Self-Preferencing at Amazon: Evidence from Search Rankings[†]

By Chiara Farronato, Andrey Fradkin, and Alexander MacKay*

Many firms, from retailers to investment management companies, offer their own products alongside products sold by competitors. While this practice has largely been accepted in many settings (see, e.g., Dubé 2022), it has faced substantial scrutiny in online markets, especially for large digital platforms such as Google and Amazon.[1] Regulators are especially concerned that digital platforms may give preferential treatment to their own products—for example, Google Maps or Amazon Basics—over those of third-party sellers, a practice referred to as self-preferencing.

Self-preferencing can be good or bad for consumers. When digital platforms introduce and promote new products, this can increase variety and generate competition that lowers prices and increases quality. However, if the favorable treatment of vertically owned products makes it difficult for consumers to find their preferred options, consumers could be harmed. In addition, self-preferencing could discourage product innovation by other firms and cause some competitors to leave the market altogether.

In this paper, we explore whether Amazon engages in self-preferencing on its marketplace. Consistent with this behavior, we find that Amazon-branded products are ranked higher than observably similar products in consumer search results. To show this, we collect new micro-level consumer search data using a custom browser extension installed by a panel of study participants. Using this methodology, we observe search positions, search behavior, and product characteristics. This allows us to evaluate whether Amazon brands are ranked higher in search results, holding other observable factors constant.

Our work contributes to a recent literature on vertical integration on Amazon (Jeffries and Yin 2021; Lee and Musolff 2022; Gutiérrez 2021; Lam 2021; Chen and Tsai 2021; Raval 2022). A key advantage of our approach is that our data reflect real consumer searches, for which results, including delivery times and targeted ads, can be personalized. Our focus on search results as a venue for self-preferencing is justified by the fact that in our data, half of product pages are reached through a search conducted on Amazon. Our evidence thus points to the pivotal role of search results in the purchase decision.

I. Data

To explore Amazon's gatekeeping role in search results, we use data from two pilot studies conducted in the summer and fall of 2022.[2] We recruited participants residing in the United States from CloudResearch and Facebook (via ads) for a study to understand the costs and benefits of vertical integration in online platforms. What we describe here about self-preferencing in search results constitutes a first step toward answering the more relevant question about the effects of vertical integration for consumer welfare, an objective of our larger project.

In order to be eligible for the pilots, participants had to be frequent Amazon shoppers—that is, they purchased from Amazon at least twice a month—and primarily use the Chrome

[1] For example, the European Union recently passed the Digital Markets Act with the goal of limiting the market power of large platforms. In the United States, President Biden's Executive Order on Promoting Competition in the American Economy and the House Antitrust Subcommittee Investigation of Competition in Digital Markets discuss similar interventions.

[2] The pilot studies were both approved under Harvard IRB21-1677.

browser on their desktop computer for online shopping. These device criteria are required because we tracked users' online behavior with Webmunk, a desktop web browser extension developed for studies of this type (Farronato, Fradkin, and Karr 2023). Consenting participants installed Webmunk on their Chrome browser for six weeks,[3] allowing us to track their browsing activity.

We have data on users' searches on Amazon, which we use to study how Amazon ranks products. Search results are the most important channel for product discovery. In our sample, 46.5 percent of product pages are reached from a search result page, which is the largest referral source (the next largest referral source is links from other web domains, at 11.2 percent). Further, the order in which products are displayed appears to be important because users do not see all search results. In 72.1 percent of searches, consumers do not click past the first results page, and based on scroll position data, only half of the products on a full page of results are actually seen by consumers.

Our data contain 228,281 search results—including both sponsored and organic results—in 3,019 unique searches conducted by 184 users. On average, Amazon returns 76 results per search, with large variation across searches (standard deviation of 70). The number of items returned for a given search is a function of both product availability and the consumer's decision whether to progress across multiple results pages.

The search terms are very idiosyncratic. The most common search is for gift cards, which we do not include in our sample. We see some identical search terms across two distinct participants (e.g., "paper towels," "trail mix," and "DayQuil"), and at most we observe the same search term across three participants (for "cat food"). The vast majority of searches are unique to each user.

We generate an indicator for whether a search result is for an Amazon-branded product. We do this in two steps. First, in real time, the browser extension identifies Amazon brands by com-

paring the content of the product's HTML with a list of predetermined character strings that include the most popular Amazon brands[4] as well as whether the item is flagged by Amazon itself as an Amazon brand.[5] Second, after the data collection, we check whether the product title contains any of the same predetermined strings. Our browser extension is able to identify 98.6 percent of all Amazon-branded products in real time.

Our data show that on average, 1.3 percent of search results are Amazon-branded products. However, there is large heterogeneity across searches. Only 19.7 percent of searches return at least 1 Amazon branded product. Among those searches, on average 5.9 percent of results are Amazon brands (standard deviation of 6.4 percent).

We observe meaningful differences in product characteristics for searches that return Amazon brands compared to those that do not. On average, products in searches with Amazon brands tend to have more consumer ratings and lower prices compared to products in other searches. They are also more likely to be eligible for Amazon Prime benefits, which include faster delivery and free shipping.

Table 1 presents product-level characteristics for the subsample of searches that return at least one Amazon brand. The right column focuses on Amazon brands, while the left column includes other products. On average, Amazon brands and other products are similar in Prime eligibility and the rate at which they are sponsored but are very different across other dimensions. Amazon brands are more likely to have faster shipping and more likely to have at least one rating. Conditional on being rated, they have

---

[3] In case participants took longer than 6 weeks to complete the study, which is possible, we cap the tracking period to 60 days. For some participants, we have less than six weeks because the second pilot is still ongoing at the time of this writing.

[4] We search for the following Amazon brands: "Amazon Basic Care," "Amazon Basics," "Amazon Collection," "Amazon Commercial," "Amazon Elements," "Amazon Essentials," "206 Collective," "Amazing Baby," "Buttoned Down," "Cable Stitch," "Core 10," "Daily Ritual," "Goodthreads," "Isle Bay," "Lark and Ro," "Moon and Back by Hanna Andersson," "Mountain Falls," "P2N Peak Performance," "Pinzon," "Presto!," "Simple Joys by Carter's," "Solimo," and "Spotted Zebra."

[5] Amazon started adding the Amazon brand badge to search results before the start of our pilots. Whenever an Amazon brand is advertised, Amazon shows the Featured from our brands flag below the product image rather than the Sponsored ⓘ flag. The browser extension identifies both the phrases "Amazon brand" and "Featured from our brands" as denoting Amazon brands.

Table 1—Amazon Brands versus Other Products

Variable	Other products	Amazon brands
Share sponsored	0.232	0.248
Share Prime eligible	0.689	0.695
Share same-day delivery	0.044	0.056
Share overnight delivery	0.018	0.033
Share with no ratings	0.044	0.004
Average stars	4.484	4.521
Average num. ratings	7,644	20,134
Average price ($)	37.83	25.77
Average rank	43.02	33.12
Num. products	54,617	2,920

Notes: The table presents descriptive characteristics of Amazon-branded products (right column) compared to other products (left column). For this table, we consider only the products appearing in searches where at least one Amazon brand also appears. Out of the 3,019 searches, 594 return at least one Amazon-branded product. The two columns significantly differ from each other at the 1 percent confidence level on all dimensions except for "Share sponsored" and "Share Prime eligible."

almost three times as many customer reviews. Amazon-branded products also tend to be cheaper, with an average price of $26, compared to $38 for other products. After controlling for many observable characteristics, Amazon brands remain about 30 percent cheaper and have 68 percent more reviews than other similar products.

Finally, Table 1 shows that on average, Amazon-branded products appear more prominently in search results. The average rank for Amazon brands is 33, compared to 43 for other products. The next section focuses on product prominence in more detail.

II. Prominence in Search Results

In this section, we focus on the prominence given to products in Amazon search results, with specific interest in identifying the features correlated with prominence. There are a number of seller blogs describing the factors entering Amazon A10's algorithm,[6] although the actual weights are, of course, proprietary.

We obtain all product features appearing in the search results. The features include prices (including any discounts), quality metrics from reviews (number and average star rating),

delivery speed (including Prime eligibility), stock availability, and the number of new and used product options. Because we have the search terms and the product titles, we also create measures of product relevance by measuring the cosine similarity and Levenshtein distance between search terms and product titles. Finally, we capture whether an item is sponsored, albeit with some noise.[7] To ensure that our results are not affected by such noise, in robustness checks we exclude products for which the sponsored flag is imprecisely measured. Those products are contained in special carousels (e.g., "Highly Rated" or "Amazon's Choice" carousels) and constitute 13.2 percent of all search results.

We run OLS regressions of the following form:

$$y_{ij} = \alpha \times amazon_{ij} + \beta \times sponsored_{ij}$$

$$+ \, \boldsymbol{\gamma} \mathbf{X}_{ij} + \epsilon_{ij},$$

where i denotes a search result for search j. We use the rank of the product in the search results page as main outcome y_{ij}. To compute the rank, we assign rank 1 to the first product shown on the upper-left side of the page, and we then sequentially allocate rank from left to right and top to bottom, like one would read a book. We use position data of each product on the web page to construct this rank. For the majority of products, a rank value can be extracted from HTML tags, which we use to validate our outcome variable.[8]

We are interested in whether the dummy for Amazon brands ($amazon_{ij}$) predicts prominence in search results. To make the size of the coefficient estimate interpretable, we compare it to the size of the coefficient for the sponsored dummy ($sponsored_{ij}$), which we expect should increase prominence.

---

[6] For example, https://eva.guru/blog/amazon-a10-algorithm/.

[7] The sponsored flag on Amazon is usually right next to the product. Sometimes, however, when the product is nested in a carousel of multiple results, the sponsored flag is outside of the HTML that we capture. In those cases, we check for other HTML tags that indicate sponsoring.

[8] We remove fewer than 1 percent of search results for which the HTML does not contain the rank set by Amazon (because, for example, it is in a special carousel) and for which the browser extension fails to record the position on the web page.

TABLE 2—FEATURES CORRELATED WITH PRODUCT PROMINENCE

	Rank (1)	Rank (2)	Rank (3)	Rank (4)	Rank (5)	In top ten (6)
Amazon brand	−6.156 (1.023)	−3.844 (0.975)	−3.874 (0.975)	−3.170 (0.758)	−2.904 (0.992)	0.028 (0.010)
Sponsored	−7.086 (0.256)	−6.546 (0.282)	−6.549 (0.282)	−5.971 (0.294)	−8.316 (0.476)	0.098 (0.003)
Major brand			−0.662 (0.602)			
Search spell FE	Yes	Yes	Yes	Yes	Yes	Yes
Additional controls		Yes	Yes	Yes	Yes	Yes
R^2	0.489	0.516	0.516	0.520	0.468	0.148
Observations	228,281	228,281	228,281	198,116	57,537	228,281
Mean of Y	38.9	38.9	38.9	34.5	42.5	0.179
Sample	All	All	All	No carousels	Search has Amazon brand	All

Notes: The table presents estimates of OLS regressions of two measures of product prominence on product characteristics. Columns 1 through 5 show results where the outcome is rank (rank 1 corresponds to the top-left result); column 6 shows estimates for a dummy for whether the product is in the top ten results. Column 1 through 5 apply various checks to the rank regression: column 1 only has search spell fixed effects; column 2 adds observable product characteristics derived from the search results page; column 3 adds the major brand dummy; column 4 excludes products in special carousels (such as "Highly Rated" and "Amazon's Choice"); and column 5 restricts attention to searches returning at least one Amazon brand. Throughout, standard errors are clustered at the search level.

The vector $\mathbf{X}_{ij}$ contains additional controls. First, we include search spell fixed effects to account for differences in the types of products that appear across different searches. Second, we add a large set of covariates that can be broadly grouped into five groups: product relevance with the search performed (proxied by similarity measures between search terms and product titles), product availability (both in terms of stock and used/new options available), product quality (as proxied by consumer reviews), price, and delivery speed and fees.

Table 2 presents the coefficient estimates for the Amazon brand dummy and the sponsored dummy. Column 1 confirms that within a search spell, sponsored products have lower rank—that is, they are closer to the top of the page. In particular, being sponsored pushes a product up by seven positions on average, an 18 percent increase in prominence. Amazon brands are also given more prominence. Without controlling for observable characteristics, the Amazon brand coefficient is nearly as large as the effect of sponsoring, and the two coefficients cannot statistically be distinguished from each other.

Column 2 reports the results after including all characteristics observable to us as controls. The coefficient on the sponsored dummy only slightly changes, whereas the Amazon brand coefficient gets closer to 0, changing from −6.156 to −3.844. This change is consistent with the fact that Amazon-branded products tend to have characteristics, such as high ratings and low prices, that organically push them to the top of the page. Nonetheless, the Amazon brand dummy remains a strong predictor of rank, roughly 60 percent as large in magnitude as the sponsored coefficient. Estimates for other features are not presented but typically go in the expected direction: low prices, fast delivery, and high ratings tend to increase product prominence.

The results so far suggest that Amazon brands are given additional prominence in search results that cannot be explained by other observable features such as prices, ratings, or delivery times. However, this may simply be due to other characteristics that we cannot observe. One possibility is that consumers like recognizable brands. To control for that, we generate an indicator for whether a product is a "major brand." We use

a simple procedure to manually code recognizable brands that appear frequently in our search data,[9] which detects 3.9 percent of search results as carrying a major brand. We end up with 86 major brands, which include well-known brands like Adidas, Band-Aid, Duracell, Heinz, Oral-B, Pampers, and Ziploc. Column 3 in Table 2 adds the major brand dummy to our baseline regression. The major brand coefficient is of the expected sign, but only a fraction of the size of the Amazon brand coefficient and statistically indistinguishable from zero.

We also consider specifications that exclude products in special carousels (column 4) and constrain searches to those returning at least one Amazon brand (column 5). Finally, in column 6, we use a dummy variable for whether the product is shown in the top ten positions as the outcome of interest. We choose the top ten positions as an outcome of interest since they roughly correspond to the sponsored banner at the top of the page and the first two rows of standard search results. All specifications shown, as well as a number of additional checks, including specifications with interaction terms and machine learning approaches, indicate that carrying an Amazon brand is a meaningful predictor of greater prominence in search. The effect of Amazon brands tends to be 30 percent to 60 percent as large as the effect of sponsoring.

III. Conclusion

Our results, based on actual consumer searches, confirm existing (anecdotal or audit-based) evidence that Amazon brands are more prominently displayed in search results, above and beyond observable characteristics such as delivery speed and ratings, and above and beyond other major brands sold on Amazon.[10] Although we observe a rich set of product characteristics, there are other factors we do not observe, such as click and purchase rates, that may justify the higher ranking that Amazon brands receive. Finally, our findings do not necessarily imply that consumers are hurt by Amazon brands' position in search results. We are currently conducting further research to study the effects of Amazon brands and ranking policies on consumer welfare.

REFERENCES

Chen, Nan, and Hsin-Tien Tsai. 2021. "Steering via Algorithmic Recommendations." Unpublished.

Dubé, Jean-Pierre. 2022. "Amazon Private Brands: Self-Preferencing vs Traditional Retailing." Unpublished.

Farronato, Chiara, Andrey Fradkin, and Chris J. Karr. 2023. "Webmunk: A New Tool for Studying Online Consumer Behavior." Unpublished.

Gutiérrez, Germán. 2021. "The Welfare Consequences of Regulating Amazon." Unpublished.

Jeffries, Adrianne, and Leon Yin. 2021. "Amazon Puts Its Own 'Brands' First above Better-Rated Products." *The Markup*, October 14. https://themarkup.org/amazons-advantage/2021/10/14/amazon-puts-its-own-brands-first-above-better-rated-products.

Lam, H. Tai. 2021. "Platform Search Design and Market Power." Unpublished.

Lee, Kwok Hao, and Leon Musolff. 2022. "Entry into Two-Sided Markets Shaped by Platform-Guided Search." Unpublished.

Raval, Devesh. 2022. "Steering in One Click: Platform Self-Preferencing in the Amazon Buy Box." Unpublished.

---

[9] Specifically, we focus attention on the first word of product names that appear in at least 25 search results or at least 10 searches. Three research assistants were asked to flag each word as a major brand that they recognize. If two out of the three research assistants flagged a word, we consider it a major brand. We manually reviewed the list to validate the brands.

[10] Our findings of self-preferencing are conditional on the endogenous set of search results that are obtained by users. We do not investigate other channels of self-preferencing that may occur when the platform chooses whether products are presented in a given search.