

In Google we trust?

Roberto Burguet*, Ramon Caminal*, Matthew Ellman*[†]

January 2014

Abstract

We develop a model of the interacting markets for online content and offline products. We portray content providers and the search engine as competing platforms that intermediate in the product market (a horizontal relation), while also vertically related in the content market. Explicitly modeling both markets allows us to characterize the substitutability (and manipulability) of search and display advertising, and its effect on the incentives to distort organic search results as well as spillovers on the reliability of sponsored search results. Specifically, improvements in the technology for targeting display ads increases this substitutability and the threat of organic search distortions. Integration of the search engine that results in full monopolization of the display ad market improves search reliability and raises consumer and total welfare, if content providers are similar. However, partial integration, or full integration when content providers differ in their ad effectiveness, introduce additional incentives for distortion and may reduce consumer and total welfare.

Keywords: Search engine bias, internet economics, vertical integration, two-sided markets, antitrust. *JEL Classifications:* L13, L41, L82, 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.

[†]We thank Konstantina Bania, Michael Baye, Stefan Bechtold, Jan Boone, Emilio Calvano, Lisa George, Shane Greenstein, Andrei Hagiu, Sjaak Hurkens, Michael Katz, Yossi Spiegel, Catherine Tucker, Alexander White, Joel Waldfogel and seminar participants at several conferences (Évora ICT, 2013, Tilburg, TILEC III, 2013, Chicago Searle Internet Conference, 2013, Paris SAET, 2013 for useful comments. We also acknowledge financial support from the Spanish Ministry of Science and Innovation (Grant: ECO2011-29663).

“[W]e expect that advertising funded search engines will be inherently biased towards the advertisers and away from the needs of the consumers.” - Brin and Page (1998)

1 Introduction

This paper investigates the strategic interaction of a dominant search engine, which we denote by G , with online content providers and other internet agents. We characterize G 's incentives as a traffic manager, and analyze how these incentives and welfare outcomes change when G expands into content provision and ad intermediation.¹

Our research starts by explicitly modeling a search engine as a platform that intermediates two different but interacting markets. The search engine helps consumers match with providers in the market for online content (e.g., news, entertainment, dictionaries). Also, it helps consumers match with sellers in the market for offline products (e.g., apparel, electronics, holiday packages). These markets are not independent. The key interdependence is that content providers, which we call publishers, are not only customers of the search engine in the market for online content, but also competing platforms in the market for offline products. Concretely, sellers of offline products, which we call merchants, can reach consumers on G 's search results page (via sponsored search links) *or* on publishers' content pages (via display ads). Therefore, when G manages traffic in the market for online content, it does also affect the traffic to competing platforms in the market for offline products.

We propose a stylized model of these two markets and their interaction. Consumers decide whether to use G , and those who do, enter queries for both offline products and online content.² G answers the queries with (paid-for) sponsored and (non-paid-for) organic results. G chooses the reliability of both these types of result, possibly biasing them in the sense of distorting away from what is optimal for consumers. In the model, consumers click on sponsored results when seeking products and click on organic results when seeking content.³ While consumers always need G to find relevant content, they may find products via display advertising on publishers' content websites *instead* of via G .

Figure 1 depicts the two interacting markets, with dashed arrows for content and solid arrows for products. The downward arrows indicate a consumer visiting the two-sided platforms (shown as boxes) of G and publishers, and the upward arrows depict providers (shown as triangles) successfully posting links on these platforms. In particular, the merchant shown can reach the consumer by posting a

¹With market shares exceeding 90% in most European countries and a global average above 80%, Google arguably dominates online search in most of the world (State-of-search, 2012). Google has recently expanded into ad intermediation (AdSense and AdWords) and content provision (e.g., Google Finance, YouTube and Zagat). Although search engines may be “just one click away” as is often argued, indirect network externalities, combined with small switching costs, habit effects and delays in detecting reduced search quality, interfere with competitive forces.

²In Broder's (2002) taxonomy, these correspond, respectively, to “informational” and “transactional” searches.

³We will argue that this separation is empirically founded. We also offer sufficient conditions. While highly stylized, this keeps the model simple and helps highlight how the two markets interact.

sponsored search link on G but also by posting a display ad on the publisher shown; this illustrates the substitutability of the search and display ad channels.

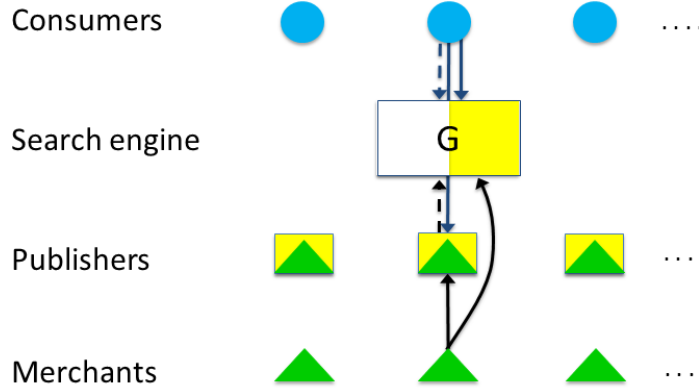


Figure 1: The two interacting markets

The interdependence of these two markets makes G and the publishers both vertically and horizontally related. By increasing the reliability of its results, G increases consumer participation, which also benefits publishers (a vertical externality). However, increasing the reliability of organic results also raises the effectiveness of display advertising on publisher platforms at the expense of sponsored search advertising on G 's search platform (a horizontal externality). So G has an incentive to instead distort organic search towards publishers that are less effective for display advertising, whether intrinsically, by choice (as for Wikipedia which carries no ads), or by poor match (consumers tend to spend less time on less relevant publishers' webs, exposing them less to display ads).⁴

Thus, G 's incentives to bias organic search lie in the substitutability between G 's (sponsored) search ads and publishers' display ads. We characterize this substitutability, identifying the role of technologies for targeting merchants' display ads to publishers' content.⁵

G 's incentives to bias sponsored search lie in a potential conflict of interest between consumers and merchants over ranking sponsored links: the merchants most willing to pay for a top position are not the best option for consumers whenever net margins and net consumer values are not aligned. In our model, merchants buy (search) ads from G and (display) ads from publishers, with all prices set by auction. G 's resulting share of merchant profits gives it an incentive to sell its search ads to high margin merchants by underweighting merchants' relevance or value to consumers in the scoring auction for G 's sponsored links.

⁴In line with this, Google AdSense recommend that publishers seek relevant consumers as a way to raise their display ad revenues, for a given content quality. For direct empirical support, see the growing body of eye-tracking-based search studies; e.g., Lorigo et al., (2008) and in particular, Wanga and Day (2007). Ellman and Germano (2009) and Wilbur (2008) offer further evidence and richer views on the relationship between content and advertising effectiveness.

⁵Substitutability lies at the heart of regulatory debates (see FTC 071-0170 and the contrasting EU report, M.4731).

Tempered by the need to attract consumers to search on G , these incentives to bias organic and sponsored search, interact as imperfect substitutes for G . By explicitly modeling both markets, we are able to conduct comparative statics to predict how the two search reliabilities respond to market characteristics and shifts in technological parameters. Thus we find that sponsored search results worsen but organic search improves when the conflict of interest between consumers and merchants rises. Meanwhile, fixing the conflict of interest and increasing the relative profitability of the offline product market worsens organic search reliability and improves sponsored search. Also importantly, the improvements in display ad targeting that have been enabled by recent technological innovations, increase the substitutability of display and sponsored advertising, which induces G to offer less reliable organic search results but more reliable sponsored search results.

Having characterized G 's incentives in the two markets, we study the effects of G 's integrations with publishers and ad intermediaries. Integration induces G to, at least partially, internalize both horizontal and vertical externalities. First, with symmetric publishers, integration that fully monopolizes ad intermediation or publishing improves the reliability of G 's search results, because G fully internalizes profits from display advertising. As usual, the internalization of the vertical externality on publishers increases G 's value of total demand (consumer participation), which here requires more reliable search results. Usually internalization of the horizontal externality points in the opposite direction, but here it complements the vertical internalization in raising total demand. This is because a non-integrated G "steals" business from publishers, not by lowering price, but by distorting organic search to make publishers' display ads less effective; eliminating business-stealing therefore raises search reliability and raises consumer participation.

Second, integrations involving only a fraction of publishers bias G 's organic search in favor of G -owned or G -affiliated publishers. Such partial integrations actually raise G 's incentives to steal business from independent publishers. We identify conditions under which this effect dominates the effect discussed in the previous paragraph, and so partial integration results in lower total welfare and lower consumer surplus than non-integration.

Third, even integrations involving all publishers (full monopolization) can have negative consequences for consumer and total surplus when publishers are asymmetric. Specifically, when publishers vary in their effectiveness as platforms for display advertising, integration creates an incentive for G to divert traffic from less to more effective publishers. This incentive is stronger than the converse incentive (to divert traffic from more to less effective publishers) under non-integration. Again, we identify conditions under which the net effect of integration is lower consumer and total welfare.

An additional effect of integration into ad-intermediation is a reduction in publisher rents. As a

result, integration may have an adverse effect on the quality of publishers' content.⁶

Summarizing, our contribution is to develop a model of the interacting markets for online content and offline products. We portray publishers and the search engine as competing platforms that intermediate in the product market (a horizontal relation), while also vertically related in the content market. Explicitly modeling both markets allows us to characterize the substitutability (and manipulability) of search and display advertising. This, in turn, allows us to explore G 's incentives to distort organic and sponsored search results, and to analyze how market and technology characteristics affect these incentives. In particular, we show that improvements in the technology for targeting display ads increase this substitutability and with it, the risk of organic search distortion. Regarding integration into content provision (publishing) and ad intermediation, we show that the consequences for search reliability and consumer and total welfare depend critically on publisher heterogeneity (in display ad effectiveness), and on the fraction of publishers that become owned by or affiliated with the search engine.

Our analysis relates to work on platform competition with multi-homing consumers and gatekeeping, as well as work on organic and sponsored search. A novel feature compared to recent studies of multi-homing,⁷ is that one platform (the search engine) is a gatekeeper whose control over links to publishers affects the degree of multi-homing.⁸

A number of recent papers study either organic or sponsored search. White (2013) presents the insight that a non-integrated search engine may bias organic search to reduce downstream merchant competition and thereby extract higher rents from sponsored search, but he does not model the content search nor the publisher platforms that compete for merchants' ad custom.⁹ White and Jain (2012) postulate reduced-form ad revenues and introduce ad-nuisance effects in a model where search engine and publishers are complements. Thus, coordination between platforms allows them to internalize the vertical externality, as in our model. Even if we interpreted ad-nuisance as a metaphor for advertising competition, horizontal externalities would be absent from their paper. Also, in independent research, de Cornière and Taylor (2013) study search bias and integration. They model search for content but, as in White and Jain (2012), they use a reduced-form approach without search for products and with ad-nuisance. Instead of complementarity between platforms, they assume exogenous substitutability between advertising on G 's and publishers' websites, which motivates an organic search bias. They also study the impact of integration with one of two asymmetric publishers under this assumption.

By instead explicitly considering micro-foundations of both the market for content and the market for products, we can endogenize substitutability between display ads and sponsored ads, and also

⁶This is of particular concern for ad-financed independent journalism on the internet.

⁷See e.g., Athey et al., (2013), Anderson et al., (2013), Ambrus et al., (2013), White and Jain (2012).

⁸Baye and Morgan's (2001) early model captures an internet gatekeeper controlling links to merchants but has no competing platforms.

⁹In White (2013), merchants cannot choose between display and search advertising; the focus is on downstream pricing.

substitutability between organic and sponsored distortions from the engine’s point of view. Moreover, this permits comparative statics exercises, including the effect of market conditions and technological parameters on these substitutabilities.

One element of our model is the potential existence of a conflict of interest in the product market. The auction literature on product search has focused on this issue and the role of the mechanisms used to allocate sponsored links. Early work characterized generalized second-price position auctions and investigated auction design.¹⁰ Athey and Ellison (2011) and Chen and He (2011) show how search auctions generate advantageous selection of those privately-informed merchants most attractive to consumers; essentially, consumers seek merchants that will attract them to click-through and buy. In some of these studies, conflicts of interest arise from private information of merchants about their value per click (e.g., in Athey and Ellison, 2011, engines distort rankings to extract merchants’ rent) or to private consumer ability to evaluate product relevance (as in Athey and Ellison, 2011, or Hagiu and Jullien, 2011, where engines distort results to induce consumers to search more, generating a positive externality on merchants). In other studies, engines distort search to relax downstream product market competition, thereby raising merchants’ value of advertising (see e.g., Chen and He, 2011, Eliaz and Spiegler, 2011, Hagiu and Jullien, 2011, White, 2013, Xu et al., 2010 and 2011).

Xu et al., (2012) and Taylor (2013) consider a model with only product search, and focus on the fact that reliable organic results constitute “self-cannibalization” by G , since organic links give merchants a free substitute to sponsoring links on G . This is also present in our model. In fact, removing the unexplained restrictions considered in these papers, organic results become superfluous.¹¹

Search bias is not so easy to detect, but Tarantino (2013) and especially Edelman and Lai (2013) provide convincing empirical evidence of bias of Google towards integrated publishers and merchants.¹²

The paper is organized as follows. Section 2 presents the baseline model and characterizes the social optimum, before analyzing the equilibrium of the game with full separation in section 3, with monopolizing 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 analyze welfare and surplus implications. Section 8 discusses assumptions and extensions and we conclude in section 9, gathering proofs in the Appendix.

¹⁰See e.g., Lehaie (2006), Edelman et al., (2007), Varian (2007), Börgers et al., (2007), and on design, Edelman and Schwarz (2010), Liu et al., (2010) and Athey and Ellison (2011). Paid-for prominence (Armstrong and Zhou, 2011) and paid-for priority (Choi and Kim, 2010) are also related.

¹¹In Xu et al., (2012), consumer clicks on organic and sponsored links are independent of link reliabilities. In Taylor (2013), search engines suboptimally offer only one sponsored link, and sponsored results implicitly cannot use information used in organic results.

¹²Tarantino (2013) adds the point that a general search engine integrated into vertical search would divert traffic towards its vertical search engines, which he models as merchants (travel intermediaries often sell travel products), omitting the 2-sided platform aspect.

2 The baseline model

We first present the essential elements of the baseline model, justifying the key assumptions at the end of the section; section 8 discusses further extensions. We start by describing the two markets, for online content and for offline products; next, we present the two types of platforms that operate in these markets. We then explain the different platforms' pricing rules, followed by the timing of decisions.

Online content and offline products. A mass one of consumers, indexed by i , value specific varieties of online goods, called “content”, available on publishers' websites, and produced at zero cost. Each of N publishers 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” content, $n(i)$, that generates net utility, $u > 0$, while any other content, $n \neq n(i)$, generates zero net utility and further units imply a net loss. Consumers do not know the identity of the publishers offering their preferred content and require the help of the search engine.

Each consumer also has a unit demand for offline goods, called “products.” Products vary by category $j \in \{1, 2, \dots, J\}$ and type $k \in \{1, 2\}$, giving $2J$ products defined by pairs (j, k) . Consumer i only values one category, denoted $j(i)$, with one unit of product $(j(i), k)$ giving i a net benefit v_k where $0 < v_2 < v_1$; so $(j(i), 1)$ and $(j(i), 2)$ are i 's “best-match” and “second-best” products, respectively. All other products or additional units imply a net loss. As with content, consumers do not know the identity of the merchants offering their preferred offline goods.

Merchants each sell one product, which they make available on their websites. We call a merchant type k if its product is type k ; such merchants earn a unit margin m_k . Two merchants produce each product, implying $4J$ merchants and ensuring competition for all advertising opportunities.¹³

We treat the symmetric case where each product category and each publisher's content interests the same fraction of consumers, $\frac{1}{J}$ and $\frac{1}{N}$, respectively.

Platforms. G intermediates the market for content. Consumers, being unaware of the identity of their preferred publisher, must turn to G to facilitate their search. They type in a query describing their desired content. G can perfectly interpret each query and identify the relevant publisher, $n(i)$. G responds by providing a single link to a publisher's website. For now, we assume this link is organic (non-paid for) and leads to some publisher's content. G may or may not provide the link to the consumer's favorite content. In particular, with some probability, denoted r^O , the link leads to the site of the consumer's best-match publisher, $n(i)$, and otherwise to some other publisher's site. G chooses $r^O \in [0, 1]$.

G also intermediates the product market. Consumer i may conduct a product search in G by typing

¹³This competition is not critical but simplifies the analysis as it implies zero merchant profits in equilibrium.

in a query describing i 's favorite product category, which G can interpret perfectly. G can also identify the four relevant merchants producing goods in category $j(i)$ and their types. G answers with a single sponsored result and no (relevant) organic result. We again rationalize this behavior below. G uses a weighted second-price auction to allocate the single sponsored result, as described in detail shortly.

Unlike in the content market (where only G intermediates), in the product market, publishers offer additional intermediating platforms. Each publisher n offers one “display ad” slot for a link to a merchant’s website, whenever a consumer i visits n ’s website to consume its content and i notices the ad with probability $\alpha \in (0, 1]$ if i is visiting her best-match publisher $n(i)$ and probability $\alpha\beta$ if i is on any other publisher’s web, where $\beta < 1$.¹⁴ Display ad targeting is less accurate than for search ads (which are informed by consumer queries). Publishers, aided by ad intermediary targeting technologies, only observe a signal $s(i) \in \{1, 2, \dots, J\}$ of each visiting consumer’s product category interest. This signal correctly indicates i ’s preferred product category with probability σ : formally, $s(i) = j(i)$, with probability σ and otherwise points to a random distinct category.¹⁵

Summarizing, as shown in Figure 1, publishers provide online content to consumers but they are also platforms that bring together consumers and merchants. At the same time, the search engine is a platform that intermediates in both markets. By responding to content queries, G is the only platform for matching consumers with content publishers, which are themselves match-makers in the market for offline products. By responding to product queries, G directly matches consumers and merchants.

Auctions for ads. When a consumer i enters a query for a product in G ’s website, all merchants “observe” the query and bid for the sponsored link. Bidding determines the “pay per click” price, denoted “PPC”. G neglects any bid by merchants selling products in category $j \neq j(i)$, and partially discounts the bids of merchants selling product $(j(i), 2)$ by a factor $\mu \leq 1$, which is an endogenous variable selected by G . The winner of the auction is determined by comparing the weighted bids.¹⁶ The PPC rate is set equal to the second highest bid. In case of a tie, if the set of winners include both types of bidders, then G selects a type 1 merchant with probability r^S and a type 2 merchant with probability $1 - r^S$. Also G always assigns equal probabilities to any tying merchants of the same type. In equilibrium, merchants of the same type submit the same bid. Moreover, we will see that G chooses μ so that the types always tie (the four relevant merchants tie in each auction). Hence, r^S represents the “reliability” of sponsored search results just as r^O represents the “reliability” of organic results.

Publishers also auction the ad space on their websites for each consumer visit using a standard second price auction (a weighted auction offers them no advantage). That is, they conduct an auction

¹⁴Typically, $\alpha < 1$ since display links must distract consumers from the publisher’s content to generate click-through.

¹⁵Targeting precision satisfies $\frac{1}{J} \leq \sigma \leq 1$ since $\sigma = \frac{1}{J}$ if the ad intermediary has no information on consumer preferences and $\sigma = 1$ if it can identify the consumer’s best-match category with probability one. $n(i)$ and $s(i)$ are independent.

¹⁶Let b_k^{\max} denote the maximum bid of merchants offering product $(j(i), k)$. If $b_1^{\max} > \mu b_2^{\max}$, a type 1 merchant wins. Conversely, if $b_1^{\max} < \mu b_2^{\max}$, a type 2 merchant wins.

for the PPC for displaying an ad, contingent on the signal $s(i)$ of the consumer's preferred product category.¹⁷ Merchants submit bids and the publisher displays one highest bidding merchant's ad to this consumer. Publishers have no costs except possible charges for ad intermediation, which we can ignore for the time being.¹⁸

Consumer search. We model consumers' participation in using G 's search engine as a single decision that depends on the overall reliability of sponsored and organic results. Each consumer i has a joint cost c_i of using the search engine for both content and product queries; this includes the foregone expected benefits from alternative search. 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 $F(c)$, such that the reverse hazard rate, $H(c) = \frac{f(c)}{F(c)}$, is decreasing.¹⁹ Each consumer observes her cost c_i privately and prior to deciding participation. Gross of her direct plus opportunity costs given by c_i , i 's expected gain from participating in online search is the sum of expected gains from consuming online content plus offline products, found via a display ad during content search or via a search ad during product search. The highest possible such gain is $u + v_1$. We assume $c_H > u + v_1$ so that consumer participation in online search is interior in any equilibrium.

Conflict of interest. Intuitively, the interests of consumers and merchants could be aligned for some products but misaligned for others. That is, m_1 could be higher than m_2 , whence, given $v_1 > v_2$, the type of goods preferred by consumers would also be the type that generates the highest profits. But as we justify below, consumers and merchants' interests may also be misaligned: $m_1 < m_2$. In general, interests will be misaligned for some fraction of searches (some fraction of product categories). All we need is that this fraction is non-zero, but for simplicity we assume the fraction is one, that is, $m_1 < m_2$; nothing qualitative is affected. Moreover, to simplify the welfare analysis, we focus on the case $m_1 + v_1 > m_2 + v_2$, so the social optimum has only type 1 transactions.

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

¹⁷Rather than reveal their signals to merchants in real-time, publishers and G in its search auctions, make their information available for interaction with merchants' bidding strategies in automated auctions. But it is equivalent to describe as if merchants bid after observing the information.

¹⁸These charges are zero in the case without integration, because intermediaries have zero costs and we assume Bertrand competition. So we need only model ad intermediaries explicitly when the search engine owns one.

¹⁹This ensures that G faces an increasing 'marginal factor cost' of attracting the consumer base it "sells" on to merchants.

a product query and can visit the website of the merchant that appears in the sponsored search result and can then buy that merchant’s product. Merchants, publishers, and consumers always observe the outcomes of previous stages.

None of our results depend on consumers always conducting their searches in this exact order (first content and then product). All we need is that a positive fraction of consumers buy from merchants that they find via publishers’ display ads, thereby reducing their need for G in the product market.

2.1 Discussion

We now show how several features of the model that might appear simplistic or far-fetched are actually equivalent to richer and more realistic representations.

Search results. We have assumed that G provides exclusively organic results for content queries and sponsored results for product queries. While clearly extreme, this split is consistent with empirical studies finding that people use sponsored links more than organic results when conducting product searches or “e-commerce search queries” (Greenspan, 2004, Jansen, 2007), but place more trust in organic results when seeking content (Hotchkiss et al., 2005, Jansen and Resnick, 2006).

In terms of the logic of the model, it is very intuitive that, in the case of product queries, merchants are willing to pay to appear prominently in the list of the search engine’s results. Also, it is in G ’s interest to withhold useful organic links in order to avoid cannibalization, as we explain in the introduction. Clearly, nothing changes in our model if G offers organic results that could be valuable for consumers but that are dominated by sponsored results, since consumers will never click on them.²⁰

In the case of content queries, in principle nothing in our model would prevent publishers from participating in position auctions. In the real world, they seem not to do so. Multiple factors may explain this behavior. A full explanation goes beyond the scope of this paper, but we mention two reasons. One is that publishers may have negligible willingness to pay on each individual query, owing to transaction costs. Having to constantly adapt bidding strategies, over a myriad of potentially relevant keywords, to changes in content and query patterns could be very costly for publishers with highly dynamic content, such as news websites. This dissuades publishers from bidding. Another possible reason has to do with the informational nature of publishers’ content, which makes observability and therefore reputation-building particularly difficult. In some cases, publishers may attempt to build

²⁰In fact, in our model, G suffers no loss if the link to the winner of the position auction is displayed twice, once marked as an sponsored result and the second time marked as an organic result. Consumers would then be indifferent and would presumably click on the organic result with some probability γ . Merchants would not pay directly when their organic results are clicked, but their willingness to pay for a sponsored click would be multiplied by $\frac{1}{\gamma}$. G would still need to suppress organic links to merchants that do not bid on G ’s sponsored auctions (some merchants have accused G of this but we are not aware of any proven evidence). All our results would remain unchanged. Yang and Ghose (2010) and Blake et al. (2013) provide mixed field study evidence on whether Google actually removes organic links to merchants that stop bidding for sponsored slots, based on experiments in which merchants stopped bidding in position auctions; note that the latter paper studies Ebay and may be explained by the reputation discussion below.

a reputation for providing reliable content by minimizing the visibility of their commercial interests. Bidding on sponsored links to expand their audiences may hurt their reputation by making their profit motive more salient.

We have assumed that G offers a single result, both for content and product searches. Within the model this is in fact optimal since G knows all the relevant information about consumers' needs, and consumers have a unit demand for both online and offline goods. In reality, most searches generate multiple results, which makes sense for consumers seeking multiple goods, but only if some of the goods demanded are sufficiently similar for a single query to identify them together. A more important reason for multiple results is that queries are often ambiguous and consumers retain private information or ability to evaluate results.

Joint participation constraint. We model consumer search participation as a single decision, for both product and content searches. This simplification captures the fact that high quality results in either type of search tend to spill over into improving G 's *overall* popularity or reputation as a reliable search engine. One explanation for this spillover is that consumers tend to develop a habit of using a fixed engine, rather than adapting each search to the specific search need of the moment. An independent but complementary explanation builds on observability. Consumers tend to learn about the quality of search from friends, from media reports and from their own experience. Such learning tends to be coarse, rather than fully contingent on each type of search, because communication is limited and memory and aggregative skills may be limited too. In this context, even if consumers take independent participation decisions in content and product searches, if G changes the reliability of one type of result (say, organic) this would affect its overall reputation, and hence affect consumer participation in product, as well as content searches. Modeling consumer participation as a single decision captures such spillovers in a simple and tractable fashion.

Conflict of interest. We are assuming that consumers' and merchants' always conflict, but as already discussed, we only need that they sometimes conflict. A simple example of such conflict arises when all merchants have the same cost and consumers gain the same gross (of price) utility from either relevant merchant's product, but the two types of products differ in price. Price differences can arise endogenously, remaining orthogonal to the online search for products, if merchants also sell to local markets of offline consumers whose elasticities differ and if regulatory or reputational concerns prevent merchants from price discriminating between online and offline consumers.

The scoring auction. The weighted position auction for determining PPCs seeks to capture, in a simplified framework, the mechanism that Google claims to use in reality.²¹ The outcome of the auction depends not only on merchants' bids but also on their quality scores, which capture the relevance of

²¹See Hal Varian's Youtube video on Google's scoring auctions, or Varian (2007).

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 described. Consumers would like G to position the producers of the type 1 good as the top sponsored result, 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. Google's claims would suggest that $r^S = 1$, with μ purely serving to prevent type 2 merchants from winning, but our theory and Brin and Page's (1998) early remarks suggest that this may not be the case.

Reputation and commitment. In the model, G sets μ , r^S and r^O and then consumers observe this before deciding whether or not to participate. Implicitly, this assumes that G can commit perfectly to any distortion strategy, or reliabilities r^O and r^S , that it wishes to adopt. The commitment assumption captures in a static model the idea that, over time, G can build a reputation for reliable search results. A key constraint on reputation building is the difficulty consumers have in observing reliability. Perfect observability requires consumers to know the quality of the results that G *could* have provided, as well as seeing those it actually provides. This extreme case is implausible, but consumers can certainly evaluate their own experiences and may learn from each others' experiences. So, commitment is feasible indirectly through reputation. As discussed above, this reputation and social learning interpretation can justify the consumer joint participation assumption and the role of organic results.²²

2.2 The social planner's problem

Before analyzing the model, we describe the benchmark optimal outcome for a social planner who cares about the sum of all agents' surpluses and can control how the search engine matches consumers with merchants and publishers, how publishers allocate display ads among merchants, and which consumers participate in online search. For any given participation level, the best possible outcome from this total surplus perspective is 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, $n(i)$, and can send any product-searching consumer to a best-match merchant, that is, one producing $(j(i), 1)$. This is also necessary for efficiency. First, content search must be undistorted since consumers have no alternative way to find content. Second, product search must be undistorted, because the alternative channel, display advertising, at best permits consumers to find their best-match products with probability $\sigma\alpha < 1$.²³ Abusing notation in anticipation of the equilibria where merchant

²²Even naïve consumers would detect bias if G 's organic results on a product search were empty or entirely useless, or suppress obviously relevant merchants (such as Ebay on a search query for "Ebay camera"). Similarly, regulators may be able to punish such blatant forms of bias. Indeed, recent FTC and EU investigations into Google's search results centered on bias towards Google-owned publisher where evidence is sharpest.

²³In the baseline model, display ads are redundant because product search has no imperfections, nor added costs given participation; see sections 6 and 7 for an extension. Display ads are also harmless because consumers ignore irrelevant or type 2 display ads, but see 8 on compulsive consumers.

types tie, 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 merchants – a type 1, relevant merchant. So we have,

Proposition 1 *Total surplus maximization requires the search engine to allocate traffic with no distortion from the consumer’s ideal; in the first-best, $r^O = r^S = 1$.*

This proposition also holds in the constrained scenario where the planner cannot control consumer behavior (search participation and trade). In fact, the need to attract these consumers, who neglect the positive externality of their participation on producer surplus, only reinforces the planner’s incentives to not distort search.

3 Equilibrium analysis under vertical separation

Throughout the paper, we solve for subgame perfect equilibria in undominated strategies. Consumers make many decisions but most are immediate once stated. As already noted, consumers always click on the single link after entering either type of query; that is, they follow G ’s “recommendation.” Similarly, consumers only ever buy a product from their category of interest and they never buy more than one unit overall. In the product search stage, consumers have no subsequent chance to find attractive products, so they buy the advertised product whether type 1 or type 2, provided it is from the relevant category, which it is in all equilibria of the baseline model. Anticipating this, in the prior content search stage, consumers only buy a displayed product if it is type 1 as well as relevant.²⁴ Consumers omit the product search if a display ad satisfies their demand for products since they only demand one unit, but they always gain from content search, if they paid their cost of participating in online search. Participation is the only remaining non-trivial consumer decision and is characterized below.

Search auctions. We begin by studying the optimal design of sponsored search auctions and merchants’ equilibrium bidding strategies. Consumers who did not purchase offline goods through display advertising enter their product query in the search engine. Four merchants, two of each type k , are potentially relevant and correctly anticipate that every click on their search ad leads to a purchase, so type k ’s are willing to pay a PPC of m_k to appear in the single slot of G ’s sponsored search results; merchants’ bids have no impact on their alternative sales options. As in unweighted second-price auctions, each merchant’s unique weakly dominant strategy is to bid her willingness to pay, $b_k = m_k$:

Lemma 1 *For any μ , r^O , r^S , the strategy profile $(b_1, b_2) = (m_1, m_2)$ is the unique equilibrium.*

²⁴A type 2 display ad is less attractive than continuing to a product search which offers some chance of a relevant type 1 product; $r^S v_1 + (1 - r^S) v_2 \geq v_2$. If $r^S = 0$, consumers would also be willing to buy type 2 products via display ads, but this cannot occur in equilibrium; the expectation of type 2 display purchases would lead G to set r^S slightly above 0 in order to induce type 1 display purchases, thereby raising consumer participation.

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 two types of merchant's effective willingness to pay, allowing G to use its tie-breaking rule r^S to fine-tune the probability, then equal to r^S , that a type 1 merchant wins the position auction. G need only consider this last case, since G can always set $r^S = 1$ and 0 to generate outcomes equivalent to higher and lower μ , respectively. Competition among merchants implies that, in equilibrium, the winner always pays its willingness to pay. Summarizing:

Lemma 2 *G optimally sets auction weight $\mu = \frac{m_1}{m_2}$, so that a relevant type 1 merchant wins the auction, paying m_1 per click, with G 's chosen probability r^S , and a relevant type 2 merchant wins, paying m_2 per click, with probability $1 - r^S$; G 's average revenue equals the average margin, denoted $M(r^S)$, on sponsored-search-mediated sales: $M(r^S) = r^S m_1 + (1 - r^S) m_2$.*

Display auctions. We now turn attention to the second-price auctions for display advertising that take place whenever a participating consumer lands on a publisher's webpage, during her first stage of search, her content search. Merchants compete by bidding their willingness to pay per click. They anticipate zero rents from any product searches, so merchants have no opportunity cost of winning a display ad, nor any indirect benefit. Clearly, type 2 merchants cannot gain from bidding for display ads, given that consumers only ever buy type 1 products in the content stage. The merchants indicated by the targeting signal – those selling $(s(i), 1)$ – anticipate that if they win a display ad, a fraction σ of clicks will yield sales, so they bid σm_1 , while other type 1 merchants only bid $\frac{1-\sigma}{J-1} m_1$. The merchants selling $(s(i), 1)$ always win, but their sales per click rate, σ , is still strictly below the full unit rate in search advertising.²⁵ So we have,

Lemma 3 *Publishers optimally conduct unweighted auctions and the type 1 merchants indicated as most relevant by the targeting technology bid σm_1 per click and one of them wins.*

Recall that a consumer visiting a publisher website is attracted by a display ad with probability α if visiting her favorite publisher and $\alpha\beta$ if visiting any other publisher. So in equilibrium, the average fraction of clicks on display ads is $\alpha e(r^O)$, where $e(r^O) = r^O + (1 - r^O)\beta$ and is increasing in r^O since $\beta < 1$. A fraction σ of these clicks end up in a transaction, so the fraction, η , of participating consumers who buy via display advertising (always their best-match product) is given by,

$$\eta = \sigma \alpha e(r^O) = \sigma \alpha (r^O + (1 - r^O)\beta).$$

²⁵Publishers have no incentives to use a weighted auction to allocate their display ad slots, since type 2 merchants never bid. In any case, publishers would internalize little of the consumer participation benefit from type 1 display ads since N is usually large.

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

Consumer participation and continuation equilibrium. Consumer participation is determined by the expected benefit from search participation, which we denote by \bar{c} ; all consumers with a lower cost c_i will participate. So $X = F(\bar{c})$, where \bar{c} is the sum of three expected net utilities from i consuming, respectively, online content (net value u or 0), her best-match product (net value v_1) and her second-best product (net value v_2):

$$\bar{c} = r^O u + v_1 [\eta + (1 - \eta) r^S] + v_2 (1 - \eta) (1 - r^S). \quad (1)$$

The probability in brackets of finding her best-match product sums the probabilities η via display advertising and $(1 - \eta) r^S$ via search advertising, whereas she only ever consumes her second-best product via search advertising, in the probability $(1 - \eta) (1 - r^S)$ event that she neglects display ads during content search and G shows a type 2 merchant. Note that \bar{c} increases with r^O , r^S and σ . So if G distorts traffic, by setting low values of r^O and/or r^S , it pays the cost of reduced consumer participation. The two instruments, r^O and r^S , play a similar role in encouraging participation and a high value of one reduces the sensitivity of participation to the other; that is $\frac{\partial^2 \bar{c}}{\partial r^S \partial r^O} < 0$. Drawing all this together, we have:

Proposition 2 *In the unique continuation equilibrium following any (r^O, r^S) , merchants make zero profits, each publisher earns $\Pi_n = \frac{F(\bar{c})}{N} \eta m_1$, and the search engine earns $\Pi^G = F(\bar{c}) (1 - \eta) M(r^S)$.*

Traffic management. Of the three positive factors constituting G 's profit, r^O increases the first (participation, $F(\bar{c})$) and decreases the second (the fraction who product search, $1 - \eta$), while r^S increases the first and decreases the third (the average margin of search-based trades, $M(r^S)$). So there is a simple trade-off for organic search: lowering r^O shifts trades from display to search platform, raising $1 - \eta$, but increasing r^O attracts search participation, $F(\bar{c})$. Similarly, there is a trade-off between raising r^S to attract participation, and lowering r^S to raise the average margin of search-based trades. Mathematically, the first-order conditions for maximizing G 's profits with respect to r^O and r^S are:

$$\frac{\partial \Pi^G}{\partial r^O} \frac{1}{M(r^S)} = f(\bar{c}) \frac{\partial \bar{c}}{\partial r^O} (1 - \eta) - F(\bar{c}) \frac{d\eta}{dr^O} = 0, \quad (2)$$

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

which can be rewritten using the reverse hazard rate $H(\cdot)$ as:

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

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

There is no equilibrium at $r^S = r^O = 1$ if the left hand side (LHS) of either (4) or (5) is then less than 1. So a sufficient condition for distortions is,

$$H(u+v_1)(1-\sigma\alpha)m_1\min\left\{\frac{u}{\sigma\alpha(1-\beta)m_1}, \frac{v_1-v_2}{m_2-m_1}\right\} < 1, \quad (6)$$

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 (6) holds.*

G 's choices (r^S, r^O) generate both vertical and horizontal externalities on publishers, whose aggregate profits are $F(\bar{c})\sigma\alpha e(r^O)m_1$: (i) G has a vertical externality because both r^S and r^O raise consumer participation, $F(\bar{c})$; (ii) G has a horizontal externality, because r^O raises the “effectiveness” of display advertising, $e(r^O)$. In other words, G 's policy determines both the mass of consumers engaged in online search where they demand offline products, and how this demand translates into purchases, via either display or search advertising. In the limiting case of $\beta = 1$ with the baseline assumption of symmetric publishers, G cannot reduce the effectiveness of display advertising, and G only affects publishers vertically. But if $\beta < 1$, G also affects the degree of substitutability between display and search advertising. Concretely, G can accomplish business-stealing from publishers' competing display ad platforms by distorting organic traffic, because this reduces the effectiveness of publishers' websites as advertising outlets. The distortion hurts consumers as well as publishers, even if $r^S = 1$. So, in contrast to typical horizontal relations where competition is healthy, the horizontal externality results in lower consumer, and total, welfare.

The minimum expression in (6) is instructive: G has stronger incentives to distort organic search than to distort sponsored search if $\frac{u}{\sigma\alpha(1-\beta)m_1} < \frac{v_1-v_2}{m_2-m_1}$, and conversely if the inequality is reversed. These two terms represent the respective cost-benefit ratios from marginally distorting organic and sponsored searches from $r^S = r^O = 1$:²⁶ distorting organic search reduces consumer surplus at the rate u (product trade values are fixed at v_1 since $r^S = 1$) while raising G 's ad revenues at the rate $\sigma\alpha(1-\beta)m_1$, as search-based trades substitute for display-based trades; meanwhile, distorting instead sponsored search (reducing r^S) reduces consumer surplus at the rate $v_1 - v_2$ while raising the value of G 's sponsored ads at the rate $m_2 - m_1$, both per product-searching consumer. To have r^S and r^O both interior in $(0, 1)$ requires exact equality of their respective cost-benefit ratios, otherwise the first-order

²⁶More generally, for any values of r^S and r^O , the inequality determines which cost-benefit ratio is larger.

conditions cannot hold simultaneously. So generically, at most one will be interior and interior solutions for both variables will obtain when

$$\frac{u}{\sigma\alpha(1-\beta)m_1} = \frac{v_1 - v_2}{m_2 - m_1}. \quad (7)$$

By contrast, the solution for r^O or r^S is interior for a set of parameter values with a non-empty interior.²⁷ So we can conduct relevant comparative statics.

Comparative statics and instrument substitutability.

The two instruments, r^O and r^S , are imperfect substitutes from G 's perspective. Indeed, G 's profit function satisfies the standard definition of substitutability: $\frac{\partial^2 \Pi^G}{\partial r^S \partial r^O} < 0$. As we have seen, in our stark model, G 's equilibrium strategy is typically to set one of r^O or r^S at 0 or 1. Marginal changes in parameter values then induce a change in only one of these instruments. The exception is when the change in the parameter crosses the threshold defined in (7). That is, when the change in the parameter induces a change in the relative size of the cost-benefit ratios of organic and sponsored search distortions.

Consider, first the effect of an increase in the accuracy of the targeting technology, σ . Define σ^* as the value of this parameter that, given the rest of parameters, makes (7) hold. For values of $\sigma < \sigma^*$, the left hand side of (4) is larger than the left hand side of (5) so that $r^S \leq r^O$, and for values of $\sigma > \sigma^*$ the opposite is true. Consider the case of interest where the parameter values are such that, for σ close to σ^* , $r^S \in (0, 1)$ or $r^O \in (0, 1)$.²⁸ Note that Π^G is jointly continuous in r^S , r^O , and σ . Then, from the maximum theorem, the equilibrium correspondence is upper hemicontinuous. Also, evaluated at σ^* , (4) is the same equation as (5), so that there is a continuum of solutions interior for both variables. Since $\frac{\partial^2 \Pi^G}{\partial r^S \partial r^O} < 0$, all these solutions lie on a downward sloping curve in r^S , r^O space. From upper hemicontinuity, we then conclude that:

Remark 1 *An improvement in the accuracy of the targeting technology so that σ crosses from below to above σ^* increases r^S and decreases r^O .*

The result is intuitive. As we discussed above, the two sides of (7) represent the cost-benefit ratio of distortions in organic and sponsored search, respectively. For the same participation and the same fraction of trades through display ads, η , the cost-benefit ratio from distorting organic search (left hand side of (7)) is lower with higher values of σ . This pushes r^O down and, consequently, r^S up. Thus,

²⁷The LHS of (4) and (5) are continuous functions of parameters and endogenous variables. For instance, if $\frac{v_1 - v_2}{m_2 - m_1} < \frac{u}{\sigma\alpha(1-\beta)m_1}$ and (5) holds at $r^S = r^O = 1$, a small parameter change that decreases the LHS of (5) induces r^S to fall strictly and no change in r^O .

²⁸A sufficient condition for r^S and/or r^O below 1, and r^S and/or r^O above 0, is that (6) holds and also,

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

when an improvement of targeting technologies makes display advertising a closer substitute for search advertising, G 's reaction is not only to fight that by reducing traffic to relevant publishers, but also to substitute organic search distortion for sponsored search distortion. As a result, sponsored search becomes more reliable.

The strength of the conflict of interest affects G 's choice of policy in a similar way. The simplest way to see this is by assuming that m_1 remains constant, but m_2 increases. This time, it is the cost-benefit ratio from distorting sponsored search that is reduced. Consequently, an increase in m_2 from a value where the right hand side of (7) is larger than the left hand side to a value where the opposite is true, reduces r^S and, consequently, increases r^O . That is, the sharper conflict of interest leads to less reliable sponsored search but more reliable organic search.

However, for a given strength of the conflict of interest, an increase in the profitability of the offline products market has the opposite effect. In particular, consider an increase in m_i with $m_1 - m_2$ kept constant, $i = 1, 2$. As with an increase in σ , this change reduces the cost-benefit ratio from distorting organic search (left hand side of (7)) without affecting the cost-benefit ratio of sponsored search distortions. Thus, when the increase in m_i crosses the threshold (7), G responds by reducing r^O and, consequently, increasing r^S . That is, by reducing the reliability of organic search and increasing the reliability of sponsored search.

An increase in u , and so in the relative importance of online content for consumers, increases the cost-benefit ratio of distortions in organic search, and so when the increase in u crosses the threshold (7), G will respond by improving the reliability of organic search and reducing that of sponsored search.

So far we have focused on G 's incentives to substitute r^O for r^S or vice versa, which are dominant when parameter values cross the threshold. When changes in parameter values have no effect on the ranking of cost-benefit ratios of distortions in organic and sponsored search, G 's incentives to adjust its reliability come from the impact of parameter values on $H(\bar{c})(1 - \eta)m_1$, the common terms in (4) and (5). For instance, a larger σ implies a larger \bar{c} and larger η , and so lower $H(\bar{c})(1 - \eta)m_1$. That is, when σ increases without crossing the threshold (7), G responds by reducing reliability of either sponsored or organic search, r^S or r^O . By contrast, an increase in m_i increases $H(\bar{c})(1 - \eta)m_1$. That is, a larger return per consumer gives G incentives to foster consumer participation, which is achieved by raising reliability. Similarly, larger u implies greater participation \bar{c} , and so lower value of $H(\bar{c})$, which is an additional incentive for higher distortions in either sponsored or organic search.

4 The effect of integration with full monopolization

In this section, we examine the effects of integration in the baseline model, beginning with an idealized scenario that isolates the positive sides of integration. Here and in sections 5 and 6, we study integration

into ad intermediation, deferring integration into publishing to section 7.

We consider a merger between G and one of the ad intermediaries and we refer to publishers that pay G 's ad intermediary to run their display advertising as "affiliates" (of G) and the rest as "non-affiliates" (of G). Given the absence of regulatory supervision, G can set different values of organic search reliability for affiliates and non-affiliates. We also need to explicitly describe the actions of intermediaries. So the baseline model changes in two ways. In the first stage, G now announces $(\mu, r^S, r_G^O, r_{NG}^O)$, where r_G^O and r_{NG}^O are the reliability of organic results to affiliated and non-affiliated websites, respectively. In between the first and second stages of the baseline model, each intermediary simultaneously announces its tariff T for publisher services, and then publishers respond simultaneously.

G 's ad intermediary can handle an unlimited number of publishers, so it can capture the entire surplus from display advertising. In equilibrium, all other intermediaries offer $T = 0$, as was left implicit in the previous section, and G sets $r_{NG}^O = 0$ to extract all publisher rents by charging $T_G = \frac{F(\bar{c})}{N} \eta m_1$, which each publisher accepts. In this equilibrium, the average reliability of content search, r^O , satisfies $r^O = r_G^O$. Nothing changes beyond the transfer of rents to G . G 's profits are now:²⁹

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

Relative to non-integration, G 's per-consumer profits rise by the rents ηm_1 extracted from publishers. That is, G now gains from consumer participation through display as well as sponsored search auctions. So G has stronger incentives to attract consumer participation and this encourages increased reliability of both organic and sponsored search. In addition, since η increases with r^O , G can increase these new rents by raising the reliability of organic search, which makes display ads more effective. In consequence, r^S and r^O both increase weakly, as we now prove in detail. Mathematically, the first-order conditions for maximizing G 's profits with respect to r^O and r^S , respectively, shift from conditions (4) and (5) to:

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

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

which can be rewritten as:

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

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

²⁹Here, we assume G cannot influence the type or quantity of display advertising of publishers affiliating with G . See also sections 7 and 8.

The LHS is higher for (10) than (11), for (10) than (4) and for (11) than (5), which indicate respectively, (i) $r^O \geq r^S$ with at most one being interior, (ii) given r^S , integration encourages more reliable organic search, and (iii) given r^O , integration encourages more reliable sponsored search. In fact, we can show that interaction between these two instruments does not change these insights. In particular, in the Appendix we prove:

Proposition 4 *Integration with full monopolization improves the reliability of the search engine and increases both consumer and total surplus; in particular, moving from no integration to full integration weakly raises r^S and r^O , and raises one or both strictly, unless initially r^S and r^O at a corner solution.*

Under integration, G monopolizes the display advertising intermediation market and appropriates all publisher rents, which improves G 's incentives to allocate traffic correctly, from the consumer and total surplus perspectives, for both types of queries. G internalizes both the vertical and horizontal externalities discussed above. The incentives to distort organic search are lower because the business-stealing effect disappears – G internalizes the horizontal externality. In addition, G takes into account the effect of higher consumer participation on publishers' rents – G internalizes the vertical externality from attracting consumers with higher reliability of both organic and sponsored search.³⁰

This proposition ignores the distributional consequences of vertical integration. In particular, publishers get zero profits which would, in a model with costly content, affect the quality of online content. Another potential drawback of integration lies in G 's incentives to discriminate against publishers that do not deal with G 's ad intermediaries or are otherwise less productive for G . This effect is absent in the above, extreme case of full monopolization where publishers are symmetric and all deal with G in equilibrium. In the next two sections, we illustrate how discrimination comes into play in a less extreme market structure.

5 Integration with partial monopolization

The assumptions in our baseline model, in particular constant returns to scale in ad intermediation and the absence of regulation, result in full monopolization when G enters the ad intermediation market. That is, in equilibrium, all publishers affiliate with G 's ad intermediary. So G 's policy of discriminating against the publishers not dealing with G 's ad intermediary did not translate into any actual 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 start examining the consequences of partial monopolization of the display advertising

³⁰The only remaining incentive to distort organic search derives from the fact that the integrated entity gets higher per consumer profits via search than via display ads – $M(r^S) \geq m_1$; cf., section 7. As with non-integration, improved targeting technologies increase remaining incentives to distort traffic.

market. Rather than model the exact obstacles that prevent full monopolization, we suppose that G 's ad intermediary can handle the advertising business of at most a fraction γ of publishers.³¹

As in the previous section, we let G treat affiliated and non-affiliated websites differently, by setting r_G^O and r_{NG}^O , and again, G offers ad intermediation in exchange for a tariff, T_G , seeking to attract the maximal feasible fraction, γ , of publishers. We assume consumers do not know whether their favorite publisher will be affiliated with G when they decide on participation.

Clearly, G will find it optimal to send any diverted traffic to a publisher in the G network of affiliates.³² So extending the effective visit notation, $e(r^O)$, from previous sections, by letting e_G and e_{NG} denote the average aggregate “effectiveness” of visits to publishers inside and outside the G network, respectively, we have,

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

$$e_{NG}(r_{NG}^O) = r_{NG}^O,$$

assuming that indeed the fraction γ of publishers accept G 's offer. The fraction of trades occurring on G 's affiliated and non-affiliated publishers is then $\eta_G = \sigma \alpha e_G$ and $\eta_{NG} = \sigma \alpha e_{NG}$, respectively, so the overall fraction of trades occurring via display advertising, η , is given by:

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

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

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

Participation is still determined by (1) using these averages. Lemma 3 still holds, both for publishers affiliated with G and for the non-affiliated: type 1 merchants still bid a PPC of σm_1 in both types of publishers' auctions. It only remains to determine how G sets T_G in the continuation game: G 's optimal choice is a tariff equal to a publisher's willingness to pay to be part of the G network. Expecting a fraction γ to affiliate with G , this value is:

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

G 's expected profits from these affiliated publishers sum to $\gamma N T_G$, giving overall expected profits:

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

³¹One possible interpretation of this exogenous constraint is that an increase in G 's market share above γ might trigger an unwanted investigation by the regulatory agency.

³²Distorting 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 regardless of whether the *destination* publisher is affiliated with G ; in equilibrium, all publishers have the same quantity and type of display ads. Thus, the destination of diverted traffic is irrelevant from the cost point of view, but not from the benefit point of view: G can charge its affiliated publishers a higher tariff if it sends them all the additional traffic diverted from non-affiliated sites.

where we use the fact that $\eta_G - \eta_{NG} = \frac{\eta - \eta_{NG}}{\gamma}$. The effects of r_G^O and r^S on G 's profits are analogous to those discussed in the case of full monopolization, equations (8) and (9).

In equilibrium, all publishers earn a rent equal to what they could get by refusing G 's offer. So publishers jointly appropriate $\eta_{NG}m_1 = r_{NG}^O\alpha\sigma m_1$. Clearly, r_{NG}^O increases this publisher surplus. In addition, for any $r^S < 1$, since $M(r^S) > m_1$, r_{NG}^O reduces the total producer surplus, by increasing the share η of display-mediated trades (that generates a margin m_1), and reduce search-mediated trades (with average margin, $M(r^S)$). Since G appropriates the producer minus the publisher surplus, G has an incentive to reduce r_{NG}^O for both these reasons. In the full monopolization case, G always minimizes r_{NG}^O at zero to extract all publisher rent. However, consumer participation is now increasing in r_{NG}^O , because organic distortions affecting non-affiliated publishers now affect a non-zero fraction, $1 - \gamma$, of consumers in equilibrium. This moderates G 's incentives to reduce r_{NG}^O and the optimal r_{NG}^O may be positive, but for exactly the same reason, any such distortions ($r_{NG}^O < 1$) now impose equilibrium inefficiency: they reduce consumer and social surplus. We now demonstrate how these harms can dominate the positive side of integration so that partial integration decreases overall social surplus, relative to non-integration.

As just explained, the main distinctive feature of partial integration is reflected in the first-order condition with respect to r_{NG}^O .³³

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

To emphasize the role of the fraction of affiliated publishers, γ , we rewrite this as:

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

where we embed the first and second effects in $\Psi(r_G^O, r_{NG}^O, r^S)$ defined by:

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

As γ increases towards 1, since Ψ converges to a finite number, the first term of (13) converges to 0, but the second remains strictly negative in the limit. So for sufficiently large γ , G has incentives to distort traffic to non-affiliates.

As with full monopolization, partial integration leads G to internalize the horizontal and vertical externalities on the affiliated publishers. The vertical internalization is partial, because publishers retain some rents if $r_{NG}^O > 0$. The horizontal internalization is also partial, because it only applies to affiliated publishers. Nonetheless, these effects promote reliable search results for sponsored search and organic

³³The three terms correspond to the three effects just described, now in inverse order: raising r_{NG}^O raises participation, lowers producer surplus and raises publisher surplus.

searches where an affiliate publisher is relevant; that is, r_G^O , r^S are higher. At the same time, partial integration exacerbates G 's incentives to steal business from non-affiliated publishers, because distorting organic searches that should be directed to non-affiliated publishers not only substitutes search-mediated trades for display-mediated trades (as occurred with non-integration) but also transfers the remaining display-mediated trades from non-affiliated to affiliated publishers. This exacerbated business-stealing can make consumers worse off under partial than non-integration. For a simple illustration, suppose condition (6) is broken so that G sets $r^O = r^S = 1$ under non-integration. As just explained above, under partial integration with γ sufficiently close to 1, G sets $r_{NG}^O < 1$. Both consumer and total surplus are then lower than under non-integration. Summarizing:

Proposition 5 *There exists a region of parameter values for which vertical integration with partial monopolization reduces both consumer and total surplus.*

The illustration identifies the converse of (6) and γ close to 1 as sufficient conditions for integration to cause a welfare reduction. This might suggest relatively small welfare losses, given that only a fraction, $1 - \gamma$, of publishers face distortion. However, in the Appendix, we show that γ can be arbitrarily low and the consumer and total surplus losses occasioned by integration can be substantial.

6 Asymmetric publishers

In this section, we analyze how differences in publishers' effectiveness for display advertising affect search bias incentives and welfare outcomes. It is important to also allow for imperfections in the product search channel for offline trade. Otherwise, as shown above in the baseline model, publisher asymmetries are irrelevant to an integrated monopolist, because display advertising is redundant, exactly as in the first-best. As already noted, that redundancy was an artifact of the assumption of frictionless product search. Introducing product search imperfections gives the integrated monopolist a natural motive to distort organic search to raise the effectiveness of display advertising. While the baseline model only allowed for distortions that reduced display effectiveness (via $\beta < 1$), publisher heterogeneity in display effectiveness now makes display-enhancing distortions feasible too. We demonstrate that this new distortion incentive can make integrated monopoly worse for consumers and social welfare than non-integration.

We model heterogeneity in publisher ad effectiveness in the simplest possible way: a proportion ρ of publishers are type H , characterized by a higher baseline ad effectiveness α_H than the rest, which are type L and have effectiveness $\alpha_L < \alpha_H$. We denote $\alpha = \rho\alpha_H + (1 - \rho)\alpha_L$. We assume $\alpha_L < \beta\alpha_H$ to focus on the interesting, high asymmetry case where display advertising is always less effective on type L than type H publishers, even comparing best-match visitors on the type L publishers with worst-match visitors on type H publishers. To introduce imperfection in product search, we now assume that, while

a proportion ϕ of participating consumers behave as in the baseline model, always able to search for products at no added cost after searching for content, the remaining proportion, $1 - \phi$, can only conduct a content search. 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 and before responding to display ads during content search.³⁴ Notice that during their content search, a fraction $1 - \phi$ of consumers are now willing to buy when faced with a display ad of the relevant type 2 product. But the other fraction ϕ would still wait to conduct a product search, so if ϕ is reasonably large, type 1 merchants still outbid type 2 merchants when competing for display ads. We assume $m_1 > (1 - \phi)m_2$, to indeed ensure that display ads are all of type 1 products, as in the preceding analysis.³⁵

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 by 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 $d_{a,b}$ represent the fraction of traffic diverted from type a publishers that is directed to type b publishers, for $a, b \in \{L, H\}$; of course, $d_{aH} + d_{aL} = 1$ for each $a \in \{L, H\}$. Now η becomes,

$$\begin{aligned} \eta = & \sigma \rho [r_H^O \alpha_H + (1 - r_H^O) \beta (d_{HL} \alpha_L + (1 - d_{HL}) \alpha_H)] \\ & + \sigma (1 - \rho) [r_L^O \alpha_L + (1 - r_L^O) \beta (d_{LH} \alpha_H + (1 - d_{LH}) \alpha_L)]. \end{aligned} \quad (14)$$

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

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

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

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

For a given level of sponsored search reliability r^S , if one can fix the fraction of display-mediated trades, η , G 's profits are increasing in r^O since this raises participation \bar{c} without changing G 's profit per consumer. Since η is increasing in r_H^O and decreasing in d_{HL} , if these two values are interior, one can raise d_{HL} and r_H^O such that η remains unchanged but r^O indeed increases. Thus:

Lemma 4 *Under non-integration, if $r_H^O < 1$, then $d_{HL} = 1$.*

³⁴One interpretation is that, after deciding to participate, a fraction ϕ of consumers will discover a product need and form a plan to search for and buy it, but the other fraction, $1 - \phi$, are unaware of their latent product demand and only consider buying offline products if exposed to an ad.

³⁵This inequality is unnecessary if consumers always discover how to conduct a product search upon seeing a display ad of the relevant product, of either type.

The intuition is simple. Under separation, G benefits from reducing offline trade mediated through publisher display ads, but wants to let consumers consume their preferred content as far as possible. Distorting traffic from type H to type L publishers maximizes the reduction in display-mediated trade for a given reduction in the reliability of content search.

To evaluate the possibility of a non-distorted equilibrium, we suppose $r_L^O = r^S = 1$ and focus on the incentives over r_H^O . In this case, $r^O = \rho r_H^O + (1 - \rho)$. Considering the first-order conditions for G 's profit maximization evaluated at $r^S = r_H^O = 1$ and $\eta = \sigma\alpha$, we can write the conditions for an equilibrium with no distortions as:

$$H(\bar{c})(1 - \sigma\alpha) \left(\frac{u}{\sigma[\alpha_L - \beta(d_{LH}\alpha_H + (1 - d_{LH})\alpha_L)]} + v_1(1 - \phi) \right) \geq 1,$$

$$H(\bar{c})(1 - \sigma\alpha) \left(\frac{u}{\sigma(\alpha_H - \beta\alpha_L)} + v_1(1 - \phi) \right) \geq 1, \quad (15)$$

$$\phi H(\bar{c})(1 - \sigma\alpha) \frac{v_1 - v_2}{m_2 - m_1} m_1 \geq 1, \quad (16)$$

where $\bar{c} = u + (\sigma\alpha(1 - \phi) + \phi)v_1$. The first inequality is implied by the second. Note that for $\phi = 1$ and $\alpha_H = \alpha_L$, this replicates the conditions in section 3. As in that section, these conditions are also sufficient; the LHS of both (15) and (16) are decreasing both 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 seek to demonstrate parameter values for which G will distort under integration with full monopolization, but not under separation. So suppose that G sets $r^S = 1 = r_H^O$. In this case, $r^O = \rho + (1 - \rho)r_L^O$, $\bar{c} = r^O u + \phi v_1 + \eta(1 - \phi)v_1$ and

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

Clearly, Π^G is now increasing in η as well as \bar{c} , which is itself increasing in η , for any given value of r^O . (G has no cost from increasing η when integrated, given G 's search ads are selling type 1 products just like the display ads.) It follows immediately that $d_{LH} = 1$ is optimal since d_{LH} increases η . Since $\frac{\partial \eta}{\partial r_L^O} = \sigma(1 - \rho)(\alpha_L - \beta\alpha_H)$ is negative by our assumption of substantial publisher heterogeneity, G now has a motive to decrease r_L^O , diverting search to more ad-effective publishers.³⁶ Of course, there is a trade-off, because search accuracy raises participation. These effects are captured in the first-order derivative which we evaluate at $r_L^O = 1$,

$$\frac{\partial \Pi^G}{\partial r_L^O} \frac{1}{m_1} = f(\bar{c}) \frac{\partial \bar{c}}{\partial r_L^O} [\phi + \eta(1 - \phi)] + F(\bar{c}) \frac{\partial \eta}{\partial r_L^O} (1 - \phi), \quad (17)$$

³⁶These distortions by the fully integrated G are akin to a monopoly media outlet distorting by picking news content that makes the media outlet's display ads more effective (see Ellman and Germano, 2009).

Notice that if $\phi = 1$, the second term representing the increase in display effectiveness is nullified since display is redundant, as explained earlier.³⁷ Now a sufficient condition for the existence of distortions under vertical integration is that,

$$H(\bar{c}) \frac{\sigma\alpha + \phi(1 - \sigma\alpha)}{1 - \phi} \left(\frac{u}{\sigma(\beta\alpha_H - \alpha_L)} - v_1(1 - \phi) \right) < 1, \quad (18)$$

evaluated at $\bar{c} = u + (\sigma\alpha + \phi(1 - \sigma\alpha))v_1$ as in (15) and (16). It is a simple exercise to see that condition (18) is compatible with (15) and (16), and so integration may lead to an increase in distortions in organic search. This increased distortion necessarily reduces both consumer and total surplus when $u > (m_1 + v_1)(1 - \phi)\sigma(\beta\alpha_H - \alpha_L)$. A consumer interested in an L type publisher may be attracted by a display ad with higher probability if sent to an H type publisher. The probability increase is $\sigma(\beta\alpha_H - \alpha_L)$, and with probability $1 - \phi$, this is the consumer's only chance for consuming an offline product. However, given the inequality, this potential gain in surplus does not compensate the consumer's direct utility loss from not consuming her preferred online content. This sufficient condition for distortions to result in consumer and total surplus losses is also compatible with (18), (15), and (16).³⁸ Thus, we conclude,

Proposition 6 *With asymmetric publishers integration with full monopolization can reduce consumer and total surplus.*

As in the case of symmetric publishers, integration with full monopolization induces G to internalize the vertical and horizontal externalities imposed on publishers. The difference is that G 's internalization of the vertical externality may no longer be in the interest of consumers. Indeed, under the assumptions of this section, aggregate publisher revenue is larger if traffic is distorted from publishers with low to high ad effectiveness. Moreover, an integrated G cares about this display ad revenue, which is not redundant given the friction, $\phi < 1$, in product search and search-mediated advertising. So G has a new incentive to distort traffic, this time from less to more ad-effective publishers.

We could have also considered the case where G only deals with type H publishers, perhaps because it cannot price-discriminate among publishers and prefers to set a tariff too high for type L publishers

³⁷As is already implicit, publisher asymmetry is also crucial: the key factor in the display ad effectiveness derivative, $\alpha_L - \beta\alpha_H$, is obviously negative if $\alpha_L = \alpha_H$. The distortion arises because with $\phi < 1$, display is not redundant and publisher asymmetry gives G an instrument that increases η : G can direct consumers whose preferred publisher is type L to a type H publisher. With symmetry, the effect of r^O on Π^G , for $\phi = 1$ represented by (8), is always positive if $r^S = 1$. Introducing $\phi < 1$ to the symmetric case, $\frac{\partial \Pi^G}{\partial r^O}$ is given by (17) with r^O substituting for r_L^O . Since participation and profits per consumer would then both increase with η and participation and η would still increase in r^O , $\frac{\partial \Pi^G}{\partial r^O}$ would still be positive.

³⁸For $u = (m_1 + v_1)(1 - \phi)\sigma(\beta\alpha_H - \alpha_L)$ and $\beta = 1$, the parenthesis in (18) takes the value $m_1(1 - \phi)$, and the parenthesis in (15) takes the value $(2v_1 + m_1)(1 - \phi)$. So for v_1 sufficiently large, the left hand side of (18) is smaller than the left hand side of (15), and the value of $H(\bar{c})$ can be adjusted so that the two expressions lie on the corresponding sides of 1. Finally, a small enough value of $m_2 - m_1$ guarantees that (16) holds.

to accept. In this case, both the incentives to distort traffic away from non-affiliates and away from low ad-effective publishers would coincide. The partial internalization of both vertical and horizontal externalities would then combine against the interest of consumers, reinforcing the possibility that integration lowers consumer and total surplus.

7 Integration with publishers

So far, we have considered integration of the search engine with an intermediary from the display advertising market. Integration with publishers results in similar incentives for G . Indeed, if the integrated entity did not modify how it handled display advertising, sections 4 and 5 would continue to describe the effects of integration on G 's policies. However, integration with publishers may have slightly different consequences, since this integration is likely to facilitate manipulation of the supply of advertising and the coordination of pricing strategies. In fact, there is no role for price coordination in our stylized model, since prices already extracted all merchant rents under non-integration. But by manipulating the supply of different types of advertising, an integrated entity might raise its profits beyond our section 4 and 5 predictions. To show this, we first consider the simplest and extreme case, where the search engine owns all publishers, but has them set display advertising as in section 4. For the parameter values that gave $r^S = 0$ and $r^O < 1$ in section 4, this would imply the profits derived there as,

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

with $\bar{c} = r^O u + \eta v_1 + (1 - \eta) v_2$. The reason for distorting r^O was that display ads, being of type 1, restricted G 's ability to maximize type 2 offline trades; the combined entity distorted organic results to transfer advertiser attention to its search platform where it could better exploit that attention. However, now the integrated entity could simply choose to eliminate display advertising from publishers' websites, inducing all consumers to conduct product searches and removing the motive for distorting organic search. G would then set $r^O = 1$ and could replicate its previous level of per-consumer profits from display and search ads by setting r^S equal to the prior level of η , which also replicates consumers' offline trade distribution. Since the increase in r^O raises consumer participation, the integrated entity gains strictly by removing display advertising. Alternatively, it could restrict display advertising to high margin, that is type 2 merchants, and again adjust its results strategies as just described. In the baseline setting, this is equivalent, since no one buys via type 2 display ads given any $r^S > 0$. But in the model with $\phi < 1$, the alternative of restricting a fraction of display advertising to type 2 merchants is strictly preferable to shutting down display ads, because the display ad channel is not redundant for the fraction $1 - \phi$ of participating consumers who cannot be reached by search advertising.³⁹

³⁹In either case, with *full* integration into publishing, G would have no incentive to distort organic search.

Similarly, when the search engine owns a fraction γ of publishers and G 's publishers maintain their display ads, 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 $r_G^O = 1$, leaving r_{NG}^O fixed, and setting $r^S = \frac{\gamma\eta_G}{1-(1-\gamma)\eta_{NG}}$. With this policy, the integrated entity makes the same profits per consumer and induces higher consumer participation. Again, simply shutting down display ads is an optimal strategy if $\phi = 1$, but more generally, G would only want to restrict against low margin display ads. Summarizing,

Proposition 7 *Integration with publishers has essentially the same effects as integration into ad intermediation but may differ by reducing the supply of display advertising or restricting display to high margin products.*

This minor difference would disappear if a G owned ad intermediary were also able to exert influence over display advertising. For instance, G might exert influence by committing to divert traffic away from those publishers that refuse to adjust the quantity and content of their display ads to G 's request.⁴⁰

8 Discussion

The model presented in this paper is stylized and parsimonious given its multiple objectives. In section 2, we explained and motivated our main modeling choices. Here, we consider additional extensions.

Compulsive consumers. In the baseline model, the role of advertising is purely informative: consumers are fully rational and advertising merely enables consumers to locate merchants – advertising neither persuades nor tempts people to consume, nor does it complement consumption. This is the most common view of advertising in economic models and Blake et al., (2013) provide supportive evidence for this view in the case of search advertising. But we now consider the possibility that display advertising is persuasive as well as informative. This opens the door to more negative views of the welfare implications of advertising. The analysis readily extends, with few changes.

Concretely, suppose consumers are compulsive and always consume either type of relevant product when tempted by a display ad. In this scenario, type 2 instead of type 1 merchants buy all the display ad slots, but the main trade-offs for G are very similar. The novelty is that type 2 display ads lead compulsive consumers to go against their better interest, so that \bar{c} is now decreasing in $\alpha\sigma$.⁴¹ In this case,

⁴⁰Notice that an analogy of the endogenous affiliation process described for a G ad intermediary in section 4 could, in principle, lead G to monopolize the market for publishing, diverting traffic away from publishers that reject a G buyout. However, this seems less plausible than the, already extreme, case of a monopolizing intermediary. First, unlike ownership transfers, ad contracts are regularly renewed, which facilitates repeated game effects. Second, publishers are more visible to consumers and plurality of ownership is recognized by regulators as a fundamental value in the media context.

⁴¹Given this, compulsives might actually prefer organic search to be distorted, in which case G would distort organic search fully: there would be no participation/profit trade-off. But a tradeoff remains whenever consumers still prefer undistorted search, which holds if $u > (v_1 - v_2)\sigma\alpha(1 - \beta)$.

the qualitative trade-offs for search distortions are essentially the same as with rational, self-controlled consumers. Inducing participation is somewhat more difficult, since consumers anticipate suffering from the temptation of type 2 display ads, but for the same reason, distorting organic search has a lower participation cost.

Alternatives to auctions and non-necessity of the perfect competition assumption. The second-price auction is equivalent to a mechanism where G simply sets take-it-or-leave-it offers to type 1 and type 2 merchants with probabilities r^S and $1 - r^S$, respectively. Notice that this posted-price alternative works equally well in the absence of competition between merchants of each product.⁴² So this competition is not essential for our results.⁴³ Similarly, the display-ad auctions could be replaced by posted-prices.⁴⁴ Also, in our setup, if G sets a stochastic μ satisfying $\mu > \frac{m_1}{m_2}$ with probability r^S and otherwise $\mu < \frac{m_1}{m_2}$, the outcome is the same and G 's optimal strategy is unaffected.⁴⁵

Merchant heterogeneity in CTR or CR. The parameter α might also vary by merchant. We abstracted from such asymmetries, but they are analogous to variations in click-through rates, CTR, (the probability a consumer clicks on a merchant's link) and conversion rates, CR, (probability that consumers arriving at a merchant webpage then buy) as studied by Athey and Ellison (2011) and Chen and He (2011). As discussed above, these studies are essentially compatible with ours. In our model of product search, both CTR and CR equal unity for any relevant product and only the product of CTR and CR is relevant. More generally, search engines claim to take account of these factors in weighting sponsorship bids. In particular, Google was the first to introduce click-weighted auctions in 2003, successfully reducing the prevalence of ads with low CTR, such as mobile phone ringtone and porn-site ads. We could extend our model along the lines of the aforementioned papers. Advantageous self-selection, as obtained in these papers, means no added conflict of interest, at least for the case with a single sponsored link.⁴⁶

PPC versus price per impression, PPI. Throughout the paper, we assumed sale of ad space based on price per click, PPC. The CTR for sponsored links and also for display ads, conditional on

⁴²Also all these designs for selling advertising are optimal in our setting as they extract full merchant rent.

⁴³Of course, bias in sponsored search requires two types of merchant to be interested in a common search query.

⁴⁴As noted just prior to Lemma 3, publishers optimally forego the possibility of weighting bids in the display ad auction because merchants and publishers have a common interest when consumers buy only type 1 products through display ads. With compulsive consumers, publishers' and merchants' preferences would also be aligned if publishers are numerous enough that each publisher's display ad strategy has a negligible impact on consumer participation.

⁴⁵Auctions are typically better instruments than posted prices when G faces uncertainty about the merchants' margins. Note that, if the distribution of these margins is atomless, then tie-breaking is irrelevant. The weight μ would still determine a probability r^S that a type 1 merchant wins the sponsored position.

⁴⁶Concretely, suppose each consumer only values one of the two merchants offering each product and can costlessly learn where a sponsored link leads, by observing a snippet. If we further characterize these two merchants by different probabilities of being the desired seller, this determines and equals that merchant's CTR. In the one slot setting, there would be no conflict of interest over this dimension, with both consumers and G favoring the highest CTR. Assuming G observes the CTR, G would weight the merchant bids by CTR since merchants with low CTR and high expected value per click would otherwise win too often in a PPC auction. Consumers would benefit. This would be similar to the modeling approach of Athey and Ellison (2011) who provide a number of interesting further results.

attracting consumer attention, are equal to one. In general, a lower CTR value would simply scale down the merchant’s willingness to pay for sponsored positions and display ads under PPI bidding; meanwhile bidding on PPC is not affected by CTR in our setup. This modeling difference would not affect our analysis (the distinction matters in richer environments such as that described under heterogeneity). PPC is slightly simpler to explain and more realistic.

The substitutability of search and display ads. Our framework assumes that both search and display advertising are purely informative. This implies that merchants view search and display ads as partial substitutes, which is critical to the organic search distortion results when publishers and ad intermediaries are symmetric. As discussed for compulsive consumers, the evidence in Blake et al., (2013) suggesting informative search advertising may not apply for display. More generally, there may be some persuasive brand advertising on the internet. If both display and search advertising were persuasive in a similar way, then even if advertisers wished to reach consumers multiple times, decreasing returns could generate substitutability. But in principle, some merchants may specialize into using just one advertising channel and they might separate complementary branches of their advertising strategy between the search and display advertising channels.⁴⁷ So long as this extreme separation is not the norm, search and display advertising will be substitutes from the perspective of merchants.

Empirical evidence is limited. An important, related study by Goldfarb and Tucker (2011) used a natural experiment based on “ambulance-chaser” laws restricting postal advertising by law firms to show that law firms do consider offline and online advertising to be substitutes. Hahn and Singer (2008) provide a survey that points to substitutability of the display and search advertising channels. The FTC (2007) took the controversial view that the markets are essentially independent,⁴⁸ but the EU (2008) antitrust authority notes a trend towards substitutability for advertisers in large part as the technologies used for search and display advertising were, and are becoming increasingly, similar (see point 52 of EU, 2008).⁴⁹ In particular, the EU (2008) found that some survey respondents considered search and display ads to differ only in terms of the “triggering mechanism” which we call targeting. Our model demonstrates formally how targeting of display ads raises substitutability.

Endogenous entry. Search distortions hinge on the assumption that search engine competition is ineffective (as justified by network effects in the models of Argenton and PrÄ yfer, 2012, and Etro, 2012). If instead this competition were effective, consumers could choose to use whichever engine managed to

⁴⁷If e.g., display ads complement search ads, G would wish to encourage display advertising with or without integration; this would push in the direction of removing organic search distortions in our baseline case, and induce distortions towards high α publishers in the extension of section 6.

⁴⁸Ratliff and Rubinfeld (2010) question the FTC (2007) claims, citing market research studies that suggest that search and display advertising are increasingly used for similar types of marketing; initially display ads may have been used more often for building brand awareness while search ads were arguably preferred for direct-response online sales.

⁴⁹The legal debates surrounding Google reported there are multi-faceted and many extend beyond the scope of our paper; see van Loon (2012) for a legal discussion.

build a reputation for reliability.

Endogenous entry of publishers would generate new incentives for G . Consumer and publisher entry would mutually reinforce, since consumers would then have access to a broader range of online content. Hence, under separation, G would tend to be more reliable in organic search in order to foster content variety and so consumer participation. For the same reason, if G fully monopolizes ad intermediation, G may wish to commit to a relatively low charge for ad intermediation.

9 Concluding remarks

In this paper, we constructed a model that explicitly describes the workings of markets underlying both search and display advertising. The two modes of advertising are imperfect substitutes for merchants. A monopoly search engine mediates between consumers and merchants by addressing consumers to merchants through search advertising, but content providers present an alternative, competing channel by which merchants can also reach consumers. So, in the market for offline products, content providers and the search engine are competitors. What makes the relationship between content providers and the engine unusual and complicated is the fact that the engine also mediates a related market, that for online content, where content providers meet consumers. We have shown how, in consequence, a monopoly search engine has incentives to distort organic search to make display advertising less effective, thereby increasing the value of sponsored search. At the same time, the search engine has incentives to distort sponsored search in favor of merchants with high margins. We have characterized how these incentives interact and how they depend on market characteristics, such as the power of targeting technologies.

We have also investigated the effects of integration. We showed how a monopolist in the search engine market may monopolize the entire advertising market by buying a single intermediary in the display advertising market.⁵⁰ With symmetric publishers, such integration reduces incentives to distort organic search, as well as sponsored search. However, when the engine integrates with, or affiliates, a fraction of publishers, new motives to distort organic search results emerge. Partial integration reinforces the engine's incentives to steal business from independent publishers, diverting traffic to affiliated or owned publishers. In fact, this effect may outweigh the positive effect of internalizing the rents from affiliated or owned publishers and result in lower consumer and total welfare.

Full integration does avoid the engine's incentive to divert traffic away from independent publishers, in that none remain, but when publishers' websites vary in their effectiveness as ad platforms, internalizing externalities among publishers creates new motives for distortion. Under full integration, the engine has an incentive to divert traffic from less to more effective publishers to increase the aggregate

⁵⁰Consistent with this, since buying DoubleClick in 2007 and AdMob in 2009 and Admeld in 2011, Google has overtaken Yahoo in the market for display advertising which was Yahoo's remaining strong point; see Learmonth (2011).

value of display advertising. Again, this new, negative effect may dominate the positive effects of rent internalization, so that integration results in lower consumer and total surplus.

Ultimately, predicting which of these competing effects will dominate is an empirical question. As explained above, our model provides a framework for estimating the degree of substitutability between display and search advertising. Substitutability plays an important role in determining the risk of organic search bias. We have conducted comparative statics exercises with respect to several parameters that affect this substitutability: targeting quality by publishers, the relative profitability of offline product markets, and the conflict of interest between merchants and consumers in that market.

Some relevant issues that we have not analyzed in this paper are amenable for analysis along the lines of our model. For instance, if consumers' utility from online content depends on costly publisher investments, then it seems straightforward that the reduction in publisher revenues associated with integration of the monopoly search engine into ad intermediation would lead to lower levels of investment with negative consequences for consumer and total surplus. Given the value of informative, entertaining, social and educational web content, explicit consideration of this new channel merits attention. Also, the integration of the search engine with ad intermediaries or publishers may facilitate behavioral targeting where publishers use a consumers' past search queries to target display ads more effectively. If so, integration will increase heterogeneity of ad effectiveness between an engine's affiliated and non-affiliated publishers, as well as increase average display ad effectiveness. Simple extensions of our model could be used to evaluate the entry, reliability and welfare consequences of these additional effects.

10 APPENDIX

10.1 Proof of Proposition 4

The first-order conditions with respect to r^O and r^S can respectively be written as:

$$H(\bar{c}) \frac{(1-\eta)M(r^S) + \eta\delta m_1}{M(r^S) - \delta m_1} \left[\frac{u}{\sigma\alpha(1-\beta)} + (v_1 - v_2)(1 - r^S) \right] \gtrless 1, \quad (19)$$

$$H(\bar{c}) \frac{v_1 - v_2}{m_2 - m_1} [(1-\eta)M(r^S) + \eta\delta m_1] \gtrless 1, \quad (20)$$

where $\delta = 0$ for the case with no integration and $\delta = 1$ for the case with integration. Note that for both equations, the LHS is higher when $\delta = 1$ than when $\delta = 0$. Also when $\delta = 1$, the LHS of (19) is higher than for (20) even if $u = 0$.

Suppose that under no integration $r^S = r^O = 1$. That requires that the LHS of both equations (19) and (20) are higher than 1 when evaluated at $\delta = 0$ and $r^S = r^O = 1$. So, they are also higher than 1 at $r^S = r^O = 1$ when $\delta = 1$. Consequently, $r^S = r^O = 1$ is a candidate solution under integration. In fact, it is the only candidate. In any other alternative, either r^S or both r^S and r^O are lower than

one, and this is inconsistent with the fact that the LHS of (20) is decreasing in both r^S and r^O . So the solution is unchanged.

Suppose that $0 < r^S < 1$ and $r^O = 1$ under no integration. Then (20) holds with equality for $\delta = 0$. If under integration, $r^O < 1$ then $r^S = 0$, since the LHS of (19) is higher than for (20). But again, this is inconsistent with the fact that the LHS of (20) decreases with r^S and r^O . So $r^O = 1$ under integration. So r^O is unchanged, which, given that the LHS of (20) is higher with $\delta = 1$ than with $\delta = 0$, implies r^S must be higher than under 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 remains at 0 under integration, then since the LHS of (19) under $\delta = 1$ is higher than under $\delta = 0$, r^O must be higher than under no integration.

Suppose that $r^S = 0$ and $r^O = 1$ under no integration. Similar to the last case, r^S either rises with r^O staying at 1, or r^S remains at 0, only now with r^O remaining at 1.

Finally, suppose that $r^S = r^O = 0$ under no integration. This policy could remain optimal under integration, but not if the LHS of (19) evaluated at $\delta = 1$ and $r^S = r^O = 0$ is higher than 1, i.e., if

$$H(v_2) \frac{m_2}{m_2 - m_1} \left[\frac{u}{\sigma\alpha(1 - \beta)} + (v_1 - v_2) \right] > 1 \quad (21)$$

because r^O is then necessarily positive under integration. So (21) is a sufficient condition to rule out the possibility that $r^S = r^O = 0$ remains an optimal policy. Since the LHS of (6) is higher than the LHS of (21), this proves that there exists a non-empty set of parameter values defined by conditions (6) and (21), for which integration strictly improves the reliability of the search engine in terms of one or both types of search.

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

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

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

Under partial integration, we study the derivative of Π^G with respect to r_{NG}^O expressed in equation (13). Evaluating at $r^S = r_G^O = r_{NG}^O = 1$, the derivative will be negative (and r_{NG}^O distorted) if and only if:

$$(1 - \gamma) H(u + v_1) \frac{u(1 - \sigma\alpha)}{\sigma\alpha} < 1. \quad (22)$$

We consider two extreme cases. First, suppose that $\sigma\alpha$ is very small. Then, condition (22) 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 $\frac{u}{\sigma\alpha(1 - \beta)} < \frac{v_1 - v_2}{m_2 - m_1} m_1$ and $H(u + v_1) (1 - \sigma\alpha) \frac{u}{\sigma\alpha(1 - \beta)}$ is very close to 1. Then, under non-integration G is close to indifferent between setting $r^O = 1$ and a value slightly below. In this case, any value of $\gamma > \beta$ will satisfy (22), so that, unless $1 - \beta$ is very small, the impact of traffic distortion can be substantial.

References

- [1] Argenton, C., and J. Prüfer (2012), “Search engine competition with network externalities.” *Journal of Competition Law & Economics* 8(1), 73-105.
- [2] Anderson, S., Ø. Foros, and H. Kind (2013), “Competition for advertisers in media markets.” mimeo.
- [3] Ambrus, A., E. Calvano, and M. Reisinger (2013), “Either or Both Competition: A ‘Two-sided’ Theory of Advertising with Overlapping Viewerships.” mimeo, Duke University.
- [4] Armstrong, M. and J. Zhou (2011), “Paying for Prominence.” *Economic Journal*, 121(556), 368-95.
- [5] Athey, S., and G. Ellison (2011), “Position auctions with consumer search.” *The Quarterly Journal of Economics* 126(3), 1213-1270.
- [6] Athey, S., E. Calvano, and J. S. Gans (2013) “The Impact of the Internet on Advertising Markets for News Media.” Rotman School of Management Working Paper No. 2180851.
- [7] Baye, M., and J. Morgan (2001), “Information Gatekeepers and the Competitiveness of Homogeneous Product Markets.” *American Economic Review* 91, 454-474.
- [8] Blake, T., C. Nosko and S. Tadelis (2013), “Consumer Heterogeneity and Paid Search Effectiveness: A Large Scale Field Experiment.” NBER Mimeo.
- [9] Börgers, T., I. Cox, M. Pesendorfer, and V. Petricek (2007), “Equilibrium Bids in Sponsored Search Auctions: Theory and Evidence.” mimeo.
- [10] Brin, S. and L. Page (1998), “The Anatomy of a Large-Scale Hypertextual Web Search Engine.” *Computer Networks and ISDN Systems* 30(1-7), 107-117.
- [11] Broder, A. (2002), “A taxonomy of web search.” *ACM SIGIR Forum* 36 (2).
- [12] Chen, Y. and C. He (2011), “Paid Placement: Advertising and Search on the Internet.” *The Economic Journal*, 121(556), F309–F328.
- [13] Choi, J. P., and B.-C. Kim (2010), “Net neutrality and investment incentives,” *Rand Journal of Economics*, 41(3), 446–471.
- [14] de Cornière, A., and G. Taylor (2013), “Integration and Search Engine Bias.” Mimeo.
- [15] Edelman, B. and Z. Lai (2013), “Exclusive Preferential Placement as Search Diversion: Evidence from Flight Search.” Harvard Business School NOM Unit Working Paper No. 13-087.
- [16] Edelman, B., M. Ostrovsky, and M. Schwarz (2007), “Internet Advertising and the Generalized Second-Price Auction: Selling Billions of Dollars Worth of Keywords.” *American Economic Review* 97, 242-259.
- [17] Edelman, B. and M. Schwarz (2010), “Optimal Auction Design and Equilibrium Selection in Sponsored Search Auctions.” *American Economic Review* 100, 597-602.
- [18] Eliaz, K., and R. Spiegler (2011), “A Simple Model of Search Engine Pricing.” *The Economic Journal* 121, 329-339.
- [19] Ellman, M., and F. Germano (2009) “What do the Papers Sell? A Model of Advertising and Media Bias.” *The Economic Journal* 119, 680-704.
- [20] Etro, F. (2012), “Advertising and Search Engines. A Model of Leadership in Search Advertising.” *Research in Economics* 67 (1), 25-38.
- [21] European Union, “EU (2008),” Article 8, COMP/M.4731 - Google/DoubleClick. Regulation (EC) No. 139/2004 Merger Procedure.
- [22] Federal Trade Commission, “FTC (2007),” Statement of Federal Trade Commission Concerning Google/DoubleClick, (FTC File No. 071-0170) 3 and 7.
- [23] Goldfarb, A., and C. Tucker (2011), “Search Engine Advertising: Channel Substitution When Pricing Ads to Context.” *Management Science* 57, 458-470.
- [24] Greenspan, R. (2004), “Searching for balance.” *ClickZ* (April 30), <http://www.clickz.com/3348071>.
- [25] Hagiu, A., and B. Jullien (2011), “Why Do Intermediaries Divert Search?” *Rand Journal of Economics* 42, 337-362.
- [26] Hahn, R., and H. Singer (2008), “An Antitrust Analysis of Google’s Proposed Acquisition of DoubleClick.” Working Paper.
- [27] Hotchkiss, G., M. Garrison, S. Jensen (2005), “Search engine usage in North America.” White paper, Enquiro, Kelowna, BC, Canada.

- [28] Jansen, B. (2007), "The comparative effectiveness of sponsored and non-sponsored links for Web e-commerce queries." *ACM Trans. Web* 1(1) Article 3.
- [29] Jansen, B., and M. Resnick (2006), "An Examination of Searchers; Perceptions of Nonsponsored and Sponsored Links During E-commerce Web Searching." *Journal of the American Society for Information Science and Technology*, 1949-1961.
- [30] Learmonth, A. (2011), "Google Just Blew Past Yahoo in Online Display Advertising; Amid a Fast-Growing Display Market, Search Giant Is Gaining Share." *AdAge Digital*, May 26.
- [31] Lehaie, S. (2006), "An Analysis of Alternative Slot Auction Designs for Sponsored Search." EC06.
- [32] Liu, D., J. Chen, and A. Whinston (2010), "Ex-Ante Information and the Design of Keyword Auctions." *Information Systems Research* 21, 133-153.
- [33] Lorigo, L., M. Haridasan, H. Brynjarsdóttir, L. Xia, T. Joachims, G. Gay, L. Granka, F. Pellacini, and B. Pan (2008), "Eye Tracking and Online Search: Lessons Learned and Challenges Ahead." *Journal of the American Society for Information Science and Technology* 59 (7), 1041-1052.
- [34] Ratliff, J., and D. Rubinfeld (2010), "Online Advertising: Defining Relevant Markets." *Journal of Competition Law and Economics* 6 (3), 653-686.
- [35] 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.
- [36] Tarantino, E. (2013), "A Simple Model of Vertical Search Engines Foreclosure." *Telecommunications Policy* 37(1), 1-12.
- [37] Taylor, G. (2013), "Search Quality and Revenue Cannibalization by Competing Search Engines." *Journal of Economics and Management Strategy*, 22(3), 445-467.
- [38] Van Loon, S. (2012), "The Power of Google: First Mover Advantage or Abuse of a Dominant Position?" In López-Tarruella A. (ed.) *Google and the Law: Empirical Approaches to Legal Aspects of Knowledge*, The Hague: T.M.C. Asher Press.
- [39] Varian, H. R. (2007), "Position Auctions." *International Journal of Industrial Organization* 25, 1163-1178.
- [40] Wang, J., and R. Day (2007), "The effects of attention inertia on advertisements on the WWW." *Computers in Human Behavior* 23(3), 1390-1407.
- [41] White, A. (2013), "Search Engines: Left Side Quality versus Right Side Profits." *International Journal of Industrial Organization*, Forthcoming.
- [42] White, A., and Jain, K. (2012), "The Attention Economy of Search and Web Advertisement." Mimeo, Tsinghua University.
- [43] Wilbur, K. (2008), "A Two-Sided, Empirical Model of Television Advertising and Viewing Markets." *Marketing Science* 27, 356-378.
- [44] Xu, L., J. Chen, and A. Whinston (2010), "Oligopolistic Pricing with Online Search." *Journal of Management Information Systems* 27(3), 111-141.
- [45] Xu, L., J. Chen, and A. Whinston (2011), "Price Competition and Endogenous Valuation in Search Advertising." *Journal of Marketing Research* 48(3), 566-586.
- [46] Xu, L., J. Chen, and A. Whinston (2012), "Effects of the Presence of Organic Listing in Search Advertising." *Information Systems Research* 23(4), 1284-1302.
- [47] Yang, S. and A. Ghose (2010), "Analyzing the Relationship Between Organic and Sponsored Search Advertising: Positive, Negative, or Zero Interdependence?" *Marketing Science*, 30(1).