Integration and Search Engine Bias *

Alexandre de Cornière[†] and Greg Taylor[‡]

August 5, 2013

Abstract

Competition authorities all over the world worry that integration between search engines (mainly Google) and publishers could lead to abuses of dominant position. In particular, one concern is that of own-content bias, meaning that Google would bias its rankings in favor of the publishers it owns or has an interest in, to the detriment of competitors and users. In order to investigate this issue, we develop a theoretical framework in which the search engine (i) allocates users across publishers, and (ii) competes with publishers to attract advertisers. We show that the search engine is biased against publishers that display many ads—even without integration. Although integration may lead to own-content bias, it can also reduce bias by increasing the value of a marginal consumer to the search engine. Integration also has a positive effect on users by reducing the nuisance costs due to excessive advertising. Its net effect is therefore ambiguous in general, and we provide sufficient conditions for it to be desirable or not.

Keywords: search engine, integration, advertising, bias.

JEL Classification: L1, L4, L86.

1   Introduction

The importance of Internet search engines is indisputable. They process billions of queries each day worldwide and, to a large extent, create order in the web out of a clutter of information. The business model of most search engines relies on two pillars, namely organic results (generally displayed to the left) and paid results (or sponsored links, to the right and on top). Both kinds of results are relevant to queries entered by users, the main difference being that the search engine does not receive payments from publishers who appear on organic listings, whereas sponsored links are bought by advertisers. Whilst organic results receive the majority of clicks,[1] paid results have proved to be very profitable, generating billions of dollars of annual advertising revenue.

The commercial nature of search engines, coupled with their essential role as information gate-keepers, has unsurprisingly attracted scrutiny from competition authorities all over the world. Google has been the main target of investigations, in part due to its dominant position in the search market (market share of over 90% in Europe, 65% in the US).[2] One of the most salient concerns expressed by competition authorities is the idea that Google, which owns many specialized publishers (for example, Google maps, Youtube, and Zagat), displays organic results in such a way as to favor these publishers, to the detriment of its competitors and, more importantly, to that of users. For instance, in 2013 the US Federal Trade Commission concluded a lengthy investigation of Google's business practices in which the regulator acknowledged that Google has introduced changes to its search algorithm that privileges its own content—ammounting to what is "commonly known as search bias" (FTC, 2013). Despite this observation, the FTC determined that the changes in question are likely to improve the experience of users and therefore do not warrant legal intervention. The debate, though, is far from settled. The European Commission—which explicitly listed own-content bias as one of its four concerns related to potential abuses of dominant market position by Google—has been more hawkish and is currently negotiating a settlement with the search engine that is likely to include measures aimed at affording more prominence to Google's rivals.

The existence of such a "bias" is difficult to establish empirically. In an attempt to measure it, Edelman and Lockwood (2011) compare the results of several search engines for a list of queries, and argue that Google tends to favor its own-content more than other major search engines. On

---

[1] 94% of clicks versus 6% for sponsored links, for Google and Bing, according to a study by eConsultancy.
[2] Manne and Wright (2007) provide a legal analysis of the antitrust issues facing Google.

the other hand, Wright (2011), after extending Edelman and Lockwood's initial study to other keywords, argues that the evidence "do[es] not support claims that own-content bias is of the nature, quality, or magnitude to generate plausible antitrust concerns".[3]

While the current debate mostly focuses on whether Google actually provides an unfair advantage to its own services, this paper tackles the broader question of the effects of integration between a dominant search engine and a publisher, taking into account, on top of the issue of own-content bias, the potential implications that such a move would have on the advertising market. To do so, we build a model that incorporates some key features of the search industry. In the model, consumers may use a monopolistic search engine in order to find some content that is relevant to their query. Upon receiving the query, the search engine, through its organic results, can direct users towards one of two publishers, which are initially unknown to consumers. We refer to the mapping from a query to a result as an allocation rule. Both publishers as well as the search engine are strictly ad-financed, and advertisers view space on these three media platforms as imperfect substitutes. Consumers dislike ads but cannot observe how many of them are displayed on each publisher prior to visiting it.[4]

We consider three cases. In the unintegrated case, the publishers are independent with respect to the search engine. Given that publishers do not receive payments from users, they display too many ads from a consumer surplus point of view. The provision of advertising is nevertheless insufficient from advertisers' point of view, because of publishers' market power. Explicitly modeling the three sides of the market (consumers, publishers and advertisers) allows us to shed light on the forces governing the optimal allocation of traffic, both from consumers' perspective and from the search engine's. While consumers would value an allocation rule that takes into account relevance and the nuisance generated by excessive levels of advertising, the search engine must also factor in the effect of its allocation rule on the price at which it sells its sponsored links. This implies in particular that the search engine will penalize publishers that display many ads more than consumers would wish.

We next turn to the study of the case in which the search engine and one publisher enter in a contractual arrangement by which the search engine serves the advertisements on the publisher's site and keeps a share of the revenues, while the publisher remains free of choosing how many

---

[3]See also Tarantino (2011) for a discussion of the issue of foreclosure in the vertical search engine industry, as well as for some evidence.

[4]We also consider the case of ad-loving consumers.

advertising slots to sell. We refer to this arrangement as partial integration (without control). Partial integration creates a monetary incentive for the search engine to direct more users towards that specific publisher, but we show that there exist conditions under which it leads to a reduction in bias, in particular when user participation is very sensitive to the quality of the results. The main reason for this is that partial integration increases the search engine's marginal return to attracting users on its platform.

Finally, we consider a market structure in which the search engine is fully integrated with a publisher, and therefore also chooses how many ads will be displayed on the publisher's site. The fact that advertisers view ads on the search engine and on websites as imperfect substitutes implies that such a merger leads the integrated structure to lower its supply of ad space, to the detriment of advertisers and to the benefit of users, who are more satisfied by visiting the search engine's own site compared to the unintegrated case.

As a result, we see that the effects of integration on consumers' welfare are ambiguous in general. By focusing on some key parameters, we derive some policy implications of the model.

Related literature

Much of the interest in search engines has been focused on the auctions used to sell advertisements (see, for example, Athey and Ellison, 2011; Chen and He, 2011; Edelman, Ostrovsky, and Schwarz, 2007; Varian, 2007). A more recent literature considers search engines' platform design decisions, and addresses the question of whether search engines will always wish to maximize result quality. For examples of such papers see de Cornière (2011), Eliaz and Spiegler (2011), Taylor (2013b), and White (forthcoming). Unlike these papers, where result quality is a purely vertical notion, we introduce 'real' bias by adding a horizontal component to search engines' algorithm design decision: in our paper, when a search engine privileges some publisher, it necessarily does so at the expense of another.

Burguet, Caminal, and Ellman (2013) also look at how ad market incentives affect the quality of organic and sponsored links on a search engine. In contrast to their work, we abstract from the design of the auction mechanism but endogenize publishers' ad supply decisions. Doing so allows us to study how the relative attractiveness of publishers to consumers is endogenous to the market structure, which has a variety of interesting theoretical and practical implications.

We operationalize bias as a decision by the search engine to direct some consumers away from

their ideal publisher. In this sense our work is similar in spirit to that of Hagiu and Jullien (2011), Armstrong and Zhou (2011) and Inderst and Ottaviani (2012), who consider the incentives of intermediaries to bias their advice in favor of firms from which they receive larger payments. The structure of the model is also reminiscent of some papers on network neutrality (Choi and Kim (2010) for instance), in which "bias" takes the form of priority rules in favor of one publisher. Our paper mainly differs from the preceding ones in two aspects: firstly, search engines in the baseline version of our model do not receive payments from the publishers to which they link and therefore do not have a direct financial incentive to bias their results. Secondly, search engines compete with those publishers to sell advertisements. The latter point introduces consumer multi-homing in the analysis, and thus complements a nascent literature on this dimension of media markets (Athey, Calvano, and Gans (2010), Anderson, Foros, and Kind (2011), Ambrus, Calvano, and Reisinger (2012)). Indeed, since in our paper users visit the search engine and a publisher, they can be reached by advertisers on both "platforms", which, as in the above papers, reduces the ability of each platform to extract revenue from advertisers, all the more so that consumers are exposed to many ads on the other outlet they visit. The novelty in our approach is that the search engine is able to influence which publisher consumers will visit, so that it can alleviate the multi-homing externality by directing users towards publishers with low level of advertising.[5] In related spirit, Taylor (2013a) studies how a media platform can affect the extent of multi-homing through the quality of its content. Ellman and Germano (2009) consider the incentives of a newspaper to bias its coverage against content detrimental to advertisers.

White and Jain (2010) also study a model in which both the search engine and publishers make money through advertising. However, they do not model the advertising side of the market, nor do they model the way the search engine allocates traffic among publishers, and instead focus on the distraction imposed on users.

2   Description of the model

Consider a market composed of one search engine $(i = 0)$ and two publishers $(i = 1, 2)$, a continuum of mass one of consumers, and a representative advertiser.

---

[5]See footnote 19 for a brief discussion of some differences in our modeling approaches.

2.1 The search for content

Publishers 1 and 2 produce some horizontally differentiated content. We model this differentiation by assuming that they are located at the ends of a Hotelling segment of unit length: publisher 1 to the left, publisher 2 to the right.

Consumers are uniformly distributed along the segment, and the utility they derive from the consumption of a publisher's content is a decreasing function of the distance between their position, $x \in [0;1]$, and that of the publisher. If the distance is d, the utility from reading the content is $v(d)$, v being a decreasing function. However, the market is characterized by incomplete information: consumers know what their tastes are, but they are ex ante ignorant of the mapping between their tastes and their position on the line, and are therefore unable to identify which publisher is the more relevant.

The only way for consumers to access content is to use the search engine. When a consumer decides to do so, he enters a qualitative query that reveals his taste. The search engine is uniquely endowed with knowledge of the function that maps a user's query into a latent location on the Hotelling segment, and it directs him towards one of the two publishers accordingly. We refer to the search engine's decision on which publisher to promote as an allocation rule, and we focus on allocation rules that have a threshold structure, as made explicit below:

Definition 1. *A search engine's allocation rule is a threshold $\bar{x}$, such that all users with $x < \bar{x}$ are directed toward publisher 1 and all those with $x \geq \bar{x}$ are directed toward publisher 2.*

We allow the search engine to choose any $\bar{x}$ in $[0;1]$, so that it has perfect control over the allocation of traffic between the publishers. This simplifying assumption is broadly consistent with Edelman and Lai (2013), who exploit a natural experiment involving variations in the layout of Google search results pages to show that relatively minor changes in design can have a large effect on the traffic flowing from a search engine.[6]

Consumers incur a search cost when they use the search engine. This search cost θ is distributed across the population according to a strictly increasing (on its support) c.d.f. F, and is independent of consumers' position. Write f for the associated density.

---

[6]One could easily amend the model so as to restrict the search engine to choose $\bar{x} \in [x_L, x_H]$, with $0 < x_L < x_H < 1$, but such a specification does not deliver additional insights.

2.2 Advertising

Both publishers and the search engine are financed exclusively through advertising. The quantity of advertisements displayed on website i is denoted q_i. We will refer to ads on Publishers 1 and 2 as *banners*, whereas ads on the search engine are *sponsored links*.[7] We assume throughout the paper that the quantity of sponsored links is exogenously fixed at q_0. This allows us to focus on the allocation rule as the main strategic tool of the search engine, and therefore clarifies the exposition of our original insights. The case in which q_0 is endogenously chosen by the search engine, for which the main results still hold, is relegated to the appendix.

Although there are no monetary transfers between consumers and publishers, we assume that consumers dislike ads, so that they "pay" with their attention. If a user is directed towards publisher i, he suffers a disutility $\delta_i(q_i)$.[8],[9] The utility of a user who is directed toward publisher i, which is at a distance d from his ideal location, is therefore

$$u(d, q_i) = v(d) - \delta_i(q_i) - \theta. \tag{1}$$

On the other side of the market are advertisers who wish to reach users. For simplicity, we focus on the case of a representative advertiser who wishes to send several messages to each consumer. It can reach consumers either when they are on the search engine page, or when they visit a publisher. Given an allocation rule $\bar{x}$, the expected per-user revenue of the advertiser, if it buys q_0 ad slots on the search engine and q_1 and q_2 slots on publishers' websites, is

$$R(Q, \bar{x}) = r_0(q_0) + \bar{x} r_1(q_0, q_1) + (1 - \bar{x}) r_2(q_0, q_2), \tag{2}$$

with $r_0(0) = r_1(q_0, 0) = r_2(q_0, 0) = 0$. The r_0 function represents the revenue from contacting users on the search engine, whilst r_1 and r_2 capture the additional revenue for advertisers generated by a user who is directed towards publisher 1 and 2 respectively. In Appendix A we provide microfoundations for this revenue function, as well as an alternative interpretation with multiple heterogenous advertisers buying at most one slot.

As suggested by our notation, for $i \in \{1, 2\}$, r_i depends on both q_0 and q_i. However, since a

---

[7]We use the term "banner" for expositional simplicity. Some publishers are actually vertical search engines with sponsored links. We discuss this point in section 6.

[8]See section 6 for a discussion of the case of ad-loving consumers.

[9]Since q_0 is exogenous, it does not matter whether we write $\delta_i(q_0, q_i)$ or $\delta_i(q_i)$.

consumer cannot visit both publisher 1 and publisher 2, r_1 does not depend on q_2, and vice-versa. Notice also that r_1 and r_2 do not depend on $\bar{x}$. The underlying assumption is that the location of a user in the content space is independent of its attitude towards advertisers.[10]

The key assumption of our model is that advertisers view ad space on the search engine and on publishers as imperfect substitutes: sponsored links and banners are two ways to reach consumers, even if they are not equivalent. Substitutability means that the marginal value of a type of ads (sponsored link or banner) decreases as the number of ads of the other type increases. Formally, we have

$$\frac{\partial^2 r_i(q_0, q_i)}{\partial q_0 \partial q_i} \leq 0 \tag{3}$$

for $i = 1, 2$. There is some disagreement about the strength of substitutability: Hahn and Singer (2008) provide evidence for strict substitutability, whilst the FTC (2007) took the position that the two ad markets are independent. Our model admits both possibilities and allows us to comment on the role of the strength substitutability in determining the outcome of integration. In particular, the limiting case with independent demands corresponds to the conditions in (3) holding with equality. The degree of substitutability is likely to differ substantially across content sectors and website types. For instance, a vertical search site (such as Expedia) seems like it might be particularly well-placed to substitute for a search engine in providing ads that are carefully targeted at a consumer's search query. Indeed, on a more general level, improvements in publishers' ability to target advertisements are likely to make the two forms of ads closer substitutes because display ads that are well-targeted can be used to stimulate a direct consumer response—a function that has typically been associated with search ads.

Note that $r_i(q_0, 0) = 0$ implies

$$r_i(q_0, q_i) = \int_0^{q_i} \frac{\partial r_i(q_0, q_i)}{\partial q_i} dq_i$$

so that, using (3),

$$\frac{\partial r_i(q_0, q_i)}{\partial q_0} = \int_0^{q_i} \frac{\partial^2 r_i(q_0, q_i)}{\partial q_i \partial q_0} dq_i \leq 0. \tag{4}$$

In words: holding constant the number of ads on publisher i, an increase in the supply of search ads causes the advertiser revenues flowing from ads on i to fall.

---

[10]This assumption is discussed in Appendix A.

We also assume that r_i is concave in q_i for all i and, to focus on interior solutions, that $\frac{\partial r_i}{\partial q_i}\big|_{q_i=0} > 0$. Concavity may result from the marginal ads being inherently less valuable, or from a decrease in the time spent by the user on the publisher due to the extra nuisance implied by the marginal ad, making it less likely that he notices a product he is interested in.

Finally, we assume that payments to publishers are made on a per-impression basis so that the expected per-user profit of advertisers is

$$\pi_a = R(Q, \bar{x}) - q_0 P_0 - \bar{x} q_1 P_1 - (1 - \bar{x}) q_2 P_2, \tag{5}$$

where P_i is the price of an advertisement on i.[11] Ad prices are determined so as to equalize supply (the q_i's offered by publishers) and demand by advertisers.

Figure 1 represents the market structure.

A remark on sponsored links In this paper we consider the case in which publishers can only attract traffic through organic links. It turns out that allowing publishers to buy sponsored links by participating in the auction would not significantly affect our results. To see this formally, consider the following variant of our model. Assume that publishers can participate in an auction for the sponsored links. Consumers, when facing different (sponsored or organic) links, are able to distinguish links that direct towards a publisher from links that direct towards a merchant (the advertisers in our model). If a consumer faces two links to publishers, one organic and one sponsored, let α be the (endogenous) probability with which he clicks on the organic link.

Suppose that $\alpha < 1$. In the auction, the identity of the winner depends on publishers' expected per-impression revenues $\alpha q_i P_i$ (since this is how much they are willing to pay per-impression). Notice that this is independent of the position of the user on the Hotelling line, so that the expected relevance of a sponsored link publisher is that of a random draw. However, if $\bar{x} \in (0; 1)$, the expected relevance of the organic link is strictly higher than that, and so the user optimally sets $\alpha = 1$. In other words, publishers' sponsored links are never clicked on in equilibrium.[12]

---

[11]Per-click pricing for sponsored links would not alter our qualitative results, as long as there is substitutability between sponsored links and banners.

[12]Notice that this logic does not imply that all sponsored links are never clicked on, since it does not cover links to merchant websites.
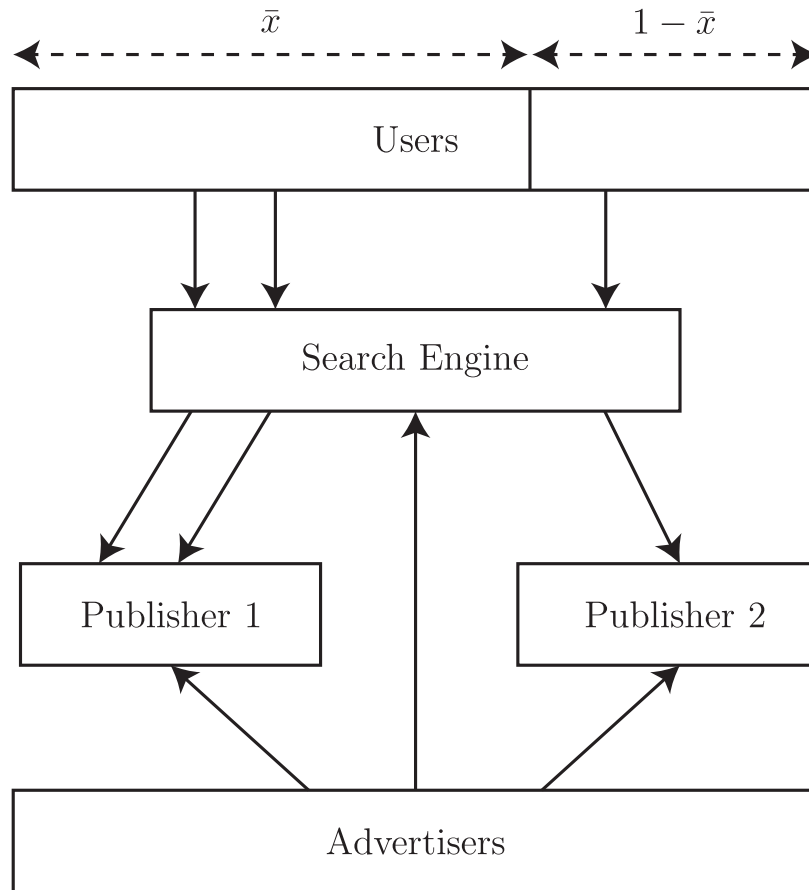
Figure 1: Market structure

2.3   Information and timing

The timing of the game is the following:

1. Publishers 1 and 2 simultaneously choose their advertising level q_1 and q_2. The search engine chooses an allocation rule $\bar{x}$.[13]

2. Advertisers observe $\bar{x}$, q_0, q_1 and q_2. The advertising market clears: P_0, P_1 and P_2 equalize supply and demand for each publisher.

3. Consumers observe $\bar{x}$. They decide whether to use the search engine or not by comparing the surplus they expect to gain from doing so to their idiosyncratic search cost.

4. Each user enters a query, and visits the publisher suggested by the search engine.

The assumption that users do not observe q_1 and q_2 is consistent with the fact that they have no prior knowledge of the content offered by the two publishers. The assumption that $\bar{x}$ is observed by consumers may seem strange if one takes the model at face value. Indeed, consumers certainly cannot observe the algorithm used by search engines. Our interpretation of this assumption of the model is, in a broad sense, that a lower quality of results will eventually drive some users away from the search engine, either because they hear about it in the media or they experience it themselves. For completeness, we analyze the case in which $\bar{x}$ is not observed by consumers in the appendix.

We look for Nash equilibria in which (i) consumers' participation strategies are optimal for every $\bar{x}$ given (correct) expectations about q_1 and q_2, (ii) taking the consumers' participation strategy as given, the choices of q_1, q_2, and $\bar{x}$ are mutual best responses, and (iii) ad prices clear the market. We first solve backwards steps 2 to 4. Then we study three cases in order to solve step 1: in the first case, the publishers are independent from the search engine; in the second case the search engine partially integrates with a publisher; in the third case the search engine takes full ownership control of a publisher. The difference between partial and full integration is that

---

[13]One could find arguments in favor of a different timing, either to have $\bar{x}$ chosen before the q_i's, or after. For example, if the search engine were to program its algorithm to condition its results upon the ad levels on sites 1 and 2 then the game would be formally equivalent to one in which the search engine moves after the publishers. Conversely publishers often tailor their design decisions to game the search algorithm and optimize their search ranking. We are agnostic on this point; in effect, search engines and publishers appear to be stuck in a kind of Red Queen's race in which neither side is able to definitively act as a leader. Simultaneous actions also turn out to be more tractable.

in the former case the search engine obtains a share of the publisher's revenue but cannot control its choice of q_i, whilst in the latter we allow the search engine to exercise full control over the integrated publisher's policy.

3   Consumer participation and ad market clearing.

3.1   Consumer participation

Suppose that the search engine has chosen the allocation rule $\bar{x}$, and that consumers expect publishers 1 and 2 to choose ad levels q_1^e and q_2^e respectively. Then the expected disutility (or *nuisance*) from advertising is

$$\Delta(\bar{x}, Q^e) \equiv \Delta^e \equiv \bar{x}\delta_1(q_1^e) + (1 - \bar{x})\delta_2(q_2^e).$$

For a given allocation rule $\bar{x}$, the expected value from reading the content is

$$\rho(\bar{x}) \equiv E[v(d)|\bar{x}] = \int_0^{\bar{x}} v(z)dz + \int_{\bar{x}}^1 v(1 - z)dz.$$

We will refer to ρ as the *relevance* of the search engine. Its second derivative is $\rho''(\bar{x}) = v'(\bar{x}) + v'(1 - \bar{x}) < 0$, hence ρ is a concave function of $\bar{x}$, maximized at $\bar{x} = 1/2$. For given values of the nuisance Δ^e and relevance ρ, the expected utility of a consumer is therefore

$$U(\rho, \Delta^e, \theta) = \rho - \Delta^e - \theta.$$

A consumer will engage in search if and only if $U(\rho, \Delta^e, \theta) \geq 0$, which gives us the participation rate $F(\rho - \Delta^e)$

3.2   Definition of bias

The expression for user participation makes it clear that consumers care about two things when they consider using the search engine: they want the content to be relevant, but also for it to be accompanied by as few advertisements as possible. This leads us to the following important observation: *consumers may wish to sacrifice relevance in order to reduce the nuisance from advertising.*

So we are naturally led to ask: what is the optimal allocation rule from consumers' point of view?

Definition 2. *For a given expected level of ads, Q^e, the user-optimal allocation rule is defined as*

$$x^*(Q^e) \equiv \underset{\bar{x}}{\operatorname{argmax}} \, F(\rho - \Delta^e).$$

Notice that x^* is well defined since $\rho - \Delta^e$ is a strictly concave function of $\bar{x}$ (Δ^e is linear in $\bar{x}$), which implies that it is strictly quasi-concave; the c.d.f. F being increasing implies that $F(\rho - \Delta^e)$ is also strictly quasi-concave.

Because users dislike ads, we have the following result:

Lemma 1. *The user-optimal cut-off, x^*, is decreasing in q_1^e and increasing in q_2^e.*

Proof. The first order condition for determination of x^* is

$$\frac{dF(\rho - \Delta^e)}{d\bar{x}} = f(\rho - \Delta^e) \left[\rho'(\bar{x}) - \frac{\partial \Delta^e}{\partial \bar{x}} \right] = 0,$$

which can be simplified to

$$\rho'(\bar{x}) = \delta_1(q_0, q_1^e) - \delta_2(q_0, q_2^e).$$

The right hand side is increasing (decreasing) in q_1 (in q_2), and hence so must be the left hand side. That $\rho''(\bar{x}) < 0$ completes the proof. ■

Example 1. *A simple example utility function consistent with the above is a Hotelling model of content preference with linear transport cost, t, paired with linear ad nuisance costs:*

$$U(\rho, \Delta) = \underbrace{1 - \int_0^{\bar{x}} tx \, dx - \int_{\bar{x}}^1 t\,(1 - x) \, dx}_{\rho} - \underbrace{[\delta_0 q_0 + \bar{x}\delta_1 q_1^e + (1 - \bar{x})\delta_2 q_2^e]}_{\Delta^e},$$

which yields

$$x^* = \frac{1}{2} + \frac{\delta_2 q_2^e - \delta_1 q_1^e}{2t}$$

Note that this example exhibits the more general property that x^* need not equal $1/2$ because consumers hope to maximize U rather than ρ. With this in mind, we do not view $x \neq 1/2$ as constituting bias per se. Rather, we describe a search engine as being biased in favor of publisher i when it sends to i users who would obtain higher surplus from visiting i's rival. More formally,

Definition 3. *We say that the search engine is biased in favor of publisher 1 when $\bar{x} > x^*$ and in favor of publisher 2 when $\bar{x} < x^*$. The search engine is unbiased if $\bar{x} = x^*$. We refer to $|x^* - \bar{x}|$ as the absolute magnitude of bias.*

3.3 Ad market clearing

The supply of advertising space is given by q_0, q_1 and q_2. On the demand side, as is standard in Cournot models, the representative advertiser acts as a price-taker. Given the observed allocation rule $\bar{x}$ and a price vector (P_0, P_1, P_2), the advertiser's profit is given by (5). The advertiser is indifferent between buying and not buying an extra advertising slot on website i if and only if its marginal revenue $\left(\frac{\partial R(Q,\bar{x})}{\partial q_i}\right)$ is equal to the price of the slot (P_i). The demand for advertising space (q_0^d, q_1^d, q_2^d) is thus implicitly given by the following conditions:

$$P_0 = \frac{\partial r_0(q_0^d)}{\partial q_0} + \bar{x}\frac{\partial r_1(q_0^d, q_1^d)}{\partial q_0} + (1 - \bar{x})\frac{\partial r_2(q_0^d, q_2^d)}{\partial q_0}, \tag{6}$$

$$P_1 = \frac{\partial r_1(q_0^d, q_1^d)}{\partial q_1}, \tag{7}$$

$$P_2 = \frac{\partial r_2(q_0^d, q_2^d)}{\partial q_2}. \tag{8}$$

For the market to clear, prices are set so that $q_i^d = q_i$ for $i \in \{0, 1, 2\}$. This gives us the inverse demand functions $P_0(Q, \bar{x})$, $P_1(Q)$ and $P_2(Q)$.

From (3), substitutability implies that P_1 and P_2 are non-increasing functions of q_0, and P_0 is non-increasing in q_1 and q_2.

Example 1 (Continued). *One of the simplest specifications that satisfies the above properties is the following:*

$$r_0(q_0) = \alpha_0 q_0 - \frac{\beta_0}{2}q_0^2,$$

$$r_i(q_0, q_i) = \alpha_i q_i - \frac{\beta_i}{2}q_i^2 - \gamma_i q_0 q_i,$$

which leads to

$$P_0 = \alpha_0 - \beta_0 q_0 - \bar{x}\gamma_1 q_1 - (1 - \bar{x})\gamma_2 q_2,$$

$$P_1 = \alpha_1 - \beta_1 q_1 - \gamma_1 q_0,$$

$$P_2 = \alpha_2 - \beta_2 q_2 - \gamma_2 q_0.$$

In this specification, the degree of substitutability between ads on the search engine and on publisher i is measured by γ_i. When γ_i is large, an increase in the number of banners (q_i) causes a large decrease in the price of sponsored links (P_0), and reciprocally. The β parameters control the concavity of advertiser revenue. A high β_i may be interpreted as a strong sensitivity of users to ad levels: an increase in q_i makes users spend much less time on publisher i, so that the marginal revenue of ads decreases quickly. When we use this linear specification later in the paper, we assume that $\alpha_i, \beta_i, \gamma_i > 0$ for all i and, to ensure that the optimization problem has a well-defined interior solution, that $\beta_0 \beta_i - \gamma_i^2 > 0$ for $i \in \{1, 2\}$.

4 Equilibrium in the non-integrated case

As a benchmark, let us look at the situation without vertical integration—indicated by an NI superscript (for 'no integration'). This section is of independent interest, since it highlights the key incentives of a search engine when it comes to choosing how to allocate traffic among independent publishers. Before turning to this question, let us look at the advertising decisions.

The profits of the three websites are

$$\pi_0 = F(\rho - \Delta^e)q_0 P_0; \qquad \pi_1 = \bar{x}F(\rho - \Delta^e)q_1 P_1; \qquad \pi_2 = (1 - \bar{x})F(\rho - \Delta^e)q_2 P_2. \qquad (9)$$

Since ad levels are unobserved, the number of active users, $F(\rho - \Delta^e)$, does not depend on the actual values of q_1 and q_2 but only on their expected value. The first-order conditions for publishers 1 and 2 is thus the standard formula:

$$\frac{\partial \pi_i^{NI}}{\partial q_i} = 0 \iff \frac{\partial P_i}{\partial q_i}q_i + P_i = 0, \qquad (10)$$

which implicitly defines the equilibrium values of q_1^{NI} and q_2^{NI}. Each publisher raises the quantity of advertisements it sells until the point at which the marginal benefit (the extra ad sold) equals the infra-marginal losses (the lower price on all the ads sold so far).

The search engine chooses its allocation rule, $\bar{x}$, as a best response to the publishers' ad volumes

as implicitly defined by (10). The first-order condition for an interior solution $\bar{x}$ is thus

$$0 = \frac{\partial \pi_0^{NI}}{\partial \bar{x}} = \frac{dF(\rho - \Delta^e)}{d\bar{x}} q_0 P_0 + F(\rho - \Delta^e) q_0 \frac{\partial P_0}{\partial \bar{x}}. \tag{11}$$

We therefore obtain the following expression, which captures the fundamental trade-off faced by the search engine:

$$\underbrace{f(\rho - \Delta^e)\left(\rho'(\bar{x}) - \frac{\partial \Delta}{\partial \bar{x}}\right) q_0 P_0}_{\text{RQ}} + \underbrace{F(\rho - \Delta^e) q_0 \frac{\partial P_0}{\partial \bar{x}}}_{\text{IIB}} = 0. \tag{12}$$

The search engine must balance three effects in order to maximize revenue. The first one, captured by the term $f(\rho - \Delta^e)\rho'(\bar{x})$, is related to the relevance of results. Other things being equal, users enjoy relevance, and improving it helps to increase their number.

However, relevance is not the only characteristic that users enjoy. They may prefer to be shown a less relevant link when this results in them being directed to sites with fewer advertisements. The term $-f(\rho - \Delta^e)\frac{\partial \Delta}{\partial \bar{x}}$ in (12) reflects this concern. The desire to attract users explains why Google, for example, publically announces measures to penalize advertising spam sites in its ranking algorithm.[14] These spam sites are the ones that cause a large disutility to users, i.e such that $\delta_i(q_i)$ is relatively large. These two effects constitute the *quality* of the results provided by the search engine. Quality is maximal when $\frac{dF(\rho - \Delta^e)}{d\bar{x}} = 0$, i.e at $\bar{x} = x^*$. The first term on the left-hand side of (12) thus measures the *Return to Quality* (RQ) for the search engine.

The second term of (12) embodies the more subtle observation that the optimal bias from the search engine's point of view is partially determined by strategic considerations over the advertising market. Here, the search engine will use $\bar{x}$ as an instrument to maintain a high price P_0 for its own sponsored links. Note that, from (6), we have

$$\frac{\partial P_0}{\partial \bar{x}} = \frac{\partial r_1}{\partial q_0} - \frac{\partial r_2}{\partial q_0}.$$

The term $\frac{\partial r_i}{\partial q_0}$ (which, from Equation 4, is negative) is the effect of the marginal ad on the search engine upon the incremental revenue from advertising at i, which will typically be a larger negative number when ads are closer substitutes. In order to maintain a high price for its ads, the search

---

[14]See, for example, http://insidesearch.blogspot.co.uk/2012/01/page-layout-algorithm-improvement.html, accessed 19th October 2012.

engine thus has an incentive to bias results towards the publisher i for which $\frac{\partial r_i}{\partial q_0}$ is less negative. Since $\frac{\partial^2 r_i}{\partial q_0 \partial q_i} \leq 0$ (see Equation 3), the search engine will bias its results against a publisher with a relatively high number of ads. By doing so, it ensures that advertisers have fewer opportunities to advertise their products, which drives up the price of the sponsored links.[15] The strength of this effect depends upon the size of $\left| \frac{\partial^2 r_i}{\partial q_0 \partial q_i} \right|$: if this is large so that i's ads are particularly close substitutes for sponsored links then the search engine will have an especially strong incentive to direct users away from i. Because the bias does not result from direct monetary incentives, but operates through a price mechanism, we label this effect the *Indirect Incentive to Bias* (IIB). Summarizing:

Proposition 1. *In the non-integrated case, the search engine biases its rankings against publishers that display many ads or whose ads are close substitutes for its own sponsored links. Too many consumers are directed to site 1 if $\frac{\partial r_1}{\partial q_0} > \frac{\partial r_2}{\partial q_0}$, and too many consumers are directed to site 2 if $\frac{\partial r_1}{\partial q_0} < \frac{\partial r_2}{\partial q_0}$.*

For instance, if we use the linear demand specification introduced in Example 1, we have $\frac{\partial r_1}{\partial q_0} - \frac{\partial r_2}{\partial q_0} = \gamma_2 q_2 - \gamma_1 q_1$. If $\gamma_2 q_2 \geq \gamma_1 q_1$, the first order condition (11) will be satisfied for $\frac{dn}{d\bar{x}} < 0$, i.e $\bar{x} > x^*$. This example clearly illustrates that the incentive to bias against publisher 2 not only depends on the expected level of advertising on each publisher, but also on the relative degrees of substitutability, γ_1 and γ_2. In the particular case in which $\gamma_1 = \gamma_2$, we see that the search engine puts too much weight on the level of ads compared to consumers' ideal.

It follows straightforwardly from Proposition 1 that the search engine will implement unbiased results if the publishers are symmetric.

Definition 4. *Publishers are symmetric if $\delta_1(q_0, \cdot) = \delta_2(q_0, \cdot)$ and $r_1(q_0, \cdot) = r_2(q_0, \cdot)$.*

Corollary 1. *In the non-integrated market structure, when publishers 1 and 2 are symmetric the search engine implements the user-optimal allocation rule: $\bar{x} = x^* = 1/2$.*

Proof. Note from (7), (8), and (10) that symmetry in r_i implies that both publishers will choose the same ad level:

$$r_1(q_0, \cdot) = r_2(q_0, \cdot) \Rightarrow P_1(q_0, \cdot) = P_2(q_0, \cdot) \Rightarrow q_1^{NI} = q_2^{NI}.$$

---

[15]Observers have, for example, accused search engines of systematically biasing their results in Wikipedia's favor in the knowledge that the latter will (for exogenous reasons) show no advertisements.

Therefore we have $\partial r_1(q_0, q_1^{NI})/\partial q_0 = \partial r_2(q_0, q_2^{NI})/\partial q_0$, so that the second term in (11) is zero. Equation (11) reduces to $dF(\rho - \Delta^e)/d\bar{x} = 0$ so that $\bar{x} = x^*(Q)$. If $\delta_1(q_0, \cdot) = \delta_2(q_0, \cdot)$ and $q_1 = q_2$, then $\frac{dF(\rho - \Delta^e)}{d\bar{x}} = f(\rho - \Delta^e)\rho'(\bar{x})$, which equals zero when $\bar{x} = 1/2$. ∎

In the unintegrated case, the incentive for bias comes from the search engine's desire to relax competition in the advertising market by directing consumers to the site which represents less of a competitive threat. Corollary 1 holds because symmetry ensures that neither site is less of a threat then the other.

Similarly, in the special case that ad demands are independent across the search engine and publishers, the IIB effect has no bite and the unintegrated search engine's problem collapses to maximization of its quality. Again, the result is that the search engine will implement zero bias:

Corollary 2. *If ad demands are independent (if (3) holds with equality) then, in the non-integrated market structure, the search engine implements the user-optimal allocation rule, $\bar{x} = x^*$, and there is no bias.*

Proof. From (4), if ad demands are independent then $\partial r_1(q_0, q_1^{NI})/\partial q_0 = \partial r_2(q_0, q_2^{NI})/\partial q_0 = 0$— implying that the last term of (11) (the IIB effect) can be neglected. Equation 11 thus reduces to $dF(\rho - \Delta^e)/d\bar{x} = 0$ so that the search engine's FOC demands $\bar{x} = x^*(Q)$. ∎

5   Integration

5.1   Partial integration without control

Before analysing the case of full integration, it is interesting to consider the intermediate arrangement whereby publisher 1 shares a fraction ϕ_1 of its revenue with the search engine but retains full-autonomy—a case we refer to as partial integration, and denote with a P superscript.[16] The partial integration model is practically relevant for two reasons: firstly, Google is already vertically integrated into the ad platform industry and sells ad technology services to otherwise independent publishers through subsiduaries such as DoubleClick.[17] Thus, the search engine is able to appro-

---

[16]On a related note, Salop and O'Brien (2000) discuss partial ownership and show that it can either exacerbate or mitigate the price effects of integration depending on the level of corporate control involved, whilst Gans and Wolak (2012) provide an analysis in the context of the Australian electricity market.

[17]DoubleClick provides technology to aid publishers in determining which ad(s) to impress on each user, but leaves the publishers in control of their supply of ad inventory. We abstract from the efficiency effects of such technology.

priate a share of advertising revenues generated on such publishers. Secondly, even if the search engine and publisher 1 are formally-integrated, it may be the case that the internal structure of the business is such that the publisher is independently managed and thus retains a degree of autonomy in setting q_1—an arrangement that can be represented by $\phi_1 \to 1$ in this partial integration model.

Site 1's profit then becomes $\pi_1 = (1 - \phi_1)F(\rho - \Delta^e)\bar{x}q_1P_1$. Since ϕ_1 enters as a multiplicative constant, the first order condition is unchanged and hence $q_1^P = q_i^{NI}$.[18] Note that this also implies that x^* is invariant to ϕ_1: consumers' preferences over publishers are not affected by partial integration. The search engine's profits are now

$$\pi_0^P = F(\rho - \Delta^e)\left[q_0 P_0 + \phi_1 \bar{x}q_1 P_1\right],$$

so its first order condition is

$$\frac{\partial \pi_0}{\partial \bar{x}} = \frac{\partial F(\rho - \Delta^e)}{\partial \bar{x}}\left[q_0 P_0 + \underbrace{\phi_1 \bar{x}q_1 P_1}_{\text{IRQ}}\right] + F(\rho - \Delta^e)\left[q_0 \frac{\partial P_0}{\partial \bar{x}} + \underbrace{\phi_1 q_1 P_1}_{\text{DIB}}\right] = 0. \qquad (13)$$

Setting $\phi_1 = 0$ corresponds to the unintegrated case considered above. Relative to that case, we see in the first term that the search engine now cares more about the number of users it attracts because each user is associated with an additional revenue stream (this corresponds to an *Increased Return to Quality* (IRQ)). All else equal, this effect will compel the search engine to implement higher quality (less biased) results the higher is its revenue share from the content-bearing site. There is also a second effect: that the search engine has a *Direct Incentive to Bias* its results in favour of site 1 (term DIB).

In order to determine whether partial integration increases or alleviates search engine bias, we need to study two cases separately.

Case 1: In the first case, $\bar{x}^{NI} > x^*$. In other words, the unintegrated search engine is biased towards publisher 1. It is clear that partial integration cannot then lead the search engine to set $\bar{x}^P < x^*$, since a deviation to $\bar{x}^P = x^*$ would increase both the number of users and the per-user revenue. Therefore, partial integration will lead to less bias if $\bar{x}^P < \bar{x}^{NI}$, and to more bias if

---

[18]If ϕ_1 was a per-unit tax on advertising imposed by the search engine, this equality would no longer hold. Instead, publisher 1 would reduce its supply of advertising space compared to the NI case, which would make it more attractive to users. We discuss this effect in section 5.2, although it occurs for different reasons there.

$\bar{x}^P > \bar{x}^{NI}$.

In the unintegrated case, the optimal allocation rule is characterised by $\partial \pi_0 / \partial \bar{x} = 0$. Standard comparative statics then imply that a sufficient condition for partial integration to reduce $\bar{x}$ (and hence bias) is $\partial^2 \pi_0 / \partial \bar{x} \partial \phi_1 < 0$. Similarly, bias will increase in this case if profits are supermodular, viz. if $\partial^2 \pi_0 / \partial \bar{x} \partial \phi_1 > 0$. Moreover, from (13), we know that

$$\frac{\partial^2 \pi_0}{\partial \bar{x} \partial \phi_1} = q_1 P_1 \left(\bar{x} \frac{\partial F(\rho - \Delta^e)}{\partial \bar{x}} + F(\rho - \Delta^e) \right) > 0 \Leftrightarrow \eta_{\bar{x}} < 1,$$

where $\eta_{\bar{x}} \equiv -\frac{F(\rho - \Delta^e)/\bar{x}}{\partial F(\rho - \Delta^e)/\partial \bar{x}}$ is the elasticity of user participation with respect to the results quality. Conversely, having $\eta_{\bar{x}} > 1$ implies that bias will decrease with partial integration.

Case 2: The second case is that in which the non-integrated search engine is biased against publisher 1 (when $\bar{x}^{NI} < x^*$ or, equivalently, $\frac{\partial P_0}{\partial \bar{x}} < 0$). Partial integration then necessarily leads to more traffic directed to publisher 1 (to an increase in $\bar{x}$). However, at least some of this shift in site 1's favor works to undo the preexisiting bias. Whether the net result is an increase or decrease in the overall level of bias therefore depends upon the magnitude of the shift. Since the IRQ term in (13) always works to reduce bias, a sufficient condition for bias to decrease is that the second term in (13) to becomes smaller: $|q_0 \frac{\partial P_0}{\partial \bar{x}}| > |q_0 \frac{\partial P_0}{\partial \bar{x}} + \phi_1 q_1 P_1|$, i.e

$$\phi_1 < \phi^* \equiv -2 \frac{q_0}{q_1 P_1} \frac{\partial P_0}{\partial \bar{x}}.$$

In particular, this implies that (provided $\bar{x}^{NI} < x^*$) bias can always be reduced relative to the no integration case by implementing some positive degree of partial integration.

The discussion of these two cases is summarized in the following proposition.

Proposition 2. *When the search engine and publisher 1 are partially integrated,*

1. *if $\bar{x}^{NI} > x^*$ and $\eta_{\bar{x}} > 1$, the level of bias decreases,*

2. *if $\bar{x}^{NI} > x^*$ and $\eta_{\bar{x}} < 1$, the level of bias increases,*

3. *if $\bar{x}^{NI} < x^*$ and $\phi_1 < \phi^*$, the level of bias decreases.*

Notice that point 3 follows a different logic from point 1. In point 3, the search engine is initially biased against publisher 1; partial integration causes the search engine to direct more traffic to site 1 (to increase $\bar{x}$), but this serves to offset the preexisting bias so that the overall

level of bias falls. In contrast, partial integration leads the search engine to allocate *less* traffic to publisher 1 in point 1. This is because demand is highly elastic with respect to quality and each extra user attracted to the search engine through a higher quality is now associated with an additional revenue stream. The search engine thus finds it worthwhile to reduce bias, even though this means directing consumers away from a site in which it has a financial interest.

If the search engine were to receive a share of both publishers' revenues (with the share from 2 being denoted ϕ_2) then the above logic extends in a natural way: (13) becomes

$$
\frac{\partial \pi_0}{\partial \bar{x}} = \frac{\partial F(\rho - \Delta^e)}{\partial \bar{x}} \left[q_0 P_0 + \underbrace{\phi_1 \bar{x} q_1 P_1 + \phi_2 (1 - \bar{x}) q_2 P_2}_{IRQ} \right]
$$

$$
+ F(\rho - \Delta^e) \left[q_0 \frac{\partial P_0}{\partial \bar{x}} + \underbrace{\phi_1 q_1 P_1 - \phi_2 q_2 P_2}_{DIB} \right] = 0.
$$

Relative to the case with $\phi_2 = 0$, we see, firstly, that the increased returns to quality effect becomes stornger because the search engine now has two additional sources of per-user revenue rather than one—and therefore an even stronger incentive to attract additional users. Secondly, the search engine now has a direct incentive to bias only to the extent that one of the sites yields to it a higher per-user flow of revenues than does the other. If the per-user revenues that the search engine receives from the two sites are roughly equal then the direct incentive to bias will be small. Only the IRQ effect remains significant in such instances and partial integration cannot then cause bias to increase.

5.2 Full integration

Suppose, now, that the search engine is integrated with publisher 1—a case we denote with an I superscript. This corresponds to setting $\phi_1 = 1$ *and* allowing the search engine to determine q_1. Notice that such integration has both vertical and horizontal features. To the extent that the search engine provides publishers with inputs (eyeballs) that can be sold to advertisers, one may be tempted to speak about vertical integration. However, the particular structure of the online advertising market is such that the search engine: (i) does not receive payments from publishers (these are organic results) and (ii) directly competes with them when it comes to selling eyeballs to advertisers, both points being more characteristic of horizontal integration.

5.2.1   Effects of integration on advertising levels

The integrated firm's profit is now

$$\pi_{01}^I = F(\rho - \Delta^e)(q_0 P_0 + \bar{x} q_1 P_1). \tag{14}$$

Since q_0 is fixed, we can focus on the effects of integration on q_1. The optimal quantity of ads displayed on publisher 1 is given by the first order condition:

$$\frac{\partial \pi_{01}^I}{\partial q_1} = \bar{x}\left(\frac{\partial P_1}{\partial q_1}q_1 + P_1\right) + \underbrace{q_0 \frac{\partial P_0}{\partial q_1}}_{\text{CPE}} = 0. \tag{15}$$

The standard logic of horizontal mergers applies in this case. Because publisher 1 now internalizes the effect of its advertising strategy on the price of the search engine's sponsored links (the *Cross-Price Effect*, CPE), it will choose a smaller q_1 than in the non-integrated case (equation (10)). On the other hand, publisher 2 does not internalize the profit of the search engine, and therefore its advertising level remains unchanged.

Proposition 3. *Integration between the search engine and publisher 1 leads to a decrease in the quantity of ads displayed by publisher 1: $q_1^I \leq q_1^{NI}$. Publisher 2 offers the same quantity of ads under both regimes: $q_2^I = q_2^{NI}$.*[19]

For example, under the specification of Example 1, it is straightforward to establish from (10) and (15) that the ad quantities are[20]

$$q_1^I = \frac{\alpha_1 - 2\gamma_1 q_0}{2\beta_1} < q_1^{NI} = \frac{\alpha_1 - \gamma_1 q_0}{2\beta_1}; \ q_2^I = q_2^{NI} = \frac{\alpha_2 - \gamma_2 q_0}{2\beta_2}. \tag{16}$$

Ignoring—for now—the changes in the allocation rule $\bar{x}$, we can make two remarks on the effects of integration. First, integration alleviates the negative externality that advertisers exert on users, by restricting the supply of advertising space. Notice that this effect is only driven by the

---

[19] The fact that advertising levels go down after integration depends on our assumption of linear pricing. In a different game in which publishers could choose T and p in a two part tariff $T + pQ$, they would always set $p = 0$ irrespective of whether there is integration, so that the quantity of ads would remain the same even after integration. For more on this "neutrality" result with non-linear pricing schemes, see Ambrus, Calvano, and Reisinger (2012). In our view, the linear pricing assumption fits well how advertisements are sold online.

[20] In this simple example, (15) becomes $\bar{x}(\alpha_1 - 2\beta_1 q_1 - \gamma_1 q_0) - \gamma_1 \bar{x} q_0 = 0$ so that q_1^I, like q_i^{NI}, does not depend on $\bar{x}$.

improved market power of the integrated structure with respect to advertisers, and has nothing to do with a desire to attract more users through a decrease in q_1, which could be possible only if q_1 were ex-ante observable.

Second, the asymmetric reaction of the two publishers following integration changes the relative preferences of the users who were indifferent between the two, or who slightly preferred publisher 2 before integration. Indeed, in light of Lemma 1, an immediate corollary to Proposition 3 is the following:

Corollary 3. *Following integration, publisher 1 becomes relatively more attractive to users: $x^*(Q^I) \geq x^*(Q^{NI})$.*

Anticipating the next subsection, this observation implies that even if integration leads to an increase in the traffic to publisher 1 (an increase in $\bar{x}$), such an increase is not necessarily detrimental to users since they now enjoy fewer ads on the publisher they visit. Thus a search engine that directs more traffic to its own content (increases $\bar{x}$) need not be viewed as biased per se.

Another corollary of Proposition 3 is the following:

Corollary 4. *For advertisers to benefit from integration, it is necessary, but not sufficient, that consumers also benefit from it.*

Indeed, the fall in ad volumes following integration (Proposition 3) implies that ad prices must rise and per-user advertiser revenue must fall. Advertiser profits can therefore only increase if the number of users increases.

5.2.2 Effects of integration on the allocation of traffic

The effect of integration on the search engine's decision of how to allocate traffic somewhat parallels the partial integration case. In particular, the FOC can be found from (13) by setting $\phi_1 = 1$ (and recalling that integration will result in a fall in q_1). Thus, as in the partial integration model, the first term in (13) reflects the additional weight that the search engine puts on attracting users with high quality results—an effect that unambiguously acts to reduce bias post-integration. In particular, we again have that

Proposition 4. *Integration may lead to a reduction in bias, $|x^* - \bar{x}|$.*
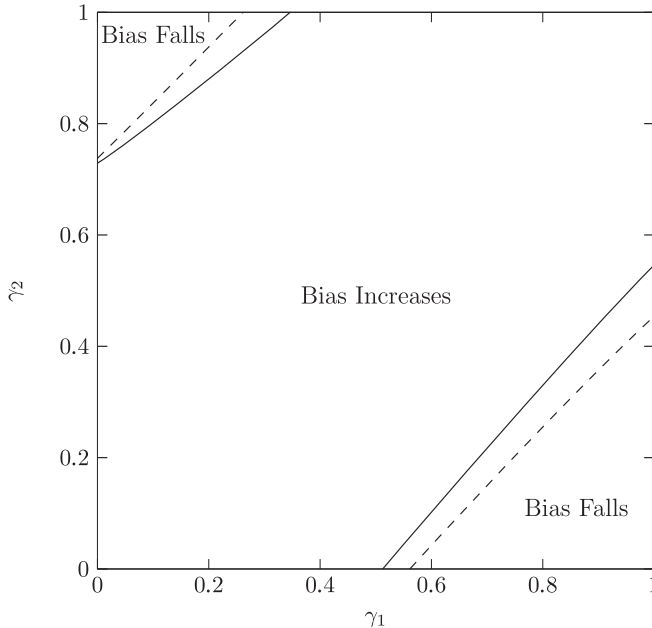
Figure 2: Plot of the changes in the absolute level of bias, b, after integration relative to the non-integrated case. The figure is plotted for the specification of Example 1 with $\alpha_i = \beta_i = 1, \delta_i = t = q_0 = 1/4$, and $\phi_1 = 1$. The solid black lines trace the loci of points along which full integration does not cause a change in bias ($b^I = b^{NI}$); the dashed lines serve a similar purpose for partial integration ($b^P = b^{NI}$).

As an illustration of this result, Figure 2 represents the change in bias implied by integration, under the specification of Example 1 for a particular parameterisation. Relative to the no-integration benchmark, we see that the integrated search engine provides less biased results for sufficient asymmetry in publishers' ad substitutability with the search engine. In this example, when $\gamma_1 \gg \gamma_2$, the unintegrated search engine heavily biases its results in favor of site 2 in order to soften advertising market competition. Integrating with site 1 provides an obvious incentive to undo this bias with the result that the overall level of bias falls. If, on the other hand, $\gamma_1 \ll \gamma_2$ then the unintegrated search engine is already heavily biased in favor of site 1 so that it has little to gain from further increases in $\bar{x}$. That x^* increases after integration then implies that the overall level of bias falls.

Since Figure 2 is plotted for $\phi_1 = 1$ we can interpret the difference between the full and partial integration cases shown there as representing the incremental effect of allowing the search engine to choose q_1.

To gain further insight, let us again restrict our attention to a symmetric setting (see Definition 4). Recall that, from Corollary 1, symmetry implies that there is no bias in the non-integrated

case. This is no longer the case with integration, as is suggested by figure 2:

Lemma 2. *When publishers 1 and 2 are symmetric, integration leads the search engine to bias its results in favor of publisher 1:* $\bar{x}^I > x^*(Q^I) \geq x^*(Q^{NI}) = \bar{x}^{NI} = 1/2$.

Proof. The last two equalities follow from Corollary 1, and the second inequality is given by Corollary 3. The only inequality that remains to be proven is the first one. To do so, take the first-order condition (13) evaluated at $x^*(Q^I)$. By definition, $\frac{dF(\rho - \Delta^e)}{d\bar{x}}|_{\bar{x}=x^*} = 0$. Since $r_1(q_0, \cdot) = r_2(q_0, \cdot)$, we also have $\frac{\partial P_0}{\partial \bar{x}} = 0$, so that $\frac{\partial \pi_{01}^I}{\partial \bar{x}}|_{\bar{x}=x^*} > 0$. The optimal value of $\bar{x}$ for the search engine is thus larger than x^*. ∎

We see that the symmetry assumption is useful, because the trade-off induced by integration is then clear: *under symmetry, integration between the search engine and a publisher leads to biased results in favor of this publisher, but also to a lower level of advertising, which benefits users.*

In the next section, we try to draw policy implications from the previous analysis.

6   Policy implications

Although the intuitive incentive to introduce own-content bias following integration is present in the model, our analysis underlines two new effects that might offset or mitigate it. The first one is the increased return to attracting an extra user (the Increased Return to Quality term in (13)), which can lead to better quality results if bias initially exists, and otherwise softens the incentive to introduce bias. The second one is the reduction in the advertising distraction incurred by users (the Cross-Price Effect in (15)). The magnitude of these effects is of course an empirical matter, outside the scope of this paper. However, the model may offer some guidance regarding the circumstances under which different effects should prevail. In this section we focus on three important parameters of the model, namely the substitutability between sponsored links and ads on the publishers' websites, consumers' attitudes towards advertising, and content differentiation. Throughout, we assume that publishers are symmetric in order to focus on the main trade-offs identified above.

6.1   Substitutability between sponsored links and banners

In practice, the extent to which advertisers regard these ads as substitutes is subject to disagreement. Whilst Hahn and Singer (2008) present survey evidence suggesting that sponsored links and banner ads are indeed substitutes, the FTC found that the relevant market for Google was the search advertising market, thereby implying that display ads showed on the various publishers are not substitutes to the sponsored links on search engines' results pages (FTC, 2007). Our view is that substitutability most likely differs across segments. It is probably rather strong between general and vertical search engines, which both display sponsored links related to a given query, and weak between a general search engine and a publisher whose content and design are more amenable to brand advertising.

The degree of substitutability is an important parameter for two key reasons. First, from a practical point of view, the extent to which advertisers view search engines and other publishers as substitutes will help to determine the relevant market, a necessary step should the European Commission or another competition authority decide to sue Google for abuse of dominant position. Secondly, in the integrated case, the degree of substitutability will affect the magnitude of the change in ad levels. Consider the first-order condition (15). It is clear that the change in q_1 with respect to the non-integrated case will be more important when substitutability is high (when $\partial P_0 / \partial q_1$ is large and negative). In the extreme case of independence (which implies $\frac{\partial P_0}{\partial q_1} = 0$), the last term of (15) disappears and one cannot expect that users would benefit from an improved surfing experience due to fewer ads after integration. In the case of independence it is also clear, from equation (13), that the search engine's new incentives will lead it to be biased in favor of publisher 1 (recall from Corollary 2 that there is no bias in the unintegrated market when ad demands are independent). Since relevance decreases and ad volumes do not change, users are made worse-off by integration. This causes them to reduce their usage of the search engine, which, along with the increase in ad prices, makes advertisers worse-off. This analysis is summarized below:

Proposition 5. *Under symmetry, when there is no ad substitutability between the search engine and the publishers, integration reduces consumers' and advertisers' surplus.*

We emphasize that the above result implies that, under symmetry, some degree of substitutability between display and search ads (so that (3) holds as a strict inequality) is a necessary
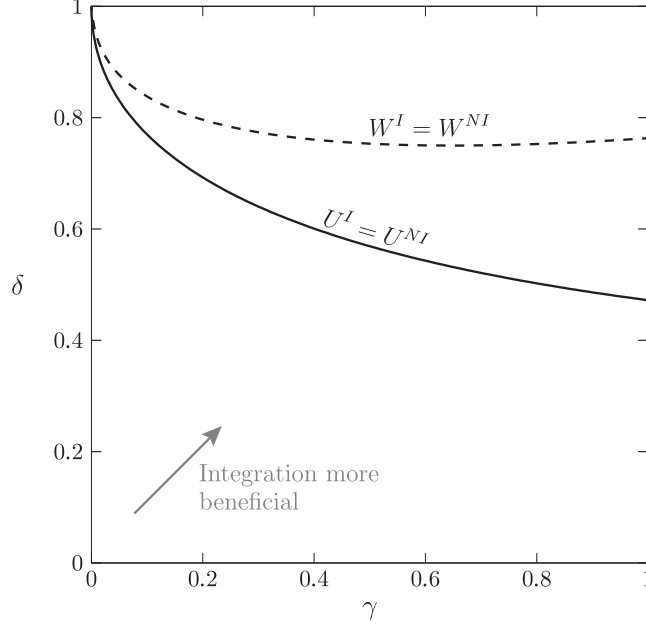
Figure 3: The effect of integration on the utility of users $(U^I - U^{NI})$ and total welfare $(W^I - W^{NI})$ for $\alpha = \beta = t = 1, q_0 = 1/4$. The solid contour indicates the locus of points satisfying $U^I = U^{NI}$ (where consumers obtain the same expected utility both before and after integration); the dashed curve serves the same purpose for overall welfare. Consumers and society are made better-off by integration everywhere above the respective curve, and worse-off everywhere below.

but, as we shall see, not sufficient condition for either advertisers or consumers to benefit from integration. With imperfect substitutability, the results are a priori more ambiguous, since users benefit from fewer ads but may suffer from biased results. In order to gain further intuition we invoke the linear Cournot-Hotelling model introduced in Example 1, with consumer search costs uniformly distributed on $[0, 1]$ so that $F(\rho - \Delta^e) = U(\rho, \Delta)$. With this specification, equilibrium ad quantities are as given in (16), $\bar{x}^{NI} = 1/2$ (by Corollary 1), and $\bar{x}^I$ can be straightforwardly calculated according to (13).

Figure 3 shows how the size of the user base (or, equivalently, the expected utility of a user) is affected by integration. Integration is more likely to be beneficial if γ is large: making search engine and publisher ads closer substitutes (increasing γ) increases the size of the competitive externality that the search engine and publisher 1 exert on each other. It follows that, when these two firms integrate, the incentive to cut quantity will be strongest when γ is large, so that consumers can then expect to benefit from a larger fall in ad volumes. Since a necessary condition for such a fall to be beneficial is that consumers dislike ads, the effect of changes γ is most strongly felt when δ is large.

As corollary 4 indicates, we see in Figure 3 that the requirement that integration increase social welfare ($W = U + \pi_0 + \pi_1 + \pi_2 + \pi_a$) is more demanding than is the requirement that user surplus increase.

6.2   Consumers' attitude towards advertising

Figure 3 also shows how, all else equal, increasing the nuisance cost of advertisements makes integration more attractive for users and more beneficial for society at large. Intuitively, we know that ad volumes fall after integration and consumers benefit more from this fall when δ is large.

In all our analysis so far, advertising has been regarded as a pure nuisance for users. That is, a user's potential surplus from trade with advertisers is more than offset by the distortion entailed by advertising. However, although the "pure nuisance" assumption is standard in the literature (see, for example, Anderson and Coate, 2005), it may be ill-suited to study some situations. In practice, for a number of services, users may derive utility from the presence of advertisers. This is the case for instance if the publishers are business directories, used by consumers to find local shops. Examples include yellow pages or map services. It turns out that our analysis is amenable to positive externalities from advertising—that is, to the assumption that $\delta_i(.)$ is non-increasing in q_j for every i, j.

In this case, if we assume symmetry between publishers, the effects of integration are rather unambiguous. Integration leads to results that are more biased and to lower levels of advertising. Both effects leave users strictly worse-off and, because the drop in *per-user* advertisers' surplus is accompanied by a decrease in user participation, advertisers also lose from integration.

Proposition 6. *Under symmetry, and when advertising exerts a positive externality on users, both consumer and advertiser surplus go down after integration.*

6.3   Content differentiation

Content differentiation between publishers 1 and 2 is also an important determinant of the desirability of integration. To see this clearly, suppose that both publishers are identical in the sense that they are not only symmetric, but also offer the same or equivalent content so that consumers are ex post indifferent across sites' content.[21] Sites for which this is likely to be a good fit include

---

[21]Note that this does not render the search engine useless because consumers may not know the content sites' addresses and still rely on the search engine as a tool to navigate to one of the sites over which they are indifferent.

pure information services such as financial data providers, weather services, or directions providers. Formally, we represent this indifference by $\rho'(\bar{x}) = 0$ for every $\bar{x}$. Given that consumers are indifferent between the two publishers, integration will lead to a complete foreclosure of publisher 2, that is $\bar{x} = 1$. Since the advertising level is lower on publisher 1 after integration (Proposition 3), consumers strictly benefit from it.

Proposition 7. *When publishers 1 and 2 are symmetric and offer undifferentiated content, integration leads to the complete foreclosure of publisher 2 ($\bar{x}^I = 1$), and users are strictly better-off than in the non-integrated case.*

Proof. Symmetry along with Proposition 3 implies that $q_1^I < q_2^I = q_i^{NI}$. Two properties follow: (i) all consumers prefer site 1 under integration in equilibrium so that $dF(\rho - \Delta^e)/d\bar{x} > 0$, and (ii) $\partial P_0/\partial \bar{x} > 0$. From (13) (with $\phi_1 = 1$) it is clear that the integrated firm's profit is increasing in $\bar{x}$:

$$\frac{\partial \pi_{01}^I}{\partial \bar{x}} = f(\rho - \Delta^e) \left[\underbrace{\rho'(\bar{x})}_{=0} - \underbrace{\frac{d\Delta}{d\bar{x}}}_{<0} \right] (q_0 P_0 + \bar{x} q_1 P_1) + F(\rho - \Delta^e) \left(\underbrace{q_0 \frac{\partial P_0}{\partial \bar{x}}}_{>0} + q_1 P_1 \right) > 0,$$

so that $\bar{x} = 1$ in equilibrium. The effect of integration on user utility is $U(\rho(1), \Delta(Q^I)) - U(\rho(\bar{x}^{NI}), \Delta(\bar{x}, Q^{NI}))$, which is positive since $\rho(1) = \rho(\bar{x}^{NI})$ and $\Delta(\bar{x}, Q^{NI}) > \Delta(\bar{x}, Q^I)$. ∎

The fall in ad supply implies that advertisers' *per-user* profit goes down with integration, but this effect may be compensated by the increase in the total number of users.

More generally, as the degree of content differentiation goes up, two opposite forces are at play: on the one hand, *for a given level of bias*, consumers expect a larger drop in utility from the increase in bias. On the other hand, the increased return to quality (IRQ) effect identified in (13) induces the integrated search engine to reduce its bias as differentiation increases (because consumers care more about being sent to the right site when sites are highly differentiated, making participation more sensitive to result quality). Whether integration is more beneficial as content becomes more differentiated is thus ambiguous in general. Figure 4 shows that, in the context of Example 1, the IRQ effect dominates for larger values of t.
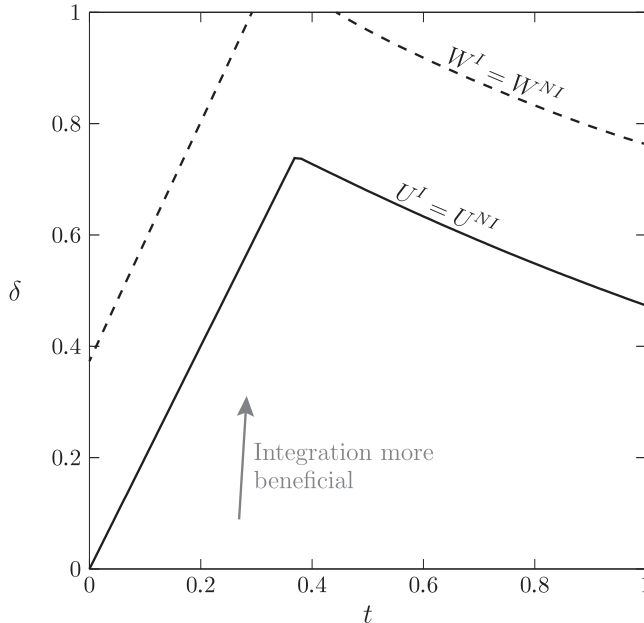
Figure 4: The effect of integration on the utility of users $(U^I - U^{NI})$ and total welfare $(W^I - W^{NI})$ for $\alpha = \beta = \gamma = 1, q_0 = 1/4$. The solid contour indicates the locus of points satisfying $U^I = U^{NI}$ (where consumers obtain the same expected utility both before and after integration); the dashed curve serves the same purpose for overall welfare. Consumers and society are made better-off by integration everywhere above the respective curve, and worse-off everywhere below.

7   Conclusion

The central importance of search engines has made them the focus of recent antitrust scrutiny, with particular attention being paid to the issue of own-content bias. In this paper we develop a theoretical model in which a search engine allocates user traffic among two publishers whilst competing with them to attract advertisers. We compare several market structures: independence of publishers as well as partial and full integration between one publisher and the search engine.

Our model delivers several insights. Firstly, we show that, even when the search engine does not have a financial interest in either publisher, it has, in general, *indirect* monetary incentives to favor publishers that exert the least competitive pressure in the advertising market. These are the ones that display few ads, or whose ads are not regarded as close substitutes to search engine sponsored links by advertisers.

Secondly, when the search engine integrates with some publisher, there exist *direct* monetary incentives for the search engine to channel traffic towards its own-content. We show, however, that there are several reasons why the level of bias in the results need not increase: (i) the new, own-

content bias may offset the initial, advertising-related bias; (ii) by increasing the expected per-user revenue of the search engine, integration increases the incentive to provide high-quality search results, which further disciplines the search engine; (iii) the integrated publisher has an incentive to lower its supply of advertising space, which benefits users and justifies making this website more prominent. Indeed, when bias is defined as a deviation from the user surplus-maximizing results, the no-bias benchmark will typically have a majority of users directed towards the integrated publisher.

The above remarks notwithstanding, search engines' and users' interests are not perfectly aligned: search engines' natural incentive to favor own-content sites will often result in bias even after consumers' preferences are fully accounted for. In this respect, our contribution has been to explore the kinds of conditions that are most conducive to welfare-improving integration. Users are most likely to benefit from integration when the integrated content site and search engine are close substitutes in the advertising market, provided users have a strong enough aversion to advertisements, in which case any change in search result quality is likely to be dominated by the equilibrium fall in ad volumes that follows integration. A similar point holds for the case in which content sites are largely homogeneous: if users are indifferent between publishers then they are relatively unaffected by bias and the most important effects are those favorable ones taking place in the ad market. However, when consumers benefit from advertising, integration is more likely to harm them. Advertisers typically fare worse than consumers after integration because integration leads to an increase in ad prices and a fall in ad volumes. The requirement that integration improves overall social welfare is therefore generally more demanding.

Our paper, of course, does not tell the whole story. In particular, we have left aside considerations of efficiency gains as well as dynamic aspects of the situation. On the first point, advocates of integration between Google and publishers argue that it will improve the user experience, by making results more easily accessible.[22] One should also take into account that major search engines, and Google in particular, have developed very efficient tools to serve ads, which may justify integration, independently of our present arguments.

Regarding dynamic aspects, one needs to take into account future incentives to invest in quality. There are at least two dimensions worth looking into. The first one concerns the search engine's

---

[22]See, for instance, the testimony of Eric Schmidt, Google chairman, before the U.S Senate Judiciary Committee in September 2011.

incentives to invest in order to improve its algorithm. While we ignore such investments in this paper, the argument according to which integration may lead to less bias through the increased return on attracting extra users can also be applied to quality dimensions that are unrelated to bias. The other dimension to consider is that of the publishers' incentives to produce high-quality content. Following integration, will a higher market share boost the marginal incentive to invest in content production, or will it lead the integrated publisher to rest on its laurels?

More generally, we have focused on the private consumer welfare implications of bias, but in some instances there may be a broader social interest in the efficacy of search, owing to externalities in the information consumption process. One obvious example is the consumption of political information, where society at large may wish a search engine to deliver clear and unfiltered access to information even when consumers do not. In such instances the appropriate bias-free benchmark is likely to differ to that discussed here. This and other open questions are left for future research.

References

Ambrus, A., E. Calvano, and M. Reisinger (2012): "Either or both competition: a two-sided" theory of advertising with overlapping viewerships," Working Paper.

Anderson, S., Ø. Foros, and H. J. Kind (2011): "Competition for advertisers in media markets," Working Paper.

Anderson, S. P., and S. Coate (2005): "Market Provision of Broadcasting: A Welfare Analysis," *Review of Economic Studies*, 72(4), 947–972.

Armstrong, M., and J. Zhou (2011): "Paying for Prominence," *Economic Journal*, 121(556), F368–F395.

Athey, S., E. Calvano, and J. Gans (2010): "Will the Internet Destroy the News Media? or Can Online Advertising Markets Save the Media?," Working Paper.

Athey, S., and G. Ellison (2011): "Position Auctions with Consumer Search," *Quarterly Journal of Economics*, 126(3), 1213–1270.

Burguet, R., R. Caminal, and M. Ellman (2013): "In Google We Trust?," Working Paper.

Chen, Y., and C. He (2011): "Paid Placement: Advertising and Search on the Internet," *Economic Journal*, 121(556), F309–F328.

Choi, J. P., and B.-C. Kim (2010): "Net neutrality and investment incentives," *RAND Journal of Economics*, 41(3), 446–471.

de Cornière, A. (2011): "Search Advertising," Working Paper.

Edelman, B., and Z. Lai (2013): "Exclusive Preferential Placement as Search Diversion: Evidence from Flight Search," Working Paper.

Edelman, B., and B. Lockwood (2011): "Measuring Bias in "Organic" Web Search," Working Paper available at http://www.benedelman.org/searchbias/.

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(1), 242–259.

Eliaz, K., and R. Spiegler (2011): "A Simple Model of Search Engine Pricing," *Economic Journal*, 121(556), F329–F339.

Ellman, M., and F. Germano (2009): "What do the Papers Sell? A Model of Advertising and Media Bias," *Economic Journal*, 119(537), 680–704.

FTC (2007): "Statement of Federal Trade Commission Concerning Google/DoubleClick FTC," FTC File No. 071-0170.

——— (2013): "Google Agrees to Change Its Business Practices to Resolve FTC Competition Concerns In the Markets for Devices Like Smart Phones, Games and Tablets, and in Online Search," FTC press release available at http://www.ftc.gov/opa/2013/01/google.shtm, accessed 9th January 2013.

Gans, J. S., and F. A. Wolak (2012): "A Comparison of Ex Ante versus Ex Post Vertical Market Power: Evidence from the Electricity Supply Industry," Working Paper.

Hagiu, A., and B. Jullien (2011): "Why Do Intermediaries Divert Search?," *RAND Journal of Economics*, 42(2), 337–362.

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

Inderst, R., and M. Ottaviani (2012): "Competition Through Commissions and Kickbacks," *American Economic Review*, 102(2), 780–809.

Manne, G. A., and J. D. Wright (2007): "Google and the Limits of AntiTrust: The Case Against the Antitrust Case Against Google," Working Paper.

Salop, S. C., and D. P. O'Brien (2000): "Competitive Effects of Partial Ownership: Financial Interest and Corporate Control," *Antitrust Law Journal*, 67, 559–614.

Tarantino, E. (2011): "Vertical Search Engines Foreclosure," Working Paper.

Taylor, G. (2013a): "Attention Retention: Targeted Advertising and the Ex Post Role of Media Content," Working Paper.

——— (2013b): "Search Quality and Revenue Cannibalization by Competing Search Engines," *Journal of Economics & Management Strategy*, 22(3), 445–467.

Varian, H. R. (2007): "Position auctions," *International Journal of Industrial Organization*, 25(6), 1163–1178.

White, A. (forthcoming): "Search Engines: Left Side Quality versus Right Side Profits," *International Journal of Industrial Organization*.

White, A., and K. Jain (2010): "The Attention Economy of Search and Web Advertisement," Working Paper.

Wright, J. D. (2011): "Defining and Measuring Search Bias: Some Preliminary Evidence," Working Paper.

A   Microfoundations for advertisers' revenue

In section 2.2, we assume that advertisers' revenue takes the form

$$r_0(q_0) + \bar{x} r_1(q_0, q_1) + (1 - \bar{x}) r_2(q_0, q_2).$$

Such an expression is consistent with at least two models describing the interaction between advertisers and consumers:

1. The first model is one in which there is a representative advertiser who wishes to reach consumers several times. More specifically, assume that showing q ads to a consumer is worth $S(q)$, S being increasing, concave and such that $S(0) = 0$. For a vector (q_0, q_1, q_2) of ad quantities, a consumer such that $x < \bar{x}$ is thus worth $S(q_0 + q_1)$, while consumers such that $x > \bar{x}$ are worth $S(q_0 + q_2)$. Therefore, the expected revenue generated by a search engine user is

$$\underbrace{S(q_0)}_{r_0(q_0)} + \bar{x} \underbrace{(S(q_0 + q_1) - S(q_0))}_{r_1(q_0, q_1)} + (1 - \bar{x}) \underbrace{(S(q_0 + q_2) - S(q_0))}_{r_2(q_0, q_2)}.$$

One easily checks that $r_0(0) = r_1(q_0, 0) = r_2(q_0, 0) = 0$ and that $\frac{\partial^2 r_i}{\partial q_0 \partial q_i} \leq 0$.

2. In the second model, there is a continuum of advertisers who wish to reach consumers only once. When a consumer sees the ad of a firm, that firm's expected profit is s. Assume that s is heterogenous in the population, distributed according to a cdf H. Assume also that there is a centralized market for advertising space. When a user enters a query $x < \bar{x}$, the supply of advertising space is $q_0 + q_1$. For any price p_1, the demand for advertising space equals the mass of firms with $s \geq p_1$, i.e $1 - H(p_1)$. The market-clearing price p_1 for an impression to a consumer with $x \leq \bar{x}$ is thus such that $1 - H(p_1) = q_0 + q_1$, i.e $p_1 = H^{-1}(1 - q_0 - q_1)$. The total revenue of advertisers is given by the expected profit of the advertisers who buy a slot. For $x < \bar{x}$, it is $\int_0^{q_0 + q_1} H^{-1}(1 - q)dq$, and for $x > \bar{x}$, it is $\int_0^{q_0 + q_2} H^{-1}(1 - q)dq$. Rearranging terms leads to the following expression for advertisers' revenue:

$$\underbrace{\int_0^{q_0} H^{-1}(1 - q)dq}_{r_0(q_0)} + \bar{x} \underbrace{\int_{q_0}^{q_0 + q_1} H^{-1}(1 - q)dq}_{r_1(q_0, q_1)} + (1 - \bar{x}) \underbrace{\int_{q_0}^{q_0 + q_2} H^{-1}(1 - q)dq}_{r_2(q_0, q_2)},$$

which also satisfies $r_0(0) = r_1(q_0, 0) = r_2(q_0, 0) = 0$ and $\frac{\partial^2 r_i}{\partial q_0 \partial q_i} \leq 0$. In that case, the price of an ad on publisher 1 is $p_1 = H^{-1}(1 - q_0 - q_1) = \frac{\partial r_1(q_0, q_1)}{\partial q_1}$, the price of an ad on publisher 2 is $p_2 = H^{-1}(1 - q_0 - q_2)\frac{\partial r_2(q_0, q_2)}{\partial q_2}$, and the average price of an ad on the search engine is $p_0 = \bar{x}p_1 + (1 - \bar{x})p_2 = \frac{\partial[r_0(q_0) + \bar{x}r_1(q_0, q_1) + (1 - \bar{x})r_2(q_0, q_2)]}{\partial q_0}$. This model is thus formally equivalent

to the one developed in the main text.

Another assumption that we make is that the value of a consumer to advertisers is independent of his position x. If we consider the second model above, this assumption can be justified if we assume that each advertiser a has a willingness to pay equal to $s_a(x)$ (instead of s_a). We also need to assume that the centralized advertising market is such that firms are allowed to condition their demands for ad space on the position of the consumer, which corresponds to a situation of targeted advertising. If, for every x, the number of firms with $s_a(x) \leq p$ is $H(p)$ and does not depend on x, then our results carry through.

B Endogeneity and observability of q_0 and x.

One of the main lessons of the paper is that consumers can be made better-off by integration, despite an increase in bias, thanks to the lower level of ads to which they are exposed. Assuming that q_0 is fixed allowed us to show this clearly, since only q_1 was affected by integration. Moreover, such an assumption also allowed us to focus on the allocation rule as the only strategic variable of the search engine.

If we relax the assumption and make q_0 endogenous, such a result is much harder to obtain. Indeed, the change in q_0 will also have an effect on q_2, so that the three variables change in a priori ambiguous directions.

We are led to consider two cases in turn, depending on whether q_0 is observed by consumers prior to their participation decision. Suppose first that q_0, like q_1 and q_2, is not observed by consumers. In the linear case (Example 1), we can show that integration decreases the expected number of ads that each user will be exposed to.

Proposition 8. *Suppose that the model specification is that of Example 1, with symmetry between publishers 1 and 2. Then the expected number of ads seen by a user goes down.*

Proof. In the non integrated case, q_0 maximizes $q_0 P_0 = q_0(\alpha - \beta q_0 - \gamma(x q_1 + (1-x)q_2))$, and q_1 and q_2 maximize $q_i(\alpha - \beta q_i - \gamma q_0)$. The solution to the system of first-order conditions is then $q_0^{NI} = q_1^{NI} = q_2^{NI} = \frac{\alpha}{2\beta+\gamma}$, for a total number of ads seen by a user $m^{NI} \equiv \frac{2\alpha}{2\beta+\gamma}$

In the integrated case, q_0 and q_1 are chosen so as to maximize $q_0 P_0 + \bar{x} q_1 P_1 = q_0(\alpha - \beta q_0 - \gamma(\bar{x} q_1 + (1-\bar{x})q_2)) + \bar{x} q_1(\alpha - \beta q_1 - \gamma q_0)$, while q_2 still maximizes $q_2(\alpha - \beta q_2 - \gamma q_0)$. This gives us

a vector $(q_0^I(\bar{x}), q_1^I(\bar{x}), q_2^I(\bar{x}))$:

$$q_0^I(x) = \frac{2\alpha\beta - \alpha\gamma - \bar{x}\alpha\gamma}{4\beta^2 - \gamma^2 - 3\bar{x}\gamma^2}, q_1^I(\bar{x}) = \frac{\alpha\left(4\beta^2 - 4\beta\gamma + (1 - \bar{x})\gamma^2\right)}{8\beta^3 - 2(1 + 3\bar{x})\beta\gamma^2}, q_2^I(\bar{x}) = \frac{2\alpha\beta^2 - \alpha\beta\gamma - \bar{x}\alpha\gamma^2}{\beta\left(4\beta^2 - \gamma^2 - 3\bar{x}\gamma^2\right)}.$$

The expected number of ads showed to a user will then be $m^I(x) \equiv q_0^I(x) + \bar{x}q_1^I(\bar{x}) + (1 - \bar{x})q_2^I(\bar{x})$.

It is then straightforward to show that $m^{NI} - m^I(\bar{x})$ is a decreasing function of $\bar{x}$. At $\bar{x} = 1$, it is of the same sign as $\beta - \gamma$, i.e positive. Thus, for every $\bar{x}$, $m^{NI} > m^I(\bar{x})$. This is true in particular for $\bar{x} = \bar{x}^I$. ∎

On the other hand, suppose that q_0 is observed along with $\bar{x}$. Then the optimal choice of q_0 must solve the trade-off between achieving a high per-user revenue (i.e maximizing q_0P_0) and attracting many users (maximizing $F(\rho - \Delta^e)$). The first-order condition, in the non-integrated case, is

$$-f(\rho - \Delta^e)\frac{\partial\Delta^e}{\partial q_0}q_0P_0 + F(\rho - \Delta^e)\left(P_0 + q_0\frac{\partial P_0}{\partial q_0}\right) = 0,$$

whereas that in the integrated case is

$$-f(\rho - \Delta^e)\frac{\partial\Delta^e}{\partial q_0}q_0P_0 + F(\rho - \Delta^e)\left(P_0 + q_0\frac{\partial P_0}{\partial q_0} + \bar{x}q_1\frac{\partial P_1}{\partial q_0}\right) = 0.$$

It is difficult to obtain general, tractable results for the endogenous and observable q_0 case, but equilibria can easily be computed and exhibit the same basic properties observed in the baseline model presented above. Consider, for instance, Example 1 in the symmetric case with $\alpha = \beta = t = 1$ and F uniform on $[0, 1]$. Figure 5 then shows the effect of integration upon advertising volumes. The total volume of ads seen by those consumers visiting either site falls with integration, which implies that integration results in *every* consumer seeing fewer ads in total.

It is interesting to look at the decomposition of ads amongst the integrated firm's portfolio of sites. The quantity of ads shown on the search engine itself falls (Figure 5(c)) as the firm seeks to alleviate the competitive externality implied by substitutability of advertisements. As in the baseline case, this effect is strongest when ads are close substitutes (when γ is large). One might expect the same to be true of ads on the integrated publisher. However, in this case there is another, countervailing consideration. Intuitively, the integrated firm has an incentive to move its supply of ads away from the search engine to the publisher—where the increase in ad volume is
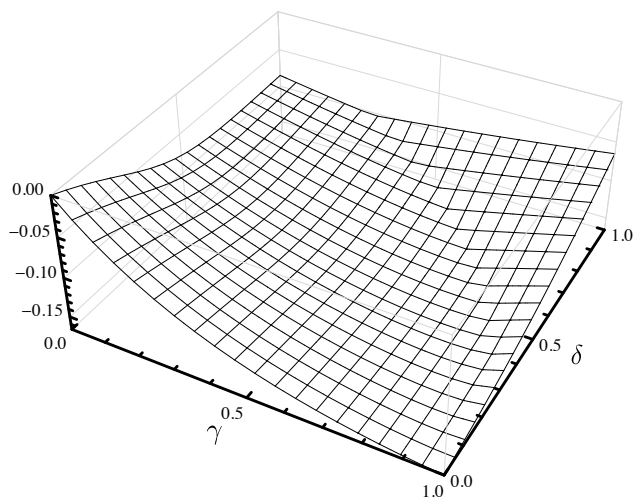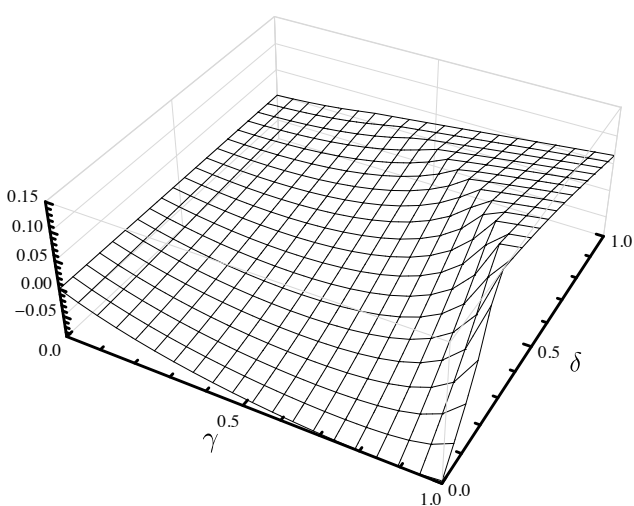
(a) Change in number of ads seen by consumers directed to site 1 $(q_0^I + q_1^I - q_0^{NI} - q_1^{NI})$.
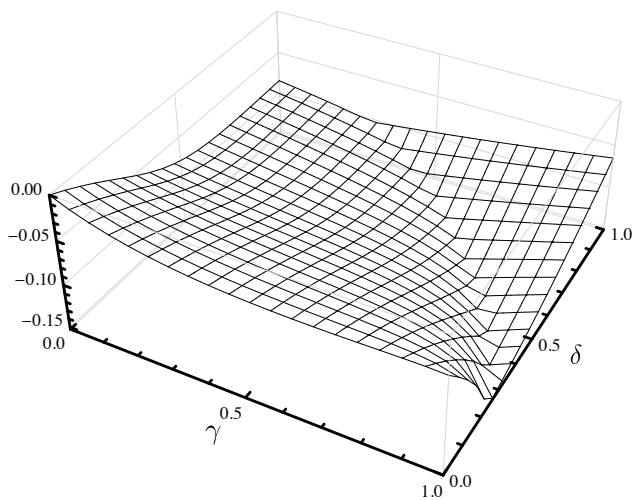
(b) Change in number of ads seen by consumers directed to site 2 $(q_0^I + q_2^I - q_0^{NI} - q_2^{NI})$.
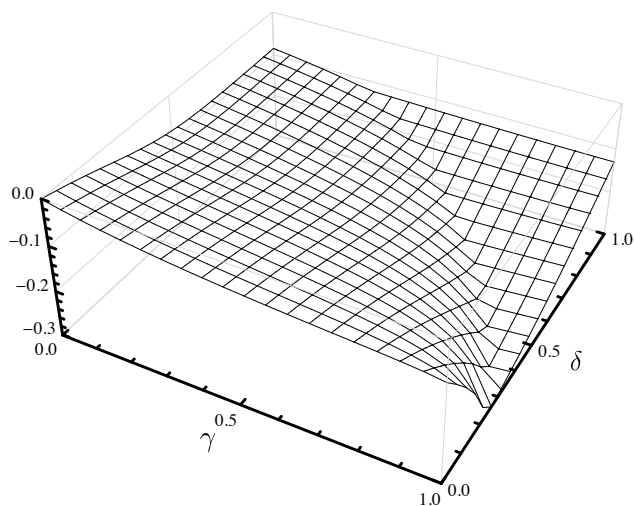
(c) Change in q_0 $(q_0^I - q_0^{NI})$.

(d) Change in q_1 $(q_1^I - q_1^{NI})$.

Figure 5: Plots of the effect of integration on ad volumes when consumers observe q_0. The kink in these surfaces stems from the fact that, for δ and γ sufficiently high, the integrated search engine sets a corner solution of $q_0 = 0$.
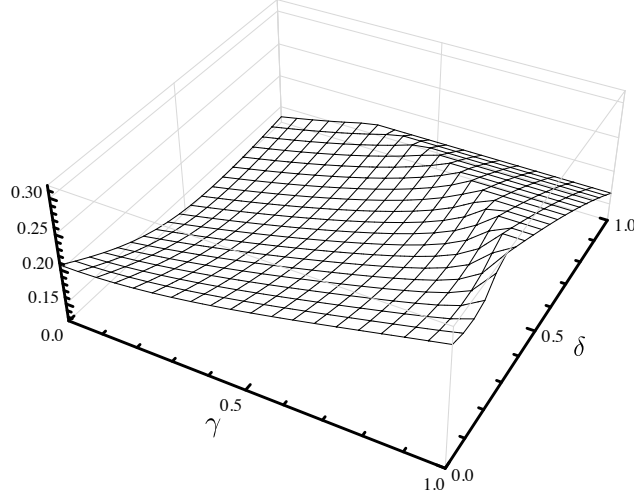
Figure 6: Plots of the post-integration bias (b^I) when consumers observe q_0.

not observed by consumers.[23] When the ad nuisance cost is sufficiently high, the possibility of such a deception is so tempting that q_1 increases in spite of the competitive externality.

Although ad levels on site 1 may increase, the offsetting effect of substitutability ensures that they fall relative to those at site 2. Indeed, this can easily be seen analytically from the first order conditions of the publishers: for a given choice of q_0, site 2 solves (10), whereas the integrated firm must set q_1 to solve (15). Since the CPE term in (15) is negative, and since $q_1^{NI} = q_2^{NI}$ under symmetry, we have

Proposition 9. *Under symmetry, integration causes site 1's ad volume to fall relative to that of site 2 so that $q_1^I < q_2^I$.*

Thus, by Lemma 1, site 1 becomes relatively more attractive to consumers after integration. This manifests as an increase in the user-optimal allocation rule, x^*. Although it is therefore in consumers' interest that additional traffic be directed to the integrated publisher, integration creates a sufficiently strong incentive for own-content bias to ensure that the integrated search engine favors its own site *too much*. The scale of the ensuing bias is shown in Figure 6 (it is easily verified that Lemma 2 carries over to the endogenous q_0 case so that pre-integration bias is zero).

Given the above discussion, it is not a priori clear whether or not consumers are harmed by integration: advertising levels fall, which reduces the advertising nuisance burden, but bias is introduced into the search results—decreasing the average content consumption utility. Figure 7

---

[23]Although this behaviour is not observed, q_1 is correctly anticipated in equilibrium.
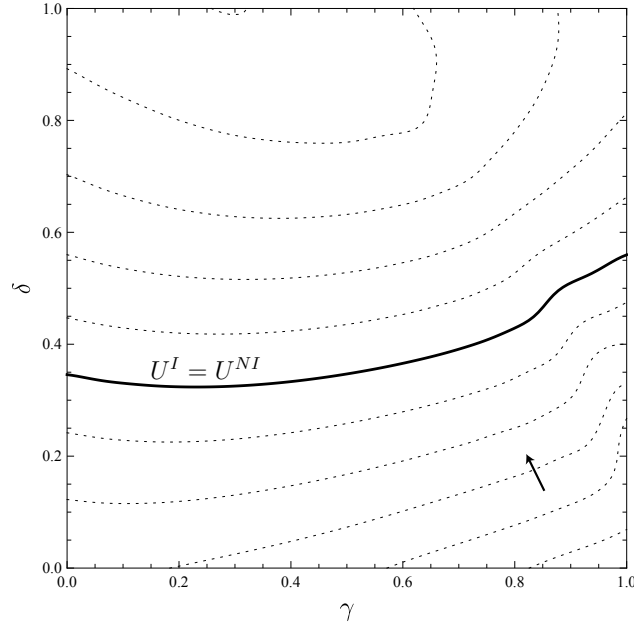
Figure 7: Contour plots of the effect of integration on consumer surplus when consumers observe q_0. Contours are drawn every 0.015 units and the difference $(U^I - U^{NI})$ is increasing in the direction indicated by the arrow. The bold, solid contour indicates the locus of points along which integration is consumer surplus neutral. Consumer surplus increases (decreases) with integration everywhere above (below) this curve.

shows the overall effect of integration on consumer surplus (with uniformly distributed search costs, consumer surplus is equal to the number of users). This mirrors the analysis of the exogenous q_0 case in that integration can benefit consumers *despite* an increase in bias. As one might expect in light of the above, users most benefit from integration when the ad nuisance is large (since they then gain relatively more from the fall in ad volumes) and are harmed otherwise.

Observability of the allocation rule. The assumption that $\bar{x}$ is observed may seem strange if one takes the model at face value. Indeed, consumers certainly cannot observe the algorithm used by search engines. Our interpretation of this assumption of the model is, in a broad sense, that a lower quality of results will eventually drive some users away from the search engine, either because they hear about it in the media or they experience it themselves.

However, it is legitimate to ask how important this assumption is. It turns out that without observability of $\bar{x}$, users would be better-off after integration, unless possibly if the publishers are very asymmetric.

To see this, consider first equation (11). Having $\bar{x}$ unobserved by consumers (but still observed

by advertisers) implies that $\frac{dF(\rho^e - \Delta^e)}{d\bar{x}} = 0$, so that

$$\frac{\partial \pi_0^{NI}}{\partial \bar{x}} = F(\rho^e - \Delta^e)q_0 \frac{\partial P_0}{\partial \bar{x}} = F(\rho^e - \Delta^e)q_0 \left(\frac{\partial r_1}{\partial q_0} - \frac{\partial r_2}{\partial q_0} \right).$$

The profit π_0^{NI} is therefore an affine function of $\bar{x}$, so that the search engine will choose either $\bar{x} = 0$ or $\bar{x} = 1$ (or be indifferent to the value of $\bar{x}$ when $\frac{\partial r_1}{\partial q_0} = \frac{\partial r_2}{\partial q_0}$). Unobservability of $\bar{x}$ thus implies that the indirect incentives to bias (IIB) are so powerful as to lead to the complete foreclosure of one of the publishers.

Consider now the integrated case. Whether $\bar{x}$ is observed by consumers does not affect the fact that q_1 is now lower than without integration. The derivative of the search engine's profit is now

$$\frac{\partial \pi_{01}^I}{\partial \bar{x}} = F(\rho^e - \Delta^e)q_0 \left(\frac{\partial r_1}{\partial q_0} - \frac{\partial r_2}{\partial q_0} + q_1 P_1 \right),$$

which, evaluated at (q_1^I, q_2), is larger than $F(\rho^e - \Delta^e)q_0 \left(\frac{\partial r_1}{\partial q_0} - \frac{\partial r_2}{\partial q_0} \right)$ evaluated at (q_1^{NI}, q_2) because $q_1^{NI} \geq q_1^I$.

Therefore we must have $\bar{x}^I \geq \bar{x}^{NI}$ and $(\bar{x}^I, \bar{x}^{NI}) \in \{0; 1\}^2$. So we have three cases:

1. $\bar{x}^I = 0, \bar{x}^{NI} = 0$,

2. $\bar{x}^I = 1, \bar{x}^{NI} = 0$,

3. $\bar{x}^I = 1, \bar{x}^{NI} = 1$.

In cases 1 and 3, the equilibrium allocation rule is not affected by integration, so that consumers clearly benefit from integration. In case 2, integration leads to a switch from a situation in which all users are directed towards publisher 2 with an ad level $q_2^{NI} = q_2^I$ to a situation in which publisher 1 receives all the traffic and displays q_1^I ads. For small levels of asymmetry between publishers 1 and 2, the positive effect of seeing fewer ads will be the dominant effect, and consumers will be better-off. The only case in which integration could be bad for consumers would be a situation in which publisher 1 displays more ads than publisher 2 even after integration, which can only happen when there is a large asymmetry in terms of the ad-revenue functions r_1 and r_2.