In Google we trust? ⋆

Roberto Burguet *, Ramon Caminal, Matthew Ellman

Institut d’Anàlisi Econòmica CSIC, Spain
Barcelona GSE, Spain

A R T I C L E   I N F O

Article history:
Received 25 January 2014
Received in revised form 8 February 2015
Accepted 15 February 2015
Available online 25 February 2015

JEL classification:
L13
L41
L82
L86

Keywords:
Search engine bias
Internet economics
Vertical integration
Two-sided markets
Antitrust

A B S T R A C T

We examine the incentives of a monopolistic search engine, funded by advertising, to provide reliable search results. We distinguish two types of search results: sponsored and organic (not-paid-for). Organic results are most important in searches for online content, while sponsored results are more important in product searches. By modeling the underlying markets for online content and offline products, we can identify the sources of distortions for each type of result, and their interaction. This explicit treatment proves crucial for understanding, not only spillovers across markets, but also fundamental policy issues, such as the welfare effects of integration. In particular, integration of the engine with a small fraction of content providers is welfare-enhancing when incentives to distort are stronger for sponsored than organic search, but welfare-reducing in the opposite case.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Search engines are indispensable for finding relevant content and products from the massive array of options available on the web. Numerous observers, including competition authorities, have questioned whether consumers and society can entrust this service to a monopolistic search engine that is funded by advertising from product sellers.1 To evaluate these concerns, it is important to identify the main sources of market failure in the markets that search engines intermediate.

This paper contributes to that task. We distinguish two types of search result, sponsored and organic (not-paid-for), and we model the separate underlying markets for online content and offline products.2 Explicit treatment of both types of search and both types of result allows us to identify the sources of distortions (away from consumers’ ideal matches) and proves crucial for analyzing spillovers across markets and fundamental policy issues, such as the welfare effects of integration. Indeed, one of the most important policy insights of this paper is that integration of the engine with a small fraction of content providers (henceforth, called publishers) is welfare-enhancing when incentives to distort sponsored search are stronger than those for distorting organic search, but welfare-reducing in the opposite case.

Before presenting our specific approach and contributions, we briefly sketch the engine’s trade-offs in choosing search reliability for each type of result separately. The trade-off is simplest for sponsored search. A search engine is a platform that connects consumers and merchants. Consumers searching for offline products (e.g., apparel, electronics, holiday packages) can use the search engine’s webpage to find links to relevant merchants’ websites. The engine typically charges merchants for well-placed links, called sponsored search links or “search ads,” but does not charge consumers. As with any two-sided platform that only

1 With 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). Sponsored search generated Google’s profits initially, but with recent expansion into ad intermediation (AdSense and AdWords) and content provision (e.g., Google Finance, YouTube and Zagat), display advertising from owned and affiliated websites accounts for an increasing share of profits (see e.g., Delo (2012)).

2 Website owners can buy (sponsor) well-positioned links to their websites in the sponsored, but not the organic, results.
charges one side, it is motivated to favor that side’s interest. Thus, an ad-funded engine might encourage consumers to visit the merchants most willing to pay for sponsored ads, even when not ideal for consumers. The need to attract consumers to search online provides a countervailing reputational incentive against excessive product search distortion.

A general search engine is also a platform that connects consumers and publishers. When consumers search for online content (e.g., news, entertainment, encyclopedias), most of the relevant results are organic (publishers have limited interest in sponsoring). But the engine may still have incentives to distort. Publishers provide content on websites with “display ads” that link consumers to merchants. So publishers are customers of the search engine in the content market, but also competing intermediaries in the product market. This gives the engine an incentive to distort content search to reduce display ad-effectiveness. Again, reputational concerns temper the distortion incentive.

Formally, our model consists of a dominant search engine, denoted by G, that interacts with a set of consumers, publishers and merchants. At its core, the search engine is a platform that intermediates two distinct but interacting markets: its search results match consumers with publishers in the content market and with merchants in the product market where publishers’ display ads also match consumers with merchants. The engine’s key instruments are the reliability of its organic and sponsored results. To make the analysis tractable, we assume that consumers use its organic results for online content search and its sponsored results for offline product search. This assumption is extreme but approximates the empirical evidence.

The model highlights three key aspects of the mechanism that generates organic search distortions. First, participating consumers use the same engine for both content and product searches. So the engine can attract users of its sponsored search results by building a reputation for reliable organic results. Indeed without such a participation spillover, the engine would not provide any organic results at all. Second, when a consumer buys a product via a display ad while searching for content, the consumer has less need to engage in product search. So the engine earns less from selling sponsored search ads. That is, display ads are substitutes for search ads.3 Third, G can make display ads less effective by diverting content-searching consumers. We consider two different mechanisms: (i) G can send consumers to less relevant publishers, frustrating contextual targeting (i.e., based on correlation between content and product taste) and inducing briefer visits and less time for ad display4; (ii) G can redirect consumers to publishers that are intrinsically less ad-effective or choose to be less ad-intensive (e.g., Wikipedia and the BBC have no ads). Thus G may distort organic search to steal business from publishers.

Our microfounded model allows us to study how organic and sponsored search distortions interact. We show that they are substitute instruments for G, owing to two mutually reinforcing effects. First, since display and search ads offer substitute channels for product-searching consumers, distorting one search type raises consumer sensitivity to the other. Second, participation spillovers imply that increased reliability of one search type relaxes G’s reputational concern when choosing the other.

We also predict how technological change and market characteristics affect equilibrium distortions. Technological innovations improving the effectiveness and targeting of display ads induce G to offer less reliable organic search results, potentially reducing consumer welfare, despite the positive direct effect of more effective display ads. Lower engine-use costs lead G to reduce the reliability of search results, worsening consumer access to offline products. Conversely, improvements in an alternative (non-strategic) search engine induce G to improve search reliability. As for market spillovers, when content search becomes more important for consumers, the engine unambiguously reduces sponsored search reliability, worsening consumer access to products. But in the converse direction, increasing the importance of product search has ambiguous effects on organic search reliability and access to content.5

Our second set of results address the question of how integration affects these distortions and welfare, and incentives to integrate. Integration may take place via direct publisher acquisition or via publishers paying a G-owned ad intermediary to serve their display ads. In either case, G internalizes a share of the publisher’s display ad rents. This has vertical and horizontal effects. The vertical effect is always beneficial to consumers: G values participation more and therefore tends to raise both sponsored and organic search reliability.

The horizontal effects are more subtle. Partial integration removes G’s incentives to engage in stealing business from publishers that become affiliates, but exacerbates G’s business-stealing incentives from independent publishers. G can now divert all distorted organic traffic to its affiliates, thereby increasing its display, as well as its search ad revenues. The fact that a small number of affiliates allows G to capture a sizable amount of diverted traffic generates a clear effect in our baseline model: when it is optimal for G to distort organic search, integrating a small number of publishers causes a discrete drop in consumer and total welfare. However, when instead G finds it optimal to distort sponsored search and not organic search, this effect is absent and the vertical internalization effect raises search reliability and welfare.

In sum, the welfare consequence of G’s integration of a small number of publishers depends critically on the relative incentives for distorting the two types of search. In addition, our derivation warns against using the pre-merger size of a given publisher to assess the risks of integration with G: G can divert traffic from non-affiliates to turn the publisher into a market leader.

The consequences of increasing the extent of integration from a non-zero level are much simpler. The detrimental horizontal effect falls with the number of non-affiliates (from which to steal business), leaving only the positive horizontal and vertical effects. So expanding integration raises search reliability and also welfare, independent of the relative distortion incentives.

Integration also affects publishers: the joint rents of G and affiliates always rise,6 but independent publishers’ rents may fall. Integration tends to decrease the traffic that G directs to independent publishers, reducing their display ad rents. This might have potentially serious consequences for content quality in the long run.

Variation in publishers’ display ad-effectiveness generates close parallels to our two sets of results. Under separation, G would then divert traffic from more to less ad-effective publishers, whereas with integration, G would divert from non-affiliated to affiliated and ad-effective publishers. So integration would lead to organic distortions that create a more ad-intensive experience for consumers.

The paper is organized as follows. We review the relevant literature in the next section. Section 3 presents the baseline model and characterizes the social optimum. Section 4 examines the equilibrium of the game with independent publishers. We study the effects of integration in Section 5 and Section 6 contains a few concluding remarks. The proofs are gathered in Appendix A.

---

3 Search ads are better targeted to product queries but display and search ads are increasingly close substitutes for merchants owing to intensive online content consumption and targeting innovations, e.g., contextual targeting and “cookies.”

4 Further empirical testing is needed, but this match effect (which requires no publisher asymmetry) is consistent with Google AdSense’s advice to publishers and eye-tracking studies such as Wang and Day (2007) and Lorigo et al. (2008).


6 We abstract from intra-organizational transaction costs. Such costs and regulatory concerns would limit the extent of consolidation but modeling them and predicting industry structure lie beyond the scope of this paper. Since non-affiliates may benefit from integration between G and other publishers, the process of bargaining over integration can become a war of attrition. The equilibria of such merger games typically involving reaching some agreement (integration) with probability one, but with delay.
2. Related literature

Our work is related to a number of recent papers, some studying organic search, others sponsored search. White’s (2013) Section 4 extension also considers both organic and sponsored results, though only for product searches. Organic and sponsored results interact because more results of either type raise product market competition, lowering the product price. The search engine’s advertising fee determines the number of sponsored results and hence the engine’s value to consumers. The number of relevant organic results is not a choice variable for the engine, but instead an unwanted consequence of efforts to improve “engine quality.”

We present a very different interaction mechanism with explicit content search and display advertising: by distorting organic search, the engine lowers the substitutability between publishers’ display ads and the engine’s sponsored search ads. The two models differ in some important predictions. For instance, in White (2013), a positive shock to engine quality (lowering user costs), can improve consumer access to products (by lowering price). This is essentially because White (2013) assumes that engine quality and the number of organic results move together, so the shock has a direct depressing effect on product price which can dominate. By contrast, we let the engine choose organic, as well as sponsored, search reliability, and as high engine quality exacerbates the engine’s incentive to distort results, the shock leads to unambiguously poorer access to products. So our modeling of organic results ensures that quality and price move in opposite directions, in contrast to the result stated in White’s Proposition 4.

In independent and contemporaneous work, de Cornière and Taylor (2014) study content search but not product search: they postulate a reduced-form for advertising, with ad nuisance costs, and with search engine and publisher ads as exogenous substitutes; that is, they assume that publisher ads lower the representative merchant’s willingness to pay for the engine’s ads. Their study and ours share the basic insight that integration between G and publishers has both horizontal and vertical effects on search distortions and consumer welfare. Our thoroughly microfounded model (with merchants optimizing both display and search advertising and consumers searching for products as well as content) lays out an explicit mechanism for the substitutability of search and display ads. This strategy allows us to study which market characteristics are most conducive to each type of search distortion. In addition, the search type distinction underlies key messages, such as the critical dependence of the welfare impacts of integration on relative distortion incentives.

Many papers have looked at (sponsored) product search. Our simple model is fully consistent with richer auction models tailored to analyze this market. The potential for conflicting interests in our product market is a mild variant on prior work where a search intermediary in a two-sided market tends to distort in favor of the paying side (merchants) against the non-paying side (consumers): in Hagiu and Jullien (2011), the intermediary distorts search to induce privately-informed consumers to search more, or to relax downstream product market competition, thereby raising merchant margins. In Athey and Ellison (2011), search bias is a side-effect of rent extraction from privately informed merchants. Xu et al. (2012) and Taylor (2013) model organic and sponsored results, but with only product search. They focus on how reliable organic results constitute “cannibalization” by G, since organic links give merchants a free substitute to sponsoring links on G (also true in our model). However, without the exogenous restrictions on link-clicking in Xu et al. (2012) and on link-setting in Taylor (2013), organic results would actually be superfluous for product search (as holds in our model).

Search bias is not easy to detect, but a Financial Times study (see Waters (2013)) suggestively identifies distortions in Google’s sponsored search results by comparing them with Google’s own Shopping search engine. For organic search, Edelman and Lai (2013) explore an instance of search diversion by Google towards integrated publishers and merchants. Empirical evidence on display and search ad substitutability is also scarce. Hahn and Singer’s (2008) survey of online retailers suggests significant substitutability; see also Ratliff and Rubinfeld (2010). More distantly, Goldfarb and Tucker (2011) present a natural experiment that indicates substitutability between online and offline search.

3. The baseline model

In this section we present a model of a monopolistic search engine that facilitates consumer search by providing both sponsored (paid) and organic (unpaid) results. The model explicitly considers two types of search goods and publishers as alternative platforms that compete with the search engine to host merchants’ ads. This setup allows us to study the incentives to distort organic as well as sponsored results, and their interaction. It also helps understand the horizontal and vertical externalities of search engine decisions on publishers and hence the engine’s incentives to integrate with publishers. Here, we present a streamlined version of the model. We discuss the key simplifying assumptions in Appendix A.

3.1. Two types of goods

We consider two different types of goods: online “content” and offline “products.” A mass one of consumers value specific varieties of each type of good. Online content is produced at zero cost and made available on publishers’ websites. Each publisher has one website with unique content. Each consumer has a favorite or “best-match” content that generates net utility, . Any other content generates zero net utility and consuming more than one unit implies a net loss.

Consumers also have a unit demand for products. Each product is unique and is produced by one of a continuum of merchants, who makes it available on its website. Each consumer has a favorite or “best-match” product that generates a high net benefit , and a “second-best-match” product that gives a low net utility , where . All other products or additional units imply a net loss. Some products generate a high margin, , for the merchant, while others generate a low margin, , where .

---

7 More broadly, we extend work on multihoming and gatekeeping: our search engine is a platform that determines multi-homing endogenously by controlling links to publisher platforms. See e.g., Athey et al. (2013), Anderson et al. (2013), Ambrus et al. (2013), White and Jain (2012) on multihoming. Baye and Morgan (2001) model an internet gatekeeper that controls links to merchants, but faces no competing platforms.

8 Despite product homogeneity, consumers implicitly click on all links so the price falls with number of results.

9 More specifically, White’s main result in Section 4 (Proposition 4) states conditions for price and quality to be complements or substitutes. Notice that they are complements in his baseline model, so it is precisely White’s modeling of organic results, introduced in this extension, that generates the effect pushing quality and price in opposite directions.

10 White and Jain’s (2012) parallel reduced-form instead assumes complementarity of the engine and publisher ads.

11 See e.g., Lehaie (2006), Edelman et al. (2007), Varian (2007), Liu et al. (2010), Athey and Ellison (2011), and Chen and He (2011). In the working paper version (Burguet et al., 2014), we explicitly model a scoring auction, that weights by merchant relevance, to allocate sponsored search links.

12 See also Chen and He (2011), Eliaz and Spiegler (2011), White (2013), and Xu et al. (2010, 2011) on the latter.

13 They also identify the possibility of advantageous selection as in Chen and He (2011) and, more generally, in settings with paid-for prominence (e.g., Armstrong and Zhou, 2011) and paid-for priority (Choi and Kim, 2010).

14 Tarantino (2013) provides further evidence and the point that a general search engine integrated into vertical search would divert traffic towards its vertical search engines (modeled as merchants, not 2-sided platforms).
3.2. Conflict of interest

Consumer and merchant interests are misaligned if the types of products most valued by consumers are not the most profitable for merchants. In general, interests will be aligned on some products and misaligned on others and our results only depend on misalignment arising for a positive fraction of searches. But for simplicity, we assume complete misalignment: all products generating high consumer value, \( \nu_H \), yield low merchant margins, \( m_L \), while all products generating low consumer value, \( \nu_L \), yield high merchant margins, \( m_H \). We also simplify the welfare analysis by focusing on the case, \( m_H + \nu_H > m_L + \nu_L \); that is, the social optimum requires that only high consumer value trades take place.

3.3. Platforms

The search engine, \( G \), intermediates the market for content. Consumers are unaware of the identity of their preferred publisher, but can use \( G \) to facilitate their search. They type in a query describing their desired content. \( G \) can perfectly interpret each query and identify the publisher providing the consumer’s best-match content. It is optimal for \( G \) to respond with a single link to a publisher’s website. We assume that this link is organic (non-paid for), \( G \) chooses the probability, denoted by \( \pi^G \), that this link leads to the site of the consumer’s best-match publisher/favorite content, otherwise linking to some other publisher’s site.

\( G \) also intermediates the product market. When consumers search for products on \( G \), they type in a query describing their desired product, which \( G \) can interpret perfectly to identify each consumer’s best-match and second-best-match products. A natural interpretation is that each query describes a “product category” characterizing these two products. \( G \) can also identify the two relevant merchants who sell them. \( G \) responds with a single sponsored result, that it sells as a search ad to a relevant merchant as described in ad prices just below; \( G \) provides no relevant organic results.

While \( G \) is the only intermediary in the content market, publishers also intermediate the product market: they tie display ads to the content on their platforms. Display ads are, like search ads, purely informative. Each publisher offers one “display ad” slot for a link to a merchant’s website (see ad prices). Whenever a consumer visits a website to consume its content, she notices the ad with probability 1 if she is visiting her best-match publisher, and probability \( \beta < 1 \) if she is on any other publisher’s site. Publishers do not perfectly identify the consumer’s favorite products, and only observe a noisy signal. With probability \( \sigma \in [0, 1] \), the signal identifies the consumer’s product category of interest (the two favorite products), and otherwise points to a random, different category. Thus, parameter \( \sigma \) represents the precision of the targeting technology for display advertising.

3.4. Ad prices

When a consumer enters a query for a product in \( G \)’s website, \( G \) identifies the consumer’s relevant category and the merchants producing the high and low (consumer) value products. \( G \) makes a take-it-or-leave-it offer to one of the merchants, offering to show its link (as the single result), at a given price per click or “PPC”; this is the search “sponsorship” price. \( G \) can (and in equilibrium, will) randomize over the two relevant merchants: with probability denoted \( \pi^G \), \( G \) offers the merchant selling the high value product to show its link for a price, and with complementary probability \( 1 - \pi^G \), \( G \) offers to link to the merchant selling the low value product, charging a possibly different price. Hence, \( \pi^G \) represents the “reliability” of sponsored search results just as \( \pi^G \) represents the “reliability” of organic results. Similarly, each publisher makes a take-it-or-leave-it offer PPC to one of the merchants after observing its targeting signal.\(^{16}\)

3.5. 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 has a joint cost \( c \) of using the search engine for both content and product queries; this includes the foregone expected benefits from alternative search instruments. Each cost \( c \) is observed privately before deciding on participation and is an independent draw from a continuous random variable on \([0, \infty)\) with density function \( f(c) \) and cumulative distribution function \( F(c) \). We assume that the reverse hazard rate, \( H(c) = \frac{f(c)}{1 - F(c)} \), is decreasing.\(^{17}\) Gross of her direct-plus–opportunity costs \( c \), the consumer’s expected gain from participating in online search is the sum of expected gains from consuming online content plus the offline products found via display or search ads. The highest possible gain is \( u + \nu_H \) and we assume \( \mathcal{T} > u + \nu_L \) so that consumer participation in online search is interior in any equilibrium.

3.6. The timing

At the beginning of the game, \( G \) announces the reliability of its search results, \( \pi^G \) and \( \pi^G \). All consumers observe this information and decide whether or not to use the search engine. If a consumer decides to search on \( G \), she types in her content query and visits the website that appears as the top organic search result. This publisher observes its signal of the consumer’s preferences on products and sells its display ad to a merchant. If the merchant accepts the offer, the consumer may click on the merchant’s display ad to the merchant’s webpage where she may buy the merchant’s product. If she buys it, she leaves the market. Otherwise, she types in her product query, \( G \) responds by selling (sponsored) search ads and the consumer visits the sponsoring merchant’s website (that appears as the top sponsored search result), where she may buy that merchant’s product.

None of our results hinge 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 found via publishers’ display ads, reducing their need to use \( G \) for product search. Finally, to simplify, we assume a continuum of publishers and merchants, with each publisher’s content and merchant’s product interesting the same fraction of consumers.

3.7. The social planner’s problem

Consider 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, since \( m_H + \nu_H > m_L + \nu_L \). This is feasible: \( G \) can send each content-searching consumer to the best-match publisher and can send any product-searching consumer to a best-match merchant. It requires undistorted search results as consumers have no other way to find content, and display ads are an imperfect alternative to sponsored search results, leading to best-match products with at most probability \( \sigma \leq 1 \). So we have,

\(^{16}\) A parameter for reduced click-through on display relative to search ads would not affect our qualitative results.

\(^{17}\) This ensures that \( G \) faces an increasing ‘marginal factor cost’ of attracting the consumer base it “sells” on to merchants.
3.8. Benchmark

Total surplus maximization requires the search engine to allocate traffic without distorting from the consumer’s ideal; in the first-best, \( r^0 = r^* = 1 \).

This also holds in the constrained scenario where the planner cannot control consumer behavior. The need to attract consumers, who neglect their positive participation externality on producer surplus, only reinforces the planner’s incentives to not distort search.

4. Equilibrium analysis with independent publishers

In this section, we consider the no-integration case where all publishers are independent of the search engine. We focus on the engine’s incentives to distort sponsored and organic results, showing how the two types of distortions are imperfect substitutes from the engine’s viewpoint. We also derive a number of interesting comparative static results. First, an improvement in the targeting technology used to allocate display ads may hurt consumers. Second, improving the search engine’s baseline attractiveness (called quality, \( s \), in White (2013)) tends to reduce consumers’ online access to products. We also study how changes in the market for content spill over to the market for offline goods and conversely. The substitutability between the two types of distortions is a critical factor in all these cases.

Throughout the paper, we solve for subgame perfect equilibria in undominated strategies. Consumers make many decisions but most are immediate once stated. Consumers always click on the single link after entering either type of query; they follow G’s “recommendation.” Consumers buy at most one product unit, accepting low value products (second-best-matches) in the product search stage (their last chance to find a product), but only buying via a display ad (in the earlier content search stage) if it is high value (best-match).

Importantly, consumers omit the product search if a display ad already satisfies their unit demand for products; they always conduct the content search if they pay their cost of participating in online search.

4.1. Pricing sponsored search ads

Consumers who did not purchase offline goods through display advertising enter their product query in the search engine. This is the last chance for merchants (and consumers) to trade and \( G \) has all the market power, so \( G \) sets a price per click equal to the merchant’s willingness to pay. Since the relevant low margin merchant implies high consumer value and vice versa, it follows that \( G \)’s average revenue per product search equals the average margin, denoted \( M(r^0) \), on sponsored-search-mediated sales: \( M(r^0) = r^0 m_L + (1 - r^0)m_H \).

4.2. Pricing display ads

In the first stage of search, content search, each consumer lands on a publisher’s website where merchants can buy display ads. These purchases have no opportunity cost for merchants since they anticipate zero rents from any subsequent product searches. High margin merchants clearly have no demand for display ads since consumers never buy low value products in the content stage. By contrast, the low margin (high value) merchant indicated by the targeting signal is willing to pay price per click \( c m_L \), since each click generates a sale with the probability \( \sigma \) of relevance. Other low margin merchants are less willing to pay since less likely relevant. So the publisher always offers the slot to the low margin merchant indicated by the signal and charges \( c m_L \) per click.

Since each consumer clicks on a display ad with probabilities \( 1 \) and \( \beta \) for favorite and other publishers, respectively, the average equilibrium fraction of such clicks is \( r^0 + (1 - r^0)\beta \). A fraction \( \sigma \) of these clicks end up in a transaction, so the fraction, \( \eta \), of participating consumers who buy a product via a display ad is given by,

\[
\eta = \sigma r^0 + (1 - \sigma)\beta.
\]

Note that \( \eta \) increases with organic search reliability, \( r^0 \), and the quality of the targeting technology, \( \alpha \).

4.3. Consumer participation and continuation equilibrium

Consumers participate if their opportunity cost \( c \) is below a cut-off, denoted by \( \hat{c} \), equal to the expected benefit from search participation:

\[
\hat{c} = r^0 u + \nu_H - (\nu_H - \nu_L)(1 - \eta)(1 - r^0).
\]

The first term is the expected payoff in the market for content: with probability \( r^0 \), a participating consumer finds her preferred content worth \( u \). The other two terms represent the expected payoff in the market for products: with probability \( (1 - \eta)(1 - r^0) \), a participating consumer does not reach a relevant display ad and product search only gives access to the second-best-match product worth \( \nu_L \), but otherwise she gets \( \nu_H \).

Note that \( \hat{c} \) increases with \( r^0, r^2, \) and \( \alpha \), and \( \frac{\partial \hat{c}}{\partial \beta} > 0 \) and \( \frac{\partial \hat{c}}{\partial \alpha} > 0 \). So distorting traffic (lowering \( r^0 \) and/or \( r^2 \)) costs \( G \) by reducing the mass of consumer participants, \( F(\hat{c}) \). Moreover, \( G \)’s two instruments, \( r^0 \) and \( r^2 \), play a similar and substitutive role in encouraging participation: a high value of one reduces the sensitivity of participation to the other.

Now a mass \( F(\hat{c}) \) of consumers trade via display advertising, leaving a mass \( (1 - \eta)F(\hat{c}) \) who conduct a product search and trade via search advertising. Drawing all this together, we have:

**Lemma 1.** In the unique continuation equilibrium after \( G \) sets \( \langle r^0, r^2 \rangle \), merchants make zero profits, publishers earn \( \Gamma^G = F(\hat{c})\eta m_L \), and the search engine earns \( \Pi^G = F(\hat{c})(1 - \eta)M(r^0) \).

4.4. Traffic management

The profit expression, \( \Pi^G \), identifies \( G \)’s two trade-offs. First, raising \( r^0 \) increases consumer participation, \( F(\hat{c}) \), but decreases the fraction of consumers who engage in product search, \( 1 - \eta \). That is, less distorted organic search raises display-mediated product trades and lowers sponsored link sales. Formally, the first-order condition with respect to \( r^0 \) of \( G \)’s optimization problem is:

\[
\frac{\partial \Pi^G}{\partial r^0} = \frac{1}{M(r^0)} \frac{\partial M(r^0)}{\partial r^0} f(\hat{c}) \frac{\partial \hat{c}}{\partial r^0} - (1 - \eta) F(\hat{c}) \frac{\partial \eta}{\partial r^0} = 0.
\] (3)

Second, raising \( r^2 \) attracts more users, \( F(\hat{c}) \), but lowers the average margin of trades via sponsored search and so the average price of sponsored links, \( M(r^0) \). Formally:

\[
\frac{\partial \Pi^G}{\partial r^2} = \frac{1}{1 - \eta} \frac{\partial F(\hat{c})}{\partial r^2} M(r^0) + F(\hat{c}) \frac{\partial M(r^0)}{\partial r^2} = 0.
\] (4)

Distorting either search type reduces \( \hat{c} \) and with it the relative sensitivity of participation to \( \hat{c} \), measured by the reverse hazard rate, \( H(\hat{c}) = \frac{1}{\partial(1 - \eta)} \frac{\partial \hat{c}}{\partial(1 - \eta)} = 1 \).

For this is a common element in both first-order conditions. To better understand these first-order conditions, we rewrite them as:

\[
H(\hat{c})(1 - \eta) \left[ -\frac{u}{\alpha(1 - \beta)} + (\nu_H - \nu_L)(1 - r^0) \right] = 1.
\] (5)
\[ H(\hat{c})(1-\eta) \left[ \frac{v_H-v_L}{m_H-m_L} + (v_H-v_L) \left( 1-r^2 \right) \right] = 1. \] (6)

We can now characterize the conditions under which G chooses to distort at least one type of result. If the left hand side (LHS) of either Eq. (5) or Eq. (6) evaluated at \( r^2 = r^2 = 1 \) is less than 1 then at least one type of result is distorted in equilibrium. A sufficient condition on the primitives is,

\[ H(u + v_H)(1-\sigma)m_H \min \left\{ \frac{u}{\alpha(1-\beta)m_L} - \frac{v_H-v_L}{m_H-m_L}, 1 \right\} < 1. \] (7)

This is also necessary, as the LHS of Eq. (5) and Eq. (6) are both decreasing in \( r^2 \) and in \( \rho^2 \). So we have:

**Proposition 1.** The search engine distorts traffic, \( r^2 < 1 \) and/or \( \rho^2 < 1 \), if and only if Eq. (7) holds.

G’s choices \( (r^2, \rho^2) \) have both vertical and horizontal external effects on publishers, whose profits are \( \Pi' = (\hat{c})'m_G \). (i) Both \( r^2 \) and \( \rho^2 \) raise consumer participation, \( F(\hat{c}) \), benefitting both G and publishers (the vertical externality); (ii) but \( \rho^2 \) raises the “effectiveness” of display advertising, \( \hat{c} \), which increases clicks on publishers’ ads and lowers clicks on sponsored links. 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. Under the baseline assumption of symmetric publishers, in the limiting case of \( \beta = 1 \), G cannot reduce the effectiveness of display advertising and G’s relationship with publishers is purely vertical.\(^{19}\) But if \( \beta < 1 \), G can accomplish business-stealing from publishers’ competing display ad platforms by distorting organic traffic to reduce the effectiveness of publishers’ websites as advertising outlets. This hurts consumers as well as publishers, even if \( r^2 = 1 \). So, in contrast to typical horizontal relations where competition benefits consumers, the horizontal externality results in lower consumer, and total, welfare.

The minimum expression in Eq. (7) identifies G’s relative incentives to distort its two instruments, \( r^2 \) and \( \rho^2 \). G has stronger incentives to distort organic search than to distort sponsored search if \( \frac{u}{\alpha(1-\beta)m_L} < \frac{v_H-v_L}{m_H-m_L} \), and conversely if the inequality is reversed. The intuition is simplest for distortions from \( r^2 = \rho^2 = 1 \). These fractions then represent the respective cost–benefit ratios from marginally distorting organic and sponsored searches: organic distortions reduce per-consumer search surplus at the rate \( u \) (all product trade values equal \( v_H \) when \( r^2 = 1 \)) while raising G’s per-searcher ad revenues at the rate \( \alpha(1-\beta)m_L \) as search-based trades substitute for display-based trades; sponsored search distortions (reducing \( \rho^2 \)) lower per-consumer-searcher consumer surplus at the rate \( v_H - v_L \) while raising G’s per-product-searcher ad revenue at the rate \( m_H - m_L \). Only when these ratios are identical can the first-order conditions hold simultaneously and both instruments \( r^2 \) and \( \rho^2 \) be interior in \( (0, 1) \).\(^{20}\)

\[ \frac{u}{\alpha(1-\beta)m_L} = \frac{v_H-v_L}{m_H-m_L}. \] (8)

4.5. The substitutability between G’s instruments

G’s profit function satisfies the standard definition of substitutability: \( \hat{c}^{u', v_H} < 0 \). A shock that raises one of the instruments will tend to reduce the other. For instance, more reliable sponsored results (higher \( \rho^2 \)) foster consumer participation, \( F(\hat{c}) \), and reduce the sensitivity of consumer participation to changes in \( \rho^2 \) (lower \( \rho^2 \)), both of which exacerbate G’s incentives to distort organic search (lower \( r^2 \)).

---

19 Even if \( \beta = 1 \), G may distort organic search if publishers vary in display-ad-effectiveness (see Section 5).

20 This is non-generic but as Eqs. (5) and (6) have continuous LHS’s, \( r^2 \) or \( \rho^2 \) is interior for a positive measure of parameters.

### 4.6. Improved display ad targeting

An increase in \( \sigma \) allows better targeting of publisher’s display ads. This reduces the LHS of both Eqs. (5) and (6). So, from an equilibrium with only \( r^2 \) interior, increased \( \sigma \) reduces G’s optimal level of \( r^2 \) and, when only \( r^2 \) is interior, it reduces \( r^2 \). These negative indirect effects on consumer welfare mitigate the positive direct effect of improved targeting (which raises consumer access to best-match products). Indeed, for organic distortions, the negative effect dominates.\(^{21}\) For sponsored search distortions, either effect may dominate, so consumer welfare may rise or fall:

#### 4.6.1. Comparative static 1

Enhanced display ad targeting technologies have a positive direct effect on consumer welfare by facilitating access to best-match products, but the search engine reacts by reducing search reliability so that consumer welfare and publisher rents may fall.

A parameter change can simultaneously shift both instruments, but only at the non-generic points where Eq. (8) holds. For instance, fixing all parameters other than \( \sigma \) and raising \( \sigma \) from below to above the value \( \sigma^* \) at which Eq. (8) holds, causes the LHS of Eq. (5) to fall from below to above the LHS of Eq. (6), so that G switches from strictly preferring sponsored over organic distortions to the converse. If \( r^2 \in (0, 1) \) and/or \( \rho^2 \in (0, 1) \) at \( \sigma^* \), this provokes an upward jump in \( r^2 \) and a downward jump in \( \rho^2 \).\(^{22,23}\)

### 4.7. Search engine attractiveness

We briefly examine the consequences of exogenous changes in the engine’s attractiveness (at given reliability choices). For instance, consumers’ alternatives to using G might worsen or G might see how to make the engine easier to use. Following White (2013), we define engine attractiveness, \( s \), as a leftward shift in the distribution \( F(\hat{c}; s) \) of consumers’ participation costs.\(^{24}\) Participation, \( F(\hat{c}; s) \), now increases with \( s \) as well as \( \hat{c} \), while the reverse hazard rate, \( h(\hat{c}; s) \), decreases with both, and Eq. (8) is unaffected. So raising \( s \) leads G to lower search reliability, \( r^2 \) or \( \rho^2 \) (whichever is interior) and this lowers the utility that consumers gain from offline products:

#### 4.7.1. Comparative static 2

An increase in search engine attractiveness leads to less accessible products.

This is in contrast to White (2013)’s ambiguous result. The main reason for the discrepancy is that he assumes that organic result quality is tied to search engine attractiveness, whereas we let the engine choose the quality of organic results optimally, and it opts to unambiguously reduce result quality when its attractiveness \( s \) rises.

### 4.8. Cross-market interactions

The interlinkage between content and product markets via the substitutability of display and search advertising generates cross-market spillovers. For instance, suppose online content becomes more important for consumers, in that their value from best-match content, \( u \), rises. If only \( r^2 \) is interior, the rise in \( u \) relaxes the participation constraint, and this leads to greater sponsored search distortions; that is, \( r^2 \) falls. So increasing...

---

21 In particular, if \( r^2 = 1 \), \( \sigma \) has no direct effect as consumers can anyway find high value products via search ads, so raising \( \sigma \) clearly generates lower consumer welfare, by provoking reductions in \( r^2 \) or \( \rho^2 \), and this generalizes.

22 A sufficient condition is: \( H(u + v_H)(1 - \sigma^*)u < \sigma^*(1 - \beta) < H(v_H + \sigma^*(v_H - v_L))(1 - \alpha^*)u \).

23 Consumer payoffs remain continuous because they are indifferent among G’s optimal distortion set at \( \sigma = \sigma^* \).

24 White (2013) calls s “quality” but note that it is distinct from the engine’s search reliability choices.
the content market’s importance can reduce accessibility of offline products. Suppose now that offline products become more important for consumers, so that \( v_p \) increases, and consider its effect on the content market. Relaxation of the consumer participation constraint similarly induces \( G \) to reduce \( r^D \), but there is an additional effect: the sensitivity of consumer participation to changes in \( r^D \) also increases and this reduces the incentives to distort organic search. The net effect is now ambiguous. The reason for this asymmetry in the spillovers between content and products is simple: \( r^D \) has no direct effect on consumers’ access to content, whereas \( r^P \) directly affects their access to products by enhancing display advertising.25

5. Integration

We now examine the consequences of \( G \)'s integration with a fraction of publishers. We initially take integration with a publisher to mean that \( G \) obtains a share of the publisher’s display ad profits, but no control over its display ad supply. This is appropriate for agreements in which affiliated publishers deal exclusively with an ad intermediary owned by \( G \) or, equivalently, acquisitions where the business model maintains publishers as separate decision units. In both cases, we refer to the publishers involved as G-affiliates. At the end of the section, we discuss the alternative scenario in which \( G \) also controls integrated publishers’ display ad policies.

We have already seen how \( G \) may use organic search distortions to steal business from non-affiliated publishers (all were non-affiliated in the last section). Integration heightens this motive since, by diverting to its affiliates, \( G \) now gains in display, as well as search, ad revenues. On the other hand, these affiliates’ display revenues raise \( G \)'s revenues per participating consumer, so integration also encourages \( G \) to raise both search and display reliabilities.

To analyze the consequences of these changes, this section introduces two extensions of the game examined so far. First, \( G \) affiliates a fraction \( \gamma \) of publishers before it fixes its search reliability. For simplicity, we assume that \( G \) becomes the residual claimant on all affiliated display ad revenues.26 Now \( G \) may want to treat affiliates and non-affiliates differently. Hence the second change: given \( \gamma \), but simultaneous with setting \( r^D \), \( G \) sets two possibly different values, \( r^D_{\text{aff}} \) and \( r^D_{\text{non-aff}} \), for the reliability of organic results on queries where the best-match publisher is, respectively, affiliated and not.

Consistent with earlier assumptions, consumers do not know whether their best-match publisher is affiliated with \( G \) or not. So their participation decisions are based on average organic search reliability and we can retain the notation \( r^D \) and \( \eta \), where Eq. (1) and Eq. (2) still define \( \eta \) and \( \tilde{\gamma} \) with now,

\[
r^D = \eta r^D_{\text{aff}} + (1-\gamma) r^D_{\text{non-aff}}.
\]

Given that \( G \) internalizes the display ad profits of its affiliates, it is immediate that \( G \) diverts any distorted organic traffic to a \( G \)-affiliate. So non-affiliated publishers only receive visits when their content is best-match for the content query and \( G \) answers the query without distorting. Thus, the average of display-mediated product trades (per participating consumer) on each non-affiliate’s site is:

\[
\eta_{\text{non-aff}} = \sigma r^D_{\text{non-aff}}.
\]

whereas the average of product trades on each \( G \)-affiliate’s site is:

\[
\eta_{\text{aff}} = \sigma r^D_{\text{aff}} + \left(1 - r^D_{\text{aff}}\right) \beta + \frac{1 - \gamma}{\gamma} \left(1 - r^D_{\text{non-aff}}\right) \beta.
\]

Total display-mediated trades (per participating consumer) are:

\[
\eta = \eta_{\text{aff}} + (1-\gamma) \eta_{\text{non-aff}}.
\]

All publishers sell their display ads to low-margin merchants at \( c_{mL} \). So each non-affiliated publisher earns rents \( \Pi_{\text{non-aff}}^D = F(\tilde{\gamma}) c_{mL} m_{L} \). Finally, \( G \)'s profits are:

\[
\Pi_{\text{G}}^D = F(\tilde{\gamma}) \left[ (1-\eta) M(r^D) + \gamma \eta_{\text{aff}} m_{L} \right].
\]

The interpretation of Eq. (9) is simple. As in Lemma 1, each of the \( F(\tilde{\gamma}) \) participating consumers clicks on and trades via a sponsored link with probability \( (1 - \tilde{\gamma}) \) and, on average, each click results in a merchant payment of \( M(r^D) \) to \( G \). Each participating consumer also trades via a \( G \)-affiliate’s display ad with probability \( \gamma \eta_{\text{aff}} \), yielding \( G \) an average display ad revenue per participant of \( \gamma \eta_{\text{aff}} m_{L} \).

To simplify the presentation, we restrict attention to scenarios where, under separation, only one instrument, \( r^D \) or \( r^P \), is distorted and it takes a strictly positive value (i.e., one is interior, the other undistorted); Appendix A2 gives necessary and sufficient conditions.27 Since the effects of integration prove critically dependent on the relative incentives to distort sponsored versus organic search results, we distinguish two parameter regions: the inequality (explained in A2),

\[
\frac{u - \gamma_{\text{aff}} - v_{L}}{m_{L} - m_{L}} > 0
\]

defines region SD where \( G \)'s incentives to distort sponsored search relative to organic search are strong (the label SD indicates a tendency towards sponsored distortions); the reverse inequality defines region OD (organic distortions) where it is \( G \)'s organic distortion incentives that are relatively strong. As well as determining distortions under integration, inequality (10) guarantees that \( r^D > r^P \) and only sponsored search is distorted under separation, on SD.28 By contrast, \( r^P \) could be distorted instead of \( r^D \) under separation on OD.

5.1. Toeholds

We begin by examining the consequences for \( G \) of entering the display advertising market with just a toehold, i.e., integrating with a small fraction of publishers \((\gamma > 0 \text{ but small})\). Integration has three effects. First, the display ad losses from diverting traffic away from affiliated publishers encourage \( G \) to raise \( r^D \). Second, internalizing vertical externalities on affiliated publishers encourages \( G \) to raise \( r^D \) and \( r^P \) to raise participation. But these two positive effects are small when the fraction \( \gamma \) of affiliated publishers is small. By contrast, a third, negative effect may be large: \( G \) will now divert any distorted organic traffic to its own affiliates, capturing a discrete fraction of display ad rents if \( r^D_{\text{aff}} < r^P \).29 So, on region OD, a small integration has a discrete negative effect on \( r^P \) and on welfare. On SD, \( r^P \) remains undistorted and vertical internalization raises \( r^D \). In consequence (proof in Appendix A):

Proposition 2. \( G \)'s integration with a small fraction of publishers affects consumer and total surplus: positively on SD and negatively on OD. On SD, sponsored reliability rises and organic reliability is unchanged, while

25 Changes in product market profitability have simpler spillover effects, since they do not directly influence participation. For instance, increasing \( m_{L} \) raises incentives to distort sponsored search. If \( r^D \) is at a corner solution and \( r^P \) > 0, this has no effect on either market, until \( m_{L} \) crosses the threshold (8) and then \( G \) responds by increasing \( r^D \) (and reducing \( r^P \). Thus, an increase in product market profitability can facilitate consumer access to content.

26 Other corner solutions complicate without adding much economic insight, except in one case: see footnote 30.

27 Eq. (10) implies LHS - RHS in Eq. (8); integration raises the relative attractiveness of organic over sponsored distortions.

28 G’s per consumer profits from its affiliated publishers, \( \gamma \eta_{\text{aff}} m_{L} \), converges to \( \sigma(1 - r^P) / m_{L} > 0 \), as \( \gamma \) goes to 0.
on OD, sponsored reliability may rise but organic reliability falls and dominates.

Proposition 2 shows that the consequences of integration for welfare depend heavily on G’s relative preferences for using the two types of distortions. When G strongly prefers to distort sponsored search but is well motivated to offer reliable organic search results, integration necessarily reduces distortions: on SD, it remains too costly to distort organic search, and vertical internalization raises sponsored search reliability. In contrast, when G prefers to distort organic search, integration with any publishers exacerbates that distortion. Even though integration may reduce sponsored distortions, the negative effect on organic distortions dominates and consumer welfare decreases.\textsuperscript{30}

**Remark.** Note the subtle antitrust consequences of these results. Pre-merger size may be a poor guide for predicting the consequences of integration between G and a given publisher. Even if pre-merger, G had no incursion into publishing or display advertising, a merger with a minor player may have large effects, because G has incentives to turn an affiliated minor player into a leader in the market, and in the process reduces the reliability of search and consumer welfare.\textsuperscript{31}

5.2. Integration expansions

Turning now to expansions in integration, raising $\gamma$ from any non-zero level $\gamma > 0$, the vertical incentives for reliability continue to increase. On region SD, this immediately implies that sponsored reliability improves (and organic search remains undistorted). On region OD, $r_{OD}$ or even $r^p$ may fall, but only in response to, and never dominating, the direct positive effect of increased $\gamma$ on average organic reliability and consumer welfare. The positive direct effect is simple: increasing integration raises the fraction of consumers whose best-match content is on a G-affiliate, raising average organic reliability as $r_{OC} < r_{CG}$. Summarizing (again proven in Appendix A):\textsuperscript{32}

**Proposition 3.** For $\gamma > 0$, expanding the integration level $\gamma$ raises average organic search reliability as well as consumer and total welfare: sponsored search reliability increases on SD and decreases on OD.

Overall, consumer and total welfare increase continuously with integration as $\gamma$ rises from 0 upwards when G has relatively strong incentives to distort sponsored results (SD). When the incentives to distort organic results are dominant (OD), integration initially causes a discrete fall in consumer and total welfare, but welfare then increases with expanding integration (and may eventually exceed the welfare level under separation).

5.3. Endogenizing integration

So far, we have discussed the consequences of exogenous changes in integration. To endogenize the level of integration would require an explicit consideration of the gains from integration, but also the (organizational and regulatory) costs of integration, and the bargaining process that determines how these gains are shared. Although a full analysis lies beyond the scope of this paper (see footnote 6 on the costs and on bargaining intricacies), a couple of remarks are in order. First, on the potential gains, integration allows G to internalize its vertical externality on affiliated publishers, and to steal business more effectively from non-affiliates (via G-affiliates’ display ads, as well as search ads). Both effects result in a larger joint profit for the conglomerate. So there are always potential gains from expanding the level of integration, but the magnitude of these gains varies: on SD, the gains emerge from the vertical internalization alone, whereas on OD, integration sometimes allows G to also benefit from more intense business-stealing from non-affiliates that amplifies the return to integration. Second, on the nature of integration agreements, having G as residual claimant on display ad revenues is in fact optimal for conglomerate profits. So, independent of bargaining power, any privately efficient agreement would generate exactly the results of Propositions 2 and 3.\textsuperscript{33}

5.4. Publisher rents

Integration also affects the revenues of non-affiliated publishers. The effect is positive on SD, since non-affiliates benefit from increased consumer participation and organic reliability stays fixed. However, a small integration is bound to hurt non-affiliates on OD, since both consumer participation and the reliability of organic results then decrease. With further integration, consumer participation recovers but traffic to non-affiliates remains lower than under separation if organic distortion incentives are strong enough. These effects of integration on publisher revenue could be important in a more general model where publishers choose the quality of their content (raising $u$) or whether to enter. Thus, integration could have serious consequences for content quality and, given scale economies in publishing, for diversity.

5.5. Publisher heterogeneity

So far, we assumed all publishers to be equally effective and equally intensive as advertising platforms. Given homogeneity, diverting organic traffic away from consumers’ preferred publishers permitted stealing business only under our assumption ($\beta < 1$) that consumers spend less time on less preferred websites, or that organic diversions frustrate contextual targeting. Even without this assumption (i.e., if $\beta = 1$), the same effects would obtain if some publishers are more effective for display advertising (inducing higher click-through rates) than others, or simply more ad intensive. Then, under separation, G would prefer to divert traffic from more to less ad-effective (or intensive) publishers. Under integration, the optimal diversion would change direction: G would prefer to divert traffic from non-affiliated to affiliated, ad-intensive publishers. Thus, integration would lead to organic distortions that create a more ad-effective or ad-intensive experience for consumers.\textsuperscript{34}

5.6. Control over display advertising

We have assumed throughout that publishers maintain control over the quantity and characteristics of their display ads after integration. When G integrates directly with publishers, rather than only affiliating them via a G-owned ad intermediary, the integrated structure may control the supply and nature of its publishers’ display ads. The above qualitative results still stand, but the integrated firm would conduct less organic distortion. Instead of ever distorting organic search among affiliates to reduce cannibalization of high margin search ads by lower

\textsuperscript{30} Moving beyond our restriction to settings where only one instrument is partially distorted under separation, there is one new case with results of a different flavor: if under separation, $1 > r^p / r^f > 0$, then integration with fewer publishers may result in lower organic reliability and lower consumer welfare even on SD. So Proposition 2 adjusts, but only when distortion incentives are so high that $r^f = 0$ and stays there.

\textsuperscript{31} In a more general model where publishers are not all equally imperfect substitutes, each minor player would only become the champion of its specific content sector, but our analysis would readily generalize.

\textsuperscript{32} It may come as a surprise that sufficient integration does not eventually eliminate sponsored distortions (given publisher homogeneity). The reason why full integration does not guarantee search reliability is that G cannot control display advertising, so G can ensure a higher margin per product trade mediated via a search ad than by an affiliates’ display ad.

\textsuperscript{33} Even if merging partners cannot achieve full maximization of joint profits, our qualitative results still hold if partners grant G a fixed fraction of display-ad rents.

\textsuperscript{34} This is reminiscent of de Cornière and Taylor’s (2014) results on ad intensity, but in their work, the consumer loses owe to an assumed distaste for ads. In our model, display ads do not hurt consumers, rather the opposite: a consumer benefit from ads (be they search or display ads) that link them to merchant webs underlies the substitutability of display and search ads that drives our results (and those of related papers). Instead, here the consumer welfare loss results directly from poorer access to relevant content.
margin affiliated display-ads, it would directly eliminate those display ads or restrict them to high margin products.

6. Concluding remarks

This paper has considered a monopolistic search engine’s incentives to distort sponsored and organic search. Our microfounded model of the engine as information gatekeeper in the markets for online content and offline products allowed us to explicitly explore the causes and consequences of display and search ad substitutability. The model readily generated comparative statics on key industry parameters. We were also able to identify the distortion, distribution and welfare effects of integration into content provision or display advertising. Further empirical work on the relative importance of these two types of distortions is a vital complement of this research. Google’s apparent Midas touch in turning small and moderate companies in a number of content sectors into large or dominant players (Youtube, Blogger and Google Books, Health, Finance, Flight Search and Maps all spring to mind) is certainly consistent with our model but not a proper test.35

We have investigated an environment where sponsored and display advertising are substitutes. This is in line with the view of the European Commission, and contrary to that of the FTC, which judged them not to be substitutes (see EU (2008) and FTC (2007)). Online advertising, and not only search advertising, may then constitute a relevant market from the perspective of competition policy. Building on this view, our model offers a new angle on the debate over more recent concerns raised by the European Commission. In November 2010, the Commission initiated investigations into alleged malpractices by the engine as information gatekeeper in the markets for online content provision or display advertising. Further empirical work on the relative importance of these two types of distortions is a vital complement of this research. Google’s apparent Midas touch in turning small and moderate companies in a number of content sectors into large or dominant players (Youtube, Blogger and Google Books, Health, Finance, Flight Search and Maps all spring to mind) is certainly consistent with our model but not a proper test.35

We have investigated an environment where sponsored and display advertising are substitutes. This is in line with the view of the European Commission, and contrary to that of the FTC, which judged them not to be substitutes (see EU (2008) and FTC (2007)). Online advertising, and not only search advertising, may then constitute a relevant market from the perspective of competition policy. Building on this view, our model offers a new angle on the debate over more recent concerns raised by the European Commission. In November 2010, the Commission initiated investigations into alleged malpractices by the engine as information gatekeeper in the markets for online content provision or display advertising. Further empirical work on the relative importance of these two types of distortions is a vital complement of this research. Google’s apparent Midas touch in turning small and moderate companies in a number of content sectors into large or dominant players (Youtube, Blogger and Google Books, Health, Finance, Flight Search and Maps all spring to mind) is certainly consistent with our model but not a proper test.35

Further regulatory insights could be developed by extending the present analysis. Endogenizing content quality via costly investment and entry and exit of publishers is particularly important; pursuing our publisher rents discussion (Section 5) may uncover substantial long-run welfare effects of integration. Other extensions worth pursuing are endogenizing integration along the lines discussed in Section 5, enriching the analysis of targeting, and modeling investment in search engine precision.

Appendix A

A.1. Discussion of the benchmark model

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.

35. Empirical work should distinguish how far their successes owe to the acquired companies’ nature and their nurture by Google, as opposed to “nepotism” where Google diverts traffic to favor companies affiliated into the Google “family.”

A.1.1. Reputation and commitment

In the model, G sets \( r^p \) 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, the reliabilities \( r^p \) and \( r^o \), that it wishes to adopt. This commitment assumption captures in a static model the idea that G can gradually build up a reputation for reliable search results. A key constraint on reputation-building derives from imperfect observation of 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. We now use this reputation and social learning interpretation to justify the consumer joint participation assumption and the role of organic results.

A.1.2. 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. Then, even if consumers decide participation in content and product searches independently, if G changes the reliability of one type of result (say, organic) this would affect its overall reputation, and hence affect consumer participation on G for both types of search. Modeling consumer participation as a single decision captures such spillovers in a simple and tractable fashion.

A.1.3. 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 model, it is very intuitive that merchants are willing to pay to appear in the search engine’s results for product queries and willing to pay more if G withholds useful organic links: G avoids cannibalization by removing organic links on product searches. Clearly, nothing changes if G offers organic results so long as they are dominated by sponsored results implying that consumers would never click on them.36

In the case of content queries, in principle nothing in our model would prevent publishers from paying for sponsored links. In the real world, they seem not to do so. Multiple factors may explain this behavior. A full explanation is beyond the scope of this paper, but we mention

36. In fact, in our model, G suffers no loss if the link to the sponsored ad is also displayed as an organic result. Consumers would then be indifferent and would presumably click on the sponsored result with some probability \( \kappa \). Merchants would not pay directly when their organic results are clicked, but their willingness to pay for a sponsored click would be multiplied by \( 1/\kappa \). All our results would remain unchanged. G would still need to suppress organic links to other merchants (some merchants have accused G of this but we have not seen documentary evidence). 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; as the latter studies eBay, it may be explained by the above discussion of reputation.
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 sites. This dissuades publishers from bidding. Another possible reason is that the informational nature of publisher content reduces observability and makes reputation-building particularly difficult. Some publishers may attempt to build a reputation for 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 products, 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.

A1.4. Pricing sponsored ads

We assumed $G$ makes take-it-or-leave-it offers of the PPC to the relevant merchants in any particular product query, selecting a merchant that supplies the best match with probability $\rho$, and a second-best merchant with probability $1 - \rho$. In the working paper (Burguet et al., 2014), we obtain the same results with a second-price auction replacing the take-it-or-leave it offers, provided that multiple merchants produce each variety. The auction must be weighted, as follows: $G$ invites bids by merchants in the relevant category, partially discounting the bids of second-best variety producers by a factor $\mu$.

In equilibrium, merchants of the same type submit the same bid, equal to their willingness to pay. Moreover, $G$ chooses $\mu$ so that the two types tie in equilibrium. The resulting prices and transactions are identical to those of the benchmark model.

This weighted position auction for determining PPCs seeks to capture, in a simplified framework, the mechanism that Google claims to use in reality.14 The outcome of the auction depends not only on merchants’ bids but also on their quality scores, which capture the relevance of merchant products to consumer demands, as inferred from queries, as well as factors such as website quality. In so doing, Google recognizes the possibility of conflicting interests between consumers and merchants that we have described. The choice of $\mu$ and $\rho$ reflects $G$’s compromise between linking consumers to best-match products and promoting the products of merchants with highest willingness to pay. Google’s claims would suggest that $\rho = 1$, with $\mu$ purely serving to prevent merchants with less preferred goods from winning, but our theory suggest that this may not be the case.

A1.5. Conflict of interest

As already noted, we only need to assume that consumer and merchant interests sometimes conflict, but for a simple example, we derive an extreme case of full conflict. Suppose 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 product differ in price. Price differences can arise endogenously, yet orthogonal to $G$’s online product search choices, if merchants also sell to distinct local markets of fully offline consumers with differing elasticities and if regulatory or reputational concerns prevent merchants from price discriminating between online and offline consumers; merchants facing inelastic offline markets set higher prices so they earn higher margins $m_h$ while consumers gain lower net utility ($v_H$). In this example, $m_H - m_L = v_H - v_L$.

A2. Proofs: preliminaries

Under integration, and for a given $\gamma > 0$, the first-order conditions of $G$’s optimization problem (maximizing Eq. (9)) with respect to $r_H, r_{NG}$, and $r_O$ are given by:

$$H(c) \frac{dc}{d\gamma} \left[ (1-\gamma)M(r_H^\gamma) + \gamma v_L m_L - \beta M(r_O^\gamma) - m_L \right] = 0. \quad (11)$$

$$H(c) \frac{dc}{d\gamma} \left[ (1-\gamma)M(r_H^\gamma) + \gamma v_L m_L - \gamma \beta m_L + (1-\beta)M(r_O^\gamma) \right] = 0. \quad (12)$$

$$H(c) \frac{dc}{d\gamma} \left[ (1-\gamma)M(r_H^\gamma) + \gamma v_L m_L - (1-\gamma)(m_H - m_L) \right] = 0. \quad (13)$$

Evaluated at the same values of $r_H, r_{NG}, r_O$, the left hand side, LHS, of Eq. (11) exceeds that of Eq. (12). If inequality (10) is reversed, then $r_H^\gamma$ is lower than $r_O^\gamma$ (and only one can be interior). This implies $r_{NG}^\gamma > r_O^\gamma$ on region $SD$ and $r_{NG}^\gamma \leq r_O^\gamma$ on $OD$ with at most one being interior in each case.

Also note that under separation, the first-order condition for $r_O$ is

$$H(c) \frac{dc}{d\gamma} \left[ (1-\gamma)M(r_O^\gamma)^\gamma \right] - \alpha(1-\beta)M(r_O^\gamma)^\gamma = 0. \quad (14)$$

Evaluated at $\gamma = 0$, and for the same values of $r_H$ and $r_{NG} = r_O$, the LHS of Eq. (12) is lower than that of Eq. (14). So, when interior, $r_O$ as a function of $\gamma$ is discontinuous at $\gamma = 0$. Finally, under separation the first-order condition that determines $r_O$ is Eq. (13) evaluated at $\gamma = 0$; that is, $r_O$ as a function of $\gamma$ is continuous at $\gamma = 0$.

A necessary and sufficient condition for only one instrument to be distorted and optimally be strictly positive under separation is that Eq. (III) holds and either Eq. (I) or Eq. (II) holds.

$$\frac{\mu u_m m_H}{m_H - m_L} > \frac{1}{H(u + v_L + \omega(v_H - v_L))(1-\alpha)^{\gamma}}. \quad (I)$$

$$\frac{v_H - v_L}{m_H - m_L} > \frac{\mu}{\alpha(1-\beta)} \frac{1}{H(v_H)(1-\alpha)^{\gamma}}. \quad (II)$$

$$\min \left\{ \frac{v_H - v_L}{m_H - m_L} \frac{\mu}{\alpha(1-\beta)} \frac{1}{H(u + v_L)(1-\alpha)^{\gamma}} \right\} > 0. \quad (III)$$

Where Eq. (1) ensures that $r^\gamma \neq 0$ where $r^\gamma$ is lower and Eq. (II) similarly ensures $r_O$ cannot be zero where $r_O$ is lower, and Eq. (III) ensures $r^\gamma = r^\gamma_1 = 1$ is not a solution. Necessity follows from the fact that $\frac{\mu v H}{u + v L} < 0$. It is readily verified that the set of parameters that satisfy these conditions is an open set.

A3. Proof of Proposition 2

Consider the region $SD$, so that $r_O^\gamma \geq r^\gamma$ under separation, and $r_{NG}^\gamma \geq r^\gamma$ under integration. Since only one instrument is partially distorted under separation, then $r_O^\gamma = 1 > r^\gamma > 0$. Note that, for $\gamma = 0$, Eq. (13) is also the first-order condition for $r^\gamma$ under separation. Now, if $r_{NG}^\gamma = r_O^\gamma = 1$ then $r_O^\gamma$ is larger under integration, as the LHS of Eq. (13) is increasing in $\gamma$. Also,

---

14 Notice that $G$ never gains by simply suppressing all publisher links, because we implicitly assume that getting no content link at all gives consumers a sufficiently low payoff.

15 See Hal Varian’s Youtube video on Google’s scoring auctions, or Varian (2007).
as \( \frac{d_1^B}{\sigma_{NC}} < 0 \), \( r_{NC}^2 \geq r^2 \), and they cannot be both interior, we conclude that indeed \( r_{NC}^2 = 1 \), and hence, \( r^2 \) is larger under integration.

Since search reliability improves with \( \gamma \), consumer surplus also increases. Producer surplus can be written as:

\[
PS = F(\bar{c}) \left[ (1 - \eta)M^*(r^2) + \eta m_1 \right].
\]

Since \( r^2 = 1 \) both before and after integration, \( \gamma \) only affects PS through \( r^2 \). Also, the derivative of PS with respect to \( r^2 \) is larger than the derivative of \( IF^2 \) with respect to \( r^2 \) if the effect on returns per participating consumer is the same in both cases, but PS per consumer is larger than \( G \)'s profits per consumer. Thus, at \( G \)'s profit-maximizing \( r^2 \), PS is increasing in \( r^2 \), and so an increase in \( r^2 \) caused by a higher level of integration raises PS and hence total welfare too.

Consider now the region OD. Let us first deal with the case \( \frac{m_H - m_L}{m_H - m_L} > \frac{1}{\tau_C} \). Under separation \( 0 < r^2 < r^2 \) is 1 and, similarly, \( r_{NC}^2 \leq r^2 \) under integration with at most one interior solution. Thus, Eq. (14) holds, and so, at that same value of \( r^2 \) and for \( r_{NC}^2 = 1 \), Eq. (12) is strictly negative. Also, since the LHS of Eq. (13) increases with \( \gamma \), and since \( \frac{d_1^B}{\sigma_{NC}} < 0 \), we conclude that under integration \( r^2 > 1 \) and \( r_{NC}^2 \leq r^2 \). Hence, since \( r^2 \rightarrow r_{NC}^2 \) as \( \gamma \rightarrow 0 \), \( r_{NC}^2 \) falls discretely after integration. Therefore, consumer welfare also falls. Finally, both under separation and integration, PS evaluated at \( r^2 = 1 \) is \( PS = F(\bar{c})m_1 \). Hence, since \( \bar{c} \) is lower under integration, so is producer surplus (and therefore total welfare).

Suppose now that \( \frac{m_H - m_L}{m_H - m_L} > \frac{1}{\tau_C} \). Thus, under separation, \( 1 = r^2 > r^2 > 0 \), and under integration \( r_{NC}^2 \leq r^2 \), and only one can be interior. As in the previous case, \( r^2 \) increases with integration. Indeed, suppose that the optimal solution under integration is \( r_{NC}^2 = 1 \). In this case, since Eq. (13) evaluated at \( \gamma = 0 \) holds, \( \frac{m_H - m_L}{m_H - m_L} > \frac{1}{\tau_C} \) then the LHS of Eq. (12) would be strictly negative, and hence we reach a contradiction.

The welfare assessment of the latter case is more elaborate. First, suppose that after integration \( r^2 = 0 \) and \( r^2 > 0 \) is higher than under separation. Hence, \( \eta = \sigma \) under separation and \( \eta = \sigma \beta \) under integration. Under separation, Eq. (13) can then be written as:

\[
H(\bar{c}_{sep})(v_H - v_L)(1 - \sigma)M^*(r_{sep}^2) - (m_H - m_L) = 0.
\]

(15)

and under integration (evaluated at \( \gamma = 0 \)) the same first-order condition (for \( r^2 = 1 \)) is:

\[
H(\bar{c}_{int})(v_H - v_L)(1 - \sigma \beta)M^*(r_{int}^2) - (m_H - m_L) \geq 0.
\]

(16)

Suppose that consumer welfare is higher under integration, that is, \( \bar{c}_{int} > \bar{c}_{sep} \). This means that \( H(\bar{c}_{int}) < H(\bar{c}_{sep}) \). Substituting for \( \bar{c} \), \( \bar{c}_{int} > \bar{c}_{sep} \) also implies

\[
(1 - \sigma \beta)(1 - r_{int}^2)(1 - \sigma)(1 - r_{sep}^2) < \frac{\eta}{v_H - v_L}.
\]

and taking into account that \( \frac{d_1^B}{\sigma_{NC}} > \frac{m_H - m_L}{m_H - m_L} m_1 \), we should then have

\[
(1 - \sigma \beta)M^*(r_{int}^2) < (1 - \sigma)M^*(r_{sep}^2).
\]

This is a contradiction, since the above inequality and \( H(\bar{c}_{int}) < H(\bar{c}_{sep}) \) then implies that the LHS of Eq. (15) is larger than the LHS of Eq. (16). Thus, \( r_{NC}^2 \) can increase or decrease with \( \gamma \). In particular, if \( r_{NC}^2 \) is sufficiently close to 1, then the first term is small and \( r_{NC}^2 \) increases with \( \gamma \). However, if \( r_{NC}^2 \) is sufficiently small, then the second term is small and \( r_{NC}^2 \) decreases with \( \gamma \). In both cases, consumer and producer surplus increase with \( \gamma \). Indeed, on the one hand, if \( r_{NC}^2 \) increases with \( \gamma \) then \( r^2 \) also increases with \( \gamma \), and since \( r^2 \) and \( r_{NC}^2 \) are fixed and equal to 1, then \( \bar{c} \) and consumer welfare increase with \( \gamma \). Also, since \( PS_{int} = F(\bar{c})m_1 \), PS also increases with \( \gamma \). On the other hand, if \( r_{NC}^2 \) decreases with \( \gamma \), then \( 1 - (1 - \gamma)\sigma r_{NC}^2 \) also increases with \( \gamma \). Since Eq. (20) must
hold, $H(u^2 + v_1)$ must fall; that is, $r^2$ and $c$ both increase with $\gamma$, and so consumer and producer surplus increase with $\gamma$.

Now consider the case $r^2_{NC} = 1$ for a given $\gamma$. As we have just discussed, for $r^2_{NC} = 1$ the slope of $IT^{\gamma}$ with respect to $r^2_{NC}$ increases with $\gamma$. Therefore $r^2_{NC} = r^2 = 1$ is still a solution for larger $\gamma$.

Consider the case $r^2_{NC} = 0$ for given $\gamma$ but $r^2 > 0$. Now, $r^2$ is given by Eq. (13) with $r^2 = 1$ and $\theta = 0$. That is:

$$H(c)(v_H - v_L)\left[1 - \eta M(r^2) + \eta m_L\right] - (m_H - m_L) = 0. \quad (21)$$

where $\eta = \alpha (\gamma + (1 - \gamma)\beta)$. Using the implicit function theorem, we have $\frac{dc}{d\gamma} < 0$. Hence, an increase in $\gamma$ raises $r^2$ but lowers $\gamma$. We show that the net effect on $c^*$ is positive. Indeed, suppose that

$$\frac{dc}{d\gamma} = u + (v_H - v_L)(1 - r^2)\alpha(1 - \beta) + (v_H - v_L)(1 - \eta) \frac{dr^2}{d\gamma}, \quad (22)$$

is negative, and so $H(c)$ increases with the larger $\gamma$. Since Eq. (21) still holds, this means that $(1 - \eta)M(r^2) + \eta m_L$ decreases, that is,

$$-(m_H - m_L)\left[1 - r^2\right] \alpha(1 - \beta) + (1 - \eta) \frac{dr^2}{d\gamma} < 0,$$

so that $\frac{dc}{d\gamma} = -\left(1 - r^2\right)\alpha(1 - \beta)$. Substituting in Eq. (22), we obtain a contradiction that shows that $\frac{dc}{d\gamma} = 0$.

Note that if $r^2_{NC} = 0$, profits of non-affiliated publishers are also zero, and hence G's profits coincide with $P$. Thus, as $\gamma$ increases then both $c^*$ and $F(c^*)$ increase but $H(c)$ falls. Hence, from Eq. (21) $\frac{dc}{d\gamma} = 0$ and $\eta m^2_{NC}$ increases. Consequently, $P$ increases with $\gamma$.

Finally, suppose that for some value of $\gamma$, $r^2 = 0$ (and so $r^2_{NC} = 0$). Then, the case $G's$ profit (and $P$) only depend on $r^2 = \gamma r^2_{NC}$. If Eq. (11) evaluated at $r^2 = 1$ is strictly positive, then optimally $r^2 = 1$ after a (small) increase in $\gamma$. That raises $r^2$ as well as consumer and producer surplus. If Eq. (11) is satisfied with equality for $r^2 = 1$ and some value of $\gamma$, then for any $\gamma > \tilde{\gamma}$, the optimal value of $r^2_{NC} = \frac{\gamma}{\tilde{\gamma}}$ and an increase in $\gamma$ will not affect consumer or producer surplus.

References