

“Integration and Search Engine Bias” by Alexandre de Cornière and Greg Taylor

Summary by Ariel Hsieh

A search engine is a two-sided platform that connects users, publishers, and advertisers. On one side, users come to the search engine to access publisher content. On the other side, advertisers seek to reach users via ads on publisher sites and the search engine itself. The search engine uses algorithms to decide what links are displayed on the results page, which makes it possible for the search engine to bias towards its own content.

Consider a search engine that directs users to one of two publishers. For example, a user comes to Google and enters “Seattle Seahawks super bowl” in the search box. Google can direct the user to ESPN or SportsMag, a hypothetical site that publishes sports news. After entering her search term, the user sees either an ESPN link or a SportsMag link on the results page. She does not know that Google specifically chose between ESPN and SportsMag, but her preferences for content, which Google knows, are more aligned with what ESPN offers (e.g. she likes videos and ESPN has better quality videos).

This paper studies the incentives for a search engine to bias results when the search engine integrates with a publisher. If Google integrates with SportsMag, is Google incentivized to bias towards SportsMag? Intuitively, Google will want to bias towards its own publisher because more clicks on SportsMag means more revenue for Google. However, if the user cares about Google’s quality (e.g. she will not use Google if she knows Google is biased), then Google has an incentive to direct her to content she prefers, which is ESPN.

The authors model a series of decisions made by users, advertisers, a search engine, and publishers to study this tradeoff. In particular, their model shows what happens to the level of search engine bias, as well as the quantity of ads on publisher sites, when there is no integration, partial integration, and full integration between a dominant search engine and a publisher.

Setting up the model

The model involves four parties: many users, a representative advertiser, one search engine, and two publishers.

Users are consumers who want access to content, but can only do so through a search engine that will direct them to one publisher. Their utility depends on the benefit received from consuming content, the cost of consuming content (e.g. time spent reading an article), and, assuming they dislike ads, the disutility from viewing ads.

Users have preferences for content. Figure 1 shows that there is an optimal allocation of users who prefer publisher 1 and users who prefer publisher 2. Users are aware of their own preferences, but have no prior knowledge of publisher content. Before a user decides to search, she observes the quality of a search engine by comparing her own preferences to the known preferences of the search engine. More users participate in search when the search engine is less biased and publishers display few ads.

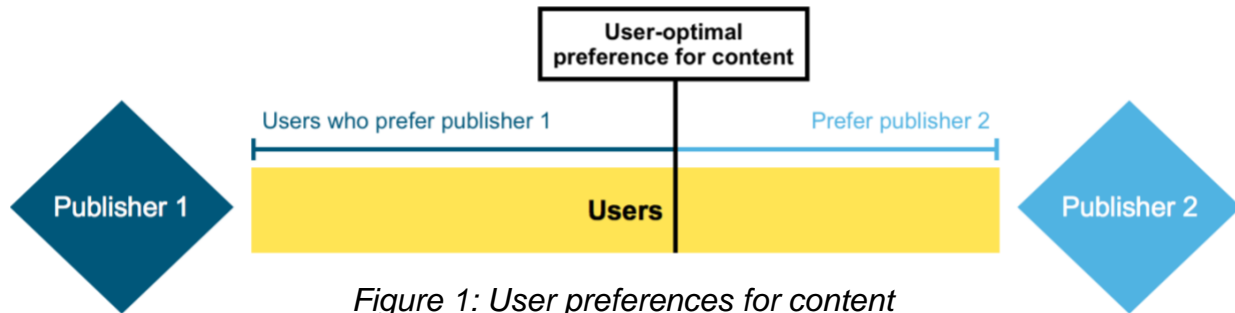


Figure 1: User preferences for content

A representative **advertiser** reaches users by showing ads on the search engine (sponsored links) and ads on the publisher sites (banner ads).¹ The advertiser views sponsored links and banner ads on publisher 1 as imperfect substitutes: sponsored links and banner ads on publisher 2 are substitutable to some other degree. The advertiser’s per-user profit is the revenue from reaching users on each publisher and the search engine (e.g. users purchase an advertiser’s product after seeing an ad), less the costs of advertising. Maximizing profit turns out to be a trade-off: the advertiser minimizes costs when ad prices and quantities are low, but maximizes revenue by buying more ad space to reach many users. While the effect of ad quantity on profit may be ambiguous, the advertiser is unambiguously better off with low ad prices and more users participating in search.

The **search engine** is an ad-financed platform that directs users to one publisher and offers sponsored links to advertisers as ad space. Since advertisers view sponsored links and banner ads as substitutes, the search engine competes with publishers to sell ad space. The search engine’s profit depends on how many users search and the quantity and price of sponsored links. As an intermediary between users and publishers, the search engine directs users according to an allocation rule. Below a certain threshold, users are directed to publisher 1; at and above that threshold, users are directed to publisher 2. Figure 2 shows that bias exists when the search engine’s allocation rule differs from users’ preferences.

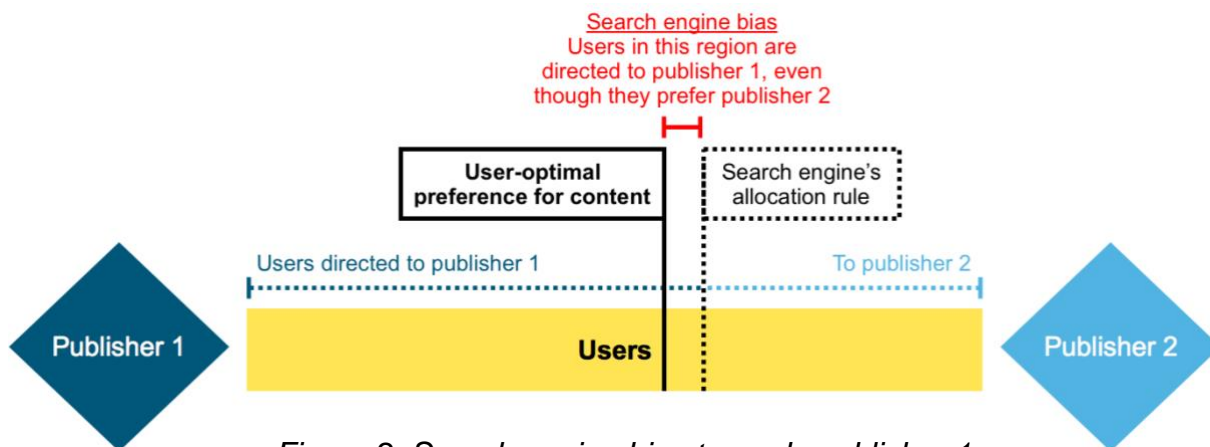


Figure 2: Search engine bias towards publisher 1

¹ The model is formally equivalent for a scenario with one advertiser who wants to send messages to each of many users, and a scenario with many advertisers who each want to reach users only once.

If the user-optimal preference for content is less than the search engine's allocation rule, as in Figure 2, then the search engine is biased towards publisher 1. If the user-optimal preference is greater than the allocation rule, then the search engine is biased towards publisher 2. If there is no difference, then the search engine is unbiased.

Absent integration, the search engine has no direct financial incentive to bias towards a publisher. Since the search engine wants to attract as many users as possible to maximize its profit, it is indirectly incentivized to present less biased results and direct users to publishers with few ads. In their main analysis, the authors assume the quantity of sponsored links is exogenously fixed, so that the search engine has one primary tool to maximize its profit: choosing an allocation rule.² A higher price for sponsored links also increases profit, but this price is determined by competition in the market for ad space. The search engine is aware of users' preferences, perhaps from tracking user search history, and is thus aware of its own bias.

The two **publishers** are ad-financed, offer banner ads to advertisers, and compete with the search engine to sell ad space. They publish horizontally differentiated content and receive views from users through organic links displayed on the search engine's results page.³ A publisher's profit depends on how many users visit its site and the quantity and price of its banner ads. Each publisher has one primary tool to maximize its profit: choosing the quantity of banner ads to supply to advertisers.

The search engine, publishers, and advertiser participate in the market for ad space. Substitutability defines the relevant market. As long as the advertiser views sponsored links and banner ads on each publisher as substitutes to some positive degree, the market contains the search engine, publisher 1, and publisher 2. In this market, demand is determined by the advertiser. Supply is determined by the quantity of ads that the search engine and each publisher chooses. There is sufficient competition so that prices adjust to clear the market; the prices of ads on each platform are functions that equalize supply and demand. Prices are sensitive to the substitutability between ads and the quantity of ads. For example, when the advertiser views sponsored links and banner ads on publisher 1 as close substitutes, an increase in the quantity of banner ads leads to a decrease in the price of sponsored links. A decrease in banner ads leads to an increase in the price of sponsored links.

² If the quantity of sponsored links is not exogenously fixed, then the search engine has an additional tool to maximize profit. The model yields the same results for an endogenous quantity of sponsored links: overall publisher ad levels fall after the search engine integrates with publisher 1, with fewer ads on publisher 1 than on publisher 2. Users are more attracted to publisher 1 because it has fewer ads, but the level of search engine bias is ambiguous and depends on how much users dislike ads.

³ In this model, publishers can also buy sponsored links through an ad auction. However, users can distinguish between organic and sponsored links (e.g. sponsored links are marked as "Ad"), and will always click on organic publisher links in equilibrium. This is because the expected relevance of an organic link (which users know by observing a search engine's level of bias) is strictly better than the random expected relevance of a sponsored link.

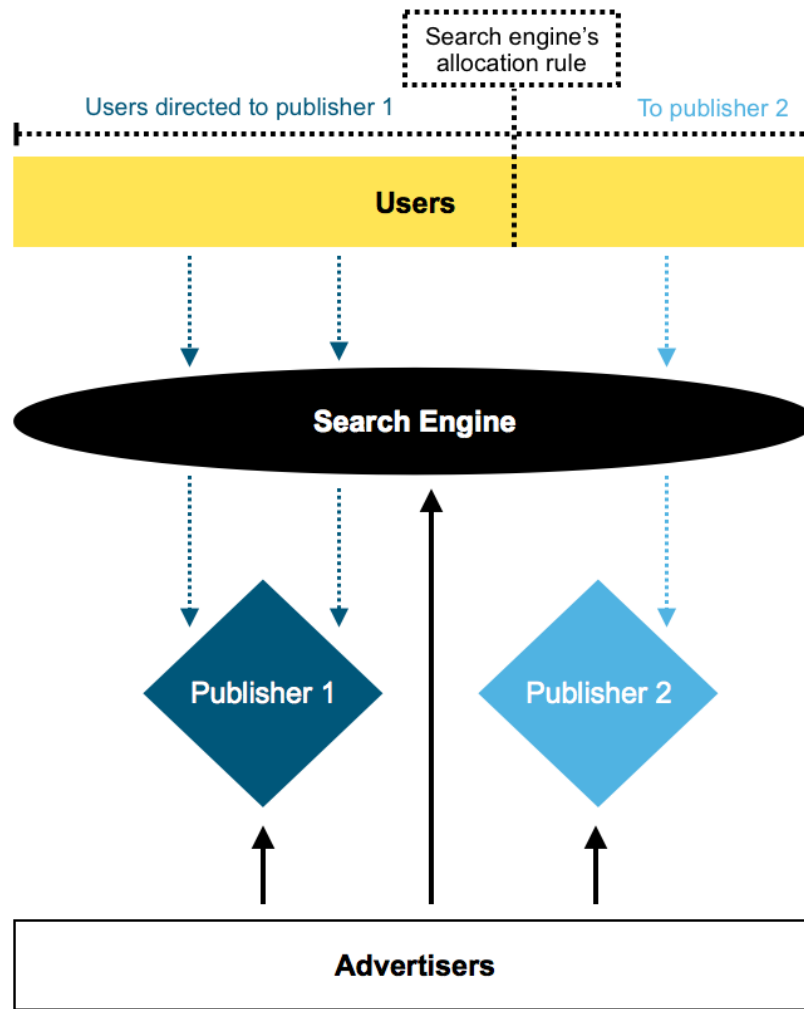


Figure 3: How the parties interact (adapted from Figure 1 in the original paper)

Structure of the model

The users, search engine, two publishers, and representative advertiser interact in a game with four stages. In stage one, the publishers simultaneously choose their supply of banner ads and the search engine chooses its allocation rule. In stage two, advertisers observe the supply of ads on each platform. The market for ad space clears, which yields prices for each type of ad. In stage three, users take note of the search engine's quality and decide whether to search by comparing their expected utility from searching to an outside option of zero. In stage four, users search.

The authors solve this game by backwards induction. First, they derive the user participation expression, which represents the number of active users who decide to search. Participation depends on the intrinsic benefit from consuming a publisher's content, the relevance of the search engine relative to users' own preferences, and the expected disutility from viewing ads. Then, the authors derive prices for each type of ad. Price functions are inverse demand functions for the quantity of ads demanded by

advertisers. To determine the quantity of ads supplied by each platform and the search engine's allocation rule, the authors consider three situations: no integration between the search engine and either of the publishers, partial integration between the search engine and publisher 1, and full integration between the search engine and publisher 1.

Baseline case with no integration

When there is no integration between either of the publishers and the search engine, the market for ad space contains three platforms: publisher 1, publisher 2, and the search engine. Each publisher chooses a quantity of ads that maximizes its own profit.

The search engine chooses an allocation rule that maximizes its profit, but also cares about attracting users. This trade-off is captured by two effects. One, the search engine wants to choose an allocation rule as close as possible to the user-optimal preference and favor the publisher with fewer ads because users are attracted to relevant content and dislike ads. Two, the search engine has an indirect incentive to bias towards the publisher with relatively few ads because this softens competition in the market for ad space. If the allocation rule directs users to the publisher with fewer ads, then publishers will compete for user traffic by lowering their ad supplies. The advertiser who views sponsored links and banner ads as substitutes has a higher demand for sponsored links because there are fewer opportunities to advertise on publisher sites. When prices adjust to equalize supply and demand, the price of sponsored links is higher than if the search engine did not bias. The search engine has an indirect incentive to bias towards publishers with fewer ads to drive up the price of sponsored links.

Balancing these two effects depends on publisher symmetry and substitutability between platforms. For example, if publisher 1 and publisher 2 are symmetric, i.e. they offer the same quantity of ads at the same price, then there is no incentive to bias because neither publisher is more or less of a competitive threat to the search engine. Similarly, the search engine has no incentive to bias if the advertiser does not view sponsored links and banner ads as substitutes. In both cases, the search engine will focus on improving quality and set its allocation rule equal to the user-optimal preference. ***In the baseline case with symmetric publishers and no integration, publisher 1 and publisher 2 choose the same quantity of ads and the search engine is unbiased.***

Effects of partial integration

Partial integration means that publisher 1 shares a fraction of its per-user revenue with the search engine, but retains full autonomy as a publisher. For example, Google's DoubleClick sells ad tech services to otherwise independent publishers. As in the baseline case, each publisher chooses a quantity of ads that maximizes its profit. With symmetric publishers, publisher 1 and publisher 2 supply the same quantity of ads. Even though publisher 1 shares a fraction of its profit with the search engine, it is an independent publisher and unaffected by the search engine's incentive to bias.

The search engine faces another trade-off between maximizing profit and attracting users. Relative to no integration, the search engine has an increased incentive to direct users to relevant results and favor the publisher with fewer ads. The search engine receives an extra stream of per-user revenue from publisher 1 and thus cares even more about attracting many users to the ecosystem. This extra stream of revenue

also creates a direct incentive to bias. The overall level of bias, compared to the baseline case, depends on how much users care about the search engine's quality.⁴

If users care more about quality, then the search engine chooses an allocation rule that is close to the user-optimal preference; the level of bias can be smaller than the case with no integration. If users care less about quality (i.e. they are not driven away by poor quality) or the search engine has market power (i.e. users have no good alternatives or available alternatives are inferior), then the search engine chooses an allocation rule that favors its partially integrated publisher because it can afford to do so without deterring users; the level of bias will be larger. ***With symmetric publishers and partial integration, publisher 1 and publisher 2 choose the same quantity of ads and the search engine has a direct incentive to bias towards publisher 1.***

Effects of full integration

The search engine receives all of the per-user revenue from publisher 1 with full integration. Assuming that advertisers view banner ads and sponsored links as substitutes, full integration is similar to a horizontal merger. An immediate effect of integration is that the fully integrated platform has control over the allocation rule and the quantity of ads supplied on publisher 1. In addition, there is one less platform that sells ads to advertisers, which relaxes competition in the market for ad space.

As before, publisher 2 chooses a quantity of banner ads that maximizes its profit. Publisher 1, as part of the integrated platform, wants to choose a quantity that maximizes the platform's profit. In this model where the integrated publisher chooses an action to make sponsored links more valuable than banner ads, there emerges a cross-price effect where a change in the supply of ads on publisher 1 accompanies a change in the price of the search engine's sponsored links. The integrated firm wants to raise the price of sponsored links, so publisher 1 chooses a lower quantity of ads than before to increase advertiser demand for sponsored links.⁵

Full integration means that the search engine has an even greater incentive to direct users to relevant content and favor the publisher with fewer ads. The quantity chosen by publisher 1 is less than the quantity chosen by publisher 2, so the search engine can conveniently bias towards publisher 1.

Compared to the baseline case with symmetric publishers, publisher 1 supplies less ads than publisher 2 and the search engine biases towards its own publisher. However, the search engine's ability to bias is tempered by the incentive to attract more users. This means choosing an allocation rule close to the user-optimal preference for content. Though this model shows that integration can raise

⁴ More precisely, the *user participation elasticity* is a measure for how much users care. User participation is *elastic* to the search engine's quality when an improvement in quality leads to a change in user participation that is larger than the magnitude of the improvement. User participation is *inelastic* when the change in participation is smaller than the magnitude of the improvement.

⁵ As long as advertisers view banner ads and sponsored links as substitutes, the quantity of ads on publisher 1 falls below the quantity of ads on publisher 2. If advertisers do not view sponsored links and banner ads as substitutes, then publisher 1 has no incentive to supply fewer ads. Even with integration, no substitutability means that symmetric publishers will supply the same quantity of ads.

quality because users see less ads and less biased search results, the integrated platform also raises its rival's costs by making it harder for the rival (publisher 2) to get traffic.

Conditions to improve welfare

With full integration, user and advertiser welfare depends on the substitutability of sponsored links and banner ads, how much users dislike ads, and the degree of content differentiation between publisher 1 and publisher 2. Regulation and enforcement around these conditions could make integration less harmful to users and advertisers.

In the case of symmetric publishers, users are better off when the advertiser views sponsored links and banner ads as strong substitutes because users are directed to publisher 1 and view fewer ads. Their utility increases the more they dislike ads. In fact, if users strongly dislike ads, they may shift their optimal preference for content towards publisher 1, which gives the search engine more room to bias. When publishers offer differentiated content, users care about being sent to the “right” site and put more weight on the search engine’s quality before deciding to search.

Advertisers are worse off when they view banner ads and sponsored links as strong substitutes. Integration relaxes competition in the market for ad space. The integrated platform uses its increased market power to restrict the quantity of banner ads and drive up the price of sponsored links, which increases the advertiser’s costs. In this model, the only way to increase advertiser welfare is to increase the number of users who participate in search; more traffic means more revenue for the advertiser.

Recall that each publisher has one primary choice variable, which is choosing a quantity of banner ads. In this setting, users are generally better off and advertisers are generally worse off after integration. However, publishers make all kinds of choices such as bidding on their rivals’ key words, changing the nature of content, and blending ads and content. Some of these choices may lead to different welfare implications for users and advertisers.

Caveats of the model

The authors note that their model does not include potential efficiency gains from integration. An integrated platform may have improved algorithms that allow the search engine to direct users to content faster and more interactively than without integration. The model also does not account for dynamic aspects of competition. For example, integration may affect a publisher’s future incentives to invest in high-quality content.

Policy implications

Overall, the model shows that integration can lead the integrated search engine to bias results towards its own publisher, relative to the case with no integration. Left unchecked, integration between a dominant search engine and a publisher may lead to foreclosure of rival search engines and publishers, higher barriers to entry in markets for search and digital advertising, harm to users through increased search engine bias, and harm to advertisers through higher ad prices.

Recent antitrust investigations reveal that these possibilities are more than speculative. In the EU, the European Commission has raised three cases against Google

and search engine bias was part of the investigations. In the UK, the Competition and Markets Authority has released a report investigating markets for online search, social media, and digital advertising.⁶ Among their recommendations, the CMA proposes a new committee to develop and enforce a code of conduct that governs the behavior of dominant online platforms and evolves as platforms continue to innovate. This code will be designed to establish rules that prevent harm and promote competition in digital markets.

⁶ The Competition and Market Authority's report is available at: <https://www.gov.uk/cma-cases/online-platforms-and-digital-advertising-market-study>