

Integration and Search Engine Bias ^{*}

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

Abstract

We study the effects of integration between a search engine and a publisher. In a model in which the search engine (i) allocates users across publishers and (ii) competes with publishers to attract advertisers, we find that the search engine is biased against publishers that display many ads—even without integration. Integration can (but need not) lead to own-content bias. It can also benefit consumers by reducing the nuisance costs due to excessive advertising. Advertisers are more likely to suffer from integration than consumers. On net, the welfare effects of integration are ambiguous.

— Accepted for publication in the *RAND Journal of Economics*, 2014. —

^{*}We are grateful to the editor, Benjamin Hermalin, for helping us to significantly improve the exposition of this article. We also benefited from useful comments from and discussions with Mark Armstrong, Bernard Caillaud, Jacques Crémer, Gabrielle Demange, Matthew Ellman, Andrei Hagiu, Bruno Jullien, Jens Prüfer, Régis Renault, Chris Snyder, Wing Man Wynne Lam, anonymous referees, as well as seminar participants at Cambridge, ENSAE-Crest, Manchester, and Oxford, and participants at RES 2013 (Royal Holloway), the 4th Workshop on the Economics of ICTs (Evora, 2013), the 2013 TILEC workshop on media regulation (Tilburg), the 7th bi-annual conference on The Economics of Intellectual Property, Software and the Internet (Toulouse, 2013), the Fourth Annual Searle Center Conference on Internet Search and Innovation (Northwestern, 2013), the 11th Media Economics Workshop (Tel-Aviv 2013) and CRESSE 2013 (Corfu). The first author benefited from funding from the Agence Nationale de la Recherche (ANR 2010-BLANC-1801 01).

[†]Department of Economics and Nuffield College, University of Oxford; adecorniere@gmail.com.

[‡]Oxford Internet Institute, University of Oxford; greg.taylor@oii.ox.ac.uk.

1 Introduction

The business model of most search engines relies on two pillars: organic results and paid results (or sponsored links). The main difference between the two is the search engine receives no payments for organic results, whereas sponsored links are bought by advertisers. While organic results receive the majority of clicks,¹ paid results have nevertheless generated billions of dollars of annual advertising revenue.

As crucial on-ramps to the information superhighway, it is not surprising that search engines have attracted scrutiny from the world's competition authorities. Google, reflecting its dominant position (market share of over 90% in Europe, 65% in the US), has been especially scrutinized.² A principal concern of competition authorities is that Google, which owns many specialized publishers (for example, Google maps, Youtube, and Zagat), could display organic results so as to favor these publishers, to the detriment of competitors and, ultimately, consumers. For instance, in 2013, the US Federal Trade Commission concluded that Google has introduced changes to its search algorithm that privileges its own content—amounting to what is “commonly known as search bias” (FTC, 2013). Despite this observation, the FTC determined that Google's practices likely improve consumers' experience 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 about Google's dominant position—has been more hawkish and is negotiating a settlement 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 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

¹94% of clicks versus 6% for sponsored links, for Google and Bing, according to a study by eConsultancy.

²Manne and Wright (2007) provide a legal analysis of the antitrust issues facing Google.

³See also Tarantino (2013) for a discussion of the issue of foreclosure in the vertical search engine industry, as

Reflecting this debate, this article considers the broader question of the overall implications that integration between a dominant search engine and a publisher has on the advertising market. To do so, we consider a monopoly search engine. Based on the query it receives, the search engine can direct a consumer towards one of two publishers. A mapping from a query to a result is an allocation rule. Both publishers, as well as the search engine, are strictly ad-financed; and advertisers view these three media platforms as imperfect substitutes. Consumers dislike ads, but are unaware of the number of ads a site displays prior to visiting it.⁴

We consider three cases. One is no integration. Because publishers receive no payments from users, they display more ads than would maximize consumer surplus. The provision of advertising is nevertheless insufficient from advertisers' point of view due to the publishers' market power. The search engine cares about two things: the number of users that it attracts and the price of its sponsored links. The former is maximised when results are unbiased. Biasing against a site with many ads, though, reduces the effective ad supply and thus increases the price of sponsored links. Bias can therefore arise absent integration.

In the second case, the search engine and one publisher enter into a contractual arrangement under which the search engine displays third-party advertisements on the publisher's site and keeps a share of the revenues, with the publisher remaining free to set the quantity of ads. We call this partial integration (without control). Partial integration creates an incentive for the search engine to direct users towards that specific publisher. It can, nevertheless, lead to a reduction in bias when user participation is very sensitive to the quality of the results because it increases the search engine's marginal return to attracting additional users.

In the third case, the search engine is fully integrated with a publisher; hence, the search engine decides the amount of advertising on the publisher's site. Because advertisers view the search engine and publishers as (imperfect) substitutes, the merged structure lowers the supply of advertising space. As a result, the integrated publisher becomes more attractive to consumers. The effects of integration on consumers' welfare are, therefore, ambiguous in general. Lastly, we

well as for some evidence.

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

derive some policy implications of the model.

Related literature

The literature on search engines has focused on the auctions used to sell ads (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 whether search engines will wish to maximize result quality. Examples are de Cornière (2013), Eliaz and Spiegler (2011), Taylor (2013), and White (2013). Unlike these articles, 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 article, 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. In contrast to their work, we abstract from the mechanism by which ads are sold, but endogenize publishers' ad supply decisions. This allows us to endogenize the relative attractiveness of publishers to consumers, which has a variety of interesting theoretical and practical implications.

We capture bias as the search engine's decision to direct some consumers away from their ideal publisher. In this sense, our work is similar to Hagi and Jullien (2011), Armstrong and Zhou (2011), and Inderst and Ottaviani (2012), who consider intermediaries' incentives to bias their advice in favor of firms from which they receive larger payments. The structure of our model is also reminiscent of some work on network neutrality (Choi and Kim (2010) for instance), in which "bias" takes the form of priority rules in favor of one publisher. Our article differs from this literature in two main 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 publishers to sell advertisements. The latter point introduces consumer multihoming and, as such, 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, because, in our model, users visit the search engine and a publisher, advertisers can reach them on both platforms. As in those articles, multihoming reduces the ability of each platform to extract revenue from advertisers. The novelty in our approach is that the search engine influences which publisher consumers visit; it can therefore restrict alternative opportunities for reaching consumers by directing them to a publisher with fewer ads.

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 how the lack of coordination between the search engine and publishers can result in an excessive level and an inappropriate allocation of advertising across websites.

2 The model

There are three platforms (indexed by i): a search engine ($i = 0$) and two publishers ($i = 1, 2$). Assume a continuum of consumers of mass one and a representative advertiser.

The search for content

The publishers produce horizontally differentiated content. They are located at the ends of a Hotelling line: publisher 1 at the location zero, publisher 2 at one.

Consumers are uniformly distributed along the Hotelling line. A consumer's utility from consuming a publisher's content is a decreasing function of the distance between her position, $x \in [0, 1]$, and the publisher's. If the two are distance ℓ apart then the publisher's content is worth $\theta - c(\ell)$ to the consumer, where c is a transportation cost ($c' > 0$) and θ is a consumer-specific intrinsic benefit from content. Assume θ is distributed according to a strictly increasing distribution function, F , with associated density f , and is independent from the consumer's location.

Information is incomplete: consumers know their tastes, but are ignorant of the mapping

between their tastes and their position on the line; hence they are unable to identify which is the more relevant publisher. Consequently, the only way for consumers to access content is via the search engine. The search engine knows a consumer’s location on the Hotelling line, and can direct him to one of the two publishers accordingly. Call the search engine’s referral an allocation rule. We focus on allocation rules with a threshold structure:

Definition 1. *A search engine’s allocation rule is a threshold \bar{x} : users with $x < \bar{x}$ are directed to publisher 1 and the rest to publisher 2.*

We allow the search engine to choose any \bar{x} in $[0, 1]$; hence it can perfectly allocate traffic between the publishers. This simplifying assumption is broadly consistent with results from Edelman and Lai (2013). They 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.⁵

Advertising

Both publishers and the search engine are financed exclusively through advertising. The quantity of advertisements displayed on website i is denoted q_i . Call ads on publishers 1 and 2 *banners*, and ads on the search engine *sponsored links*.⁶ We assume that the quantity of sponsored links is exogenously fixed at q_0 . This permits us to focus on the allocation rule as the search engine’s principal strategic tool. The case of an endogenous q_0 , for which the main results still hold, is relegated to the appendix.

There are no monetary transfers between consumers and publishers, but consumers dislike ads. Denote the disutility from ads by $\delta_i(q_i)$.^{7,8} The utility of a user who visits publisher i , which is at

⁵One 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.

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

⁷See section 6 for a discussion of the case of ad-loving consumers.

⁸because q_0 is exogenous, it does not matter whether we write $\delta_i(q_0, q_i)$ or $\delta_i(q_i)$.

a distance ℓ from his ideal location, is therefore

$$u(\theta, \ell, q_i) = \theta - c(\ell) - \delta_i(q_i).$$

On the other side of the market are advertisers who wish to reach users. For simplicity, we consider 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 $\mathbf{q} = (q_0, q_1, q_2)$ ads is

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

with $r_0(0) = r_1(q_0, 0) = r_2(q_0, 0) = 0$. The r_0 function captures the revenue from reaching users on the search engine, while r_i is the additional revenue from reaching them on publisher i .⁹

As suggested by our notation, r_i depends on both q_0 and q_i for $i \in \{1, 2\}$. However, because a consumer doesn't visit both publishers, r_1 does not depend on q_2 , and *vice versa*. Notice too that r_1 and r_2 do not depend on \bar{x} : a user's location is independent of his attitude towards advertisers.¹⁰

The key assumption of our model is that advertisers view ads on the search engine and the publishers as imperfect substitutes. Substitutability means that the marginal value of one type of ads decreases as the number of the other type increases that is,

$$\frac{\partial^2 r_i(q_0, q_i)}{\partial q_0 \partial q_i} \leq 0 \tag{1}$$

for $i = 1, 2$. There is, we note, disagreement about substitutability: Hahn and Singer (2008) provide survey evidence that such ads are strict substitutes, while the FTC (2007) took the position that the two ad markets are independent. We allow for both possibilities.

Because $r_i(q_0, 0) = 0$, (1) implies that $\frac{\partial r_i(q_0, q_i)}{\partial q_0} \leq 0$: holding constant the ads on publisher i , more ads on the search engine mean advertiser revenues from ads on a publisher fall.

⁹In Appendix A we provide microfoundations for this revenue function, as well as an alternative interpretation with multiple heterogeneous advertisers buying at most one slot.

¹⁰This assumption is discussed in Appendix A.

Assume r_i is concave in q_i and, to ensure interior solutions, that $\frac{\partial r_i}{\partial q_i}|_{q_i=0} > 0$. Concavity reflects diminishing marginal revenue from advertising.

Payments to publishers are made on a per-impression basis: the advertiser’s expected per-user profit is

$$\pi_a = R(\mathbf{q}, \bar{x}) - q_0 P_0 - \bar{x} q_1 P_1 - (1 - \bar{x}) q_2 P_2, \quad (2)$$

where P_i is the price of an ad on platform i .¹¹ Given the quantity of ads offered by publishers and the search engine, prices adjust to clear the market (similar to a Cournot model).

Figure 1 represents the market structure.

A remark on sponsored links: we limit attention to the case in which publishers obtain traffic through organic links only. Allowing publishers to buy sponsored links by participating in an ad auction would not significantly affect our results. To see this, assume that 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 $\alpha q_i P_i$ (because 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. If $\bar{x} \in (0, 1)$ then the expected relevance of the organic link, on the other hand, is strictly better than random, and so the user optimally sets $\alpha = 1$. In other words, *publishers’* sponsored links are never clicked on in equilibrium.

Information and timing

The timing of the game is:

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

1. The publishers simultaneously choose their advertising levels. The search engine chooses an allocation rule, \bar{x} .¹²
2. Advertisers observe \bar{x} and \mathbf{q} . The advertising market clears: P_0 , P_1 , and P_2 equalize supply and demand for each platform.
3. Consumers learn \bar{x} . They decide whether to use the search engine by comparing their expected surplus with their outside option, normalized to zero.
4. Consumers search.

The assumption that consumers do not observe q_1 and q_2 is consistent with the assumption that they have no prior knowledge of the publishers' content. Assuming that consumers know \bar{x} may seem strange, but is meant to capture the idea that consumers know about the search engine's quality. For completeness, we analyze the case in which \bar{x} is not observed by consumers in the appendix.

We solve the model by backward induction.

3 Consumer participation and ad market clearing

Consumer participation

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

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

¹²One could find arguments in favor of a different timing, either to have \bar{x} chosen before the q_i 's or after. In fact, search engines and publishers appear to be stuck in a kind of Red Queen's race in which neither side is able to definitely act as a leader. Simultaneous actions also turn out to be more tractable.

For a given allocation rule \bar{x} , the expected utility (gross of nuisance costs) for a consumer of type θ in searching is $\theta + \rho(\bar{x})$, where

$$\rho(\bar{x}) \equiv -E[c(\ell)|\bar{x}] = -\int_0^{\bar{x}} c(\ell)d\ell - \int_{\bar{x}}^1 c(1-\ell)d\ell$$

is the *relevance* of the search engine. Its second derivative is $\rho''(\bar{x}) = -c'(\bar{x}) - c'(1-\bar{x}) < 0$, 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

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

A consumer will engage in search if and only if $U(\rho, \Delta^e, \theta) \geq 0$, which gives us the participation rate $1 - F(\Delta^e - \rho)$.

Definition of bias

The expression for user participation makes it clear that consumers care about two things: they want relevant content and as few ads as possible. What is the optimal allocation rule from consumers' point of view?

Definition 2. For a given vector of ads, \mathbf{q}^e , the user-optimal allocation rule is defined as

$$x^*(\mathbf{q}^e) \equiv \operatorname{argmax}_{\bar{x}} \rho - \Delta^e.$$

Notice that x^* is uniquely defined because $\rho - \Delta^e$ is a strictly concave function of \bar{x} . Because users dislike ads:

Lemma 1. The cutoff, x^* , is decreasing in q_1^e and increasing in q_2^e .

Proof. The first-order condition defining x^* is

$$\rho'(\bar{x}) - \frac{\partial \Delta^e}{\partial \bar{x}} = 0;$$

hence, $\rho'(x^*) = \delta_1(q_0, q_1^e) - \delta_2(q_0, q_2^e)$. The right-hand side is increasing (decreasing) in q_1 (in q_2), and thus so must the-left hand side. That $\rho''(\bar{x}) < 0$ completes the proof. ■

Example 1. *Suppose a linear transport cost, t , and linear ad nuisance costs:*

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

which yields

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

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

Ad market clearing

The price of advertising adjusts to clear the market. Given the observed allocation rule \bar{x} and a price vector (P_0, P_1, P_2) , the advertiser's profit is given by (2). The advertiser is indifferent between buying and not buying the marginal ad on website i if and only if its marginal revenue ($\frac{\partial R(\mathbf{q}, \bar{x})}{\partial q_i}$) is equal to the price of the slot, P_i . The corresponding demand for advertising, (q_0^d, q_1^d, q_2^d) , is thus given by

$$P_0 = \frac{\partial r_0(q_0^d)}{\partial q_0} + \bar{x} \frac{\partial r_1(q_0^d, q_1^d)}{\partial q_0} + (1-\bar{x}) \frac{\partial r_2(q_0^d, q_2^d)}{\partial q_0}, \quad (3)$$

$$P_1 = \frac{\partial r_1(q_0^d, q_1^d)}{\partial q_1}, \quad (4)$$

$$P_2 = \frac{\partial r_2(q_0^d, q_2^d)}{\partial q_2}. \quad (5)$$

This gives us the inverse demand functions $P_0(\mathbf{q}, \bar{x})$, $P_1(\mathbf{q})$ and $P_2(\mathbf{q})$.

From (1), 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). *Suppose*

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

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

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 example, 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. When we use this linear specification later in the article, 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

Suppose no integration—indicated by an NI superscript. The profits of the three websites are

$$\begin{aligned}\pi_0^{NI} &= [1 - F(\Delta^e - \rho)] q_0 P_0, \\ \pi_1^{NI} &= \bar{x} [1 - F(\Delta^e - \rho)] q_1 P_1, \\ \pi_2^{NI} &= (1 - \bar{x}) [1 - F(\Delta^e - \rho)] q_2 P_2.\end{aligned}$$

Because ad levels are unobserved, the number of active users, $1 - F(\Delta^e - \rho)$, 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 are thus

$$\frac{\partial \pi_i^{NI}}{\partial q_i} = 0 \iff \frac{\partial P_i}{\partial q_i} q_i + P_i = 0, \quad (6)$$

which define the equilibrium values of q_1^{NI} and q_2^{NI} .

The search engine chooses its allocation rule, \bar{x} , as a best response to the publishers' ad volumes, as defined by (6). The first-order condition for its choice \bar{x} is thus

$$0 = \frac{\partial \pi_0^{NI}}{\partial \bar{x}} = \underbrace{f(\Delta^e - \rho) \left(\rho'(\bar{x}) - \frac{\partial \Delta}{\partial \bar{x}} \right) q_0 P_0}_{\text{RQ}} + \underbrace{[1 - F(\Delta^e - \rho)] q_0 \frac{\partial P_0}{\partial \bar{x}}}_{\text{IIB}}. \quad (7)$$

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

Second, users are willing to tradeoff relevance for a reduction in ads. The term $-f(\Delta^e - \rho)\frac{\partial \Delta}{\partial \bar{x}}$ in (7) reflects this. The desire to attract users explains why Google, for example, publically announces measures to penalize advertising spam sites in its ranking algorithm.¹³ These two effects together constitute the *quality* of the results provided by the search engine. Quality is maximal when $\frac{d[1-F(\Delta^e-\rho)]}{d\bar{x}} = 0$, i.e. at $\bar{x} = x^*$. The first term on the right-hand side of (7) thus measures the

¹³See, for example, <http://insidesearch.blogspot.co.uk/2012/01/page-layout-algorithm-improvement.html>, accessed 19th October 2012.

Return to Quality (RQ) for the search engine.

The third effect, the term labeled IIB, reflects the more subtle point that the optimal bias from the search engine’s point of view is partially determined by strategic considerations concerning the advertising market. That is, the search engine also uses \bar{x} as an instrument to maintain a high price, P_0 , for its own sponsored links. From (3), 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 $\frac{\partial r_i}{\partial q_0}$ is the effect of the marginal sponsored link on the revenue from advertising at i . In order to maintain a high price for its ads, the search engine has an incentive to bias results towards the publisher i for which $\frac{\partial r_i}{\partial q_0}$ is less negative. Because $\frac{\partial^2 r_i}{\partial q_0 \partial q_i} \leq 0$ (see Equation 1), 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.¹⁴ The strength of this effect also depends upon the degree of substitutability between sponsored links and banners, $\left| \frac{\partial^2 r_i}{\partial q_0 \partial q_i} \right|$: if it is large, so that i ’s ads are particularly close substitutes for sponsored links, then the search engine has 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 publisher i if $\frac{\partial r_j}{\partial q_0} > \frac{\partial r_i}{\partial q_0}$.*

For instance, 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 (7) is satisfied for $\frac{dF}{d\bar{x}} < 0$; hence $\bar{x} > x^*$. Observe the bias against publisher 2 not only depends on the amount of advertising on each publisher, but also on relative substitutability, γ_1 and γ_2 . In particular, when $\gamma_1 = \gamma_2$, the search engine puts too much weight on the level of ads compared to consumers’ ideal.

¹⁴Observers 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.

If, in Example 1, websites only differ in their degree of substitutability γ_i , (6) implies

$$q_i = \frac{\alpha - \gamma_i q_0}{2\beta}.$$

Thus, the condition $\gamma_2 q_2 > \gamma_1 q_1$ becomes $(\gamma_1 - \gamma_2)[q_0(\gamma_1 + \gamma_2) - \alpha] > 0$. If α is small relative to $q_0(\gamma_1 + \gamma_2)$, then publishers' supply of advertising space is very elastic with respect to γ_i , so that the search engine is biased *in favor* of its closest substitute (which shows fewer ads). If α is relatively large, the elasticity of q_i with respect to γ_i is low, and the search engine is biased *against* its closest substitute (whose ads exert a stronger externality on P_0).

It also follows from Proposition 1 that the search engine implements 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: $\bar{x}^{NI} = x^* = 1/2$.*

Proof. From (4), (5), and (6), both publishers will choose the same ad level: $r_1(q_0, \cdot) = r_2(q_0, \cdot) \Rightarrow P_1(q_0, \cdot) = P_2(q_0, \cdot) \Rightarrow q_1^{NI} = q_2^{NI}$. Hence, $\partial r_1(q_0, q_1^{NI})/\partial q_0 = \partial r_2(q_0, q_2^{NI})/\partial q_0$, so that the second term in (7) is zero. Consequently, $d[1 - F(\Delta^e - \rho)]/d\bar{x} = 0$, thus $\bar{x}^{NI} = x^*(\mathbf{q})$. If $\delta_1(q_0, \cdot) = \delta_2(q_0, \cdot)$ and $q_1 = q_2$, then $\frac{d[1 - F(\Delta^e - \rho)]}{d\bar{x}} = f(\Delta^e - \rho)\rho'(\bar{x})$, which equals zero when $\bar{x} = 1/2$. ■

In the unintegrated case, an incentive for bias derives 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 than the other.

Similarly, in the special case of no substitutability, the IIB effect has no bite and the unintegrated search engine will therefore maximize its quality; hence zero bias.

Corollary 2. *If ad demands are independent (if (1) holds with equality) then, in the non-integrated market structure, the search engine implements the user-optimal allocation rule, $\bar{x}^{NI} = x^*$, and there is no bias.*

Proof. Integrating (1), 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 IIB effect can be neglected. Equation (7) thus reduces to $d[1 - F(\Delta^e - \rho)]/d\bar{x} = 0$ so that the search engine’s FOC demands $\bar{x}^{NI} = x^*(\mathbf{q})$. ■

5 Integration

Partial integration without control

Next we consider the intermediate situation in which publisher 1 shares a fraction ϕ_1 of its revenue with the search engine but retains full-autonomy—call this partial integration, and denote it with a *PI* superscript. This situation is relevant because Google is vertically integrated into the ad platform industry and sells ad technology services to otherwise independent publishers through subsidiaries such as DoubleClick.¹⁵ Thus, the search engine is able to appropriate a share of advertising revenues generated on such publishers.

Now, site 1’s profit is $\pi_1 = (1 - \phi_1)[1 - F(\Delta^e - \rho)]\bar{x}q_1P_1$. Because ϕ_1 is effectively a profit tax, site 1’s quantity decision is unaffected; hence, $q_1^{PI} = q_1^{NI}$. Consequently, x^* is invariant to ϕ_1 : consumers’ preferences over publishers are not affected by partial integration. The search engine’s profits are now

$$\pi_0^{PI} = [1 - F(\Delta^e - \rho)][q_0P_0 + \phi_1\bar{x}q_1P_1];$$

hence, profit maximization entails choosing \bar{x} to solve

$$\frac{\partial \pi_0}{\partial \bar{x}} = \frac{\partial [1 - F(\Delta^e - \rho)]}{\partial \bar{x}} \left[q_0P_0 + \underbrace{\phi_1\bar{x}q_1P_1}_{\text{IRQ}} \right] + [1 - F(\Delta^e - \rho)] \left[q_0 \frac{\partial P_0}{\partial \bar{x}} + \underbrace{\phi_1q_1P_1}_{\text{DIB}} \right] = 0. \quad (8)$$

Relative to the case of no integration, the search engine cares even more about the number of users it attracts because each user is associated with an additional revenue stream (an *Increased Return to Quality*, IRQ). All else equal, this effect compels the search engine to implement higher quality

¹⁵DoubleClick 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.

(less biased) results the higher is its revenue share from the content-bearing site. There is also a second effect: the search engine has a *Direct Incentive to Bias* its results in favor of site 1 (term DIB).

To determine how partial integration affects search engine bias, we consider two cases.

Case 1: $\bar{x}^{NI} > x^*$, the non-integrated search engine is biased towards publisher 1. Partial integration cannot then lead the search engine to set $\bar{x}^{PI} < x^*$ because a deviation to $\bar{x}^{PI} = x^*$ would increase both the number of users and the per-user revenue. Therefore, partial integration leads to less bias if $\bar{x}^{PI} < \bar{x}^{NI}$, and to more if $\bar{x}^{PI} > \bar{x}^{NI}$.

In the non-integrated case, the optimal allocation rule is characterized 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 increases if $\partial^2\pi_0/\partial\bar{x}\partial\phi_1 > 0$. Moreover, from (8), we know that

$$\frac{\partial^2\pi_0}{\partial\bar{x}\partial\phi_1} = q_1P_1 \left(\bar{x} \frac{\partial[1 - F(\Delta^e - \rho)]}{\partial\bar{x}} + [1 - F(\Delta^e - \rho)] \right) > 0 \Leftrightarrow \eta_{\bar{x}} < 1,$$

where $\eta_{\bar{x}} \equiv -\frac{\partial[1 - F(\Delta^e - \rho)]/\partial\bar{x}}{[1 - F(\Delta^e - \rho)]/\bar{x}}$ is the elasticity of user participation with respect to the results quality.

Case 2: $\bar{x}^{NI} < x^*$, the non-integrated search engine is biased against publisher 1. Partial integration then necessarily leads to more traffic directed to publisher 1 (i.e., to an increase in \bar{x}). Because publisher 1 received too little traffic prior to integration, the increase in traffic to site 1, if not too large, may work to reduce bias. If the increase in traffic to site 1 is too large, bias changes sign and increases in magnitude. Because the IRQ term in (8) creates incentives to reduce bias, a sufficient condition for bias to decrease is that the second term in (8) becomes smaller: $|q_0 \frac{\partial P_0}{\partial \bar{x}}| > |q_0 \frac{\partial P_0}{\partial \bar{x}} + \phi_1 q_1 P_1|$; that is

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

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 yields:

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 is less than it is absent integration,*
2. *if $\bar{x}^{NI} > x^*$ and $\eta_{\bar{x}} < 1$, the level of bias is more than it is absent integration,*
3. *if $\bar{x}^{NI} < x^*$ and $\phi_1 < \phi^*$, the level of bias is less than it is absent integration.*

Point 3 follows a different logic than 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 offsets 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 the search engine gains 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: (8) becomes

$$\frac{\partial \pi_0}{\partial \bar{x}} = \frac{\partial [1 - F(\Delta^e - \rho)]}{\partial \bar{x}} \left[q_0 P_0 + \underbrace{\phi_1 \bar{x} q_1 P_1 + \phi_2 (1 - \bar{x}) q_2 P_2}_{IRQ} \right] + [1 - F(\Delta^e - \rho)] \left[q_0 \frac{\partial P_0}{\partial \bar{x}} + \underbrace{\phi_1 q_1 P_1 - \phi_2 q_2 P_2}_{DIB} \right] = 0.$$

Relative to the case with $\phi_2 = 0$, we see, firstly, that the increased returns to quality effect becomes stronger 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 causes bias to decrease.

Full integration

Suppose, now, that the search engine is fully integrated with publisher 1—a case we denote with an *FI* superscript. This corresponds to setting $\phi_1 = 1$ and allowing the search engine to determine q_1 .

Effects of integration on advertising levels

The integrated firm's profit is

$$\pi_{01}^{FI} = [1 - F(\Delta^e - \rho)] (q_0 P_0 + \bar{x} q_1 P_1).$$

Recall our assumption that q_0 is fixed exogenously; hence, the integrated firm's only decision vis-à-vis advertising is q_1 . The profit-maximizing number of ads displayed on publisher 1 is given by the first-order condition:

$$\frac{\partial \pi_{01}^{FI}}{\partial q_1} = \bar{x} \left(\frac{\partial P_1}{\partial q_1} q_1 + P_1 \right) + \underbrace{q_0 \frac{\partial P_0}{\partial q_1}}_{\text{CPE}} = 0. \quad (9)$$

The standard logic of horizontal mergers applies in this case. Because publisher 1 now internalizes the effect of its advertising supply 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 6). 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: $q_1^{FI} \leq q_1^{NI}$. Publisher 2 offers the same quantity of ads*

under both regimes: $q_2^{FI} = q_2^{NI}$.¹⁶

For example, under the specification of Example 1, using (6) and (9), the ad quantities are¹⁷

$$q_1^{FI} = \frac{\alpha_1 - 2\gamma_1 q_0}{2\beta_1} < q_1^{NI} = \frac{\alpha_1 - \gamma_1 q_0}{2\beta_1}; \quad q_2^{FI} = q_2^{NI} = \frac{\alpha_2 - \gamma_2 q_0}{2\beta_2}. \quad (10)$$

Ignoring—for now—the changes in the allocation rule \bar{x} , we can make two remarks about the effects of integration. First, it ameliorates the negative externality that advertisers exert on users, by restricting the supply of advertising space. This effect derives from the improved market power of the integrated structure with respect to advertisers. 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:

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

In other words, even if integration causes more traffic to be directed to publisher 1, users are not necessarily harmed because they now face fewer ads on the publisher they visit. Thus a search engine that directs more traffic to its own content 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.

¹⁶ That $q_1^{FI} \leq q_1^{NI}$ depends on our assumption of linear pricing. With two-part tariffs, integration would have no effect on the quantity of ads. (See Ambrus, Calvano, and Reisinger (2012) for more on this “neutrality” result). In our view, this assumption fits well how advertisements are sold online.

¹⁷In this simple example, (9) 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^{FI} , like q_1^{NI} , does not depend on \bar{x} .

Effects of integration on the allocation of traffic

The effect of integration on the search engine’s allocation decision parallels the partial integration case: the first-order condition can be found from (8) by setting $\phi_1 = 1$. Thus, as in the partial integration model, the first term in (8) reflects the additional weight that the search engine puts on attracting users with high quality results—an effect that unambiguously serves to reduce bias post-integration. In particular, we have that

Proposition 4. *Integration can lead to a reduction in bias, $|x^* - \bar{x}|$.*

Figure 2 illustrates the change in bias implied by integration for a particular parametrisation of Example 1. Relative to the no-integration benchmark, 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 incentive to undo this bias; hence, 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 it has little to gain from further increases in \bar{x} . That x^* increases after integration means that the overall level of bias falls.

Because 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, we again restrict attention to a symmetric setting (see Definition 4). From Corollary 1, symmetry implies that there is no bias in the non-integrated case; Figure 2 suggests that this is no longer true under integration.

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

Proof. The last two equalities follow from Corollary 1, and the second inequality from Corollary 3. It remains to prove the first one. To do so, take the first-order condition (8) evaluated at

$x^*(\mathbf{q}^{FI})$. By definition, $\frac{d[1-F(\Delta^e-\rho)]}{d\bar{x}}|_{\bar{x}=x^*} = 0$. Because $r_1(q_0, \cdot) = r_2(q_0, \cdot)$, we also have $\frac{\partial P_0}{\partial \bar{x}} = 0$, so that $\frac{\partial \pi_{01}^{FI}}{\partial \bar{x}}|_{\bar{x}=x^*} > 0$. The optimal value of \bar{x} for the search engine is thus larger than x^* . ■

Under symmetry, integration between the search engine and a publisher leads to biased results in favor of that publisher, but also to a lower level of advertising, which benefits users.

6 Policy implications

The search engine has a bias toward own content following integration, but, as we showed, two additional effects temper that effect. The first is the increased return to attracting an extra user (the IRQ term in (8)), which can lead to better quality results if bias initially exists, and otherwise softens the incentive to introduce bias. The second is the reduction in the advertising distraction incurred by users (the CPE in (9)). The magnitude of these effects is an empirical matter. However, the model may offer some insight as to when different effects should dominate. In this section, we focus on three important parameters of the model, namely the substitutability between sponsored links and banners, 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.

Substitutability between sponsored links and banners

The degree of substitutability is important for two reasons. First, the extent to which advertisers view search engines and other publishers as substitutes helps to determine the relevant market. Second, the degree of substitutability affects the magnitude of the change in ad levels between no-integration and integration (see the CPE in (9)). In the extreme case of no substitutability ($\frac{\partial P_0}{\partial q_1} = 0$), the last term of (9) disappears and, thus, users do not benefit from fewer ads after integration. In this case, equation (8) shows that integration biases the search engine in favor of publisher 1 (recall, from Corollary 2, that there is no bias in the unintegrated market when ad demands are independent). Because 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. In sum,

Proposition 5. *Under symmetry, when there is no ad substitutability between sponsored links and banners, integration reduces consumers’ and advertisers’ surplus.*

Under symmetry, some degree of substitutability between display and search ads (so that (1) holds as a strict inequality) is, therefore, a necessary condition for either advertisers or consumers to benefit from integration. It is not, however, sufficient, as we show below. With imperfect substitutability, the results are more ambiguous because users benefit from fewer ads but could suffer from biased results. Example 1 provides the relevant intuition. In this example, equilibrium ad quantities are given by (10), $\bar{x}^{NI} = 1/2$ (by Corollary 1), and \bar{x}^{FI} can be determined using (8). Figure 3 shows how the size of the user base (or, equivalently, the expected utility of a user) is affected by integration for a particular set of parameter values. Integration is more likely to be beneficial when sponsored links and banners are close substitutes because this increases the competitive externality between the search engine and publisher 1. Consumers can then expect to benefit from a larger fall in ad volumes, all the more so when the disutility from ads, δ , is large.

Consumers’ attitude towards advertising

Figure 3 also shows how, all else equal, the greater the nuisance cost of advertisements, the more attractive integration is for users and society at large: ad volumes fall after integration and consumers benefit more from this when δ is large.

So far, advertising has been considered a pure nuisance for users. Although that assumption is standard in the literature (see, for example, Anderson and Coate, 2005), users may benefit from advertising in some contexts: for example, if the publishers are business directories (online “yellow pages”). It turns out that our analysis is amenable to positive externalities from advertising—that is, to the assumption that $\delta_i(\cdot)$ is non-increasing in q_j for every i, j .

In this case, if we assume symmetry between publishers, the effects of integration are 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.*

Content differentiation

The degree to which the content of publishers 1 and 2 differs is important for assessing the desirability of integration. To see this, suppose that the publishers offered the same content. Formally, this corresponds to $\rho'(\bar{x}) = 0$. Given consumer indifference to the publishers, integration leads to publisher 2's being completely foreclosed, that is $\bar{x} = 1$. Because the advertising level is lower on publisher 1 after integration (Proposition 3), consumers strictly benefit from integration.

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

Proof. Symmetry and Proposition 3 imply that $q_1^{FI} < q_2^{FI} = q_i^{NI}$. Two properties follow: (i) all consumers prefer site 1 under integration in equilibrium, so that $d[1 - F(\Delta^e - \rho)]/d\bar{x} > 0$; and (ii) $\partial P_0/\partial \bar{x} > 0$. From (8) (with $\phi_1 = 1$), the integrated firm's profit is increasing in \bar{x} :

$$\frac{\partial \pi_{01}^{FI}}{\partial \bar{x}} = f(\Delta^e - \rho) \left[\underbrace{\rho'(\bar{x})}_{=0} - \underbrace{\frac{d\Delta}{d\bar{x}}}_{<0} \right] (q_0 P_0 + \bar{x} q_1 P_1) + [1 - F(\Delta^e - \rho)] \left(\underbrace{q_0 \frac{\partial P_0}{\partial \bar{x}}}_{>0} + q_1 P_1 \right) > 0,$$

implying $\bar{x} = 1$ in equilibrium. The change in user utility from integration is $U(\rho(1), \Delta(\mathbf{q}^{FI})) - U(\rho(\bar{x}^{NI}), \Delta(\bar{x}, \mathbf{q}^{NI}))$, which is positive because $\rho(1) = \rho(\bar{x}^{NI})$ and $\Delta(\bar{x}, \mathbf{q}^{NI}) > \Delta(\bar{x}, \mathbf{q}^{FI})$. ■

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

As the degree of content differentiation goes up, two opposite forces are relevant: on the one hand, *for a given level of bias*, consumers lose from the increase in bias. On the other hand,

the increased return to quality (IRQ) effect causes the integrated search engine to reduce 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 .

7 Conclusion

The central importance of search engines to online commerce has made them the focus of recent antitrust scrutiny; of particular concern has been the issue of own-content bias. Our model delivers several insights relevant to this debate. 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 the search engine's own sponsored links.

Secondly, if the search engine integrates with a publisher, then this creates *direct* monetary incentives for the search engine to channel traffic towards its own content. We show, however, that these new incentives need not increase the level of bias in search results: (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; and (iii) the integrated publisher has an incentive to lower its supply of advertising space, which benefits users and justifies making that publisher's site more prominent.

The above notwithstanding, search engines' and users' interests are not perfectly aligned: a search engine's natural incentive to favor its own sites can lead to 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 and users have a strong 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.

Our article, of course, does not tell the whole story. In particular, we have ignored possible efficiency gains as well as any dynamics. On the first point, Google has argued that integration will improve the user experience, by making results more easily accessible.¹⁸ 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, our analysis omits future incentives to invest in quality. There are at least two dimensions worth considering. One concerns the search engine's incentives to invest in order to improve its algorithm. The other 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?

We have assumed bias is observable to users as a simple way of capturing the idea that bias will eventually drive away some of a search engine's users. In practice, consumers learn about a search engine's quality over time so that it cannot commit to a result quality ex ante and must instead build a reputation for quality. A potentially interesting line of research is therefore to investigate how reputation dynamics weigh on the incentive (not) to bias.

¹⁸See, for instance, the testimony of Eric Schmidt, Google executive chairman, before the U.S Senate Judiciary Committee in September 2011.

A Microfoundations for advertisers' revenue

In Section 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. A representative advertiser wishes to reach consumers several times. Specifically, assume that showing q ads to a consumer is worth $S(q)$, S increasing, concave, and $S(0) = 0$. For a vector (q_0, q_1, q_2) of ad quantities, a consumer at $x < \bar{x}$ is thus worth $S(q_0 + q_1)$, a consumer at $x > \bar{x}$ worth $S(q_0 + q_2)$. Therefore, the expected revenue derived from a search engine user is

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

It is readily seen 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. There is a continuum of advertisers, each of whom wishes to reach consumers only once. When a consumer sees the ad of a firm, that firm's expected profit is s . Assume that s varies according to the distribution function 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 $s \geq p_1$; that is, $1 - H(p_1)$. The market-clearing price p_1 for an impression to a consumer with $x \leq \bar{x}$ thus satisfies $1 - H(p_1) = q_0 + q_1$. 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 our main conclusions is that consumers can be made better-off by integration, despite an increase in bias, because they suffer fewer ads. Assuming that q_0 is fixed allowed us to show this clearly, because 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 make q_0 endogenous, such a result is 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 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), integration decreases the expected number of ads to which each user will be exposed.

Proposition 8. *Under the assumptions of Example 1, with symmetry between publishers 1 and 2, the expected number of ads seen by a user goes down following integration.*

Proof. In the non-integrated case, q_0 maximizes $q_0 P_0 = q_0(\alpha - \beta q_0 - \gamma(xq_1 + (1 - x)q_2))$, 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 $(q_0^{FI}(\bar{x}), q_1^{FI}(\bar{x}), q_2^{FI}(\bar{x}))$:

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

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

It is then straightforward to show that $m^{NI} - m^{FI}(\bar{x})$ 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} , $m^{NI} > m^{FI}(\bar{x})$. This is true in particular for $\bar{x} = \bar{x}^{FI}$. ■

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 $q_0 P_0$) and attracting many users (maximizing $1 - F(\Delta^e - \rho)$). The first-order condition, in the non-integrated case, is

$$-f(\Delta^e - \rho) \frac{\partial \Delta^e}{\partial q_0} q_0 P_0 + [1 - F(\Delta^e - \rho)] \left(P_0 + q_0 \frac{\partial P_0}{\partial q_0} \right) = 0,$$

whereas, in the integrated case, it is

$$-f(\Delta^e - \rho) \frac{\partial \Delta^e}{\partial q_0} q_0 P_0 + [1 - F(\Delta^e - \rho)] \left(P_0 + q_0 \frac{\partial P_0}{\partial q_0} + \bar{x} q_1 \frac{\partial P_1}{\partial q_0} \right) = 0.$$

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.

Where does the integrated firm host ads? Ads on the search engine itself fall (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 not observed by consumers.¹⁹ 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 (6), whereas the integrated firm must set q_1 to solve (9). Because the CPE term in (9) is negative, and because $q_1^{NI} = q_2^{NI}$ 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 $q_1^{FI} < q_2^{FI}$.*

Thus, by Lemma 1, site 1 becomes relatively more attractive to consumers after integration. This entails an increase in the user-optimal allocation rule, x^* . Although it is, therefore, in consumers’ interest to have additional traffic be directed to the integrated publisher, integration creates an overly strong incentive for own-content bias; 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, whether or not consumers are harmed by integration is ambiguous: there are fewer ads, which reduces the nuisance burden, but bias is introduced into the search results. Figure 7 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 insofar as integration can benefit consumers *despite* an increase

¹⁹Although this behavior is not observed, q_1 is correctly anticipated in equilibrium.

in bias. As one might expect in light of the above, users benefit most from integration when the ad nuisance is large 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. This assumption is meant to capture the idea that low-quality results will eventually drive some users away from the search engine.

However, it is worth considering how important this assumption is. Without observability of \bar{x} , users would be better off after integration, unless possibly if the publishers are very asymmetric.

To see this, consider equation (7). Having \bar{x} unobserved by consumers (but still observed by advertisers) implies that $\frac{d[1-F(\Delta^e-\rho^e)]}{d\bar{x}} = 0$, so that

$$\frac{\partial \pi_0^{NI}}{\partial \bar{x}} = [1 - F(\Delta^e - \rho^e)] q_0 \frac{\partial P_0}{\partial \bar{x}} = [1 - F(\Delta^e - \rho^e)] q_0 \left(\frac{\partial r_1}{\partial q_0} - \frac{\partial r_2}{\partial q_0} \right).$$

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 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

$$\frac{\partial \pi_{01}^{FI}}{\partial \bar{x}} = [1 - F(\Delta^e - \rho^e)] q_0 \left(\frac{\partial r_1}{\partial q_0} - \frac{\partial r_2}{\partial q_0} + q_1 P_1 \right),$$

which, evaluated at (q_1^{FI}, q_2) , is larger than $[1 - F(\Delta^e - \rho^e)] q_0 \left(\frac{\partial r_1}{\partial q_0} - \frac{\partial r_2}{\partial q_0} \right)$ evaluated at (q_1^{NI}, q_2) because $q_1^{NI} \geq q_1^{FI}$.

Therefore we must have $\bar{x}^{FI} \geq \bar{x}^{NI}$ and $(\bar{x}^{FI}, \bar{x}^{NI}) \in \{0, 1\}^2$. So we have three cases: (i) $\bar{x}^{FI} = 0, \bar{x}^{NI} = 0$, (ii) $\bar{x}^{FI} = 1, \bar{x}^{NI} = 0$, (iii) $\bar{x}^{FI} = 1, \bar{x}^{NI} = 1$. In cases (i) and (iii), the equilibrium allocation rule is not affected by integration, so that consumers clearly benefit from

integration. In case (ii), integration leads to a switch from a situation in which all users are directed towards publisher 2 with an ad level $q_2^{NI} = q_2^{FI}$ to a situation in which publisher 1 receives all the traffic and displays q_1^{FI} 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 .

References

- Ambrus, A., Calvano, E., and Reisinger, M. “Either or both competition: a two-sided” theory of advertising with overlapping viewerships.” Mimeo, Duke University, 2012.
- Anderson, S., Foros, Ø., and Kind, H. J. “Competition for advertisers in media markets.” Mimeo, University of Virginia, 2011.
- Anderson, S. P. and Coate, S. “Market Provision of Broadcasting: A Welfare Analysis.” *Review of Economic Studies*, Vol. 72 (2005), pp. 947–972.
- Armstrong, M. and Zhou, J. “Paying for Prominence.” *Economic Journal*, Vol. 121 (2011), pp. F368–F395.
- Athey, S., Calvano, E., and Gans, J. “Will the Internet Destroy the News Media? or Can Online Advertising Markets Save the Media?” Working Paper No. 2180851, Rotman School of Management, University of Toronto, 2010.
- Athey, S. and Ellison, G. “Position Auctions with Consumer Search.” *Quarterly Journal of Economics*, Vol. 126 (2011), pp. 1213–1270.
- Burguet, R., Caminal, R., and Ellman, M. “In Google We Trust?” Mimeo, Barcelona Graduate School of Economics, 2013.
- Chen, Y. and He, C. “Paid Placement: Advertising and Search on the Internet.” *Economic Journal*, Vol. 121 (2011), pp. F309–F328.

- Choi, J. P. and Kim, B.-C. “Net neutrality and investment incentives.” *RAND Journal of Economics*, Vol. 41 (2010), pp. 446–471.
- de Cornière, A. “Search Advertising.” Discussion Paper no. 649, Department of Economics, University of Oxford, 2013.
- Edelman, B. and Lai, Z. “Exclusive Preferential Placement as Search Diversion: Evidence from Flight Search.” HBS Working Paper no. 13-087, Harvard Business School, Harvard University, 2013.
- Edelman, B. and Lockwood, B. “Measuring Bias in “Organic” Web Search.” Mimeo, Harvard University, 2011.
- Edelman, B., Ostrovsky, M., and Schwarz, M. “Internet Advertising and the Generalized Second-Price Auction: Selling Billions of Dollars Worth of Keywords.” *American Economic Review*, Vol. 97 (2007), pp. 242–259.
- Eliasz, K. and Spiegel, R. “A Simple Model of Search Engine Pricing.” *Economic Journal*, Vol. 121 (2011), pp. F329–F339.
- FTC. “Statement of Federal Trade Commission Concerning Google/DoubleClick.” FTC File No. 071-0170, 2007.
- FTC. “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>, 2013.
- Hagiu, A. and Jullien, B. “Why Do Intermediaries Divert Search?” *RAND Journal of Economics*, Vol. 42 (2011), pp. 337–362.
- Hahn, R. W. and Singer, H. J. “An Antitrust Analysis of Google’s Proposed Acquisition of DoubleClick.” Mimeo, Georgetown University, 2008.
- Inderst, R. and Ottaviani, M. “Competition Through Commissions and Kickbacks.” *American Economic Review*, Vol. 102 (2012), pp. 780–809.
- Manne, G. A. and Wright, J. D. “Google and the Limits of AntiTrust: The Case Against the Antitrust Case Against Google.” Mimeo, George Mason University, 2007.

- Tarantino, E. “A Simple Model of Vertical Search Engines Foreclosure.” *Telecommunications Policy*, Vol. 37 (2013), pp. 1–12.
- Taylor, G. “Search Quality and Revenue Cannibalization by Competing Search Engines.” *Journal of Economics & Management Strategy*, Vol. 22 (2013), pp. 445–467.
- Varian, H. R. “Position auctions.” *International Journal of Industrial Organization*, Vol. 25 (2007), pp. 1163–1178.
- White, A. “Search Engines: Left Side Quality versus Right Side Profits.” *International Journal of Industrial Organization*, Vol. 31 (2013), pp. 690–701.
- White, A. and Jain, K. “The Attention Economy of Search and Web Advertisement.” Mimeo, Tsinghua University, 2010.
- Wright, J. D. “Defining and Measuring Search Bias: Some Preliminary Evidence.” Mimeo, George Mason University, 2011.

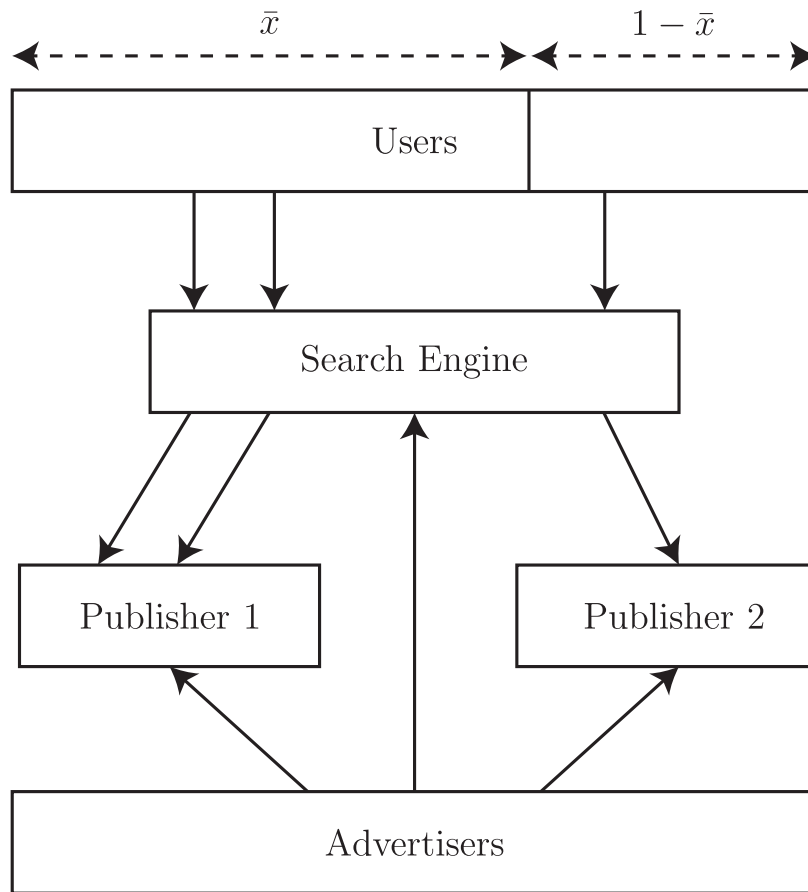


Figure 1: Market structure

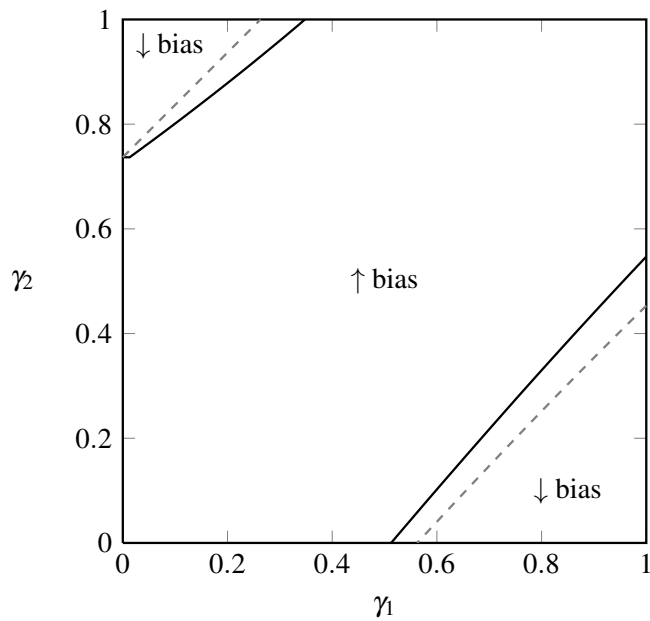


Figure 2: Plot of the changes in the absolute level of bias after integration relative to the non-integrated case. The figure is plotted for the specification of Example 1 with $\alpha_i = \beta_i = 1, \delta_i = t = q_0 = 1/4$, and $\phi_1 = 1$. The solid black lines trace the loci of points along which full integration does not cause a change in bias; the dashed lines serve a similar purpose for partial integration.

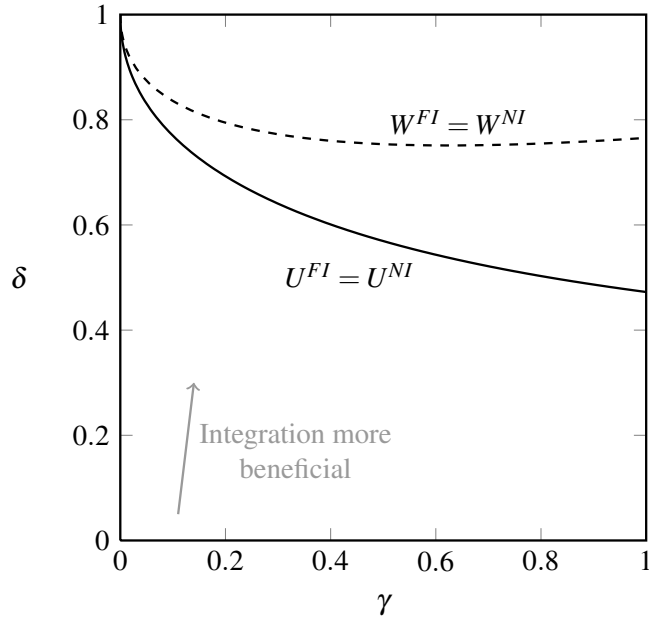


Figure 3: The effect of integration on the utility of users ($U^{FI} - U^{NI}$) and total welfare ($W^{FI} - W^{NI}$) for $\alpha = \beta = t = 1, q_0 = 1/4$, where $W = U + \pi_0 + \pi_1 + \pi_2$. The solid contour indicates the locus of points satisfying $U^{FI} = U^{NI}$ (where consumers obtain the 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.

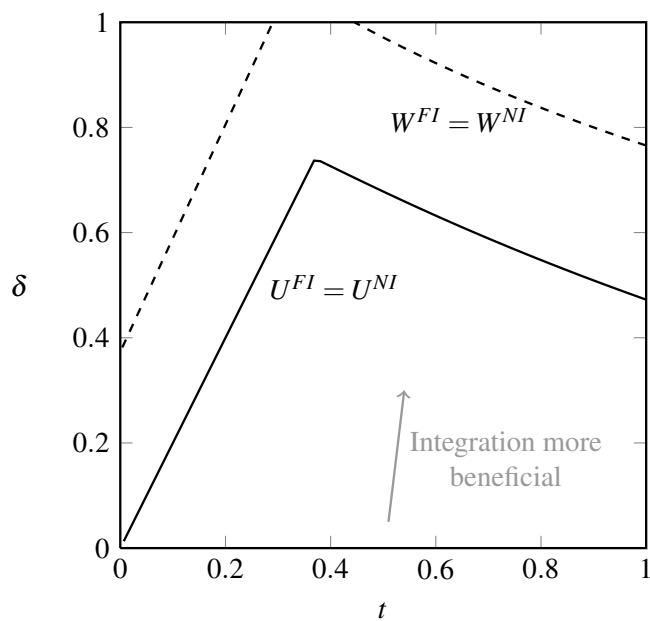
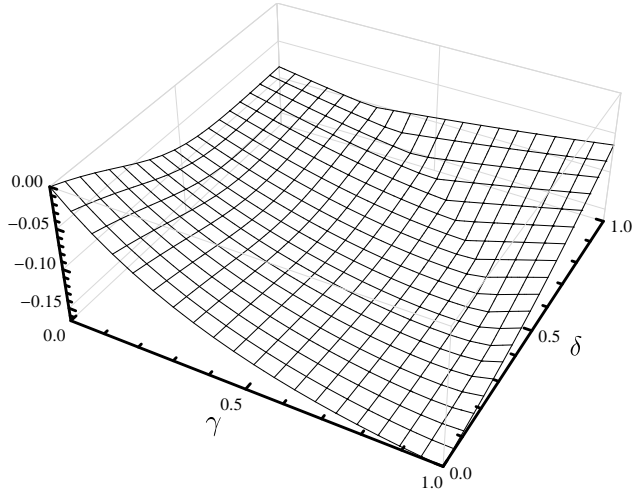
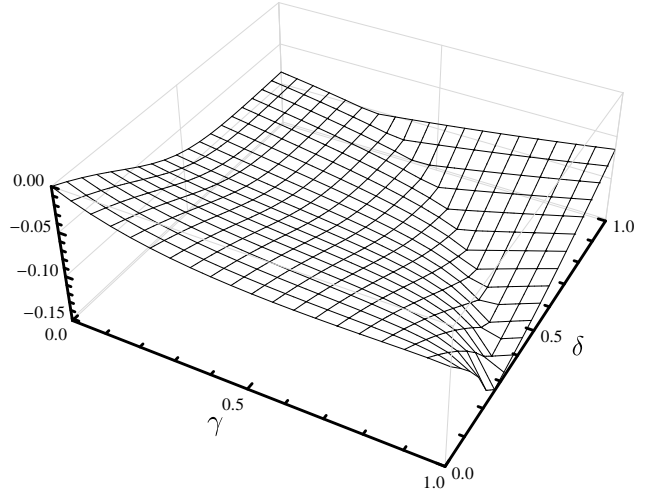


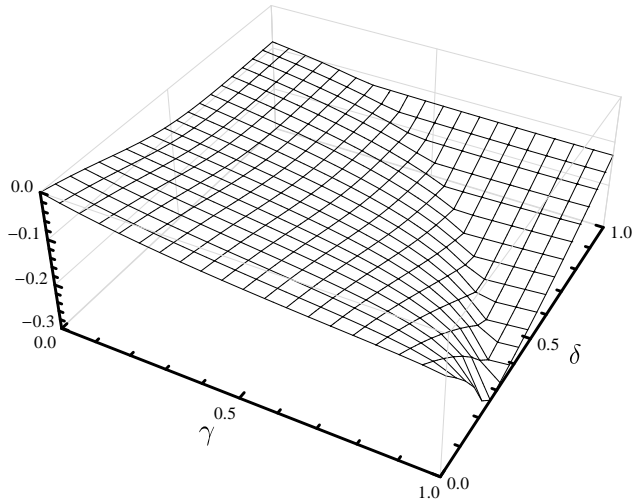
Figure 4: The effect of integration on the utility of users ($U^{FI} - U^{NI}$) and total welfare ($W^{FI} - W^{NI}$) for $\alpha = \beta = \gamma = 1, q_0 = 1/4$. The solid contour indicates the locus of points satisfying $U^{FI} = U^{NI}$ (where consumers obtain the same expected utility both before 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.



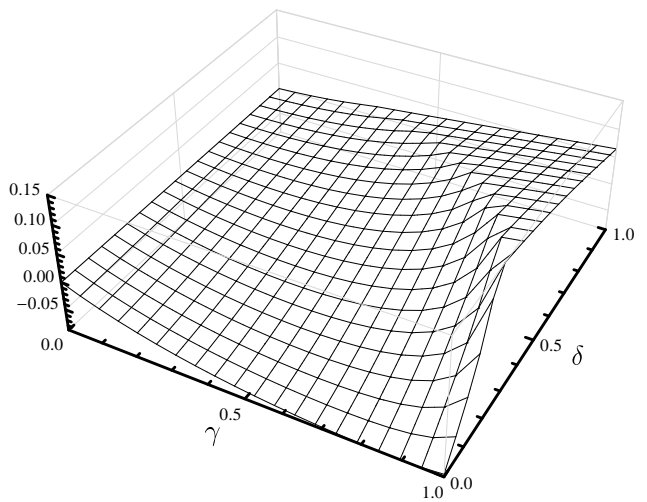
(a) Change in number of ads seen by consumers directed to site 1 ($q_0^{FI} + q_1^{FI} - q_0^{NI} - q_1^{NI}$).



(b) Change in number of ads seen by consumers directed to site 2 ($q_0^{FI} + q_2^{FI} - q_0^{NI} - q_2^{NI}$).



(c) Change in q_0 ($q_0^{FI} - q_0^{NI}$).



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

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 $q_0 = 0$.

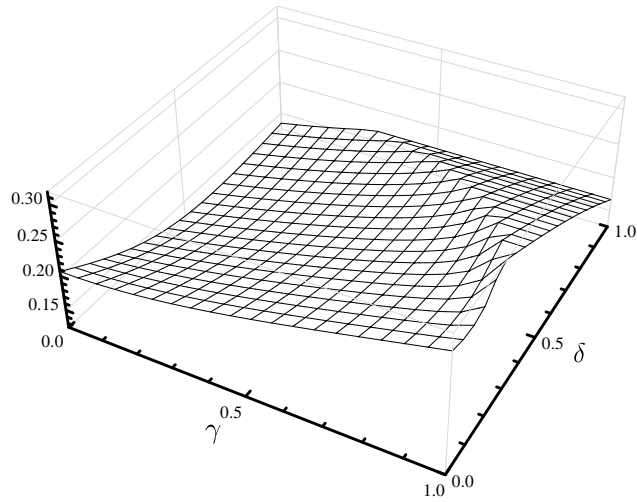


Figure 6: Plots of the post-integration magnitude of bias when consumers observe q_0 .

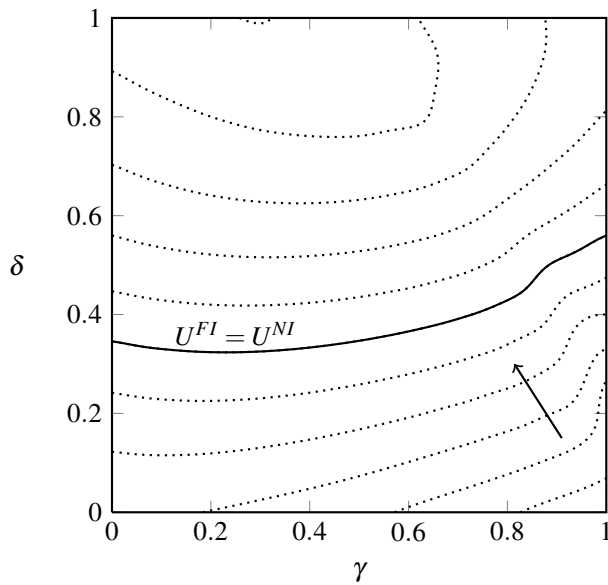


Figure 7: Contour plots of the effect of integration on consumer surplus when consumers observe q_0 . Contours are drawn every 0.015 units and the difference ($U^{FI} - U^{NI}$) is increasing in the direction indicated by the arrow. The solid contour indicates the locus of points along which integration is consumer surplus neutral. Consumer surplus increases (decreases) with integration everywhere above (below) this curve.