenton and Prufer (2012) develop a model in which a search engine with an initial advantage can monopolize the market by learning from past searcher experiences.

consume content on the publisher's website. So publishers operate in a two-sided market. On the advertising side, their display ads compete with Google's (sponsored) search ads for merchant demand.[2] On the other side, consumers seeking content may need to use Google's search engine which then intermediates as an upstream gatekeeper. The first point gives Google an incentive to interfere with display advertising; the second point may give Google the power to do so. Together, they consitute the first source of potential organic search distortions.

We make a parallel distinction of consumer searches by objective: "content searches" for online goods and "product searches" for offline goods. In principle, consumers could reach both publishers' content and merchants' products via either sponsored or organic search results, but in the equilibrium of our model consumers only use organic results when seeking content and only use sponsored results when seeking products. This simplifying split reflects two ideas.[3] First, for product searches, Google has clear incentives to distort its organic results so that consumers only use the sponsored results, thereby obliging merchants to sponsor links to get traffic. Second, as a rule, publishers do not pay for sponsored links, so consumers use organic results exclusively when conducting content searches.[4] The split captures, in a stylized fashion, the findings of Greenspan (2004) and Jansen (2007) that people use sponsored links more than organic results when conducting product searches or "e-commerce search queries", but otherwise place more trust in organic results (see Hotchkiss et al., 2005). Thanks to this split, we only need to keep track of two types of bias: bias in organic results for content searches and bias in sponsored results for product searches.[5]

The first potential source of bias for these organic (content search) results derives from the substitutability (for merchants) between display and search advertising. Display is a stronger substitute for search advertising when content consumers are more attentive, receptive and responsive to ads. The effectiveness of display advertising depends, indirectly, on organic search distortions. First, distorting organic search towards publishers that are less effective for advertising (or, like Wikipedia and the BBC, decide against displaying ads) reduces the average effectiveness of display ads. Second, even with symmetric publishers, distorting

[2]We characterize this substitutability of display and search advertising which lies at the heart of recent regulatory controversy (FTC statement, 071-0170 versus EU report, M.4731).

[3]It makes the model highly tractable. We readily account for a less extreme split in section ??.

[4]Publishers have little appetite for sponsoring links for a number of reasons. Publishers mostly sell informational goods where reputation and credibility are key, so consumers prefer to rely on independent recommendations (such as organic results); consumers are suspicious of publishers that need or choose to pay to reach them. Also sponsoring links draws attention to publishers' advertising interest which may not be otherwise salient. Another important reason is practicality: publishers may not be able to anticipate which of the myriad of potential content queries will indicate when their content provides a good match; publishers of dynamic content, such as news sites, are particularly afflicted.

[5]Throughout the paper, we define bias relative to what consumers are searching for, though we take account of producer surplus in our welfare analyses.

organic search worsens the matching between consumers and publishers, which impoverishes display advertising because consumers are less attentive to less relevant publishers.[6] The second potential source of organic bias is that when Google integrates into ad intermediation or publishing, Google internalizes the ad revenues of its own publishers and its affiliates (those publishers that pay Google for ad intermediation). This is the most heavily discussed bias in regulatory circles.

The source of incentives to bias sponsored search resides in a potential conflict of interest between consumers and merchants over ranking sponsored links: the merchants most willing to pay for a top position may not be the best option for consumers since net margins and net consumer values need not be perfectly aligned. Google's optimal scoring auction for sponsored links does factor in "quality" by discounting "less relevant" merchants whose products provide lower net consumer value, but Google internalizes a share of merchant profits so it may be tempted to underweight relevance so that less relevant merchants with higher margins can sometimes win.[7]

These temptations to bias organic search and sponsored search are naturally tempered by Google's need to attract consumers to search via Google. As a result, a profit-maximizing search engine may distort search results relative to the allocation rules that maximize consumer surplus and these distortion incentives interact. In particular, the distortions are imperfect substitutes for Google: more bias in one type of search lowers the incentive to distort the other. Display advertising is very sensitive to the ability of publishers or ad intermediaries to target consumers. Technological innovation in this field is constant and may play an important role in determining the importance of the vertical and horizontal issues that we discuss here. We model targeting as the ability of publishers or ad intermediaries to identify consumer perferences, and find that, indeed, improved targeting exacerbates the incentives for traffic distortion.

Integration into ad intermediation changes Google's distortion incentives. Notice first that Google's control over organic search gives Google the gatekeeper power to restrict the flow of traffic onto any publisher website that does not become a Google advertising affiliate. We present three main findings on integration: (1) monopolization of the intermediation market improves the reliability of Google's search results, because Google internalizes profits from display advertising; (2) on the other hand, vertical integration without full ad intermediary

[6] This claim underlies Google AdSense's basic advice to publishers that they can raise their display ad revenues by attracting relevant consumers as well as by publishing interesting content. We refer to Ellman and Germano (2009) and Wilbur (2008) for further evidence and richer views on the relationship between content and advertising effectiveness.

[7] So, the model only partially supports Google's claims about its sponsored search position auctions; see Google's chief economist, Varian (2007), for a description of how the scoring auction should weight each merchant's bid by merchant quality and relevance to the searching consumer.

monopolization biases Google's organic search to favor publishers that deal with Google as ad intermediary against those that do not. Concretely, we identify conditions under which non-integration generates higher total welfare and higher consumer surplus than does vertical integration with partial monopolization. In addition, we show that (3) even with full monopolization, integration can have negative consequences for organic search and for total surplus when publishers are asymmetric in their effectiveness as platforms for display advertising.[8]

Advertising revenues are fundamental to the business models of most web-based publishers (including the much beleaguered news media). So we characterize the sharing of surplus among all the actors involved in producing web-mediated "trades". Our main finding, that vertical integration may reduce the share of advertising surplus that publishers can appropriate, presents a serious concern: in an extended model with investments, this would discourage publishers (such as news media) from investing in quality or creating new products.

Internet trade and search engine incentives are active research fields in economics. Our unified model covers both organic and sponsored search and explicit merchant competition for advertising opportunities. Early work in the literature studied sponsored search alone: equilibrium bidding by merchants in position auctions (see e.g., Lehaie, 2006, Edelman and Ostrovsky, 2007, 2010, Edelman, Ostrovsky and Schwarz, 2007, Varian, 2007, Borgers et al., 2007) and auction design (see e.g., Liu, Chen and Whinston (2010) and Athey and Ellison (2011) who model the design of scoring auctions with reserve values and the extraction of rent from merchants). Chen and He (2011) and Athey and Ellison (2011) study position auctions in a context with asymmetric information. Auctions induce a positive self-selection effect, but in Athey and Ellison (2011) the interests of search engine and consumers are not perfectly aligned due to a search externality: consumers incur the cost of searching but do not take into account the positive effect of their search on producer surplus. This search externality is also present in Hagiu and Jullien (2011), who show that a search intermediary may bias information so as to increase consumer search, total surplus and profits. We treat a more basic conflict of interest: the imperfect alignment of merchants' margins and consumers' net benefits.[9]

For organic search bias and integration, Hahn and Singer (2008) provide a law and economics analysis of the Google-DoubleClick merger of 2007. White (2012) also considers this merger and the insight that organic results can interfere with sponsored search, but his paper

[8] We develop similar results for publisher integration.

[9] A third strand of the literature shows how search distortions can affect the determination of merchant values via product market price competition (see e.g., Chen and He, 2011, Hagiu and Jullien, 2011, White, 2012, Xu, Chen, and Whinston, 2010 and 2011).

lacks a micro-founded model of product search.[10] Our model focuses on the direct interaction in the advertising market by merchants who can choose among display and search advertising. More recently, independent work by de Corniere and Taylor (2012) also examine some of the topics we analyze. Using a reduced-form approach, they study how substitutability between display and sponsored search advertising affects organic search bias and they look at the impact of integration with one of two publishers. Under separation, Google biases organic search against publishers that are closer substitutes to its own sponsored advertising (or impose larger ad nuisance costs on consumers). Integration with this closer competitor may lead to less display advertising and to less bias. Their results are largely compatible with ours. Our explicit, micro-founded framework derives a value for the exogenous substitutability parameter in their model. We also explicitly model competition among advertisers and derive a bias in sponsored search which interacts nontrivially with organic search bias. For instance, an improvement in publishers' targeting technologies can reduce the quality of organic results (which is quite intuitive), but may instead impoverish the sponsored results (much less intuitive). There are various other modeling differences between the two papers; in particular, in our model the search engine can reduce the effectiveness of display advertising even when publishers are symmetric, so organic search distortions can arise in a symmetric environment. On the other hand, we abstract from nuisance costs and their implications for search bias.

The paper is organized as follows. Section 2 presents the baseline model and we characterize the social optimum, before analyzing the equilibrium of the game with full vertical separation in section 3, with monopolizing vertical integration into ad intermediation in section 4, and partial monopolization in section 5. In section 6, we allow for publisher asymmetries in ad effectiveness and section 7 treats integration with publishers. In all cases, we compare welfare and surplus implications. Section 8 discusses assumptions and extensions and we conclude in section 9, gathering proofs in the Appendix..

2 The baseline model

Here we present the essential elements of the baseline model, postponing to Section 8, the motivation and endogenization of our main assumptions, and some possible generalizations. There are five types of agent: consumers, merchants, publishers, intermediaries (between publishers and merchants) in the display advertising market, and a search engine, G. There are five types of agent: consumers, merchants, publishers, intermediaries (between publishers

[10]All merchants sell the same product and no merchant can choose between display and search advertising, but his "left-side merchants" (the right-side ones pay for search ads) get free advertising and affect product market prices, so they indirectly affect sponsored search revenues.

and merchants) in the display advertising market, and a search engine, G.

Products and content. A mass one of consumers, indexed by i, value specific varieties of offline and online goods, seeking one unit of each. Offline goods, called "products", vary by product category, indexed by $j \in \{1, 2, ..., J\}$, and by product type, indexed by $k \in \{1, 2\}$, giving $2J$ products, each defined by a pair (j, k). Consumer i is only interested in one category, denoted $j(i)$, and gets a net benefit v_1 from one unit of her best match product $(j(i), 1)$ and v_2 from her second-best product $(j(i), 2)$. Consuming multiple units or any other product gives a net loss. We assume $0 < v_2 < v_1$. Merchants produce one product each and make them available on their websites. Two merchants produce each product, implying $4J$ merchants. This ensures competition for all advertising opportunities.[11] A merchant is type k if its product is type k. A type k merchant earns a unit margin m_k and we assume $m_1 < m_2$. This homes in on the possible conflict of interest between consumers and merchants that arises whenever consumer values and merchant margins are imperfectly aligned. To simplify the welfare analysis, we assume $m_1 + v_1 > m_2 + v_2$, so the social optimum has only type 1 transactions.

There are N online goods, which we call "content." Publishers make content available on their websites where they can expose their visitors to display ads. Each publisher has exactly one website with unique content, so publishers, content and websites share an index, n. Each consumer i has a favorite or "best match" publisher, $n(i)$, whose content directly generates net utility u, $u > 0$, while any other publisher content, $n' \neq n(i)$, generates zero net utility; any second unit implies negative utility. Consumers also gain indirect utility if publishers display ads that lead to offline trade. Content is free so each publisher's objective is to maximize its display ad revenues. We treat the symmetric case where each product category and publisher interest the same fraction of consumers, $\frac{1}{J}$ and $\frac{1}{N}$, respectively.

Display advertising. Ad intermediaries operate between publishers and merchants. They can serve merchants' ads onto the display areas of publishers' websites, contingent on consumer visits and a signal, from their targeting technology, of the visitor i's product category interest; this signal $s(i) \in \{1, 2, ..., J\}$ is correct, $s(i) = j(i)$, with probability σ and otherwise a random distinct category.[12] When consumer i visits website n, publisher n has the space to show her up to one merchant's display ad at zero cost. Clicking on this ad leads i through to the chosen merchant's website. This click-through occurs with probability $\alpha \in (0, 1)$ if i is visiting her best match publisher $n(i)$ but falls to $\beta\alpha$ if i is on any other

[11] Having two (or more) merchants per product implies zero merchant profits in equilibrium, which greatly simplifies the analysis, but competition is not a critical assumption (see section 8).

[12] Targeting precision satisfies, $\frac{1}{J} \leq \sigma \leq 1$ as $\sigma = \frac{1}{J}$ if the ad intermediary has no information on consumer preferences and $\sigma = 1$ if the ad intermediary can identify the consumer's best match category with probability one. $n(i)$ and $s(i)$ are independent. .

publisher's web; $\beta \in (0,1)$, capturing the idea that consumers spend less time on webs with less relevant content, which makes them less likely to notice and click on an ad; section 8 discusses further. Intermediaries have zero costs and we assume free entry. So, absent integration, intermediary competition leads to delivery of contingent display ads at zero cost. We only need model ad intermediaries explicitly when the search engine owns one.

Publishers conduct second-price auctions for displaying an ad contingent on each visitor's signal $s(i)$. Merchants submit bids representing their declared willingness to pay per click (PPC); the highest bidder wins and pays the second highest PPC bid. Each merchant producing (j,k) should be willing to pay $\sigma m_k CR_k$ if $s(i) = j$ and $\frac{1-\sigma}{J-1} m_k CR_k$ if not, where CR_k is the conversion rate or probability that i would purchase (j,k) if clicking on its ad, conditional on merchant relevance, $j = j(i)$. In our equilibria, $CR_1 = 1$ and $CR_2 = 0$. Publishers have no costs except possible ad intermediation charges, which are, as noted above, zero in the baseline case.

The search engine. Consumers do not know the identity of the publishers and merchants offering their preferred online and offline goods. They can use G to search for content, first, and then for products.[13] To conduct each type of search, they type in a query consisting of a set of keywords, and G responds by providing a set of results which consist of links to publishers' or merchants' websites. These results are separated into two groups: a list of "sponsored results" (typically appearing on the right and at the top of the screen with a yellow background) where website owners can sponsor a link to their site (in that they pay for a well-placed link), and a list of unsponsored or "organic results" where G commits against adjusting results as a function of any such direct payment. As explained in the introduction, for content search, we omit sponsored results so consumers only use organic results; we describe this and some other search characteristics as mere assumptions, but discuss endogenization in section 8.

Content search. Each consumer types in a query describing her content interest. G perfectly interprets this query and locates consumer i's preferred publisher, $n(i)$. G shows a single organic result, a single link to some publisher website n; in effect, G recommends n and consumers optimally follow. G could please consumers and always show/recommend the website of the consumer's best match, but G may have incentives to distort search by showing a different website. Specifically, G recommends the consumer's best match, $n(i)$, with probability r^O and shows a randomly chosen other website with the complementary probability. So r^O represents the "reliability" of organic results in the case of content search.

Product search. Each consumer types in a query describing her product interest. G

[13] All we require is that consumers sometimes purchase, via display advertising, a product for which they might otherwise have conducted a product search.

can perfectly interpret these queries and identify i's preferred product category, $j(i)$, and the corresponding two type 1 and two type 2 merchants. In addition, merchants know which queries are relevant to them. G provides no organic results and shows only one merchant in its sponsored search list. So the reliability of product search is fully determined by the reliability of sponsored search.

Search advertising. G uses a weighted second-price auction to allocate the single sponsored result.[14] The weights take into account each merchant's "quality" or relevance to consumers: G discounts the bid of type 2 merchants by a factor μ, $\mu < 1$. Merchants who know their product is irrelevant to a particular query never bid a positive amount, so there is no need to discount them and we can restrict attention to the bids, b_1, b_2, of the merchants offering the best match and second best products, respectively; in equilibrium, both merchants selling the type k product place the same bid, b_k.

A bid represents a merchants' declared willingness to "pay per click" (PPC): every time a consumer clicks on the sponsored result, the merchant will pay G the PPC determined in the corresponding auction. The winner of the auction is determined by comparing the weighted bids, b_1 and μb_2. The PPC rate is set equal to the lowest bid that would have allowed the winner to win the auction. Since there are two merchants of each type, in equilibrium the winner of the auction will be chosen at random and pay her bid. If $b_1 > \mu b_2$, one of the type 1 merchants wins and the PPC rate is b_1. Conversely, if $b_1 < \mu b_2$, a type 2 merchants wins, paying a PPC of b_2. If the two types tie, G applies a tie-breaking rule that favors type 1 merchants with probability r^S and type 2 merchants with probability $1 - r^S$. In our equilibria, tying always occurs and in this case r^S represents the "reliability" of sponsored search.

Consumer participation. We model consumers' participation in Google as a single decision that depends on the overall reliability of sponsored and organic results.[15] Consumers have a joint cost c_i of using the search engine for both content and product queries.[16] We assume that c_i is an independent draw from a continuous random variable on $[0, c_H]$ with density function $f(c)$ and cumulative distribution function $F(c)$. We assume that the hazard rate, $H(c) = \frac{f(c)}{F(c)}$, is decreasing.[17] We normalize all consumers' expected utility from not

[14] This is often called a position, paid-for placement or sponsored keyword auction.

[15] This simplification captures the fact that high quality results in either type of search spill over into improving Google's overall popularity or reputation as a reliable search engine. This holds for two reasons: consumers develop a habit of using a fixed search engine instead of thinking whether to search online or offline for each search query; also limited memory and communication (when learning from friends and media) lead consumers to learn the quality of search results in a coarse fashion, not fully contingent on the type of searches conducted.

[16] This reflects opportunity costs including the foregone value of alternative searches as well as any direct disutilities involved in searching online.

[17] This guarantees that, thinking of G as a monopsonist facing a supply provided by consumers, G faces an

participating in online search to zero. Gross of her cost c_i, the expected gain for a consumer i from participating in online search is given by the expected gains $r^O u$ from consuming online content, plus the expected gains from offline products, both through display advertising during the content search stage and also through search advertising in the product search stage. The highest possible gain is $u + v_1$. We assume $c_H > u + v_1$, which guarantees that consumer participation in online search is always interior in equilibrium.

The timing. In the first stage, G and publishers announce their auction rules, including design variables, μ, r^S and r^O. In the second stage, merchants choose their bidding strategies in both the position auctions and display advertising auctions. In the third stage, consumers decide whether or not to use the search engine. If they do search on G, they type in their query for content and visit one publisher's website. While consuming online content, they may be attracted to click on a displayed ad through to a merchant's web where they may buy the merchant's product. Then they can leave the market or they type in a product query and visit one merchant's website where they can buy the merchant's product. Merchants, publishers, ad intermediaries, and consumers can observe the actions chosen in the previous stages.

The social planner's problem. For a benchmark, we consider a social planner that maximizes the sum of all agents' surpluses and can control how the search engine matches consumers with merchants and publishers and how publishers and intermediaries allocate display ads among merchants, as well as consumer participation in online search. For any given participation level, the best possible outcome from this total surplus perspective is for each consumer for each consumer to consume her best match content and one unit of her best match product (recall $m_1 + v_1 > m_2 + v_2$). This is feasible: G can send each consumer to the best match publisher and can send any product-searching consumer to the best match merchant. It is also necessary for efficiency: display ads linking to best match merchants can provide an alternative channel for some efficient offline trades, but not all (as $\alpha < 1$), so undistorted product search is necessary as well as sufficient.[18] Abusing notation in anticipation of the equilibri

Display ads can provide an alternative channel for matching consumers and merchants, but this is redundant in the baseline where product search has no cost or imperfection; display can do no harm either as consumers can ignore irrelevant or type 2 ads. Display advertising never suffices for efficient offline trade (as $\alpha < 1$), so it is necessary that G does not distort product search. Undistorted content search is of course necessary too. Abusing notation

increasing "marginal factor cost".

[18]This display channel is redundant in the baseline model where product search has no cost or imperfection (and consumers ignore irrelevant or type 2 display ads).

in anticipation of the equilibria with tying merchant types, we momentarily let r^O denote the probability that G sends content searchers to their best match publishers, while r^S is the probability of sending product searchers to their best match, that is, relevant type 1, merchants. So we have,

Proposition 1. *Total surplus maximization requires the search engine to allocate traffic with no distortion from the consumer's ideal; it requires* $r^O = r^S = 1$.

This proposition also holds in the constrained scenario where the planner cannot control consumer behavior (i.e., consumers set their privately optimal participation and trade decisions). In fact, the incentives not to distort traffic are reinforced: consumers neglect the positive externality of participation on producer surplus, so the social planner would, if feasible, want to stimulate participation by further improving search engine reliability.

3 Equilibrium analysis under vertical separation

We begin by studying the optimal design of sponsored search auctions and merchants' equilibrium bidding strategies. Those consumers that did not purchase offline goods through display advertising will get involved in product search by entering a query in the search engine. Four merchants (two of each type) are potentially interested in the query, so the PPC that a merchant is willing to pay to appear in the single slot of G's sponsored search results is m_k. At this point merchants have no alternative advertising options and they correctly anticipate that every click in their link implies a purchase. In a standard second-price auction, each merchant's unique weakly dominant strategy is to bid her willingness to pay: $b_k = m_k$. This is still so in our setup:

Lemma 1. *For any* μ, r^O, *and* r^S, *the strategy profile* (b_1, b_2), *where* $b_k = m_k$, *is the unique equilibrium in undominated strategies.*

Competition among merchants implies that in equilibrium the winner always pays its willingness to pay. So if $\mu < \frac{m_1}{m_2}$, type 2 bids are so discounted that a type 1 merchant always wins and traffic is efficiently allocated. Conversely, if $\mu > \frac{m_1}{m_2}$, type 2 merchants always win and product search is inefficient. Discounting type 2 merchants by the precise weight $\mu = \frac{m_1}{m_2}$ equates the effective willingness to pay of the two types of merchant, so G can use its tie-breaking rule r^S to fine-tune the probability, also then equal to r^S, that a type 1 merchant wins the position auction. G need only consider this last case, since setting $r^S = 1$ captures the high μ case and $r^S = 0$ captures the lower μ case.[19] Summarizing:

[19] In our setup, setting values $\mu > \frac{m_1}{m_2}$ with probability r^S and $\mu < \frac{m_1}{m_2}$ with converse probability generates

Lemma 2. *G optimally sets the auction weights in such a way that type one merchants win the auction and pay* m_1 *per click with probability* r^S*, and type 2 merchants win and pay* m_2 *with probability* $1 - r^S$*.*

We now turn attention to auctions for display advertising. If a consumer is not attracted by display advertising, merchants still have a chance of capturing this consumer through sponsored search advertising. However, merchants anticipate that the value of selling to a consumer in this second stage is nul since in case of winning a position auction they must pay a PPC equal to their margin. Hence, when merchants bid in publishers' auctions they only take into account the potential profits obtained by such an advertising channel. Moreover, since each publisher is too small to affect consumer participation, publishers have no incentives to give different weights to the bids of different merchant types. That is, they have no incentive to set a weight "μ" different from 1. So, publishers are willing to allocate the single slot to one of the winning merchants in the unweighted auction. However, in equilibrium only one type of merchants has incentives to participate in the auction. In principle, type 2 merchants are in a better position to win the auction. However, if consumers are exposed to type 2 ads then they will prefer not to buy from them. If they did so, their net utility would be v_2, whereas by waiting and conducting their product search they would obtain an expected net utility of $r^S v_1 + \left(1 - r^S\right) v_2 \geq v_2$.[20] So we focus on the case where consumers are only willing to purchase type 1 goods when attracted by display advertising. In this case, type 2 merchants do not have any incentives to bid and competition between identical type 1 merchants implies that they bid a PPC equal to σm_1, since they anticipate that a fraction σ of clicks yield sales.

Lemma 3. *Publishers optimally conduct unweighted auctions, in which, in the equilibrium in undominated strategies, only type* 1 *merchants participate actively, bidding a PPC equal to* σm_1*.*

Since the probability that a consumer visiting a publisher website is attracted by a display ad is α if visiting her favorite publisher and $\alpha\beta$ if visiting any other publisher, then in equilibrium the average fraction of clicks on display ads is $\alpha e\left(r^O\right)$, where $e\left(r^O\right) = r^O + \left(1 - r^O\right)\beta$. Finally, the fraction of clicks that end up in a transaction depends on the quality

an outcome equivalent to the auction design. So, the second price auction is equivalent to a mechanism where G simply sets take-it-or-leave-it offers to each type of merchant with probabilities r^O and r^S to type 1 merchants with probability $1 - r^S$, and to type 2 merchants with probability $1 - r^S$; this would work without competition. However, in an alternative setup where G is uncertain v and m about margins and values, an auction is valuable and a deterministic value of μ could determine the probability of success of each type of merchant, with ties having zero probability.

[20] In the main text we will ignore the possible equilibria where consumers are willing to purchase type 2 goods when attracted by display ads, under the expectation that $r^S = 0$, and parameter values are indeed such that G finds it optimal to set $r^S = 0$.

of the targeting technology, σ. As a result, the fraction of participating consumers that purchase one unit of their best match through display advertising, η, is given by

$$\eta = \sigma\alpha e\left(r^{O}\right).$$

Note that η increases with the quality of the targeting technology, σ, and with the reliability of organic results, r^{O}. If we let X denote consumer participation, then the mass of consumers performing product search is $X(1-\eta)$.

Let us now examine the determinants of consumer participation. If we let $\overline{c}$ be the highest value of the cost of using the search engine of all participating consumers then $X = F(\overline{c})$,where

$$\overline{c} = r^{O}u + v_1\left[\eta + (1-\eta)r^{S}\right] + v_2(1-\eta)\left(1-r^{S}\right). \tag{1}$$

The first term is the expected net utility from consuming online content. The second term is the expected net utility from consuming her best match product. Such consumption will occur with probability η through display advertising and with probability $(1-\eta)r^{S}$ through sponsored search advertising. Finally, the third term is the net expected utility from consuming her second-best product, which will occur with probability $(1-\eta)\left(1-r^{S}\right)$ as a result of sponsored search advertising. Note that $\overline{c}$ increases with r^{O}, r^{S} and σ:

$$\frac{\partial\overline{c}}{\partial r^{O}} = u + (v_1 - v_2)\left(1-r^{S}\right)\sigma\alpha(1-\beta), \qquad \frac{\partial\overline{c}}{\partial r^{S}} = (v_1 - v_2)(1-\eta),$$

$$\frac{\partial\overline{c}}{\partial\sigma} = (v_1 - v_2)\left(1-r^{S}\right)\alpha e\left(r^{O}\right).$$

If the search engine engages in traffic distortion (setting low values of r^{O} and/or r^{S}) then it will pay the cost of having lower consumer participation. In fact, these two instruments, r^{O} and r^{S}, have a similar role in encouraging participation. Also, a high value of one of these instruments reduces the sensitivity of consumer participation with respect to changes in the other instrument; that is $\frac{\partial^2\overline{c}}{\partial r^{S}\partial r^{O}} < 0$.

The average merchant margin on all sales that materialize through sponsored search is:

$$M\left(r^{S}\right) = r^{S}m_1 + \left(1-r^{S}\right)m_2$$

which is decreasing in r^{S}. The following proposition follows directly:

Proposition 2. *Given $\left(r^{O}, r^{S}\right)$ there exists a unique equilibrium in undominated strategies in which merchants make zero profits, each publisher earns $\Pi_n = \frac{F(\overline{c})}{N}\eta m_1$, and the search engine earns $\Pi^{G} = F(\overline{c})(1-\eta)M\left(r^{S}\right)$.*

We can now analyze G's incentives for traffic management. The first order condition for maximizing G's profits with respect to r^{O} is:

$$\frac{\partial\Pi^{G}}{\partial r^{O}}\frac{1}{M\left(r^{S}\right)} = f(\overline{c})\frac{\partial\overline{c}}{\partial r^{O}}(1-\eta) - F(\overline{c})\frac{d\eta}{dr^{O}} = 0. \tag{2}$$

The first term is positive and represents the increased participation in search associated with more accurate organic search for online content, and so with a higher probability that the consumer finds her best match online content. The second term is negative. Raising the accuracy of organic search increases the value of display advertising, which reduces consumers' need for product search.

The first order condition for the maximization of G's profits with respect to r^S is:

$$\frac{\partial \Pi^G}{\partial r^S}\frac{1}{1-\eta} = f(\bar{c})\frac{d\bar{c}}{dr^S}M\left(r^S\right) + F(\bar{c})\frac{dM\left(r^S\right)}{dr^S} = 0. \tag{3}$$

Again, the first term is positive. A more accurate sponsored search attracts more consumers since it results in a larger probability of finding the best match product when searching for offline products. On the other hand, an increased search accuracy reduces the profitability to G of each participating consumer. This is because G receives a higher price (PPC) whenever a type 2 merchant wins the relevant position auction over type 1 merchants. G internalizes the full share of merchant values from search advertising, given their display advertising choices, and therefore gains by distorting search towards the high margin, type 2 merchants.

It will be useful to rewrite the two first order conditions (2) and (3) as follows:

$$H(\bar{c})(1-\eta)\left[\frac{u}{\sigma\alpha(1-\beta)} + (v_1 - v_2)\left(1-r^S\right)\right] = 1. \tag{4}$$

$$H(\bar{c})(1-\eta)\left[\frac{v_1 - v_2}{m_2 - m_1}m_1 + (v_1 - v_2)\left(1-r^S\right)\right] = 1. \tag{5}$$

These two conditions can only hold simultaneously if parameter values satisfies a very specific condition:

$$\frac{u}{\sigma\alpha(1-\beta)m_1} = \frac{v_1 - v_2}{m_2 - m_1}. \tag{6}$$

So:

Lemma 4. *Generically, (6) fails and r^S and r^O cannot be both interior in* $(0,1)$.

The left hand side of (6) represents the cost/benefit ratio of distorting organic search. Higher distortion in organic search, i.e., lower r^O, may lead to a private loss u in consumers' surplus, but also a switch of demand from publishers' ads to sponsored ads of $\sigma\alpha(1-\beta)m_1$. The right hand side of (6) represents the cost/benefit ratio of distorting organic search, i.e., reducing r^S. Higher distortion in sponsored search, reduces consumers' surplus by $v_1 - v_2$ but raises the value of sponsored ads by $m_2 - m_1$. If $\frac{v_1 - v_2}{m_2 - m_1} < \frac{u}{\sigma\alpha(1-\beta)m_1}$, then distorting product search is more attractive for G, whereas distorting content search is more attractive if the inequality is reversed.

Consequently, a sufficient condition for distortions to occur is that

$$H(u + v_1)(1-\sigma\alpha)\min\left\{\frac{u}{\sigma\alpha(1-\beta)}, \frac{v_1 - v_2}{m_2 - m_1}m_1\right\} < 1, \tag{7}$$

so that the LHS of (4) and (5) evaluated at $r^S = r^O = 1$ are both larger than 1. Moreover, since the LHS of both (4) and (5) are decreasing in both r^O and r^S, this condition is also necessary, giving:

Proposition 3. *The search engine allocates traffic inefficiently ($r^S < 1$ and/or $r^O < 1$) if and only if condition (7) holds.*

The reader should notice that the solution for either r^O or r^S is interior for a set of parameter values with a non empty interior.[21] So we can conduct meaninful comparative statics.

We now examine the effect of a targeting technology improvement. Obviously, a marginal change in σ can only affect variables with interior solutions. Since we can only have one interior solution, the comparative statics are quite straighforward. If $0 < r^S < 1$ then r^O is either equal to 0 or 1 (and not affected by σ). The equilibrium value of r^S is determined by equation (5). In this case an increase in σ raises the number of transactions of offline products generated by display advertising, η, which in turn reduces the sensitivity of consumer participation to changes in r^S. So, the cost of distorting sponsored search is lower. Moreover, these additional transactions involve best matches, and since $r^S < 1$ consumers are made better off, which raises consumer participation, $\bar{c}$. Hence, the benefits from raising profits per consumer by distorting sponsored search are higher. In sum, an increase in σ implies a lower value of r^S.[22]

If $0 < r^O < 1$ then either r^S is equal to 0 or to 1. In this case the equilibrium value of r^O is given by equation (4). A higher value of σ enhances the effectiveness of organic search distortion (higher $\frac{\partial \eta}{\partial r^O}$). Hence, the same reduction in r^O achieves a higher increase in the fraction of participating consumers conducting product search. Besides this direct effect there are also two indirect effects. As discussed above, a higher value of σ has a positive impact on both $\bar{c}$ and η. The lower number of consumers conducting product searches ($1 - \eta$ is lower) reduces the incentives to attract consumers (lower cost of distorting organic search). Similarly, higher consumer participation exacerbates the incentives to increase profits per participating consumer (higher benefits from distorting organic search.) Hence, these indirect effects also induce the search engine to set a lower value of r^O.[23]

Summarizing:

[21]The LHS of (4) and (5) are continuous functions of parameters and endogenous variables. For instance, consider the case $\frac{v_1 - v_2}{m_2 - m_1} < \frac{u}{\sigma\alpha(1-\beta)m_1}$ and parameter values satisfying (5) at $r^S = r^O = 1$; a small change in a parameter that decreases the LHS of (5) implies a strict small decrease in r^S and no change in r^O.

[22]Note that the overall effect of σ on consumer participation is ambiguous: The direct effect is positive, but this may be more than compensated by the negative indirect effect (lower r^S).

[23]As in the previous case the overall effect of σ on consumer participation is ambiguous, since the direct positive effect may be more than compensated by the negative indirect effect (lower r^O). Similarly, the overall effect on η is also ambiguous.

Proposition 4. *An improvement in the targeting technology reduces the reliability of the search engine. More precisely, if either r^S or r^O are interior, a marginal increase in σ reduces the equilibrium value of that variable.*

This result suggests that r^S and r^O are imperfect substitutes from the point of view of G's profits. Indeed, $\frac{\partial^2 \Pi_G}{\partial r^S r^O} < 0$. These two instruments interact through two main channels. First, they both increase consumer participation. Second, a higher value of any of these instruments reduces the sensitivity of consumer participation with respect to the other instrument.

4 The effect of integration and full monopolization

In this section we begin to examine the effects of integration in our baseline model. This model is highly stylized, which simplifies the analysis considerably but at the same time causes somewhat extreme results. Some features of the model are particularly important. First, all publishers are symmetric. Second, G has all the bargining power vis-a-vis publishers. Third, the market is completely unregulated and hence G finds no limit to its profit-maximizing strategies. Some of these issues will be discussed later on.

Suppose that G merges with one particular ad intermediary. Such integration has both a vertical and a horizontal dimension. The search engine is essential in the production of display advertising. In order to attract the maximum number of consumers through ads posted in publishers' websites, merchants need not only the contribution of publishers and intermediaries (the effective use of the ad space and the targeting technology) but they also need the search engine's contribution sending consumers to their best match content websites. So, the acquisition of an ad intermediary by the search engine can be characterized as a vertical integration. But at the same time display advertising is an imperfect substitute of sponsored search advertising. In this sense such acquisition can be also caracterized as horizontal integration.

In the absence of regulatory supervision, G could set different values of r^O for publisher websites whose advertising business is run by G's ad intermediary, and for the rest. In fact, G may "threaten" websites with $r^O = 0$ in case they deal with a different intermediary. Suppose ad intermediaries offer their services to publishers in exchange for a tariff T, and next publishers determine the PPC through second price auctions. Then, if G has no constraint on the number of publishers that its ad intermediary can handle, it can capture the entire surplus from display advertising. Indeed, in equilibrium all other intermediaries will offer $T = 0$, G will set $r^O = 0$ for publishers dealing with any other intermediary, and G's intermediary will set $T^G = \frac{F(\bar{c})}{N} \eta m_1$. So, G's profits will now be:

$$\Pi^G = F(\bar{c}) \left[(1 - \eta) M \left(r^S \right) + \eta m_1 \right].$$

Once again, let us first examine G's optimal behavior in allocating traffic. The effect of r^O on G's profits is given now by:

$$\frac{\partial \Pi^G}{\partial r^O} = f(\bar{c}) \frac{d\bar{c}}{dr^O} \left[(1-\eta) M\left(r^S\right) + \eta m_1\right] - F(\bar{c}) \frac{d\eta}{dr^O} \left[M\left(r^S\right) - m_1\right] = 0. \tag{8}$$

Two new effects appear relative to the case of no integration (condition (4). First, the per-consumer rents, i.e., $\left[(1-\eta) M\left(r^S\right) + \eta m_1\right]$ are higher. That is, G has higher incentives to attract consumers. Second, a higher r^O increases the rents per consumer that G extracts from publishers, ηm_1. Both effects induce a higher value of r^O.

The first order condition with respect to r^S is:

$$\frac{\partial \Pi^G}{\partial r^S} = f(\bar{c}) \frac{d\bar{c}}{dr^S} \left[(1-\eta) M\left(r^S\right) + \eta m_1\right] + F(\bar{c})(1-\eta) \frac{dM\left(r^S\right)}{dr^S} = 0. \tag{9}$$

Again, with respect to the case of no integration (condition (5) we have an additional effect. Namely, the gains from consumer participation apply to a larger base (display as well as sponsored search auctions). Hence, G faces stronger incentives to set a higher value of r^S with respect to the case of no integration. [24]

Equations (8) and (9) can be rewritten as follows:

$$H(\bar{c}) \frac{(1-\eta) M\left(r^S\right) + \eta m_1}{M\left(r^S\right) - m_1} \left[\frac{u}{\sigma\alpha(1-\beta)} + (v_1 - v_2)\left(1 - r^S\right)\right] = 1, \tag{10}$$

$$H(\bar{c}) \left[\frac{v_1 - v_2}{m_2 - m_1} m_1 + (v_1 - v_2)\left(1 - r^S\right)(1-\eta)\right] = 1. \tag{11}$$

These two equations cannot hold simultaneously since the LHS of (10) is higher than the LHS of (11). Therefore, $r^O \geq r^S$ and only one of them can be an interior solution. Also, the LHS of (10) is larger than (4), which indicates that, for a given r^S, under integration G has incentives to set a higher value of r^O. Similarly, the LHS of (11) is higher than the LHS of (5). Again, for a given r^O, under integration G has incentives to set a higher value of r^S. Finally, we check that interaction between these two instruments does not change this initial insight. In particular, we prove that:

Proposition 5. *Integration with full monopolization improves the reliability of the search engine and increases both consumer and total surplus. More precisely, from an initial case of non-integration, integration weakly raises r^S and r^O, and raises one or both strictly, except possibly staying at a corner solution if initially at a corner (certainly if initially at the corner with $r^S = r^O = 1$ integration has no effect).*

Proof. See the Appendix. □

[24] Note also that as for non-integration, improved targeting technologies again increase G's incentives for traffic distortion.

Under integration G monopolizes the display advertising intermediation market and as a result incentives to allocate traffic in both types of queries are improved from the point of view of both consumer and total surplus. Now G internalizes a vertical externality. This is somewhat analogous to the case of upstream-downstream integration in markets mediated by prices. When G integrates with ad intermediaries, and then is able to appropriate the rents of publishers, it acquires incentives to increase these rents. More accurate organic search has a direct impact in this direction. Also, profits per participating consumer increase, which enhances G's incentives to attract new consumers. This can be accomplished by improving the realiability of its search results. G also internalizes a horizontal externality: G has less incentive to distort r^O, because G now benefits from this (aside from the continued indirect effect of substitution to type 1 advertising on display).

This proposition ignores many of the distributional consequences of vertical integration. In particular, publishers get zero profits which would, in a model with costly content, affect on the quality of online content. In a richer model with endogenous content quality, the type of integration we have considered in this section is likely to have an additional effect on consumers' welfare, since publishers' incentives to invest in quality would be significantly reduced.

Another potential drawback of integration has to do with G's incentives to discriminate against publishers not dealing with G's intermediaries. This effect is absent in the above, extreme case of full monopolization, since all publishers are symmetric and deal with G in equilibrium. In the next two sections we illustrate how this effect comes into play in a less extreme market structure.

5 Integration with partial monopolization

The assumptions in our baseline model, in particular publisher symmetry and constant returns to scale in the ad intermediation technology, result in full monopolization when G enters the ad intermediation market. So G's discrimination against the publishers not dealing with G's ad intermediary did not translate into any discrimination among publishers in equilibrium. In a more realistic setting, publishers would be heterogeneous and G's integration with one ad intermediary would typically result in partial monopolization. So in this section, we examine the consequences of partial monopolization of the display advertising market without specifying the obstacles that prevent full monopolization.

Suppose that G's ad intermediary can at most handle the advertising business of a fraction γ of publishers.[25] As in the previous section, we let G treat these two types of websites

[25] One possible interpretation of such exogenous constraint is that an increase of G's market share above γ

differently. G offers their potential ad intermediation customers the possibility of receiving more traffic in exchange for a tariff, T_G. We assume G cannot price discriminate so it offers the same deal to all websites. Also, we assume that customers take their participation decision without knowing whether their favorite publisher is more or less likely to be affiliated with G. So, G announces r_G^O and r_{NG}^O, the probabilities that a customer looking for content offered by a publisher is sent to that publisher when it is affiliated with G and when it is not, respectively. Clearly, G will find it optimal to send any diverted traffic to a publisher in the G system.[26]

Extending the notation from previous sections, we let e_G and e_{NG} be the fraction of "effective" visits received by publishers in the G system or outside, respectively. That is,

$$e_G\left(r_G^O, r_{NG}^O\right) = r_G^O + \left(1 - r_G^O\right)\beta + \frac{1-\gamma}{\gamma}\left(1 - r_{NG}^O\right)\beta,$$

and

$$e_{NG}\left(r_{NG}^O\right) = r_{NG}^O.$$

Accordingly, the proportion of purchases from each type of publisher is $\eta_t = \sigma\alpha e_t$ for $t = G, NG$. Given this, participation is still determined by (1), although now the average sales induced by display advertising, η, is given by:

$$\eta = \gamma\eta_G + (1-\gamma)\eta_{NG},$$

and the average accuracy of organic search results, r^O, is:

$$r^O = \gamma r_G^O + (1-\gamma) r_{NG}^O.$$

In equilibrium T_G will be such that all websites are indifferent between accepting and rejecting the deal offered by G, but it has to be the case that at least a fraction γ of websites will accept the offer and G makes a deal with exactly a fraction γ. Lemma 3 still holds for publishers affiliated with G, as well for those not affiliated. That is, merchants are still willing to bid a PPC equal to σm_1 in both types of publishers. So, publishers' willingness to pay to be part of the G system is:

$$T_G = \frac{F\left(\bar{c}\right)}{N}\left(\eta_G - \eta_{NG}\right) m_1.$$

might trigger an unwanted investigation by the regulatory agency.

[26] Distorting search for online goods away from the best match publisher has the same effect on customer participation and merchants' willingness to pay for sponsored ads whether or not the *destination* publisher is affiliated with G. Of course, each publisher's site traffic affects merchants' willingness to pay of display ads on that site, which benefits G when and only when the destination of diverted traffic is a site affiliated with G.

Since $\eta_G - \eta_{NG} = \frac{\eta - \eta_{NG}}{\gamma}$ and G's profits from affiliated publishers are equal to $\gamma N T_G$, we can write G's profits as

$$\Pi^G = F(\bar{c}) \left[(1-\eta) M\left(r^S\right) + (\eta - \eta_{NG}) m_1 \right].$$

The effects of r_G^O and r^S are analogous to the those discussed in the case of full monopolization (equations (8) and (9)). The main distinctive feature of the partial integration case is reflected in the first order condition with respect to r_{NG}^O:

$$\frac{\partial \Pi^G}{\partial r_{NG}^O} = f(\bar{c}) \frac{\partial \bar{c}}{\partial r_{NG}^O} \frac{\Pi^G}{F(\bar{c})} - F(\bar{c}) \frac{\partial \eta}{\partial r_{NG}^O} \left[M\left(r^S\right) - m_1 \right] - F(\bar{c}) \frac{\partial \eta_{NG}}{\partial r_{NG}^O} m_1. \tag{12}$$

The first term of (12) represents the effect of r_{NG}^O through consumer participation. Note that in this case, since $\gamma < 1$, r_{NG}^O positively influences consumer participation. So, in contrast to the baseline model, G's incentives to set a low r_{NG}^O are moderated and so the optimal r_{NG}^O is not always equal to zero. However, precisely because $\gamma < 1$, even if $r_{NG}^O > 0$, a value of r_{NG}^O below 1 represents traffic distortions that are realized, and negatively affect consumer and producer surpluses. The second term of (12) represents the effect of r_{NG}^O through the average rate of transactions induced by display advertising, η. If $r^S < 1$ an increase in η has a negative effect on G's profits because some transactions induced by sponsored search advertising, which on average yield a margin $M\left(r^S\right)$, are replaced by transactions induced by display advertising that yield a lower margin, $m_1 < M\left(r^S\right)$. The third term of (12) represents the effect of r_{NG}^O through the rate of transactions induced by display advertising in non-affiliated publishers, η_{NG}. This effect is also negative since a higher η_{NG} raises the profits of non-affiliated publishers, which in turn reduces the rents that G can extract from affiliated publishers. The fraction of affiliated publishers, γ, has a different impact on these three effects. Indeed, (12) can be written as

$$\frac{\partial \Pi^G}{\partial r_{NG}^O} \frac{1}{(1-\beta)\sigma\alpha} = (1-\gamma) \Psi\left(r_G^O, r_{NG}^O, r^S\right) - F(\bar{c}) \frac{m_1}{1-\beta}, \tag{13}$$

where $\Psi\left(r_G^O, r_{NG}^O, r^S\right)$ embeds both the first and the second effect; that is:

$$\Psi\left(r_G^O, r_{NG}^O, r^S\right) = f(\bar{c}) \left(\frac{u}{(1-\beta)\sigma\alpha} + (v_1 - v_2)\left(1 - r^S\right) \right) \frac{\Pi^G}{F(\bar{c})} - F(\bar{c}) \left[M\left(r^S\right) - m_1 \right].$$

As γ increases the first two effects shrink, and go to 0 as γ goes to 1. In contrast, the third effect is invariant with respect to γ and hence it remains strictly negative in the limit. Hence, if γ is sufficiently high then G will have incentives to distort traffic to non-affiliated publishers: such distortion has little impact on consumer participation but a non-negligible impact on the rents that G can collect from affiliated publishers.

As for full monopolization (the previous section) integration induces G to internalize a higher fraction of the producer rents generated by each participating consumer. Hence G has additional incentives to attract new consumers, which requires more reliable search results (higher r_G^O, r^S). However, with partial monopolization there is a new effect that works in the opposite direction: G may have incentives to distort more intensively the traffic to non-affiliated publishers. If this latter effect becomes relatively important, consumers end up being worse off under vertical integration.

We illustrate this in the case where G sets $r^O = r^S = 1$ under non-integration. As discussed above, under partial integration, for γ is sufficiently close to 1, G will set $r_{NG}^O < 1$, and both consumer and total surplus will be lower than under no integration. Summarizing:

Proposition 6. *There exists a region of parameter values for which vertical integration with partial monopolization reduces both consumer and total surplus.*

A sufficient condition is to have γ close to 1 and condition (7) failing, which might suggest relatively small welfare losses, as only traffic a fraction $(1 - \gamma)$ of publishers face distortion. However, in the Appendix, show that γ could be arbitrarily low and the losses in consumer and total surplus occasioned by integration could be substantial, owing to new distortion incentives.

6 Asymmetric publishers

In principle, differences between publishers in effectiveness for display advertising could affect search bias incentives and implications. We introduce this asymmetry in the simplest way possible: we now assume that a proportion ρ of publishers are type H, characterized by a higher baseline ad effectiveness α_H than the rest, of type L, with effectiveness $\alpha_L < \alpha_H$. We focus on the more interesting and simpler case where asymmetries are sufficiently large: we assume $\alpha_L < \beta\alpha_H$. Then even display advertising to best-match visitors of type L publishers is less effective than display advertising to worst match visitors of type H publishers.

We show shortly below that publisher asymmetries are irrelevant in our baseline model when G monopolizes the ad intermediation market, but this irrelevance is an artifact of the same simplifying assumption, frictionless product search, which made display advertising redundant in the first-best solution. G's choices do affect the relative roles of display and search advertising and the composition of offline trades, but in the baseline model, for a given participation level, distorting the effectiveness of display advertising has no effect on the quantity of offline trade; this is because each participating consumer always does a product search (at zero marginal cost) that always leads to offline trade, if they do not trade via display ads. To properly analyze publisher asymmetries, we need to allow for imperfections in the product search channel for offline trades.

Concretely, we now assume that a proportion ϕ of participating consumers behave as in our baseline model, always able to search for products at no added cost after searching for content, but the rest of consumers, $1-\phi$, can only search for content. Publishers, intermediaries, merchants, and the search engine know this but do not know which individual consumers have a viable option of product search. Consumers, by contrast, learn whether they can do a product search after choosing to participate, but before responding to display ads.[27] So a fraction ϕ of the consumers visiting a given publisher are unwilling to purchase type two goods because, as in the baseline model, they prefer to wait and search for offline goods later on, but now a fraction $1-\phi$ are willing to purchase either type.[28] Unless ϕ is very small, type 1 merchants still outbid type 2 merchants when competing for display ads. So we assume $m_1 > (1-\phi)m_2$, thereby ensuring that display ads are all of type 1 products, as in the previous sections.

G may gain by treating asymmetric publishers asymmetrically. So we distinguish the reliability of content search by whether the consumer is looking for a type H or type L publisher, denoting r_H^O and r_L^O, respectively. Moreover, G can choose to divert customers to either of the two types of publishers, so we let $d_{i,j}$ represent the fraction of diverted traffic that is diverted from type i to type j publishers, for $i, j \in \{L, H\}$. With this, η becomes,

$$\begin{aligned}\eta = \sigma\rho \left[r_H^O \alpha_H + (1-r_H^O)\beta \left(d_{HL}\alpha_L + (1-d_{HL})\alpha_H\right)\right] \\ + \sigma(1-\rho) \left[r_L^O \alpha_L + (1-r_L^O)\beta \left(d_{LH}\alpha_H + (1-d_{LH})\alpha_L\right)\right].\end{aligned}$$

Defining the average reliability of content search, $r^O = \rho r_H^O + (1-\rho) r_L^O$, participation is now given by,

$$\bar{c} = r^O u + \eta v_1 + \phi \left(1-\eta\right) \left(v_1 r^S + v_2 \left(1-r^S\right)\right).$$

We begin with the case of non-integration. G's profits are now,

$$\Pi^G = \phi F\left(\bar{c}\right)\left(1-\eta\right) M\left(r^S\right).$$

Note that, given the average number of transactions through display ads, η, and participation, $\bar{c}$, G's profits do not depend on r^O, but profits are increasing in $\bar{c}$. So, for given η and r^S, G maximizes r^O to maximize participation $\bar{c}$ to maximize profits. This implies that any diverted content search should be towards low ad effective publishers.

[27]So, when consumers choose whether to use the search engine, they anticipate that they will certainly searching for content and that, conditional on not having purchased an offline product through display advertising, they will search for an offline product with probability ϕ.

[28]This is in keeping with the interpretation of consumer participation as determined by experience. Consumers sometimes $(1-\phi)$ don't plan an offline purchase before being exposed to an ad, and in that case they may buy after clicking. But otherwise, they do plan to search for offline goods, and then would only buy if they think they found their best chance.

Lemma 5. *Under non-integration, if traffic is diverted, it is always from type H to type L publishers, but never the other way around; that is, if* $r_H^O < 1$ *then* $d_{HL} = 1$ *and if* $r_L^O < 1$ *then* $d_{LH} = 0$.

Proof. See the Appendix. □

The intuition is simple. Under separation, G benefits from reducing purchases through publishers, but prefers that consumers can access their preferred content. So if any traffic is distorted to type L publishers, which involves $r^O < 1$, and either $d_{HL} > 0$ or $d_{LH} < 1$, then G could always offer consumers a higher probability of accessing their preferred publisher, i.e., a higher r^O, and in exchange increase traffic to type L and reduce traffic to type H publishers. At root, G and the customers are not playing a zero-sum game in this case: the potential loss, from $r^S < 1$, in the likelihood of obtaining the preferred offline good when buying through sponsored search instead of display advertising, can be more than compensated by a higher likelihood of reaching the preferred content.

Now, since in equilibrium, d_{HL} is either 1 or irrelevant, and d_{LH} is either 0 or irrelevant, we can write

$$\eta = \sigma\alpha_L e(r^O) + \sigma\rho r_H^O(\alpha_H - \alpha_L).$$

Suppose $r_L^O = 1$. In this case, $r^O = \rho r_H^O + (1-\rho)$ and

$$\frac{\partial \Pi^G}{\partial r_H^O}\frac{1}{\phi M\left(r^S\right)} = f\left(\overline{c}\right)\frac{\partial \overline{c}}{\partial r_H^O}\left(1-\eta\right) - F\left(\overline{c}\right)\frac{\partial \eta}{\partial r_H^O},$$

where $\frac{\partial \eta}{\partial r_H^O} = \sigma\rho\left[\alpha_H - \beta\alpha_L\right]$. Also,

$$\frac{\partial \Pi^G}{\partial r^S}\frac{1}{\phi} = f\left(\overline{c}\right)\frac{d\overline{c}}{dr^S}M\left(r^S\right) + F\left(\overline{c}\right)\frac{dM\left(r^S\right)}{dr^S}.$$

Note that two necessary conditions for distortions not to occur are that these two expressions are negative when evaluated at $r^S = r_H^O = 1$, at which values $\eta = \sigma\alpha$, where $\alpha = \rho\alpha_H + (1-\rho)\alpha_L$. That is, the conditions are

$$H\left(\overline{c}\right)\left(1-\sigma\alpha\right)\left(\frac{u}{\sigma\left(\alpha_H - \beta\alpha_L\right)} + v_1(1-\phi)\right) \geq 1, \tag{14}$$

and

$$\phi H\left(\overline{c}\right)\left(1-\sigma\alpha\right)\frac{v_1 - v_2}{m_2 - m_1}m_1 \geq 1, \tag{15}$$

both evaluated at $\overline{c} = u + \left(\sigma\alpha(1-\phi) + \phi\right)v_1$. Note that for $\phi = 1$ and $\alpha_H = \alpha_L$, we obtain the same conditions as in Section 3. These conditions are also sufficent, for the same arguments discussed in Section 3: the LHS of both (14) and (15) are both decreasing in r^S and in r_H^O.

We now turn to the case of integration where G monopolizes the ad intermediation market. As in Section 5, we demonstrate the existence of parameter values for which G will introduce distortions under integration with full monopolization, but not under separation. So, suppose that G sets $r^S = 1 = r^O_H$, so that $\overline{c} = r^O u + \phi v_1 + \eta(1-\phi)v_1$ and

$$\Pi^G = m_1 F(\overline{c})\left[\phi + \eta(1-\phi)\right].$$

Note that Π^G is now increasing both in $\overline{c}$ and η. Moreover, $\overline{c}$ is also increasing in η, for a given value of r^O. Since η is increasing in d_{LH}, that immediately shows that $d_{LH} = 1$ is the optimal choice when $r^S = 1 = r^O_H$.[29] Finally, evaluated at $r^O_L = 1$,

$$\frac{\partial \Pi^G}{\partial r^O_L}\frac{1}{m_1} = f(\overline{c})\frac{\partial \overline{c}}{\partial r^O_L}\left[\phi + \eta(1-\phi)\right] + F(\overline{c})\frac{\partial \eta}{\partial r^O_L}(1-\phi), \tag{16}$$

where, for $d_{LH} = 1$, $\frac{\partial \eta}{\partial r^O_L} = \sigma(1-\rho)(\alpha_L - \beta\alpha_H)$. Notice the effect of publishers' asymmetry on the incentives for traffic distortion under integration. For symmetric publishers and $\phi = 1$, the effect of r^O on Π^G, $\frac{\partial \Pi^G}{\partial r^O}$ is represented in (8) and is positive if $r^S = 1$. Maintaining symmetry but introducing $\phi < 1$, $\frac{\partial \Pi^G}{\partial r^O}$ would read as (16), but with r^O substituting for r^O_L. Both participation and profits per consumer would then be increasing in η, the proportion of sales through display advertising, and both participation and η would also be increasing in r^O. So $\frac{\partial \Pi^G}{\partial r^O}$ would still be positive.

This would still the case when publishers are asymmetric if a change in r^O did not change the proportion of consumers that visit each type of publisher. However, with asymmetric publishers, G has a new instrument to increase η: addressing consumers who look for a type L publisher to a type H publisher. This is the effect of reducing r^O_L, when $d_{LH} = 1$. That is, now $\frac{\partial \eta}{\partial r^O_L} < 0$, and this gives G an incentive to reduce r^O_L.

Indeed, now a sufficient condition for the existence of distortions under vertical integration is that

$$H(\overline{c})\frac{\sigma\alpha + \phi(1-\sigma\alpha)}{1-\phi}\left(\frac{u}{\sigma(\beta\alpha_H - \alpha_L)} - (1-\phi)v_1\right) < 1, \tag{17}$$

evaluated at $\overline{c} = u + (\sigma\alpha + \phi(1-\sigma\alpha))\,v_1$. Even when (14) and (15) hold, (17) may still hold, giving,

Proposition 7. *There exists a region of parameter values for which integration with full monopolization reduces both consumer and total surplus.*

Proof. See the Appendix. □

[29]Likewise, $d_{HL} = 0$ would be the optimal choice for $r^S = 1$ and r^O_H sufficiently close to 1, although at $r^O_H = 1$ d_{HL} is irrelevant.

When $\phi = 1$, (16) is positive, since the last term in the right hand side is zero. So the problem for G would be similar to the problem with symmetric publishers, substituting the average α for the fixed value of α. As mentioned above, this was also true under non-integration; with $\phi = 1$, asymmetries have no real impact as we now explain. Consumers' participation decisions depend on the likelihood that they find their preferred publisher and their best-match offline product. Given the likelihood that consumers find their preferred publisher, determined by the average value of r^O, G's profits would also depend only on the likelihood that consumers buy their second-best, instead of their preferred, offline good. So, all that would matter would be the "average" effectiveness for the pattern of traffic. As a consequence, the analysis under integration would be equivalent to the case of symmetric publishers: the ability of G to discriminate according to ad effectiveness would not offer G effective instruments to improve upon symmetric treatment of publishers. There would be simply more degrees of freedom to attain the same relevant values: η, r^O, and r^S.

We could have also considered the case where G only deals with type H publishers, perhaps because it cannot price-discriminate among publishers and then it is in G's interest to set a tariff too high for type L publishers to accept. In this case, both the incentives to distort traffic away from non affiliates and away from low advertising efficiency publishers would coincide. This would reinforce the possibility that integration results in lower consumer and total surplus.

7 Integration with publishers

So far we have considered the integration of the search engine with one intermediary of the display advertising market. Integration with publishers would result in similar incentives for G. In particular, if the integrated entity does not modify how it handles display advertising, then sections 4 and 5 would still describe the effects of integration on G's policies. However, integration with publishers should facilitate the manipulation of the supply of different types of advertising and the coordination of their pricing strategies. The second effect is not present in our stylized model, since under separation prices are already set at the monopoly level and there is no cross price elasticity to internalize, merchants' rents already being zero. However, an integrated entity could still potentially raise its profits by manipulating the supply of different types of advertising.

Consider first the extreme, simplest case, where the search engine would own all publishers but maintains display advertising as in Section 4. There, we showed that, for some parameter values, $r^O < 1$ and $r^S = 0$, implying,

$$\Pi^G = H(\bar{c})\left[(1-\eta)\, m_2 + \eta m_1\right],$$

where $\bar{c} = r^O u + \eta v_1 + (1-\eta) v_2$. However, the integrated entity could choose to eliminate display advertising from publishers' websites. If it did so, it would have no reason to distort organic search, and so would set $r^O = 1$. Also, it could replicate the level of profits per consumer achieved with display advertising by setting r^S equal to what was then the level of η. However, in this case, consumer participation would be higher since r^O would have increased. Hence, it would be in the integrated entity's interest to remove display advertising or restrict it to high margin, that is type 2, merchants only. Notice that the second approach of restricting display ads to higher margin merchants is strictly preferable in the model with $\phi < 1$, because then a fraction $1-\phi$ of participating consumers cannot be reached through search advertising so display is valuable.

A similar argument can be made in the case the search engine owns only a fraction γ of publishers. In this case, if display ads in G's publishers are maintained then, borrowing from Section 5, there is a region of parameter values such that $r_G^O < 1$ and $r^S = 0$. A more profitable policy includes blocking type 1 display advertising in their own publishers, setting $r_G^O = 1$, keeping r_{NG}^O at the same level, and setting $r^S \in \left[\frac{\gamma(\eta_G - \eta_{NG})}{1-(1-\gamma)\eta_{NG}}, \frac{\gamma\eta_G}{1-(1-\gamma)\eta_{NG}}\right]$. With such a policy, the integrated entity makes higher profits per consumer, and induces higher consumer participation. Again, simply shutting down display ads is optimal if $\phi = 1$, but more generally, G would only want to restrict low margin display ads. Summarizing,

Proposition 8. *Integration with publishers may cause a reduction in the supply of display advertising or a restriction to high margin product displays.*

8 Discussion

One of the driving forces behind our model is a potential conflict of interest between G, who can cash on merchants' margins, and consumers, whose net utility from purchases may not be alignned with these margins. For clarity, we have chosen a extrem form of misalignment, but more general specifications can be accommodated by our model. For instance, instead of assuming that the best offline product for a consumer is a low margin one (type 1), and her second best is always a high margin one (type 2), we could assume that these could be any pair of offline products. Such more general formulation opens the door to uninteresting situations in position auctions. First, some queries would involve two type 1 or two type 2 products. In those cases, all relevant merchants would place the same bid and then G's optimal policy would consist of recommending consumers' optimal product (no conflict of interest). Second, some queries would involve a type 2 product as the best match and a type 1 product as the second-best. Once again, there would be no conflict of interest since consumers and merchants rank the two products in the same order. The existence of these segments of the market do not add any insight and have only a quantitative, not qualitative, effect on G's incentives to distort its recommendations.

We may also assume that queries do not provide enough information to identify consumers' ranking of the two products with probability 1. For instance, suppose that those consumers that value category j disagree on the ranking of the two varieties: product $(j, 1)$ is the best match for a fraction θ, and product $(j, 2)$ for a fraction $1 - \theta$. We could still assume that the product category j is all that G can observe about the consumer, but not to which group she belongs to. Clearly, only if $\theta > \frac{1}{2}$ there is a conflict of interest and the value of θ would parameterize the intensity of such conflict. Fixing $\theta = 1$, as we did in the description of the baseline model, does not alter the qualitative results.

In our model, aside from consumer preferences, offline goods differ only in the merchant's margin. This may be justified by differences in access to alternative channels for attracting consumers. For instance, merchants might realize the same margin per sale, m, in both types of goods. However, for type 1 goods, a fraction τ of consumers whose best match is $(j, 1)$ and who did not purchase good $(j, 1)$, neither through sponsored results nor as a result of being attracted by display advertisement in publishers' websites, will still purchase the good through alternative channels, perhaps offline. It is straightforward to show that in the equilibrium of this alternative formulation, the effective internet margin of types 1 and 2 products are $(1 - \tau) m$ and m, respectively. In other words, merchants' willingness to pay in position auctions and display advertising are given by the same formulas of the baseline model just letting $m_1 = (1 - \tau) m$ and $m_2 = m$.

Alternatively, margins may differ because of product prices, which we treated as exogenous. Consider the case where merchants also sell to a set of consumers who never participate in online search and cannot set a different price for offline and online consumers (perhaps regulators prevent price discrimination, perhaps merchants value a reputation for fairness). Merchants with a relatively elastic offline demand will set lower prices and have lower margins. These would be the type 1 merchants and the merchants with less elastic offline consumers would be type 2. Notice that this perspective automatically generates a conflict of interest: if online consumers are symmetric in their gross valuation of the type 1 and 2 goods, then their net value is higher for the type 1 goods which are cheaper. Indeed, $v_1 - v_2 = m_2 - m_1$ in this case.

Auction pricing. Thoughout the paper, we have assumed PPC pricing through auction. The weighted position auction for determining PPCs seeks to capture, in a simplified framework, the mechanism that Google claims to use in reality. The outcome of the auction depends not only on merchants' bids but also on the quality score of different ads, which captures the relevance of merchant products to consumer demands, as inferred from queries, as well as factors such as website quality. In so doing, Google recognizes the possibility of conflicting interests between consumers and merchants that we have just commented. Consumers would like G to position the producers of the type 1 good in the top slot of the sponsored results, but the producers of the type 2 good have a higher willingness to pay for this slot. The choice of μ and r^S reflects G's compromise between these two objectives. We have also assumed PPC pricing for ad space in publishers' websites. If, alternatively, publishers charged merchants per impression (PPI) rather than per click, then merchants' willingness to pay would simply be scaled down by the click-through rate. We also assumed that the CTR for sponsored links and CTR *conditional* on paying attention to an ad in display advertising is 1. The exact value of (a common) CTR is irrelevant under PPC but affects PPI bidding. A lower value of CTR would simply scale down the willingness to pay for positions and display ads.

We have abstracted from any difference in CTR among merchants, one of the dimensions that Google claims to take into account in the weighting of position auction bids. At least as long as we restrict to only one sponsored link, considering such asymmetries would introduce no new conflict of interest between producer surplus and consumer surplus, and could be accommodated to our baseline model. This may be done by assuming that consumer i has some probability of actually liking one and only one of the merchants offering each type $k = 1, 2$ of the category $j(i)$. Then, by observing the snippet, the consumer can costlessly learn whether the sponsored link leads to one of thoese two merchants. This would be similar to the modelling choice of Athey and Ellison (2011).

Search and results. We have modeled a particular order in the queries for online and offline goods. As already noted, our results only require some possibility that content searches result in product discoveries which dissuade product searching on the engine. If some consumers conduct their searches in the opposite order, or some other consumers only engage in one type of search then the link between the display advertising and search advertising markets would be scaled down, but the qualitative conclusions would be the same. We have introduced a slightly more complex pattern of searches there where the particular order of the baseline model was restrictive, namely in Section 6. And then only to the minimum necessary to highlight interesting results.

Our distintion between content and product searches is also extreme in that we assume no sponsored links in the former, and no informative organic links in the latter. Google usually provides relevant organic links, particularly for search queries like "buy Apple computer" which suggest the consumer has a specific merchant in mind. In such cases, Google's reputational incentives may be very strong. Notice that, already in our simplifying model, Google loses nothing by providing marginally less useful organic links to merchants that bid for sponsored links - an observation consistent with merchant complaints that their high organic rankings suddenly dove downwards when they stopped sponsoring search. Moreover, sponsored search results are free to consumers and Google could adjust its auction scoring rules to make sponsored results as reliable, in equilibrium, as any feasible organic results.[30] So, not only for the inattentive consumer who would click on whatever link is shown first, but even for the choosy one, the sponsored link may be the relevant result for a large range of queries for offline goods.

On the other hand, several factors may explain why often there are no sponsored links in the answer to queries related to online content. For instance, even though traffic to publishers has a commercial value, since then it may result in the click on a display ad, there may be too much uncertainty as to what that value may be. Indeed, the query by itself may be very much unrelated to the sort of offline good for which the consumer may have demand, if any. Such information, which is the basis for targeting, is related to previous history of the consumer, and would be difficult to define as part of the "position" offered in a position auction at Google. Therefore, a click in a sponsored link in the results of a query for content may have very little (expected) value, and so organic results are the most relevant ones in those cases. We then take the simplifying, extreme position of ignoring sponsored results. Finally, sponsored search advertising may be irrelevant when the query is for content. This is because when consumers actively seek content, G cannot distract them with ads; distraction only becomes possible once consumers find their content by which point the relevant advertising is display advertising.

We have also assumed that the list of results, organic or sponsored, contains only one relevant item. We have already comment on this assumption, which is a stylized abstraction when clicks are increasingly costlier for the consumer. For instance, if G offers a long list of organic results and commits to place the best match publisher in the top position of the list with probability r^O, and in a random position with complementary probability $1 - r^O$, and consumers incur a sufficiently high cost of making further clicks, they will find it optimal to click only on the first organic result. This would result in a model that is equivalent to the single result one. Likewise, that organic results are irrelevant in the aswer to queries for offline goods may be the consequence of G placing all merchants and internet sites but the sponsored link in nearly random order in the organic results.

Substitutability between display and search advertising. Our research is based on the view that search and display advertising are imperfect substitute services. So, our

[30] Reality is more complex; e.g., the empirical work of Yang and Ghose (2010) suggests organic and sponsored links may complement each other in generating trade (click-through and conversions).

model homes in on the different nature of targeting permitted by search and display advertising which lies at the center of the regulatory analyses, as summarized in the European Commission's case report on Google-DoubleClick: "The targeting characteristic constitutes the essential difference influencing the choices of advertisers. As explained in paragraph 12 of that report, ... [W]hile for search ads the targeting is based on the user's precisely revealed interests (via the search query), for nonsearch ads the targeting is connected with a less precise definition of the consumer's interests, determined by means of criteria such as the context of the visited web page and by its geographical location. Some respondents specified that the differentiation between search and non-search ads merely reflects the "triggering mechanism" that determines which ad to place in the inventory space and the different attitude of the viewer (more focused when using a search query), but that this does not necessarily imply the identification of a separate market." So, the incentives of Google come, partly from the nature of her position as a platform in a two sided environment, and partly from the horizontal incentives to hamper competition by an imperfect alternative for advertising represented by publishers' display advertising.

9 Concluding remarks

In this paper, we have constructed a model that explicitly describes the workings of markets for both search and display advertising. These two modes of advertising are imperfect substitutes for merchants. As a result, a monopoly search engine may have incentives to distort, not only sponsored search (in favor of merchants with higher willingness to pay but lower consumer relevance), but also organic search in order to reduce the effectiveness of display advertising and raise the value of sponsored search.

We have also shown how, in a scenario with symmetric publishers, if the monopoly search engine owns an intermediary in the display advertising market then it may be able to monopolize the entire market. In this context, the incentives to distort organic search, as well as sponsored search, are reduced. As a result, integration raises both consumer surplus and total surplus. However, if there exist any obstacles to full monopolization then the search engine finds it optimal to distort organic search from non-affiliated towards its affiliated publishers. In fact, we argue that the latter effect may dominate so that integration between the search engine and an ad intermediary results in lower consumer surplus as well as lower total welfare. Similar effects may appear when the engine monopolizes intermediation for display advertising but publishers are heterogeneous with respect to how effective advertising in their sites is. An integrated search engine will distort organic search towards more effective sites, and this distortion may well result in lower consumer and total surplus. Finally, if the search engine owns some publishers then qualitative results are also very similar, at least when the search engine cannot bundle the two types of advertising together.

The current draft of the paper neglects some important further issues. First, integration between the search engine and an ad intermediary or a publisher may facilitate behavioral targeting where publishers can use a consumer's past search queries to target display ads more effectively. Second, implications for publishers' incentives to invest in quality are already visible, but merit further investigation.

10 Appendix

10.1 Proof of Proposition 5

The first order conditions with respect to r^O and r^S for both the case of no integration and integration can respectively be written as:

$$H(\bar{c})\frac{(1-\eta)M\left(r^S\right)+\eta\delta m_1}{M\left(r^S\right)-\delta m_1}\left[\frac{u}{\sigma\alpha(1-\beta)}+(v_1-v_2)\left(1-r^S\right)\right]\gtreqless 1. \tag{18}$$

$$H(\bar{c})\frac{v_1-v_2}{m_2-m_1}\left[(1-\eta)M\left(r^S\right)+\eta\delta m_1\right]\gtreqless 1. \tag{19}$$

where $\delta = 0$ under no integration and $\delta = 1$ under integration. Note that the LHS of both equations are higher in the case $\delta = 1$ than in the case $\delta = 0$.

Suppose that under no integration $r^S = r^O = 1$. Then the LHS of both equation (18) and (19) are higher than 1 evaluated at $\delta = 0$ and $r^S = r^O = 1$. Then, they are also higher than 1 at $r^S = r^O = 1$ and $\delta = 1$. Consequently, $r^S = r^O = 1$ is also a candidate solution in case of integration. In fact, it is the only candidate. In any other alternative either $r^S < 1$ or both r^S and r^O are lower than one. But this is inconsistent with the fact that the LHS of (19) decreases with both r^S and r^O.

Suppose that $r^S < 1$, and $r^O = 1$ under no integration. Then, (19) holds with equality. If under integration $r^O < 1$ then $r^S = 0$. But again, this is inconsistent with the fact that the LHS of (19) decreases with r^S and r^O. Since $r^O = 1$ under integration, and given that the LHS of (19) is higher with $\delta = 1$ than with $\delta = 0$, then r^S must be higher than in the case of no integration.

Suppose that $r^S = 0$, and $0 < r^O < 1$ under no integration. If under integration $r^S > 0$, then once again $r^O = 1$. Instead if $r^S = 0$ under integration, then since LHS of (18) under $\delta = 1$ is higher than under $\delta = 0$ then r^O must be higher than under no integration.

Finally, suppose that $r^S = r^O = 0$ under no integration. It could also be the case that under integration $r^S = r^O = 0$ is still the optimal policy. If the LHS of (18) evaluated at

$\delta = 1$ and $r^S = r^O = 0$ is higher than 1, i.e., if

$$H(v_2)\frac{m_2}{m_2 - m_1}\left[\frac{u}{\sigma\alpha(1-\beta)} + (v_1 - v_2)\right] > 1 \tag{20}$$

then under integration r^O is positive. Hence (20) is a sufficient condition to rule out the possibility that under integration the optimal policy includes $r^S = r^O = 0$. Since the LHS of (7) is higher than the LHS of (20) then we conclude that there exist a non-empty set of parameter values defined by conditions (7) and (20), for which integration strictly improves the reliability of the search engine.

10.2 The size of γ and the potential welfare losses of partial integration

Suppose that under no integration $r^S = r^O = 1$. Hence,

$$H(u + v_1)(1 - \sigma\alpha)\min\left\{\frac{u}{\sigma\alpha(1-\beta)}, \frac{v_1 - v_2}{m_2 - m_1}m_1\right\} \geq 1$$

. If we evaluate equation (13) at $r^S = r^O_G = r^O_{NG} = 1$ then this condition will be negative if and only if:

$$(1-\gamma)H(u + v_1)\frac{u(1-\sigma\alpha)}{\sigma\alpha} < 1 \tag{21}$$

Let us consider two extreme cases. First, suppose that $\sigma\alpha$ is very small. Then, condition (21) will hold only if γ is very close to 1. In this case the loss associated with the distortion of traffic to non-affiliated publishers will be small since the fraction of non-affiliated publishers is also small. Next, suppose that we are in Region B; that is, $\frac{u}{\sigma\alpha(1-\beta)} < \frac{v_1 - v_2}{m_2 - m_1}m_1$, and moreover $H(u + v_1)(1 - \sigma\alpha)\frac{u}{\sigma\alpha(1-\beta)}$ is very close to 1. In other words, under no integration G is close to indifferent between setting $r^O = 1$ and or a value slightly below 1. In this case, any value of γ will satisfy condition (21), and hence the impact of traffic distortion can be substantial.

10.3 Proof of Proposition 7

We derive the slope of r^O_H as an implicit function of r^O_L for given r^O:

$$\left.\frac{dr^O_H}{dr^O_L}\right|_{r^O} = -\frac{1-\rho}{\rho},$$

an similarly for given value of η:

$$\left.\frac{dr^O_H}{dr^O_L}\right|_{\eta} = -\frac{1-\rho}{\rho}\frac{\alpha_L(1-\beta) - \beta d_{LH}(\alpha_H - \alpha_L)}{\alpha_H(1-\beta) + \beta d_{HL}(\alpha_H - \alpha_L)},$$

for any values of d_{ij}. The latter is larger, and therefore either $r_L^O = 1$, or $r_H^O = 0$. On the other hand, if $r_L^O < 1$, and so $r_H^O = 0$, then

$$\left.\frac{d\left(d_{HL}\right)}{dr_L^O}\right|_\eta = \frac{1-\rho}{\rho}\frac{\alpha_L(1-\beta) - \beta d_{LH}\left(\alpha_H - \alpha_L\right)}{\beta\left(\alpha_H - \alpha_L\right)},$$

and

$$\left.\frac{d\left(d_{LH}\right)}{dr_L^O}\right|_\eta = -\frac{\alpha_L(1-\beta) - \beta d_{LH}\left(\alpha_H - \alpha_L\right)}{(1-r_L^O)\beta\left(\alpha_H - \alpha_L\right)}.$$

Assume $\alpha_L(1-\beta) > \beta d_{LH}\left(\alpha_H - \alpha_L\right)$, then an increase in r_L^O together with an increase in d_{HL} may keep η constant but increase r^O. So, in this case $d_{HL} = 1$. Similarly, a decrease in d_{LH} and an increase in r_L^O will have the same effect, and so this can happen only if $d_{LH} = 0$. Contrary, assume $\alpha_L(1-\beta) < \beta d_{LH}\left(\alpha_H - \alpha_L\right)$. In this case an increase in r_L^O together with a reduction in d_{HL} may keep η constant but increase r^O. So, $d_{HL} = 0$. But

$$\left.\frac{d\left(d_{HL}\right)}{dr_H^O}\right|_\eta = \frac{\alpha_H(1-\beta) + \beta d_{HL}\left(\alpha_H - \alpha_L\right)}{(1-r_H^O)\beta\left(\alpha_H - \alpha_L\right)} > 0,$$

so that both d_{HL} and r_H^O cannot be equal to zero. This contradiction shows that if $r_L^O < 1$, then $r_H^O = d_{LH} = 0$, and $d_{HL} = 1$. Finally, if $r_L^O = 1$ then d_{LH} is irrelevant, and from the last expression we conclude that $d_{HL} < 1$ only if $r_H^O = 1$, in which case d_{HL} is also irrelevant. So, either $r_L^O = r_H^O = 1$, or $r_L^O = 1, r_H^O < 1, d_{HL} = 1$, or $r_L^O < 1, r_H^O = 0, d_{HL} = 1, d_{LH} = 0$.

10.4 Proof of Proposition 8

Conditions (14), (15), and (17) are satisfied if

$$\begin{aligned}
&((1-\sigma\alpha)\phi + \sigma\alpha)\left(\frac{u}{\sigma\left(\beta\alpha_H - \alpha_L\right)(1-\phi)} - v_1\right) < \frac{1}{H\left(\overline{c}\right)} \\
&< (1-\sigma\alpha)\left(\frac{u}{\sigma\left(\alpha_H - \beta\alpha_L\right)} + (1-\phi)v_1\right),
\end{aligned}$$

and

$$\phi H\left(\overline{c}\right)(1-\sigma\alpha)\frac{v_1 - v_2}{m_2 - m_1}m_1 \geq 1.$$

Note that given all the rest of the parameters, all three conditions are satisfied for large enough values of v_1, as long as $H\left(c\right)$ is bounded below.

11 References

Athey, Susan, and Glenn Ellison. 2012. "Position auctions with consumer search." The Quarterly Journal of Economics 126.3 (2011): 1213-1270.

Borgers, Tilman, Ingemar Cox, Martin Pesendorfer, Vaclav Petricek, "Equilibrium Bids in Sponsored Search Auctions: Theory and Evidence, mimeo, 2007.

Chen, Jianqing, Lizhen Xu, and Andrew Whinston, "Effects of the Presence of Organic Listing in Search Advertising," forthcoming in Information Systems Research, 2012

Chen, Yongmin and Chuan He, 2011, "Paid Placement: Advertising and Search on the Internet," The Economic Journal Volume 121, Issue 556, pages F309–F328, November 2011.

Economist, The, "Making news pay: Reinventing the newspaper." 2011 (Retrieved January 28, 2013, from http://www.economist.com/node/18904178)

Evans, D.S. (2008). "The Economics of the Online Advertising Industry," Review of Network Economics: 7(3), Article 2.

Edelman, Benjamin and Michael Ostrovsky, "Strategic Bidder Behavior in Sponsored Search Auctions," Decision Support Systems 43 (2007), 192-198.

Edelman, B., and B. Lockwood (2011) "Measuring Bias in 'Organic' Web Search," Mimeo.

Edelman, Benjamin, Michael Ostrovsky, and Michael Schwarz, "Internet Advertising and the Generalized Second-Price Auction: Selling Billions of Dollars Worth of Keywords," American Economic Review XCVII (2007), 242-259.

Edelman, Benjamin and Michael Schwarz, "Optimal Auction Design and Equilibrium Selection in Sponsored Search Auctions," American Economic Review C (2010), 597-602.

Ellman, M., and F. Germano (2009) "What do the Papers Sell? A Model of Advertising and Media Bias," Economic Journal, 119, 680-704.

Greenspan, R. 2004. "Searching for balance." ClickZ (April 30),
http://www.clickz.com/3348071.

Hagiu, A., and B. Jullien (2011) "Why Do Intermediaries Divert Search?," Rand Journal of Economics, 42, 337-362.

Hahn, R. W., and H. J. Singer (2008): "An Antitrust Analysis of Google's Proposed Acquisition of DoubleClick," Working Paper.

Hotchkiss, G., M. Garrison, S. Jensen. 2005. "Search engine usage in North America." White paper, Enquiro, Kelowna, BC, Canada.

Jansen, B. J. 2007. "The comparative effectiveness of sponsored and non-sponsored links for Web e-commerce queries," ACM Trans. Web 1(1) Article 3.

Lehaie, Sebastien, "An Analysis of Alternative Slot Auction Designs for Sponsored Search," EC06, 2006.

Liu, De, Jianqing Chen, and Andrew B. Whinston, "Ex-Ante Information and the Design of Keyword Auctions," Information Systems Research XXI (2010), 133-153.

State of Search, 2012, http://www.stateofsearch.com/wp-content/uploads/2010/12
/Google-market-shares-around-the-world-infographic.jpg accessed Feb. 6, 2013.

Varian, Hal R., "Position Auctions," International Journal of Industrial Organization XXV (2007), 1163-1178.

White, Alexander, 2009, "Search Engines: Left Side Quality versus Right Side Profits," Toulouse School of Economics Mimeo.

Wilbur, K.C. (2008) "A Two-Sided, Empirical Model of Television Advertising and Viewing Markets," Marketing Science, 27, 356-378.

Xu, Lizhen, Jianqing Chen, and Andrew Whinston. 2010. "Oligopolistic Pricing with Online Search" Journal of Management Information Systems, Vol. 27, No. 3, pp. 111-141.

Xu, Lizhen, Jianqing Chen, and Andrew Whinston. 2011. "Price Competition and Endogenous Valuation in Search Advertising," Journal of Marketing Research, Vol. 48, No. 3, pp. 566-586.

Yang, Sha and Anindya Ghose, 2010, "Analyzing the Relationship Between Organic and Sponsored Search Advertising: Positive, Negative, or Zero Interdependence?" Marketing Science, January/February 2011 vol. 30 no. 1.