The Welfare Effects of Sponsored Product Advertising

Chuan Yu*

April 2024

Click Here for the Latest Version

Abstract

Many retail platforms have recently expanded their advertising businesses, featuring sponsored products in search results. While sponsored product advertising can enable sellers to reveal information, it can also worsen search results and raise prices. Using data on Amazon searches, purchases, and advertising auction bids, I estimate a model incorporating consumers, sellers, and the platform to evaluate the welfare effects. Counterfactual analysis suggests that eliminating advertising could benefit consumers and sellers under a fixed commission rate, but would harm them if Amazon adjusts the rate optimally. Advertising tends to be more beneficial in markets with newer products and greater product differentiation.

*Department of Economics, Stanford University. Email: chuanyu@stanford.edu. I am particularly grateful to my advisors Liran Einav, Matthew Gentzkow, Michael Ostrovsky, and Ali Yurukoglu for their invaluable support and guidance. I thank Hunt Allcott, Claudia Allende, Lanier Benkard, Shengmao Cao, Jose Ignacio Cuesta, German Gutierrez, Brad Larsen, Pearl Li, Neale Mahoney, Adam Rosenberg, Paulo Somaini, Shoshana Vasserman, Frank Yang, and seminar participants at Boston University, Google Research, Indiana University, Purdue University, Stanford University, UNC-Chapel Hill, University of San Francisco, University of Southern California, and Yale University for helpful comments. I acknowledge Forum Brands for generously providing data. This research is supported by the SIEPR Program in Regulatory Policy and Leonard W. Ely and Shirley R. Ely Graduate Student Fellowship through a grant to the Stanford Institute for Economic Policy Research. All errors are my own.

1 Introduction

In recent years, sponsored product advertising has experienced dramatic growth on various digital retail platforms. When consumers search for products on these platforms, they typically encounter a mix of organic search results generated by the platforms' ranking algorithms and sponsored results displaying the winners of ad auctions. Amazon is the leader of this market, with its advertising revenues growing from less than \$3 billion in 2016 to \$47 billion in 2023 (Amazon 2017, 2024). Many other platforms offering a variety of goods and services, such as Walmart.com, Expedia, Instacart, and UberEats, are also rapidly expanding their advertising ventures.[1] This change represents a significant transformation in their business model.

The proliferation of sponsored product advertising has triggered a range of responses among platform users. Some consumers express frustration with sponsored products dominating search results, while others find them better suited to their needs than organic options (e.g., Weise 2019; Fowler 2022). Some sellers complain about rising selling costs, while others view advertising as an expansion opportunity (e.g., Matsakis 2022; Wang 2022). In 2023, the U.S. Federal Trade Commission (FTC) filed a lawsuit against Amazon for "illegally maintaining monopoly power," with allegations related to its advertising business.[2] Therefore, it is crucial to examine the impact of sponsored product advertising on the welfare of different market participants.

The welfare effects of sponsored product advertising are theoretically ambiguous. On one hand, sellers can utilize advertising to reveal information about their products not reflected in organic rankings, thereby enhancing efficiency by enabling high-quality or low-cost products to reach more consumers. On the other hand, the presence of low-quality sponsored products may displace higher-quality alternatives, and sellers may pass on advertising expenses to consumers, resulting in higher prices. Additionally, advertising may be a substitute for other revenue-raising instruments for the platform. If it became unavailable, the platform might resort to other instruments to raise revenues, potentially exacerbating distortions elsewhere.

In this paper, I examine the welfare effects of sponsored product advertising on Amazon, one of the world's largest retail platforms. My analysis consists of three components. First, I introduce

[1] For instance, Instacart, a grocery delivery platform whose IPO was one of the most prominent of 2023, earned \$871 million from advertising in 2023, constituting nearly 30% of its total revenues (Instacart 2024). In 2023, the Interactive Advertising Bureau and Media Rating Council proposed a series of guidelines to standardize measurement for sponsored product advertising. Over 50 platforms contributed to the development of these guidelines.

[2] The FTC asserts that Amazon "degrade[s] the customer experience by replacing relevant, organic search results with paid advertisements—and deliberately increase[s] junk ads that worsen search quality and frustrate both shoppers seeking products and sellers who are promised a return on advertising purchase" (Federal Trade Commission 2023).

a stylized example that illustrates the roles of different economic factors in shaping the welfare implications. Second, I compile extensive datasets on search results and product information on Amazon and provide descriptive evidence on the forces at play. Finally, I estimate an equilibrium model that incorporates interactions among consumers, sellers, and the platform. Using model estimates, I simulate a counterfactual scenario in which sponsored products are replaced with organic listings and quantify the welfare effects of sponsored product advertising.

The primary dataset for this paper includes a large number of search results from Amazon. I conducted web scraping for several thousand high-traffic keywords six times a day over a two-month period in 2022. These data allow me to track all organic and sponsored products on the first result page for over one million searches. I also collected information for nearly one million products that appeared in these search results, including daily prices, sales ranks, and reviews. I utilize an established method to convert daily sales ranks into daily quantities and validate this conversion using exact daily quantities for several hundred products over two years obtained from an Amazon seller. Finally, I gathered aggregate statistics on the bids submitted by auction winners for each keyword in my sample.

I begin my analysis of the data by presenting several descriptive findings on factors influencing welfare effects. First, I provide evidence of search frictions by demonstrating that a product's position in search results can impact its sales, which motivates sellers to pursue higher locations in search results. Second, by comparing organic and sponsored products in similar locations on the same result pages, I illustrate how sponsored product advertising enhances the visibility of less established products. Third, I observe that top sponsored products have prices that are 16% higher than their organic counterparts.

Next, I develop a model that incorporates interactions among market participants. In this model, consumers make purchase decisions while facing search frictions, considering only a subset of products in search results. Sellers set prices and submit bids in the ad auction to maximize expected profits, accounting for uncertainties arising from organic ranks and auction outcomes. The platform sets a commission rate, represented as a percentage of the price paid by sellers, to maximize a combination of commissions, advertising revenues, consumer surplus, and seller profits, with different weights assigned to each. By including the welfare of all participants in the objective, this approach reflects the platform's long-term considerations. The commission rate can represent a broader range of revenue-raising instruments available to the platform.

I identify the extent of search frictions by examining how daily variations in a product's organic ranks affect its sales while accounting for demand persistence. To estimate product utility, I rely on consumers' revealed preferences, explicitly incorporating heterogeneous consideration sets. In most markets, organic ranks show a moderate, though imperfect, correlation with product utility, and sponsored products exhibit lower average utility than their organic counterparts in similar locations. I derive sellers' marginal costs and advertising costs from the optimality conditions in their profit-maximization problems. Despite sponsored products often having lower marginal costs, the advertising payment reverses this cost advantage. Last, to determine the platform's objective, I assume that the observed choice of commission rate is optimal.

To evaluate the welfare implications of sponsored product advertising, I simulate a counterfactual scenario in which all sponsored slots are replaced with organic listings. I examine two cases: one where the commission rate remains unchanged at the existing level and another where the platform adjusts the commission rate optimally in response. I find that eliminating sponsored products while holding the commission rate constant results in a 12.6% increase in consumer surplus and a 14.1% rise in seller profits. These gains stem from the removal of low-utility sponsored products and the elimination of the pass-through of advertising costs to prices. The platform's total revenues drop by 14.8%, despite collecting more commissions.

When the platform can reoptimize the commission rate, the model predicts a 4.7-percentage-point increase in the commission rate following the removal of advertising. With advertising, a higher commission rate lowers sellers' margins and their willingness to pay for sponsored positions, thereby reducing the platform's advertising revenues. The elimination of sponsored positions incentivizes the platform to raise the commission rate. Transitioning from the status quo to a scenario with only organic results and a higher commission rate leads to a 4.3% reduction in consumer surplus and an 8.4% decrease in seller profits. This reversal in the welfare comparison underscores the importance of considering the platform's response when policymakers contemplate imposing restrictions on sponsored product advertising.

The effect of removing advertising on consumer surplus exhibits substantial variation across markets. While the effect is negative on average, it is positive in nearly 40% of the markets. This heterogeneity is closely related to the accuracy of organic rankings, measured as the correlation between product utility and organic ranks. An advantage of advertising is to allow sellers to reveal information about their products when organic rankings fail to reflect actual product utility. I find that advertising tends to be more beneficial in markets with fewer established products and greater product differentiation, often associated with less precise organic rankings.

As a final exercise, I explore counterfactual scenarios with varying numbers of sponsored positions. My findings suggest that while consumers and sellers generally benefit from advertising compared to a scenario with only organic results, introducing more sponsored positions can hurt consumers and sellers. Both the consumer- and seller-optimal numbers of sponsored positions are lower than the number in the status quo. For policymakers concerned about the welfare of consumers and sellers, a cap on the number of sponsored positions can be a better policy target than a complete ban on advertising. This approach preserves the benefits of sponsored product advertising while safeguarding the interests of consumers and sellers.

There are some caveats to my analysis. First, this paper examines the short-term impacts of sponsored product advertising and may not capture longer-term effects. For instance, advertising could alleviate the "cold-start" problem (Schein et al. 2002) by enhancing the visibility of new entrants, thereby encouraging more sellers to enter the market. However, the rise in advertising costs could reduce the profitability of selling on Amazon, potentially leading some sellers to leave the platform. In the appendix, I provide suggestive evidence on both effects.

Second, my estimation of product utility relies on consumers' revealed preferences, assuming that they can accurately discern product utility. However, in reality, consumers may make inferences based on a product's position and whether it is sponsored, so the perceived utility can change when advertising is removed. Additionally, consumers may exhibit aversions to sponsored products and ignore them when making purchase decisions, which could result in an underestimation of the utility of sponsored products. In the appendix, I explore several extensions that partially relax this assumption, and the findings remain qualitatively robust.

Finally, I abstract from several practices employed by Amazon aimed at enhancing the quality of sponsored results, such as the use of relevance scores to rank sellers in auctions and personalized sponsored results. By not incorporating them into my analysis, I may underestimate the benefits of advertising. Moreover, my baseline model assumes that organic rankings are fixed in counterfactual scenarios, precluding any changes by the platform to the ranking algorithm. The appendix considers an extension that allows organic rankings to change in a stylized way.

This paper contributes to several strands of literature. It builds upon extensive research in search advertising and position auctions, primarily focusing on search engines. Pioneering works by Edelman et al. (2007) and Varian (2007) investigate the theoretical properties of position auctions. Subsequent studies consider richer settings such as endogenous consumer search (Athey and Ellison 2011) and firms' pricing decisions (Chen and He 2011). Empirical studies have examined sellers' dynamic bidding incentives (Yao and Mela 2011), endogenous valuations of sponsored positions (Chan and Park 2015), and consumer heterogeneity and the effectiveness of search advertising (Blake et al. 2015). This paper adapts the canonical models to the setting of sponsored product advertising and incorporates novel features of the environment, such as price competition and the platform's endogenous response.

Some recent theoretical studies have examined the impact of sponsored product advertising on market outcomes and welfare. Ilango (2022) predicts that less prominent firms have a higher pass-through from advertising payment to prices, leading to lower competition and consumer surplus. Motta and Penta (2022) show that brand search advertising decreases welfare due to reduced competition and higher prices. In this paper, I highlight various economic factors that influence the welfare effects and quantify the magnitudes of different forces in a high-stakes setting.

My work is also connected to research on the usage of various instruments to boost platform revenues. Choi and Mela (2019) examine the trade-off between advertising revenues and commissions and how various product-ranking and ad-pricing mechanisms impact platform profits. Long et al. (2022) analyze how platforms can leverage information revealed in ad auctions to refine organic rankings and set the commission rate. My study incorporates the welfare of consumers and sellers into the platform's objective and investigates how different platform designs affect these users. A developing literature examines the welfare implications of Amazon's practices (e.g., Lee and Musolff 2023; Gutierrez 2022; Lam 2023; Farronato et al. 2023). My research focuses on advertising, a rapidly growing and highly policy-relevant sector within Amazon and numerous other platforms, which can provide insights and lessons with broader implications.

This paper is broadly related to the literature on slotting fees, which are payments from manufacturers to retailers for better shelf space. The literature has debated whether slotting fees are procompetitive (e.g., Sullivan 1997) or anticompetitive (e.g., Marx and Shaffer 2010). Sudhir and Rao (2006) find that the use of slotting fees for allocating shelf space is efficient and helps manufacturers signal private information. Sponsored product advertising is the e-commerce analog, and my paper contributes new evidence using novel data on placement and payment. I find that sponsored products deliver lower utility and can displace higher-utility alternatives. However, advertising can be beneficial in markets with fewer established products.

The remainder of the paper proceeds as follows. Section 2 describes the setting. Section 3 presents a stylized example to illustrate key economic factors at play. In Section 4, I present the data and descriptive results. I develop and estimate an equilibrium model in Section 5. Section 6 discusses identification and estimation results. In Section 7, I conduct counterfactual exercises to quantify the welfare effects of sponsored product advertising. Section 8 concludes.

2 Setting

Amazon is one of the world's largest digital retail platforms. In 2023, it generated a total sales volume of over $700 billion (Kaziukėnas 2024a). In the U.S., Amazon holds a market share of about 40% in the e-commerce retail industry (Kaziukėnas 2024b). Amazon operates both as a retailer, directly selling products to consumers, and as a marketplace where third-party sellers can offer their own products. As of 2021, the platform hosted over six million sellers worldwide, who contributed over 60% of the total sales on the platform (Kaziukėnas 2021).

When consumers visit Amazon's website or use its mobile app, they usually start by entering a keyword into the search box, which directs them to a page of search results. From there, consumers can browse through the listed products and make purchase decisions. Each listing in the search results includes a product image, a brief description, price, delivery options, total reviews, and an average consumer rating. Clicking on a listing takes consumers to a product page with more detailed information. Appendix Figure A.1 shows an example of search results.

There are two main types of search results. The majority of search results are organic, determined by Amazon's ranking algorithm based on factors like relevance to the searched keyword, sales performance, and consumer reviews. The remaining results are sponsored, where sellers pay the platform to display their products. Sponsored products are labeled onscreen as "Sponsored" to distinguish them from organic results. A product can appear as both organic and sponsored. Typically, when viewed on a desktop or laptop browser, search results are displayed in 60 positions on each page, usually organized in 15 rows and 4 columns.[3] Among these 60 positions, there are usually 12 sponsored ones, which can appear at the top, middle, or bottom of the search results.[4] On mobile devices, products are arranged vertically, and consumers can scroll down to view more products.

To advertise a product in search results, a seller selects a set of keywords and submits keyword-

[3]There are also nonstandardized positions for displaying products in a carousel layout, such as those highlighted as "editorial picks." These products are excluded from my analysis. Consumers can also navigate to later pages in the search results. According to Amazon's data, 81% of clicks happen on the first page.

[4]Appendix Figure A.2 illustrates the distribution of sponsored positions in search results. In my sample, the most common sponsored positions are 1–4, 17–18, 31–33, and 58–60. There are also less common result pages containing 22 positions in a single column, displaying larger images for products such as electronic devices.

specific bids.[5] When a consumer enters a keyword, Amazon conducts a generalized second-price auction among all sellers who bid on that keyword (Varian 2007; Edelman et al. 2007). In principle, sellers are ranked based on a combination of their bids and platform-assigned scores that measure products' relevance to the searched keyword. However, as I do not observe the relevance scores used by Amazon, I assume that products are ranked solely based on their bids. The seller with the r-th highest bid is displayed in the r-th sponsored position. Payment is made for each click, irrespective of whether it leads to a transaction. When a consumer clicks on the r-th sponsored listing, the seller pays Amazon an amount equal to the $(r+1)$-th highest bid.

Auction outcomes on Amazon can be highly dynamic. The array of products in the sponsored results may differ between two searches of the same keyword conducted within minutes of each other. This variability stems from various real-world factors, as documented and studied in the literature (e.g., Athey and Nekipelov 2011; Ostrovsky and Skrzypacz 2022).[6] To aid sellers in making informed bidding decisions, Amazon provides them with aggregate statistics on winning bids for each keyword. Specifically, it reports estimates of the median, lowest, and highest bids among recent auction winners for each keyword.

When a consumer makes a purchase on the platform, Amazon collects a commission fee from the seller, calculated as a percentage of the price. Appendix Figure A.3 illustrates that the commission rate varies across product categories, ranging from 8% to 20%. For most products in my analysis, this rate is either 15% or 17%. Additionally, Amazon may collect storage and shipping fees from sellers, which are usually fixed amounts based on product sizes. They may also collect a Prime membership fee from consumers, which includes benefits like free shipping.

As previously mentioned, products on Amazon are sold by either third-party sellers or Amazon itself. In the former case, sellers set retail prices and pay commissions to Amazon for each unit sold. In the latter case, Amazon purchases products from suppliers at wholesale prices and determines retail prices. Both third-party sellers and suppliers can advertise their products on Amazon. In my analysis, I do not differentiate between suppliers and third-party sellers. I assume that they both set retail prices, pay commissions to Amazon, and make advertising decisions.[7]

[5] Sellers can choose between automatic and manual targeting. With automatic targeting, a seller does not need to specify keywords; instead, Amazon matches the product to relevant keywords. With manual targeting, there are three match types: *broad match*, where the search term contains the keyword in any order or close variations; *phrase match*, where the search term contains the keyword; and *exact match*, where the search term exactly matches the keyword.

[6] Bidders may set limits on their expenditures, prompting the platform to distribute their participation in auctions over time. The platform may personalize advertising based on consumers' locations or purchase histories. Both bidders and the platform can experiment with different advertising campaigns or product arrangements.

[7] In cases where multiple sellers offer the same product, Amazon combines them into a single listing in search results

3 A Stylized Example

In this section, I present a stylized example to illustrate the economic factors that drive the welfare effects of sponsored product advertising. This example incorporates the essential elements: consumer search frictions, seller competition, and a profit-maximizing platform.

3.1 Setup

Consider a market with two sellers, labeled $j = 1, 2$, each offering a product with quality δ_j. The platform's search results consist of two positions, displaying either organic or sponsored results. Without loss of generality, I assume that product 1 is ranked higher in the organic ranking, even though this ranking may not necessarily reflect the actual ranking of δ_j. This discrepancy can arise when the organic ranking fails to fully capture product quality. The platform collects a commission for each unit sold, calculated as a percentage of the product's price.

A fraction of consumers consider only the product in the first position, while the rest consider products in both positions. This assumption approximates real-world search behavior, where consumers may face search costs and evaluate only a subset of products in search results. Such behavior motivates sellers to pursue a higher location in search results. Consumers have idiosyncratic preferences and may choose either a product they consider or an outside option.

I first examine a scenario where both positions display organic results. In this case, product 1 appears in the first position and is considered by all consumers, while product 2 is considered only by a subset of consumers. Each seller sets a price to maximize profits given its rank in search results. The platform chooses a commission rate to maximize total commissions.

Next, I consider a scenario where the first position is allocated through an ad auction. The auction outcome is stochastic, with the seller submitting a higher bid having a better chance of winning the auction. The auction winner is placed in the first position and pays a price for each click on the listing. The second position displays the top-ranked organic product (product 1). Each seller sets a price and submits a bid to maximize its expected profits. The platform chooses a commission rate to maximize the sum of commissions and advertising revenues. Appendix Figure A.4 visually illustrates the setup, and Appendix A provides formal definitions of consumer demand and the auction rule and derives the equilibria in both scenarios.

and selects a default seller, typically based on prices. While consumers can purchase from other sellers, about 80% of sales go through the default seller (Lee and Musolff 2023). Only the default seller can advertise the product.

3.2 Simulation Results

I assess the welfare effects of introducing a sponsored position by simulating the two cases described above and comparing welfare measures. In the simulations, both products have the same marginal cost. I hold the quality of product 1 constant while varying the quality of product 2.

Sponsored product advertising can impact welfare through various channels. To examine the contribution of each channel, I explore three assumptions regarding how sellers and the platform can respond to the introduction of a sponsored position. In the most restrictive scenario, sellers cannot adjust prices, and the platform cannot modify the commission rate; sellers determine their bids assuming fixed prices. In the second assumption, sellers have the flexibility to adjust their prices, while the commission rate remains unchanged. In the most flexible assumption, I allow both sellers to adjust their prices and the platform to select a different commission rate. For each assumption, I calculate the difference in market outcomes and welfare between the equilibrium with solely organic results and that with a sponsored result.

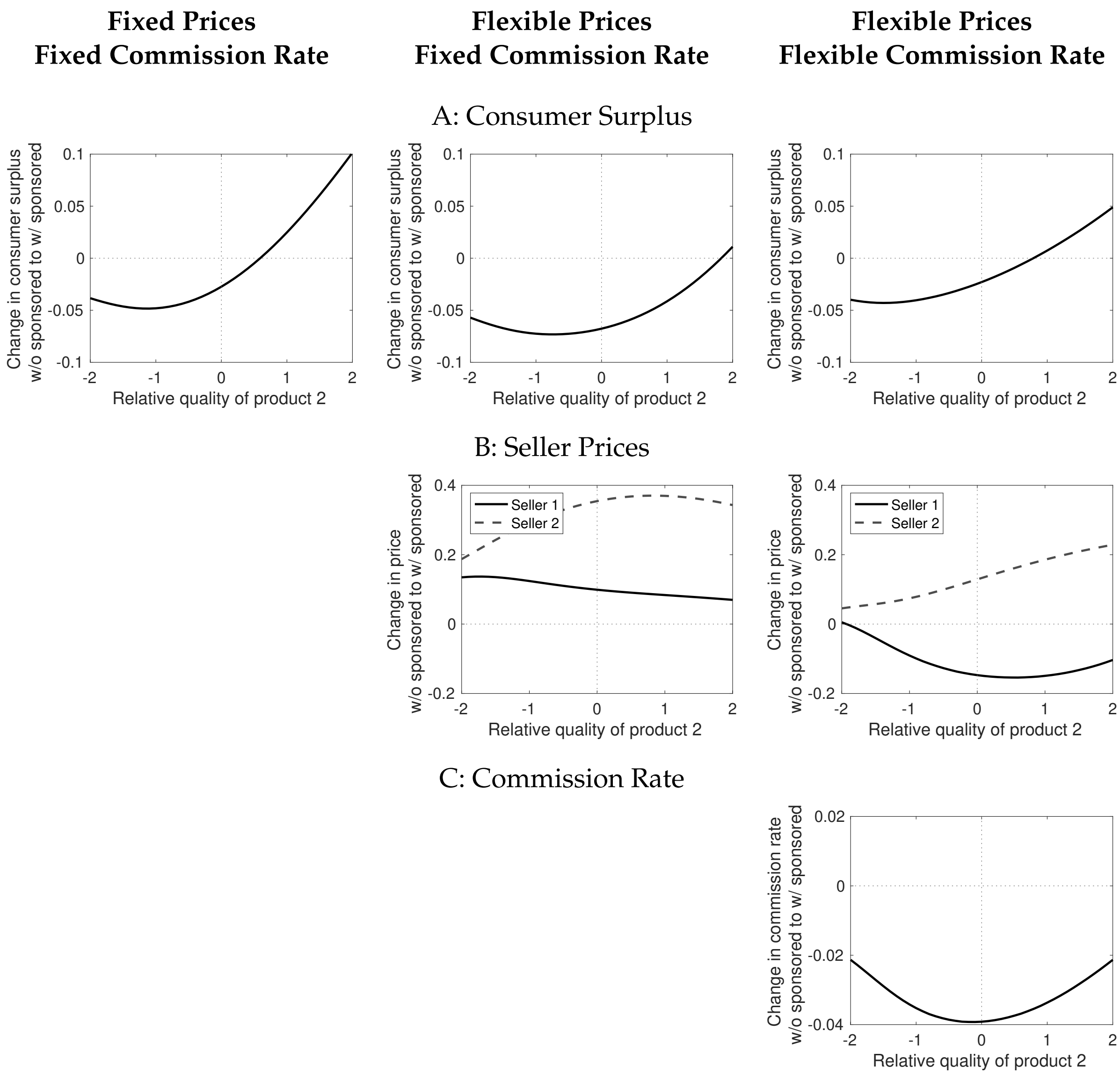
Figure 1 illustrates the change in consumer surplus following the introduction of a sponsored position.[8] The left column shows results with fixed prices and commission rates. When product 2 has lower quality than product 1, meaning that the organic ranking reflects the actual quality ranking, consumer surplus decreases regardless of which product wins the auction. If product 2 wins, it displaces the higher-quality product in the more prominent location, hurting consumers who consider only the first product. If product 1 wins, it occupies both the sponsored and organic positions, crowding out product 2 and reducing product variety for consumers who consider products in both positions. In either case, consumers are worse off.

However, when product 2 has higher quality, sponsored product advertising can enhance consumer surplus. In this scenario, product 2 often wins the sponsored position, as it anticipates greater gains from reaching more consumers. Therefore, consumers who consider only the first product gain access to a higher-quality alternative. This result suggests that in cases where the organic ranking fails to accurately reflect product quality, advertising can benefit consumers by increasing the visibility of higher-quality but lower-ranked products. This advantage is particularly relevant for products with less established records on the platform.

The second column of Figure 1 demonstrates that consumer surplus diminishes when sellers can adjust their prices after the introduction of a sponsored position. This decline occurs because

[8] Appendix Figure A.5 presents additional simulation results on seller profits, platform revenues, and total surplus.

Figure 1: Simulation Results of the Stylized Example



Notes: This figure presents the simulation results of the stylized example. For each panel, the y-axis shows the change in outcomes from the equilibrium with two ranked organic positions and no sponsored positions to the equilibrium with a sponsored position at the top and an organic position at the bottom. The x-axis represents the relative quality of product 2 compared to product 1. Each column corresponds to a different assumption regarding how sellers and the platform can respond to the introduction of a sponsored position. In the first column, neither sellers can change prices nor can the platform change the commission rate. Sellers choose bids to maximize profits taking their prices as given. In the second column, sellers can set new prices, but the platform cannot change the commission rate. In the third column, sellers can set new prices and the platform can set a new commission rate. Panels A to C display the changes in consumer surplus, the prices of both products, and the commission rate, respectively. See Section 3 for the setup of the stylized example and Appendix Figure A.5 for additional simulation results.

sellers pass on advertising expenses to consumers by raising their prices. Since prices are strategic complements, the price of the product that loses the auction also increases. The extent to which advertising expenses are reflected in prices depends on consumers' price sensitivity.

When I further allow the platform to modify its commission rate (right column of Figure 1), a notable finding is that the commission rate is lower with advertising, as illustrated in Panel C. Sellers' willingness to pay for sponsored positions is closely linked to their profit margins. A higher commission rate lowers sellers' margins and reduces their bids in the auction, resulting in lower advertising revenues for the platform. Consequently, the introduction of a sponsored position incentivizes the platform to lower its commission rate, which reduces prices and benefits consumers. The magnitude of the change in the commission rate depends on the platform's objectives, which may extend beyond short-term revenues in reality.

This example illustrates the ambiguity in the welfare effects of sponsored product advertising, which are determined by various factors. These factors include the relationship between organic rankings and product quality, the degree of overlap between sponsored and organic products, consumers' price sensitivity, and the platform's objective when determining its commission rate. In subsequent analyses, I will measure these factors for Amazon to assess the welfare effects.

4 Data and Descriptive Results

4.1 Data

The main dataset comprises a large set of scraped search results from Amazon. I selected 3,237 high-traffic keywords, which cover a variety of product categories and have a total daily search volume of 8.8 million.[9] I grouped keywords that generated search results with substantial overlap into the same market, resulting in 546 markets in total. Appendix B.1 describes the selection of keywords and the construction of markets. Appendix Table A.1 presents examples of several markets along with associated keywords. The descriptive results in this section use all keywords, while the estimation in Section 5 focuses on the highest-search-volume keyword in each market.

Between May 1 and July 5, 2022, I conducted six searches per day on Amazon for each keyword in the sample. These searches took place from 8 a.m. to 6 p.m. Pacific Time, with a two-hour interval between each search.[10] For each search, I observed all sponsored and organic results on

[9]The data on search volumes come from Jungle Scout, a prominent e-commerce intelligence service provider.

[10]To facilitate the web-scraping process, I utilized the Rainforest API tool. Each request specified a keyword and scrape time and acted as an anonymous shopper on Amazon.com. This tool routed requests through a vast network across the U.S., enabling the completion of thousands of requests within minutes. This approach provided more representative results than could be obtained through repeated searches in a single location.

the first page and their locations. I also recorded product characteristics displayed in the search results, such as price, total reviews, average consumer rating, and eligibility for free shipping. In total, I collected results from over 1.4 million searches.

I collected additional product information from Keepa.com, a platform that monitors products sold on Amazon. For each product in the collected search results, I identified its initial listing date on Amazon, the category it belonged to, and its daily Best Sellers Rank (BSR) in that category. Amazon uses the BSR system to report a product's sales performance relative to that of other products within the same category. To convert a product's daily BSRs into daily quantities, I followed an established method outlined in Appendix B.2. To validate this conversion method, I compared the estimated daily quantities based on BSRs with exact daily quantities for several hundred products over a two-year period. The exact daily quantities were obtained from a prominent Amazon seller. As shown in Appendix Figure A.6, the estimated and exact daily quantities align remarkably well, both in the cross-section and for high-frequency variation over time.

Finally, I collected the median, lowest, and highest winning bids reported by Amazon for all the keywords used in my analysis as of May 2022. These statistics were updated daily. However, upon manual examination, I found limited variation over time during my sample period. Appendix Figure A.7 presents the distribution of the median winning bids.[11]

Table 1 presents summary statistics of the data. Panel A reports the search volume and bid statistics for all keywords in the sample. The median keyword has a daily search volume of 1,088 and a median winning bid of \$1.11. The lowest and highest winning bids are \$0.72 and \$1.31, respectively. Panel B describes the characteristics of searches. On average, a result page contains 42 organic positions and 11 sponsored ones. 81% of searches yield 60 results on the first page, and my analysis focuses on these result pages. Panel C summarizes the characteristics of all the products in the search results. The median product has a price of \$28 and a daily quantity of 45 units and has been listed on Amazon for 2.8 years.

4.2 Positions in Search Results and Product Sales

The stylized example in Section 3 is built upon the assumption that a larger proportion of consumers consider the product in the top position. This assumption reflects the existence of search frictions and provides sellers with an incentive to be placed higher in search results. In this section,

[11] While I observed the median winning bids for all keywords, I observed the lowest and highest winning bids only for a subset of keywords. Panel B of Appendix Figure A.7 shows a nearly linear relationship between the median, highest, and lowest winning bids. I imputed the lowest and highest winning bids for keywords with missing data.

Table 1: Summary Statistics

	N	Mean	SD	P25	Median	P75
A: Keyword						
Daily Search Volume	3,237	2,775.28	4,050.84	534.43	1,088.23	3,522.77
Median Winning Bid ($)	3,237	1.36	0.89	0.84	1.11	1.57
Lowest Winning Bid ($)	3,237	0.88	0.58	0.54	0.72	1.01
Highest Winning Bid ($)	3,237	1.60	1.05	0.99	1.31	1.85
B: Search						
Total Positions	1,281,845	52.53	15.15	60	60	60
Organic Positions	1,281,845	41.88	12.58	48	48	48
Sponsored Positions	1,281,845	10.66	2.69	12	12	12
Total Positions = 60	1,281,845	0.81	0.39	1	1	1
C: Product						
Price	64,617,245	83.35	248.49	16.50	27.90	60.77
Daily Quantity	63,482,517	132.57	243.10	11.57	45.41	143.36
Consumer Rating (1-5)	64,434,914	4.44	0.33	4.30	4.50	4.70
Total # Reviews	64,430,254	6,798.86	18,433.52	215.00	1,355.00	5,871.00
Listed Time (Years)	46,132,028	3.68	3.54	1.11	2.76	4.93
Eligibile for Prime Shipping	66,157,677	0.86	0.35	1.00	1.00	1.00
Amazon's Choice	66,157,677	0.02	0.12	0.00	0.00	0.00
Sponsored Product	66,157,677	0.19	0.39	0.00	0.00	0.00

Notes: This table provides summary statistics for the datasets used in the analysis. Panel A presents keyword-level information. Data on search volumes and bids are collected from Jungle Scout. Panel B reports statistics at the search level. I conducted six searches a day for each keyword on Amazon from May 1 to July 5, 2022. Panel C presents product-level characteristics for all products that appeared on the first page of the search results. Daily quantities are estimated based on Amazon's Best Sellers Ranks. See Appendix B.2 for more details on the estimation method.

I analyze the impact of a product's position in search results on its sales.

Appendix Figure A.8 illustrates a negative relationship between a product's position in search results and its sales, conditional on product and keyword–day fixed effects. However, this correlation does not imply a causal relationship, as a product's position in search results can be influenced by its sales performance. Therefore, I employ a regression framework that allows me to flexibly control for unobserved demand, partially mitigating the endogeneity problem.

Let q_{jt} represent the sales quantity of product j on day t, and pos_{jt} denote its average position in the search results for a specific keyword.[12] I estimate the following regression model:

$$\log(q_{jt}) = \beta_p pos_{jt} + \sum_{\tau=1}^{3} \rho \log(q_{j,t-\tau}) + \phi_{jw} + \phi_{kt} + \varepsilon_{jt}. \tag{1}$$

Here, ϕ_{jw} represents product–week fixed effects, and ϕ_{kt} represents keyword–day fixed effects.

[12] A smaller value of pos_{jt} indicates a higher position in search results, with $pos_{jt} = 1$ representing the first position. When a product appears in the search results for multiple keywords, I retain only the keyword with the highest frequency of appearance to ensure that each product is associated with a unique keyword.

In Column 1 of Table 2, I report the estimate of β_p while including product fixed effects instead of product–week fixed effects and without accounting for lagged product sales quantities. This estimate corresponds to the slope depicted in Appendix Figure A.8.

Table 2: Relationship Between Positions and Sales

	Dep. Var. = Log(Sales)					
	(1)	(2)	(3)	(4)	(5)	(6)
Position	-0.0125	-0.0054	-0.0036	-0.0561	-0.0214	-0.0170
	(0.00010)	(0.00006)	(0.00005)	(0.00824)	(0.00340)	(0.00257)
L1. Log(Sales)			0.2534			0.3322
			(0.00204)			(0.03153)
L2. Log(Sales)			0.0035			0.0264
			(0.00130)			(0.02158)
L3. Log(Sales)			-0.0554			-0.0503
			(0.00129)			(0.02317)
N	2,725,125	2,673,123	2,598,590	19,104	18,752	18,284
Product FE	Yes	Yes	Yes	Yes	Yes	Yes
Product-Week FE	No	Yes	Yes	No	Yes	Yes
Keyword-Day FE	Yes	Yes	Yes	Yes	Yes	Yes
Products	All	All	All	Top	Top	Top
R^2	0.959	0.985	0.986	0.983	0.994	0.995

Notes: This table examines the relationship between a product's position in search results and its sales quantity. Each observation is a product on a given day. The dependent variable is the log of the product's sales quantity. The variable *Position* denotes the product's average position in the search results for a keyword, with lower values indicating higher positions. When a product appears in the search results for multiple keywords, I retain only the keyword with the highest frequency of appearance. Column 1 presents the estimate of β_p in equation (1) while including product fixed effects instead of product–week fixed effects and without accounting for lagged product sales quantities. Column 2 includes product–week fixed effects in the regression. Column 3 further includes three-period lagged product sales quantities in the regression. The sample includes products that appear in the search results for a keyword in more than 50% of the searches. Columns 4–6 replicate the regressions in columns 1-3 but restrict the analysis to products with an average position smaller than five. Standard errors are clustered at the product level and are reported in parentheses.

The endogeneity concern arises because a product's unobserved demand may simultaneously affect both its current sales and organic rank. In Column 2, I include product–week fixed effects in the regression to account for short-term fluctuations in a product's demand. In Column 3, I further include three-period lagged product sales quantities in the regression, which helps mitigate the influence of correlated demand patterns across days. With these variables included, I observe a smaller coefficient for β_p. Nevertheless, it remains statistically significant with a meaningful magnitude. The estimate in Column 3 suggests that moving up one position higher in the search results for a keyword increases a product's sales quantity by 0.36%.

These estimates aggregate products across all positions, but the impact of positions on sales may not be uniform across products. For instance, products ranked higher in search results might exhibit a more pronounced sales response to changes in positions. Columns 4–6 of Table 2 present results from the same regressions as in Columns 1–3, but restrict the analysis to products with an average position smaller than five. The coefficients for this subset of products are about four times larger, indicating a stronger impact of positions on sales for top-ranked products.

However, these regressions do not entirely address the endogeneity issue. While they include flexible controls for recent underlying demand, they do not account for the possibility of a contemporaneous demand shock that affects sales and alters the product's position simultaneously. In Section 5.1.2, I will further tackle the endogeneity problem by making structural assumptions about how demand shocks and organic ranks evolve over time.

4.3 Comparison of Sponsored and Organic Products

The example in Section 3 highlights the importance of the relationship between products' organic ranks and their quality in determining the welfare effects of sponsored product advertising. When organic rankings do not accurately reflect product quality, advertising can benefit consumers by increasing the visibility of high-quality products. However, it may harm consumers if low-quality sponsored products displace high-quality organic alternatives. Therefore, I conduct a comparison between sponsored and organic products at the top of search results.

Figure 2 compares sponsored and organic products in the top eight positions in each search, typically consisting of four sponsored and four organic products. The analysis focuses on four dimensions: the log of total reviews, years listed on Amazon, average consumer rating, and the log of price. The first two variables offer insights into a product's presence and visibility on the platform. While average consumer rating could serve as a measure of product quality, it is imperfect and susceptible to manipulation. Therefore, it should be interpreted with caution.

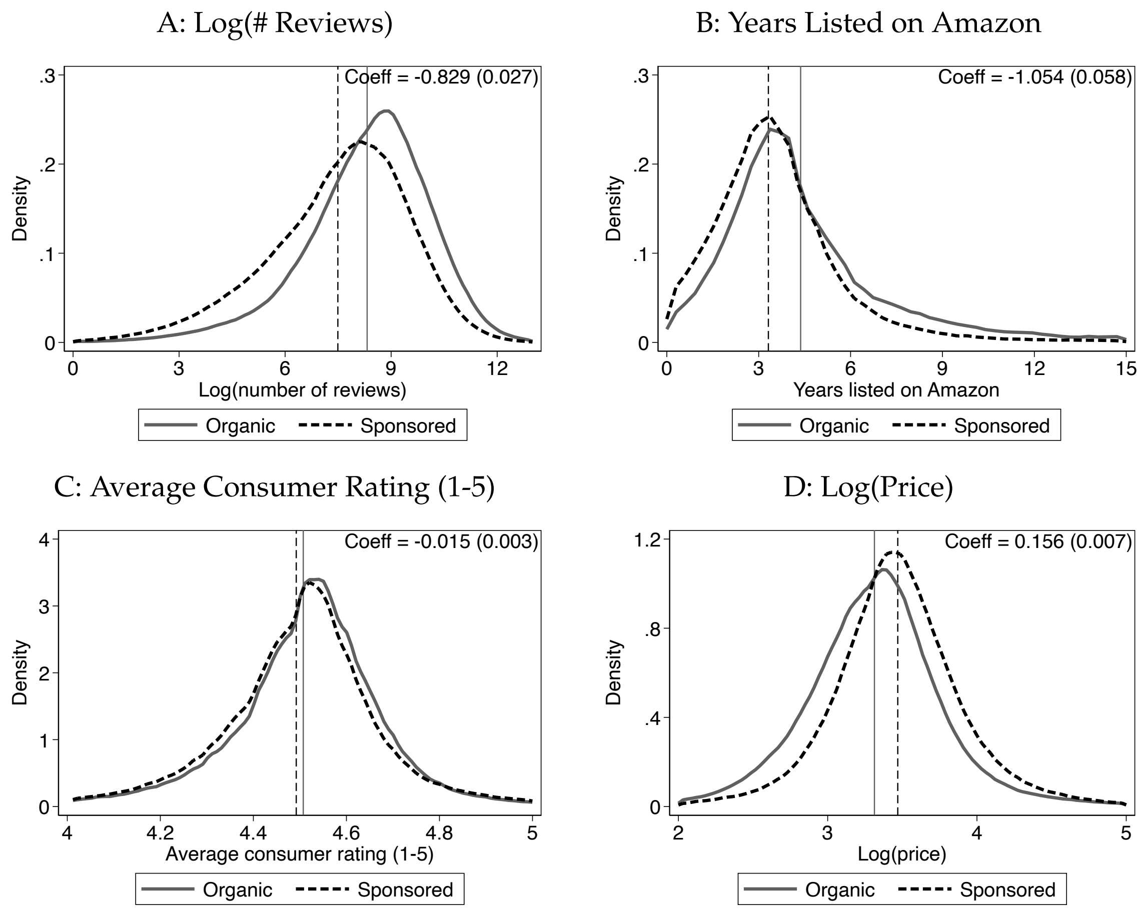
Let j index products and f index searches. I residualize each outcome y_{jf} by search fixed effects and recenter it at its sample mean, denoted as $\tilde{y}_{jf}$. Each panel of Figure 2 displays the distribution of $\tilde{y}_{jf}$ separately for sponsored and organic products, along with the estimated coefficient β_S from the following regression:

$$\tilde{y}_{jf} = \beta_0 + \beta_S Sponsored_{jf} + \epsilon_{jf}. \tag{2}$$

Here, $Sponsored_{jf}$ is an indicator for sponsored products, and β_S quantifies the average difference between sponsored and organic products in similar locations on the same result pages.

The first two panels of Figure 2 indicate that, on average, top sponsored products have 83%

Figure 2: Comparing Top Sponsored and Organic Products



Notes: This figure compares sponsored and organic products. I restrict to products located in the top eight positions in each search, typically consisting of four sponsored and four organic products. For each panel, I residualize the outcome of interest by search fixed effects and recenter it at its sample mean. This figure plots the distribution of the residualized outcome separately for sponsored and organic products, with the vertical lines indicating the corresponding means. In Panels A to D, the variables are the log of the number of reviews, years listed on Amazon, average consumer rating, and the log of price, respectively. The coefficient reported in each panel is estimated from equation (2). Standard errors are clustered at the product level and reported in parentheses.

fewer reviews and a one-year shorter listing duration compared to top organic products. These results suggest that sponsored product advertising alters the set of products consumers encounter and boosts the visibility of less established products. It is plausible that their true utility is not accurately reflected in organic rankings, and thus, they may benefit more from advertising.

Panel C of Figure 2 demonstrates that top sponsored products have only slightly lower average consumer ratings compared to their organic counterparts in similar locations. While this observation may imply that sponsored products are not significantly inferior, it is crucial to note that the average consumer rating is a noisy and potentially manipulable measure of quality. In

Section 5, I estimate product quality based on consumers' revealed preferences.

The last panel of Figure 2 illustrates that, on average, top sponsored products have prices 16% higher than their organic counterparts. This price difference could be attributed to the additional payment required for sponsored placements, raising concerns that if advertising costs lead to higher prices, sponsored product advertising could negatively impact consumers, even if it can help certain high-quality products reach more consumers.[13]

Finally, sponsored product advertising may harm consumers by causing duplication when the same product appears in both a sponsored and an organic position. To assess the extent of overlap between sponsored and organic products, I examine whether each sponsored product also appears in an organic position on the same result page. Appendix Figure A.10 illustrates that over 40% of the top four sponsored products also appear in organic positions on the same result page. This percentage decreases to 27% when considering all sponsored products.

So far, I have presented evidence of multiple factors at play. While sponsored product advertising could enhance the visibility of less established products, it could also raise prices and reduce product variety. However, these findings do not fully capture the equilibrium effects. If sponsored product advertising were eliminated, sellers would adjust their prices in response, and the platform could also respond accordingly. In the next section, I will develop and estimate an equilibrium model to account for these equilibrium effects.

5 Model and Estimation

I develop and estimate a three-stage equilibrium model. First, the platform sets a commission rate to maximize its objective. Then, sellers compete in prices and bids to maximize their expected profits. Finally, consumers facing search frictions make purchase decisions. I will describe each stage of the model in reverse order and then discuss the estimation procedure.

5.1 Demand

5.1.1 Model

I define a market as a specific keyword. Consumers enter the keyword on the platform, view the search results, and choose from the listed products or an outside option. I focus on the first page

[13] In Appendix Figure A.9, I estimate equation (2) separately for each keyword in my sample and plot the distribution of β_S across keywords. The relationships between sponsored and organic products hold for the majority of keywords.

of the search results. As I do not have access to click data, I model a consumer's purchase decision without delving into the details of the browsing process. I also abstract from the broader search process, such as which keyword to use and whether to search for another keyword.

In each market, products are indexed by $j \in \mathcal{J}$. The search results contain N positions, indexed by n. Out of these, a total of R positions, indexed by r, are designated as sponsored positions and allocated through ad auctions. The remaining positions display organic results determined by the platform's ranking algorithm. In the following estimation, $N = 60$ and $R = 12$. Let Γ denote the product arrangement, where $\Gamma_n \in \mathcal{J}$ refers to the product in the n-th position. Furthermore, let $\mathcal{J}_n(\Gamma) = \{\Gamma_1, \ldots, \Gamma_n\} \subset \mathcal{J}$ represent the set of products in the top n positions for $1 \leq n \leq N$.

Consumer i's utility from purchasing product j on day t is expressed as follows:

$$u_{ijt} = \phi_j + \psi_t + \psi_j t - \alpha p_{jt} + \xi_{jt} + \epsilon_{ijt} = \delta_{jt} + \epsilon_{ijt}. \tag{3}$$

Here, ϕ_j denotes the time-invariant product quality, and ψ_t is a day fixed effect. $\psi_j t$ represents a product-specific linear time trend, which captures a product's trajectory over time. p_{jt} is the price, and ξ_{jt} stands for an unobserved demand shock. ϵ_{ijt} represents an idiosyncratic preference shock following a type I extreme value distribution. The mean utility of the outside option, which is not buying any product on the first page of the search results, is normalized to zero. I assume that consumers can observe δ_{jt} for all products they consider; thus, they do not make inferences about product utility based on whether a product is labeled as sponsored.[14]

Consumers may have different consideration sets in the form of the top n products in search results, or $\mathcal{J}_n(\Gamma)$. Let $\lambda_n \in [0, 1]$ represent the fraction of consumers whose consideration sets contain the product in the n-th position. All consumers consider the first product, so $\lambda_1 = 1$. Only a subset of consumers consider the second product, so $\lambda_2 \leq 1$. Among consumers who consider the n-th product, only a subset consider the $(n+1)$-th product, so $\lambda_{n+1} \leq \lambda_n$. Given that choosing a product beyond the first page is part of the outside option, $\lambda_{N+1} = 0$. In general, a fraction $\lambda_n - \lambda_{n+1} \geq 0$ of consumers have the consideration set $\mathcal{J}_n(\Gamma)$.[15]

[14] The empirical evidence regarding how consumers react to sponsorship disclosure is mixed. Sahni and Nair (2020) find that disclosing a restaurant's listing as a paid ad increases calls to the restaurant by 77%, while Hui and Liu (2022) find that more salient ad disclosure prompts consumers to substitute away from advertised listings. In Appendix F.1, I explore an extension where some consumers ignore sponsored products when making purchases.

[15] This model loosely approximates a sequential search model where a consumer continues her search if and only if the expected incremental utility exceeds the search cost (e.g., Lam 2023). Consumers with varying search costs stop at different positions, resulting in different consideration sets. In my baseline model, the population shares of different consumer types, $\lambda_n - \lambda_{n+1}$, are exogenous and do not adapt to changes in prices or product arrangement. In Appendix F.2, I explore an extension where consideration sets are determined endogenously in a sequential search model.

Each consumer chooses either a product in her consideration set or the outside option. Given the prices of all sellers $\boldsymbol{p}$ and the product arrangement Γ, the market share of product j on day t is given by:

$$s_{jt}(\boldsymbol{p}, \Gamma) = \sum_{n=1}^{N} \mathbf{1}\big(j \in \mathcal{J}_n(\Gamma)\big)(\lambda_n - \lambda_{n+1}) \frac{\exp(\delta_{jt})}{1 + \sum_{j' \in \mathcal{J}_n(\Gamma)} \exp(\delta_{jt'})}. \tag{4}$$

Unlike a logit model, in equation (4), a product's market share depends not only on prices but also on the product arrangement Γ and the population shares of different consideration sets, $\lambda_n - \lambda_{n+1}$.

5.1.2 Estimation

When applying the model to the data, I introduce some additional assumptions. First, since the dataset contains multiple searches for a keyword each day, I calculate the average market share across all searches on a given day. Second, while equation (4) defines a product's market share within a specific keyword, referred to as the *focal* keyword, I observe a product's aggregate quantity on the platform. I calibrate the fraction of a product's total sales attributed to the focal keyword in Appendix C.1. In Appendix C.2, I calibrate the size of each market using industry reports. These calibrations allow me to construct a product's market share in the focal keyword.

The key primitives that I need to estimate are consumers' search frictions, captured by λ_n, and the parameter that measures consumers' price sensitivity, α. The former determines the population shares of different consideration sets, while the latter plays an important role in sellers' pricing decisions. I allow these parameters to vary across markets. To estimate α, I employ a supply-side moment, as described later in Section 5.2.2.

I parameterize the fraction of consumers who consider each position in the search results as $\lambda_n = 1/\exp\big(\beta(n-1)\big)$, where $\beta \geq 0$ is a parameter determining the magnitude of search frictions. When $\beta = 0$, there is no search friction, and a product's rank in search results does not affect its sales. A larger β implies that a greater share of consumers consider only products in top positions. In this case, all else being equal, ranking higher in search results would result in more sales. Appendix Figure A.11 provides examples of this function under different values of β.

Estimating the causal effect of a product's rank on its sales has an endogeneity problem, as the unobserved demand shock ξ_{jt} in equation (3) can be correlated with the product's organic rank r_{jt}. This correlation arises because a product's organic rank is likely affected by its sales performance. A greater demand shock leads to more sales, which could be picked up by the platform's ranking algorithm and improve the product's organic rank on the same day.

To identify the parameter β, I employ a common approach found in the literature (e.g., Grennan 2013; Sweeting 2013; Errico and Lashkari 2022). First, I assume that the unobserved demand shock ξ_{jt} follows a first-order autoregressive (AR(1)) process, where ξ_{jt} consists of a persistent component that can be predicted by the lagged demand shock and an innovation term η_{jt}:[16]

$$\xi_{jt} = \rho\xi_{j,t-1} + \eta_{jt}, \quad \eta_{jt} \sim \text{i.i.d.}, \quad \mathbb{E}(\eta_{jt}) = 0, \quad \eta_{jt} \perp \xi_{j,t-1}. \tag{5}$$

Second, I assume that:

$$r_{j,t-1} \perp \eta_{jt}. \tag{6}$$

This assumption posits that when the platform determines a product's organic rank, it does not take into account future innovations to demand. It is a reasonable assumption because, even if the platform anticipates future changes in a product's demand, it has no incentive to incorporate them into the product's current organic rank, as these changes are realized only in the future.

Last, I assume that, given the current demand shock, the current and lagged organic ranks have a nonzero correlation:

$$\text{Corr}(r_{j,t-1}, r_{jt}|\xi_{jt}) \neq 0. \tag{7}$$

This assumption is valid if the platform's ranking algorithm considers not only the current sales performance but also the sales performance over an extended period, thereby creating persistence in organic ranks.[17] Under this assumption, given the current demand shock ξ_{jt}, there exists variation in the lagged organic rank $r_{j,t-1}$ influenced by demand shocks before $t-1$. This variation carries forward to the current organic rank r_{jt} and affects a product's sales. Importantly, this variation is exogenous to the current innovation and can be used to identify the causal effect.

Given these assumptions, I estimate β and ρ for each market using the Method of Moments with the following conditions:

$$\mathbb{E}\begin{pmatrix} \eta_{jt}\xi_{j,t-1} \\ \eta_{jt}r_{j,t-1} \end{pmatrix} = 0. \tag{8}$$

To understand the identification of these two parameters, first consider a product with a constant rank over time. The fluctuations in its daily sales reveal the persistence of demand shocks, thereby identifying ρ. Then, imagine two products with identical lagged demand shocks $\xi_{j,t-1}$. Their current demand shocks would follow the same distribution given by equation (5). However, these

[16]Appendix G.1 considers alternative processes for the unobserved demand shock ξ_{jt}.

[17]Appendix Figure A.12 provides some suggestive evidence supporting this assumption.

products might have followed different trajectories before $t-1$. For example, one product may have had stable demand, while the other may have been out of stock for a period. These divergent histories lead to variations in $r_{j,t-1}$ between the two products, which then carry over to r_{jt} and result in different sales. As per equation (6), these differences in $r_{j,t-1}$ are orthogonal to the current innovation and can help identify the parameter β.[18]

5.2 Supply

5.2.1 Model

Sellers make bidding and pricing decisions every week to maximize their expected profits. In reality, a seller typically makes multiple keyword-specific bidding decisions and a single pricing decision that applies to the entire platform. However, estimating such a model is not feasible due to data and computational limitations. Therefore, I employ a partial-equilibrium approach, focusing on sellers' decisions within the focal keyword. Specifically, when seller j submits a bid b_j for this keyword and sets a price p_j, it considers its bidding decisions in other keywords as fixed but acknowledges that changing its price can impact its sales across all keywords.

When sellers make bidding and pricing decisions, they take into account the uncertainty in the realized search results on Amazon, which arises from two sources. First, organic ranks are determined in real time by Amazon's ranking algorithm, which is influenced by a variety of factors and subject to variations across different searches. I assume that the actual order of organic ranks, denoted as Γ_0, is drawn from a distribution represented as $G(\cdot)$.[19]

Second, auction outcomes can undergo frequent changes due to various real-world elements, such as sellers' budget constraints, personalized advertising, and experimentation, as discussed in Section 2. To capture this uncertainty in auction outcomes, I assume that when a seller submits a bid b_j, the realized bid $\tilde{b}_j$ is stochastic and follows $\tilde{b}_j = \omega_j b_j$, where $\omega_j > 0$ reflects the deviation of the realized bid from the submitted bid. I refer to b_j as the *targeted* bid and $\tilde{b}_j$ as the *realized* bid. The vector $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_J)$ follows an exogenous distribution denoted as $F(\cdot)$.

Let $\boldsymbol{p} = (p_1, \ldots, p_J)$ and $\boldsymbol{b} = (b_1, \ldots, b_J)$ collect the prices and targeted bids of all sellers. The realized product arrangement in the search results, denoted as $\Gamma(\boldsymbol{b}, \boldsymbol{\omega}, \Gamma_0)$, is determined by

[18] In Appendix G.2, I conduct a robustness check by exploring different values of β. The range of β I examine encompasses the value consistent with the estimate by Ursu (2018), who utilizes experimental variation to identify the impact of ranks on consumer search, as well as the value corresponding to the estimate presented in Column 6 of Table 2.

[19] In my baseline model, I assume that the distribution of organic ranks remains fixed in all counterfactual exercises. In Appendix F.3, I explore an extension where I estimate the platform's ranking algorithm as a function of past sales and prices and incorporate potential changes in the distribution of organic rankings in counterfactual scenarios.

targeted bids, deviations of realized bids, and realized organic ranks. Specifically, the seller with the r-th highest realized bid occupies the r-th sponsored position, and the remaining positions are allocated based on realized organic ranks. Seller j's expected profits in a week are given by:

$$\pi_j(\boldsymbol{p}, \boldsymbol{b}) = \left(s_j^e(\boldsymbol{p}, \boldsymbol{b}) + M_j^{other} s_j^{other}(p_j)\right)\left((1-\tau)\, p_j - c_j\right) - ad_j^e(\boldsymbol{p}, \boldsymbol{b}) - ad_j^{other}(p_j), \tag{9}$$

where

$$\begin{aligned} s_j^e(\boldsymbol{p}, \boldsymbol{b}) &= \int s_j\big(\boldsymbol{p}, \Gamma(\boldsymbol{b}, \boldsymbol{\omega}, \Gamma_0)\big) \mathrm{d}F(\boldsymbol{\omega}) \mathrm{d}G(\Gamma_0), \\ ad_j^e(\boldsymbol{p}, \boldsymbol{b}) &= \int ad_j(\boldsymbol{p}, \boldsymbol{b}, \boldsymbol{\omega}, \Gamma_0) \mathrm{d}F(\boldsymbol{\omega}) \mathrm{d}G(\Gamma_0). \end{aligned} \tag{10}$$

Here, τ denotes the commission rate charged by the platform. The marginal cost, c_j, includes production costs (or wholesale prices) and logistics expenses such as transportation costs but excludes advertising payment. $s_j^e(\boldsymbol{p}, \boldsymbol{b})$ represents the seller's expected market share within the focal keyword, and $s_j^{other}(p_j)$ denotes its aggregate market share in other keywords. M_j^{other} is the market size of other keywords relative to the focal keyword. $ad_j^e(\boldsymbol{p}, \boldsymbol{b})$ and $ad_j^{other}(p_j)$ represent the seller's expected advertising payment in the focal and other keywords, respectively.

Given the realized product arrangement $\Gamma(\boldsymbol{b}, \boldsymbol{\omega}, \Gamma_0)$, seller j's market share within the focal keyword, denoted as $s_j\big(\boldsymbol{p}, \Gamma(\boldsymbol{b}, \boldsymbol{\omega}, \Gamma_0)\big)$, is determined by equation (4). When the same product appears in both organic and sponsored positions as separate listings, I assume they receive independent preference shocks. The expected market share, $s_j^e(\boldsymbol{p}, \boldsymbol{b})$, integrates over the uncertainty in auction outcomes and organic ranks in a week and is given by equation (10).

In terms of advertising payment, when a seller wins the r-th sponsored position, its payment can be computed as the payment per click (PPC) multiplied by the number of clicks on the sponsored listing. Let $j_r(\boldsymbol{b}, \boldsymbol{\omega})$ denote the seller with the r-th highest realized bid. In a generalized second-price auction, the PPC for the seller ranked r-th is equal to the bid of the seller ranked $(r+1)$-th. Since I do not model the browsing process, I adopt a reduced-form approach to model clicks. Specifically, when product j appears in the r-th sponsored position, I define the average number of clicks per sponsored unit, or the inverse conversion rate, as γ_{jr}. A higher γ_{jr} implies that the product receives more clicks for each sponsored unit sold, resulting in higher advertising costs for the seller. Therefore, seller j's advertising payment can be expressed as:

$$ad_j(\boldsymbol{p}, \boldsymbol{b}, \boldsymbol{\omega}, \Gamma_0) = \sum_{r=1}^{R} \mathbf{1}\big(j_r(\boldsymbol{b}, \boldsymbol{\omega}) = j\big) \tilde{b}_{j_{r+1}(\boldsymbol{b}, \boldsymbol{\omega})} \gamma_{jr} s_j^S\big(\boldsymbol{p}, \Gamma(\boldsymbol{b}, \boldsymbol{\omega}, \Gamma_0)\big). \tag{11}$$

Here, $s_j^S\big(\boldsymbol{p}, \Gamma(\boldsymbol{b}, \boldsymbol{\omega}, \Gamma_0)\big)$ represents seller j's market share from a sponsored position, and $\mathbf{1}\big(j_r(\boldsymbol{b}, \boldsymbol{\omega}) = j\big)$ is an indicator of seller j winning the r-th sponsored position. The expected advertising payment, $ad_j^e(\boldsymbol{p}, \boldsymbol{b})$, is given by equation (10) and accounts for the uncertainty in search results.

In addition to the focal keyword, sellers also generate sales and incur advertising costs from other keywords. While I do not explicitly model interactions among sellers in those keywords, I approximate a seller's market share in other keywords using a logit model, where the market share is a function of its price, denoted as $s_j^{other}(p_j)$. I assume that the division between sponsored and organic sales in other keywords mirrors the observed split in the focal keyword. Similarly, a seller's advertising payment per sponsored unit in other keywords also mirrors that in the focal keyword. The relative market size, M_j^{other}, follows the same calibration outlined in Appendix C.2. Appendix D.1 provides further details on the construction of these quantities.

In equilibrium, sellers' expectations are rational. Given their expectations, each seller cannot increase its expected profits by submitting a different targeted bid or setting a different price.

5.2.2 Estimation

In this model, each seller faces two weekly decisions: the price p_j and the targeted bid b_j. There are two primitives to be estimated for each seller: the marginal cost c_j and the inverse conversion rate γ_{jr} (i.e., the average number of clicks per sponsored unit sold). I estimate these parameters by solving the first-order conditions for p_j and b_j in sellers' profit-maximization problems.

The expected profits in equation (9) rely on a seller's rational expectations regarding its market share and advertising payment. To approximate these expectations, I calculate averages across a large number of simulated search results to account for the uncertainty in auction outcomes and organic ranks. Below, I will first outline the estimation of the distributions of realized bids and organic ranks. Then, I will use these estimated distributions to simulate search results and construct sellers' expected profits. Finally, I will discuss the first-order conditions and the estimation of the price sensitivity parameter.

Bids and Organic Ranks In the model, each seller submits a targeted bid in each week, and sellers are ranked based on their stochastic realized bids. Different searches within a given week are independent realizations of the auction with the same set of targeted bids. I parameterize the realized bid as $\log(\tilde{b}_j) \overset{\text{i.i.d.}}{\sim} N\big(\log(b_j), \sigma_b^2\big)$, where σ_b quantifies the level of uncertainty in realized bids. To estimate the parameters $\boldsymbol{b}$ and σ_b in each market–week, I employ a Gibbs sampler with data augmentation (Tanner and Wong 1987). The resulting parameters can generate auction outcomes that closely match the observed realizations and align with the aggregate statistics of the winning bids reported by Amazon. Appendix D.2 describes the estimation procedure in detail.

I adopt a similar method to estimate the distribution of organic ranks. In each search, I assume that the platform's algorithm assigns each product a stochastic score drawn from a product-specific distribution, and then ranks all products based on their realized scores. These organic ranks may or may not align with the actual ranking of product utility. I parameterize the realized score as $\log(\tilde{o}_j) \overset{\text{i.i.d.}}{\sim} N(\log(o_j), \sigma_o^2)$ and use a Gibbs sampler to estimate $\boldsymbol{o} = (o_1, \ldots, o_J)$ in each market–week to match the actual realizations of organic ranks.

Expected Profits I simulate 2,000 realizations of auction outcomes and organic ranks for each market–week using the estimated distributions. For a given seller j and a specific rank $r \in \{0, 1, \ldots, R\}$, I construct search results where seller j always occupies the r-th sponsored position. When $r = 0$, the seller does not win any sponsored position. To achieve this, I hypothetically assign the r-th sponsored position to this focal seller in all realizations and allocate the remaining positions based on simulated organic ranks and the realized bids of other sellers. For each realization, I calculate the seller's market share using equation (4) and its advertising payment using equation (11). Then, I compute the average market share and advertising payment across all simulations, denoted as $s_j^r(\boldsymbol{p}, \boldsymbol{b})$ and $ad_j^r(\boldsymbol{p}, \boldsymbol{b})$, respectively.

Last, I derive seller j's probability of winning the r-th sponsored position given other sellers' bids, denoted as $q_j^r(\boldsymbol{b})$, and approximate the expected market share and advertising payment in equation (10) as:

$$s_j^e(\boldsymbol{p}, \boldsymbol{b}) = \sum_{r=0}^{R} q_j^r(\boldsymbol{b}) s_j^r(\boldsymbol{p}, \boldsymbol{b}), \;\; ad_j^e(\boldsymbol{p}, \boldsymbol{b}) = \sum_{r=1}^{R} q_j^r(\boldsymbol{b}) ad_j^r(\boldsymbol{p}, \boldsymbol{b}). \tag{12}$$

First-Order Conditions I parameterize the inverse conversion rate as $\gamma_{jr} = \exp(\gamma_j \lambda_r)$, where λ_r represents the fraction of consumers who consider the product in the r-th sponsored position, as defined in Section 5.1.1. γ_j captures product-specific conversion rates.[20] The number of clicks is at least one, as consumers who make a purchase always click on that product. I expect γ_j to be positive: a higher position in search results tends to attract more irrelevant traffic, as consumers in the exploration stage are more inclined to click on listings without completing a purchase.

[20] The parameter γ_j could also encompass sellers' dynamic incentives. For example, a seller might invest in sponsored product advertising to enhance its future organic ranks, although this possibility is beyond the scope of this paper.

Given any value of α, I estimate seller primitives, c_j and γ_j, by solving two first-order conditions. A seller's pricing decision reveals its margin and thus, the sum of its marginal cost and advertising cost. With a fixed margin, a higher bid corresponds to a higher conversion rate, and vice versa. Therefore, the seller's bidding decision helps distinguish between the two types of costs and identifies both parameters. For sellers who never appear in a sponsored position in a given week, I assume that they do not advertise and estimate only c_j for them.

Price Sensitivity I estimate the price sensitivity parameter α in equation (3) using a supply-side moment. Specifically, I utilize data from a survey conducted by Jungle Scout in 2022 (Jungle Scout 2023). This survey asked about 3,500 Amazon sellers of different sizes about their performance on the platform and found a median profit margin of 18% among these sellers. Given any value of α, I estimate seller primitives and compute each seller's profit margin in each market. Then, I search for the value of α such that the median profit margin in that market matches 18%.[21]

5.3 Platform

5.3.1 Model

The platform sets a commission rate τ to maximize a combination of short-term revenues, consumer surplus, and seller profits. Consistent with prior research on multi-sided platforms (e.g., Gutierrez 2022; Castillo 2023), Amazon can invest in user acquisition and retention to maximize its long-run profitability. Incorporating the welfare of consumers and sellers into the platform's objective approximates these long-term considerations that are not explicitly modeled.

As discussed in Section 2, Amazon also generates revenues from other sources besides commissions, such as storage and shipping fees charged to sellers. These revenue streams share a common feature: they serve as substitutes for advertising revenues. Section 3 illustrates how commissions can act as a substitute for advertising revenues, and the same intuition applies to these storage and shipping fees. An increase in these fees lowers sellers' margins and bids in the auctions, and consequently, reduces Amazon's advertising revenues. The substitution among different revenue-raising instruments is a critical aspect captured in my model. The commission rate represents just one element within a broader array of revenue-raising tools available to Amazon.

Let $CS(\tau)$ and $PS(\tau)$ represent the expected consumer surplus and seller profits under the

[21] One caveat is that the number 18% represents the median profit margin across all sellers on the platform and may not necessarily apply to each market. Unfortunately, the survey does not provide category-specific profit margins.

commission rate τ. The platform's short-term revenues include commissions $COM(\tau)$ and advertising revenues $AD(\tau)$. Appendix E.1 fleshes out these terms. Given the set of sponsored positions, the platform chooses τ to maximize the following objective function:

$$\max_{\tau} \Pi(\tau) = COM(\tau) + AD(\tau) + \mu\big(CS(\tau) + PS(\tau)\big). \tag{13}$$

Here, $\mu \geq 0$ is a nonnegative weight that reflects the platform's consideration of the welfare of other market participants relative to its short-term revenues.

In the model, I do not assume that the platform optimally selects the set of sponsored positions. The number of potential combinations of sponsored positions is vast,[22] making it implausible for the observed choice to be entirely optimal. In reality, the arrangement of sponsored positions on Amazon has been constantly evolving in recent years, likely due to ongoing experimentation.[23] In Section 7.4, I explore alternative numbers of sponsored positions and find that the platform's optimal number of sponsored positions is slightly higher than the current number.

5.3.2 Estimation

I estimate the platform's weight on the welfare of other market participants, μ, using its optimality condition in equation (13). While the model assumes a uniform commission rate, in practice, this rate varies from 8% to 20% across different markets. Let $\tau^{m,0}$ denote the observed commission rate in market m. In most markets in my sample, $\tau^{m,0}$ is either 15% or 17%. The size-weighted average commission rate across these markets, denoted as τ^0, is 15.6%. I assume that the platform chooses an average commission rate, which is then proportionally translated into market-specific commission rates. Specifically, when the platform sets an average commission rate τ, the actual commission rate in market m is calculated as $\tau^m = \frac{\tau^{m,0}}{\tau^0}\tau$.[24]

I maintain the set of sponsored positions and vary the commission rate τ. For each market–week, I solve for the new equilibrium across various values of τ and calculate consumer surplus, seller profits, commissions, and advertising revenues in the new equilibrium. To determine the weight parameter μ, I compute the derivative of the objective function in equation (13) with respect to τ and set it equal to zero at the observed rate $\tau^0 = 15.6\%$.

[22] With 60 positions in search results, there are 2^{60} possible sets of sponsored positions, leading to approximately 10^{18}.

[23] As demonstrated in Appendix E.2, there exist deviations in the set of sponsored positions that can increase consumer surplus, seller profits, and platform revenues simultaneously, benefiting the platform regardless of its objective.

[24] In Appendix F.4, I explore an extension where Amazon has the flexibility to implement market-specific commission rates. All results are qualitatively the same as the baseline results.

6 Results

6.1 Parameter Estimates

Figure 3 presents key parameter estimates. Panel A illustrates the fraction of consumers considering each position in search results. The different lines represent markets with varying degrees of search frictions.[25] In the absence of search frictions, consumers would consider all products, resulting in a flat line in the figure. However, the observed downward-sloping lines indicate a gradual decline in consumer consideration from the top position to the bottom position, which supports the presence of search frictions and is consistent with the finding in Section 4.2.

In the median market, 47% of consumers consider a product in the middle of the search results, while 22% consider the last product.[26] A median consumer has 28 products in her consideration set, corresponding to seven rows of products on a desktop or laptop browser.[27] Typically, consumers scroll down the screen multiple times, skip most products based on the information in the listings, and click on only several listings to inspect their product pages. Therefore, the number 28 should be interpreted as the upper limit on the page position that a consumer may consider rather than the number of products that she carefully examines.

Panel B of Figure 3 compares product utility across different positions in search results. When estimating product utility based on consumers' revealed preferences, it is crucial to account for the position effect. Even if products have the same utility, a product located higher in search results is considered by more consumers and captures a larger market share. A model assuming a homogeneous consideration set would overestimate the utility of products located in higher positions. The model in Section 5.1.1 explicitly incorporates consumers' heterogeneous consideration sets, thereby eliminating the mechanical correlation between positions and utility.

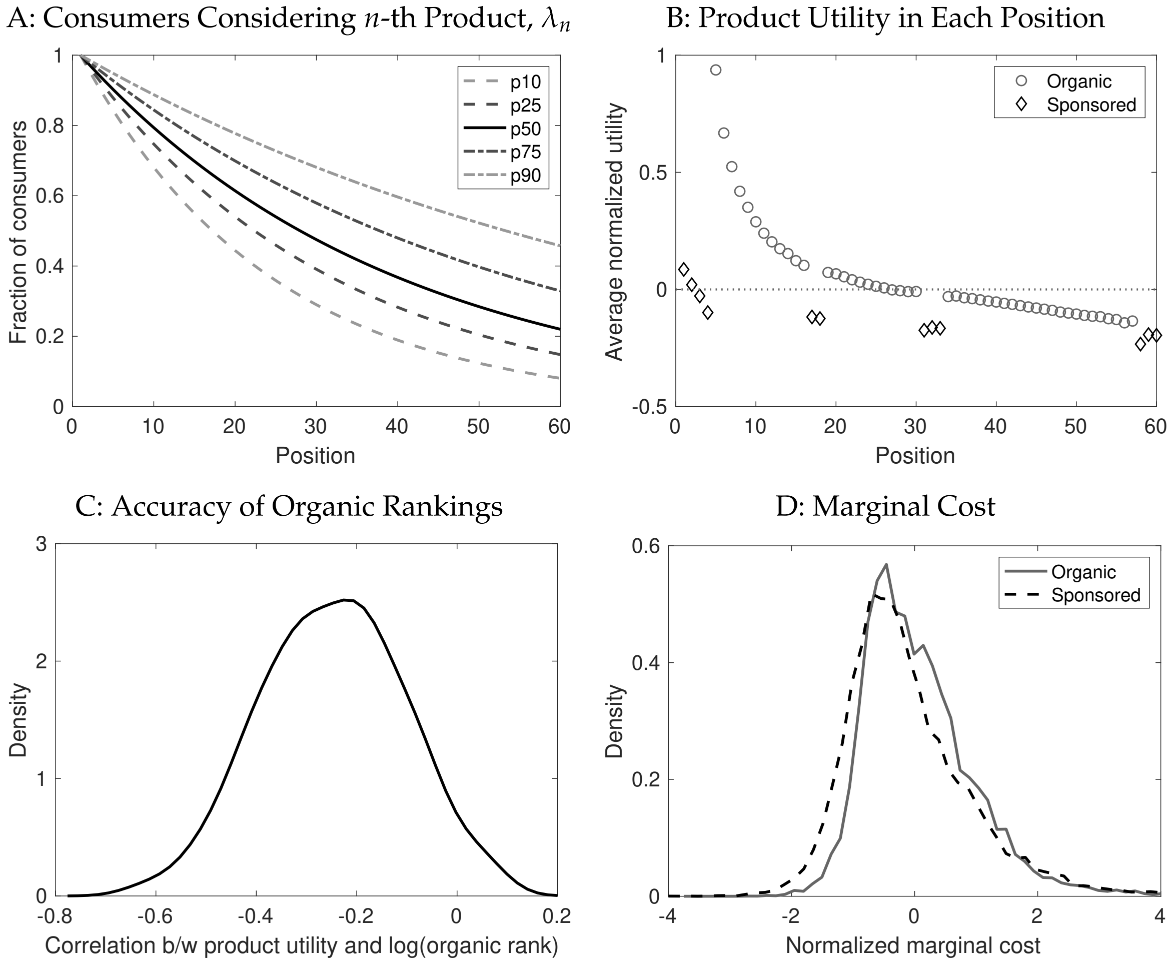
To construct this figure, I normalize the estimated product utility, $\hat{\delta}_{jt}$, to have a zero mean and unit variance across all products within a market–week. The figure illustrates the average normalized utility of products in each position, pooling searches from all market–weeks together. This comparison reveals two patterns. First, in organic positions, there exists a clear correlation be-

[25] To account for sampling errors, I utilize an empirical Bayes procedure to shrink the market-specific parameter $\hat{\beta}$ towards its mean across markets. The details of this procedure are provided in Appendix C.3.

[26] My analysis focuses on the first page of search results, with products on subsequent pages considered as part of the outside option. Some of the 22% may represent consumers who consider products beyond the first page. According to Amazon's data, 81% of clicks occur on the first page of search results (Baker 2018).

[27] Ursu (2018) utilizes an experiment on Expedia to examine the impact of ranks on consumer search and finds that a median consumer considers 34 products. Lam (2023) adopts an identification strategy similar to that in my paper and reports that a median consumer sees 15–20 products in the "Home & Kitchen" category on Amazon.

Figure 3: Estimation Results



Notes: This figure presents key estimation results. In Panel A, I calculate the fraction of consumers whose consideration sets contain the product in the n-th position in the search results, $\lambda_n = 1/\exp(\beta(n-1))$, based on the estimated parameter $\hat{\beta}$ in each market. Different lines correspond to different percentiles of λ_n across markets. In Panel B, I normalize the estimated product utility $\hat{\delta}_{jt}$, as defined in Section 5.1.1, to have a zero mean and unit variance across all products in a given market–week. The figure displays the average normalized utility of products in each position across all searches, separately for sponsored and organic positions. Panel C shows the distribution of the correlation between estimated product utility $\hat{\delta}_{jt}$ and the log of organic ranks in each market, serving as a measure for the accuracy of organic rankings. Panel D depicts the distribution of estimated marginal costs, separately for the top four sponsored and four organic products in all searches. Marginal costs are normalized to have a zero mean and unit variance across all products in a given market–week

tween higher rankings and higher average utility, and this relationship is monotonic. The average utility declines rapidly in the initial few positions. For instance, the average utility for products in the first organic position is 0.94 standard deviations higher than the average utility of all products. This difference diminishes to 0.35 standard deviations for products in the fifth organic position.

Second, sponsored products exhibit significantly lower average utility compared to organic counterparts in similar locations. For instance, products in the first sponsored position, which is also the first position on the entire page, have slightly higher utility than an average product but notably fall short of the utility of an average product in the first organic position. Sponsored products located beyond the third position have lower average utility than an average product. These observations indicate that, on average, organic rankings on Amazon effectively reflect product utility, while sponsored products tend to exhibit lower utility.

In Panel C of Figure 3, I calculate the correlation between estimated product utility and the log of organic ranks for each market. This measure helps assess the accuracy of organic rankings in each market. The figure shows a median correlation of -0.25, which suggests that while organic rankings, on average, reasonably reflect product utility, they are far from perfect. Importantly, there is considerable variation across markets, with the correlations ranging from -0.65 to 0.1. I further examine this heterogeneity in Section 7.3.

Panel D of Figure 3 depicts the distribution of normalized estimated marginal costs for the top eight products in search results, separately for sponsored and organic products. Top sponsored products have lower average marginal costs. However, this cost advantage does not translate into lower prices. As shown in Panel D of Figure 2, prices for top sponsored products are 16% higher than top organic ones. This price difference is primarily attributed to the additional advertising payment required for securing sponsored placements.

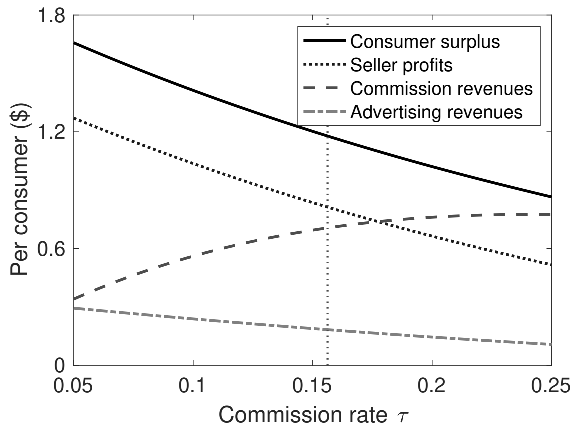
In Figure 4, I present the estimation of the platform's weight on the welfare of other market participants. As the platform increases its commission rate, consumer surplus and seller profits decrease, while commission revenues increase in a concave manner. Notably, advertising revenues decline as the commission rate rises. This is because a higher commission rate reduces sellers' willingness to pay for sponsored positions, thereby lowering advertising revenues for the platform. Using these functions, I determine the parameter μ to be 0.116 from the platform's first-order condition.[28]

6.2 Model Fit

In Figure 5, I validate the model estimates by comparing them with industry reports. Panel A focuses on three metrics commonly used in the advertising industry: (i) click-through rate, calcu-

[28] This value appears relatively low compared to estimates in the literature. For instance, Castillo (2023) finds that Uber acted as if it were maximizing rider surplus in 2017, and Gutierrez (2022) finds that Amazon placed a median weight of 1.11 on consumer welfare and 0.39 on seller welfare in 2018–2020. It is important to note that the settings are different, and the weights likely decline as a platform becomes more established.

Figure 4: Outcomes Under Varying Commission Rates



Notes: This figure displays consumer surplus, seller profits, commission revenues, and advertising revenues under different commission rates. I maintain the set of sponsored positions at the observed configuration and vary the commission rate. For each commission rate, I solve for the new equilibrium in each market–week, compute the outcome measures, and aggregate across market–weeks using market sizes.

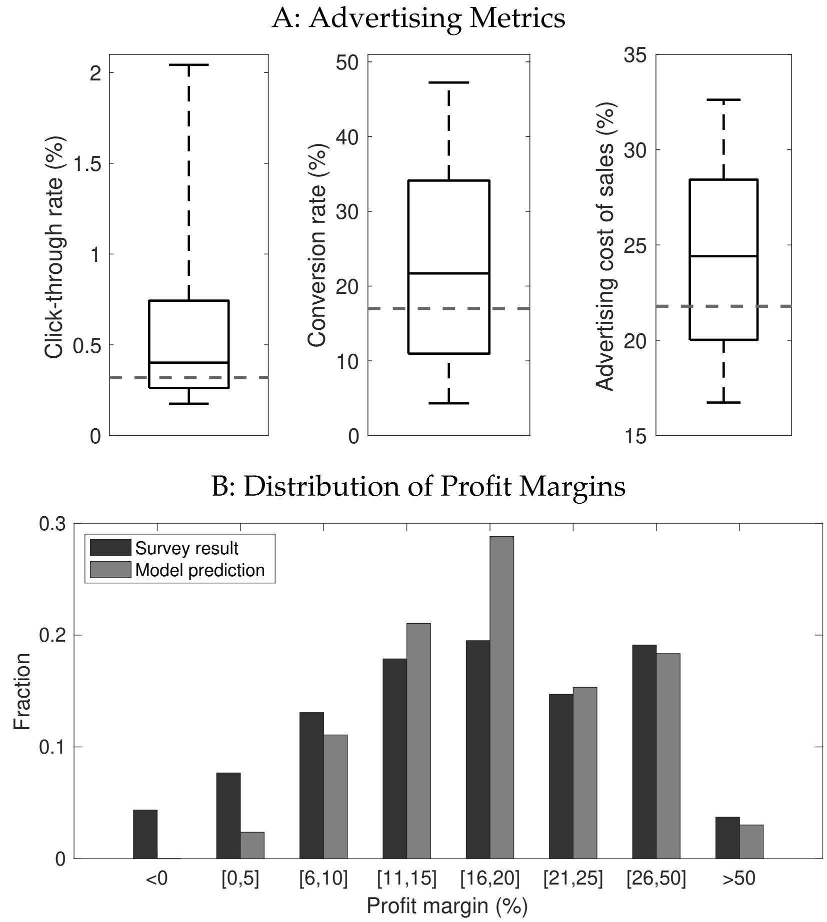
lated as total clicks divided by total impressions; (ii) conversion rate, calculated as total sponsored units sold divided by total clicks; and (iii) advertising cost of sales, calculated as the total advertising payment divided by total revenues from sponsored units sold. I compute these metrics for each market based on my model estimates and compare them to industry benchmarks. Remarkably, the model's predictions closely align with the industry reports, despite the estimation not directly targeting these metrics. For instance, the model predicts an average click-through rate of 0.40% in a median market, which closely matches the industry estimate of 0.32%.[29]

The Jungle Scout survey provides data on the distribution of sellers across eight profit margin intervals (Jungle Scout 2023).[30] Using estimates generated by my model, I calculate the profit margin for each seller and determine the distribution across these intervals. Panel B of Figure 5 compares the predicted distribution from the model with that documented in the survey. Overall, the model's predictions reasonably align with the survey results. For instance, the survey indicates that 17.9% of sellers have profit margins between 11% and 15%, while the model predicts a similar number of 21.0%. However, the survey reports that 4.4% of sellers experience negative profits, a scenario that the model precludes by construction. These sellers might have dynamic incentives, such as setting low prices to improve future organic ranks. These incentives could be reflected in the estimates of their marginal costs, which might be lower than the actual costs.

[29] In the second quarter of 2022, the average click-through rate on Amazon was 0.32%, conversion rate was 17.0%, and advertising cost of sales was 21.8% (Pacvue 2022).

[30] These intervals are <0%, 0–5%, 6–10%, 11–15%, 16–20%, 21–25%, 26–50%, and >50%.

Figure 5: Model Validation



Notes: This figure validates model estimates by comparing them with industry benchmarks. Panel A compares three metrics commonly used in the advertising industry predicted by model estimates and from industry reports: (i) click-through rate, calculated as total clicks divided by total impressions; (ii) conversion rate, calculated as total sponsored units sold divided by total clicks; and (iii) advertising cost of sales, calculated as the total advertising payment divided by total revenues from sponsored units sold. Each box in Panel A represents the distribution of one of these metrics across markets, with the middle line indicating the median, the edges of a box indicating the 25th and 75th percentiles, and the outer lines indicating the 5th and 95th percentiles. The dashed line represents the industry estimate from Pacvue in the second quarter of 2022 (Pacvue 2022). Panel B compares the distribution of profit margins for sellers. The Jungle Scout survey in 2022 provides data on the distribution of sellers across eight profit margin intervals (Jungle Scout 2023). I compute the profit margin for each seller in my sample using model estimates and calculate the corresponding distribution. The figure compares the predicted distribution from the model with that documented in the survey.

7 Welfare Effects of Sponsored Product Advertising

In this section, I quantify the welfare effects of sponsored product advertising on Amazon. I start by evaluating the overall impacts on consumers, sellers, and the platform. Next, I break down the change in consumer surplus into different components and examine the heterogeneity in the welfare effects across markets. Last, I explore alternative numbers of sponsored positions.

7.1 Overall Effects

I consider a counterfactual scenario where all sponsored positions are replaced with organic listings. To simulate this scenario, I generate 2,000 realizations of organic ranks and populate the search results with the top $N = 60$ products. Then I find the new equilibria in each market under varying commission rates and aggregate across markets using market sizes. I consider two cases that differ in how the platform reacts to the removal of sponsored positions: one where the platform maintains a fixed commission rate at 15.6%, and another where the platform adjusts its commission rate to optimize its objective. By comparing this scenario to the status quo, I can assess the welfare effects of *removing* sponsored product advertising.

Table 3 summarizes the welfare effects resulting from the removal of sponsored product advertising. The first three columns report consumer surplus, seller profits, commission revenues, and advertising revenues under the status quo and counterfactual scenario with a fixed commission rate. Eliminating sponsored positions leads to a 12.6% increase in consumer surplus and a 14.1% rise in seller profits. While there is a moderate uptick in commission revenues due to increased sales, the platform experiences a substantial 14.8% decline in total revenues because of the loss of all advertising revenues. The total surplus, encompassing consumer surplus, seller profits, and platform revenues, increases by 4.6%. These findings suggest that when the commission rate remains fixed, the presence of advertising on Amazon has an overall negative impact on consumers and results in a significant transfer from sellers to the platform.

Table 3: Welfare Effects of Removing Sponsored Product Advertising

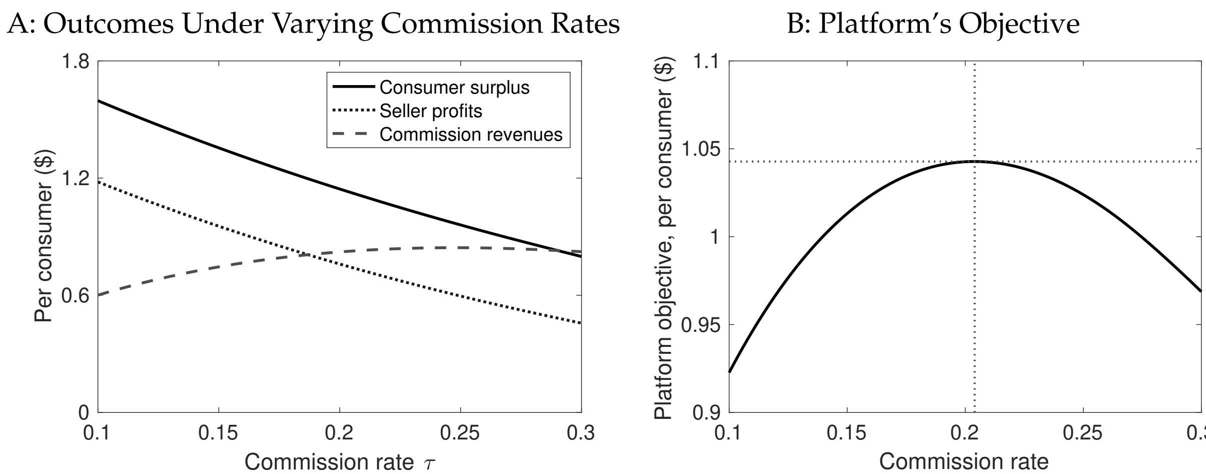
	(1)	(2)	(3)	(4)	(5)
	Status Quo	Counterfactual: Without Sponsored Product Advertising			
		Fixed Commission		Endogenous Commission	
	Level ($)	Level ($)	Change (%)	Level ($)	Change (%)
Consumer Surplus	1.18	1.33	12.6%	1.13	-4.3%
Seller Profits	0.81	0.93	14.1%	0.75	-8.4%
Platform Revenues	0.89	0.76	-14.8%	0.83	-7.2%
Commission Revenues	0.71	0.76	7.4%	0.83	16.9%
Advertising Revenues	0.18	0.00	-100.0%	0.00	-100.0%
Total Surplus	2.88	3.01	4.6%	2.70	-6.3%

Notes: This table compares outcome measures between two scenarios: the status quo with sponsored positions (column 1) and a counterfactual scenario where all sponsored positions are replaced with organic listings (columns 2-5). In columns 2 and 3, the commission rate remains fixed at the observed rate of 15.6% in the status quo. In columns 4 and 5, the platform re-optimizes the commission rate to maximize its objective. Columns 1, 2, and 4 present the outcome measures in dollars per consumer, while columns 3 and 5 calculate the percentage changes from the status quo to the respective counterfactual scenarios. The results are aggregated across market–weeks using market sizes. Figure 6 illustrates the optimal commission rate in the counterfactual scenario.

The finding that removing sponsored positions leads to an increase in consumer surplus aligns with the estimation results. As demonstrated in Panel B of Figure 3, organic ranks on Amazon can predict product utility on average. In contrast, sponsored products consistently exhibit lower utility than their organic counterparts when occupying similar positions. Despite sponsored products having lower marginal costs, the additional advertising payment effectively negates their cost advantages. Furthermore, the overlap between sponsored and organic products, as depicted in Appendix Figure A.10, exacerbates the loss of consumer surplus.[31]

In the scenario where the platform has the flexibility to adjust its commission rate following the removal of sponsored positions, Figure 6 reveals an optimal commission rate of 20.3%, a significant increase over the current rate of 15.6%. This adjustment arises because, in the presence of sponsored product advertising, a higher commission rate results in lower advertising revenues, which incentivizes the platform to maintain a relatively low commission rate. When this incentive is removed, the model predicts an increase in the commission rate.

Figure 6: Optimal Commission Rate Without Sponsored Product Advertising



Notes: This figure identifies the optimal commission rate in the counterfactual scenario where all sponsored positions are replaced with organic listings. Panel A plots consumer surplus, seller profits, and commission revenues under different commission rates. The results are aggregated across market–weeks using market sizes. Panel B constructs the platform's objective, a linear combination of revenues, consumer surplus, and seller profits, using the weight estimated in Section 6.1. The vertical dotted line indicates the commission rate that maximizes the platform's objective.

The last two columns of Table 3 illustrate the welfare effects with the new commission rate. Compared to the status quo, consumer surplus decreases by 4.3%, and seller profits decline by

[31]In the baseline analysis, I assume that the second appearance of the same product in search results does not receive an independent idiosyncratic preference shock when calculating consumer surplus. Thus, any duplication of product listings is a waste of space from consumers' perspective. In Appendix G.3, I explore an alternative assumption where the second appearance also receives its independent preference shock. The welfare results are qualitatively the same.

8.4%. While the additional commissions collected by the platform can partially offset the loss in advertising revenues, its total revenues still decrease by 7.2%. Overall, the total surplus drops by 6.3%, primarily driven by the decline in total transactions on the platform.

The reversal in the welfare comparison highlights the importance of selecting an appropriate counterfactual scenario when evaluating the effects of sponsored product advertising. Merely comparing the status quo to a scenario without advertising, while keeping all else constant, might not accurately portray Amazon's response. If advertising were prohibited, Amazon would have a strong economic motive to raise other fees. Therefore, it is crucial to compare different revenue-raising methods. My model compares introducing sponsored positions and raising the commission rate. The results indicate that the former not only raises more revenues for the platform but also exerts a smaller negative impact on consumers and sellers. As discussed in Section 5.3.1, the commission rate in my model can represent a broader array of instruments available to the platform. While Amazon might not frequently modify its commission rate in reality, it does adjust other fees, such as storage and shipping fees, on an annual basis.

How should we interpret this result? While in theory, a platform could maximize its revenues through tailored product placement and individualized seller fees, such an approach would be impractical in real-world scenarios due to the extensive information required. Instead, platforms typically have two options for allocating positions in search results: relying on existing information with sellers paying a uniform fee or conducting auctions for a subset of positions with winners paying their bids. The auction mechanism elicits private information from sellers and allows the platform to engage in price discrimination against sellers, often resulting in higher platform revenues. My finding suggests that this enhanced efficiency can also benefit other market participants. Therefore, the adoption of advertising can lead to a more efficient and mutually beneficial marketplace for all participants involved.

7.2 Breaking Down the Change in Consumer Surplus

Section 7.1 evaluates the overall impact of removing sponsored product advertising. To gain a deeper understanding of the various factors at play, I break down the overall change in consumer surplus into several components by analyzing three intermediate cases in addition to the status quo and the counterfactual scenario. These cases are defined as follows:

- **Status quo**: Commission rate is 15.6%. Sponsored products pay advertising payment.

- **Case 1 (no ad payment)**: Commission rate is 15.6%. The product arrangement is the same as in the status quo, but sponsored products do not need to pay advertising payment.
- **Case 2 (equal costs)**: Marginal costs of sponsored products are shifted by a market-specific constant so that the average marginal costs for sponsored and organic products within each market are equal. All other conditions remain the same as in Case 1.
- **Case 3 (no sponsored products)**: Commission rate is 15.6%. There are only organic results.
- **Counterfactual (endogenous commission rate)**: Commission rate is 20.3%. There are only organic results.

I denote the percentage change in consumer surplus from the status quo to Case 1 as the *no-ad-payment effect*, which reflects the impact of removing the pass-through from advertising payment to prices. Similarly, I define the change from Case 1 to Case 2 as the *cost effect*, capturing the impact of the cost differences between sponsored and organic products. Transitioning from Case 2 to Case 3 represents a shift in the set of products available to consumers. Given that any cost differences have already been eliminated, the change in consumer surplus can be interpreted as the *quality effect*. Last, the change from Case 3 to the counterfactual scenario with an endogenous commission rate is labeled the *commission effect*, quantifying the impact of a higher commission rate. Panel A of Figure 7 summarizes the definitions of these effects.

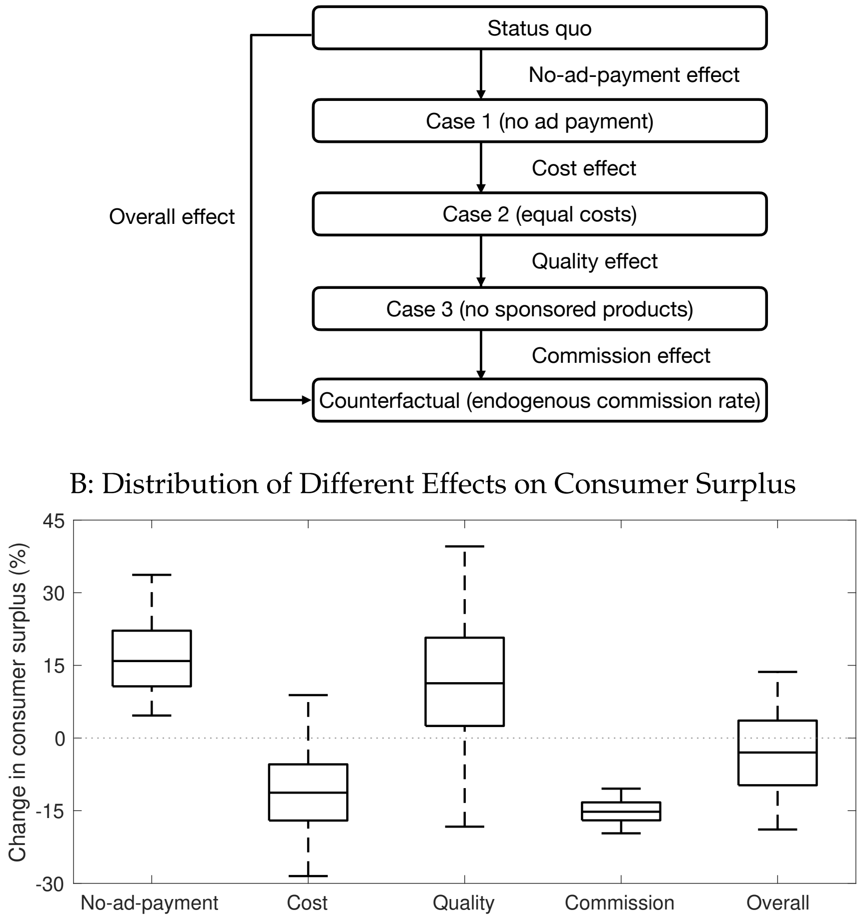
Panel B of Figure 7 illustrates the distribution of these four effects across markets.[32] The no-ad-payment effect is consistently positive, as the removal of advertising payment always lowers prices and benefits consumers. In contrast, the commission effect stands out as highly negative, indicating a significant distortion caused by a higher commission rate. Meanwhile, the cost effect tends to be negative, aligning with the finding in Panel C of Figure 3 that sponsored products often have lower marginal costs than organic products.

Notably, the quality effect is generally positive but exhibits substantial variation across markets. In about 20% of the markets, this effect turns negative. This suggests that moving from the existing mix of products to exclusively organic results may reduce the average quality of products available to consumers in some cases. When all these effects are combined, they result in a negative overall effect in the majority of markets. However, in 38% of the markets, the overall effect is positive. In the following section, I will explore this heterogeneity in more detail.

[32] Appendix Figure A.13 provides a similar breakdown for seller profits, platform revenues, and total surplus.

Figure 7: Decomposition of Welfare Effects of Removing Sponsored Product Advertising

A: Summary of All Scenarios



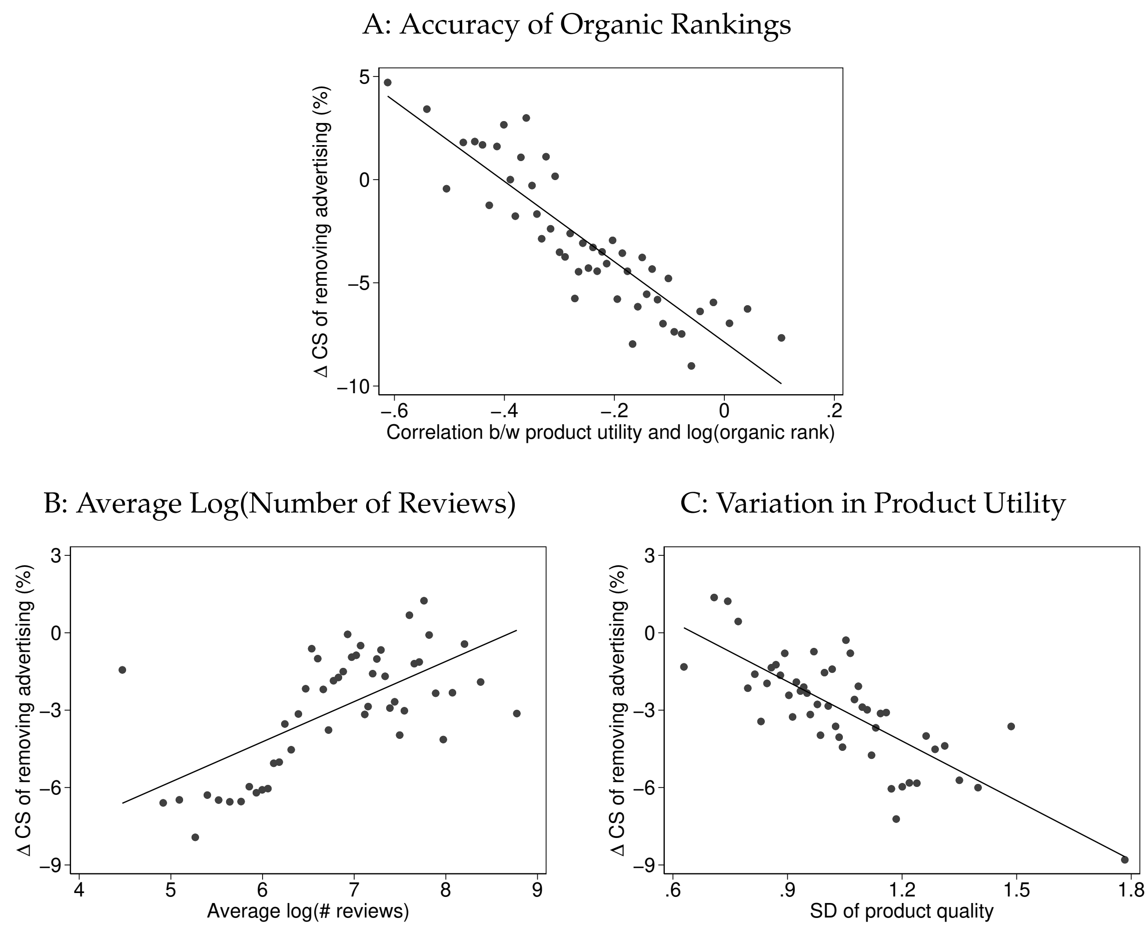
Notes: This figure decomposes the aggregate effect of removing sponsored product advertising on consumer surplus. Besides the status quo and counterfactual scenario, I consider three intermediate cases. In the first case, the commission rate and product arrangement are the same as in the status quo, but sponsored products do not need to pay advertising payment. In the second case, the marginal costs of sponsored products are shifted by a market-specific constant so that the average marginal costs for sponsored and organic products within each market are equal. All other conditions remain the same as in the first case. In the third case, the commission rate is unchanged, but all positions display organic results. The percentage change in consumer surplus from the status quo to the first case is defined as the *no-ad-payment effect*, from the first to the second case defined as the *cost effect*, from the second to the third case defined as the *quality effect*, and from the third case to the counterfactual scenario with an endogenous commission rate defined as the *commission effect*. Panel A summarizes the definitions of these effects. In Panel B, each box presents the distribution of one of the four effects and the overall effect across market–weeks, with the middle line indicating the median, the edges of a box indicating the 25th and 75th percentiles, and the outer lines indicating the 5th and 95th percentiles. Appendix Figure A.13 plots the decomposition of the aggregate changes in seller profits, platform revenues, and total surplus.

7.3 Heterogeneity Across Markets

Section 7.2 reveals substantial variation across markets in the effect of removing advertising on consumers. This section further explores the factors that contribute to this heterogeneity.

As highlighted in Section 3, a primary benefit of advertising is to enable sellers to reveal information about their products when organic rankings fail to accurately reflect product utility. Panel A of Figure 8 confirms this intuition, showing a positive correlation between the impact of eliminating advertising on consumers and the accuracy of organic rankings within a given market. The accuracy is measured as the correlation between product utility and organic ranks, as illustrated in Panel C of Figure 3. In markets where organic rankings are less accurate, removing sponsored products tends to be more harmful to consumers.

Figure 8: Heterogeneity in Welfare Effects on Consumer Surplus Across Markets



Notes: This figure examines the heterogeneity in the welfare effects on consumer surplus across markets. The y-axis represents the percentage change in consumer surplus from the status quo with sponsored positions to the counterfactual scenario without sponsored positions under endogenous commission rates, which corresponds to the "overall effect" plotted in Figure 7. In Panel A, the x-axis represents the correlation between estimated product utility and the log of organic ranks in a market, a measure of the accuracy of organic rankings. A more negative correlation indicates more accurate organic rankings. In Panel B, the x-axis represents the log of the number of reviews averaged across all products in a market. In Panel C, the x-axis represents the standard deviation of estimated product utility, $\hat{\delta}_{jt}$, across all products in a market, a measure of product differentiation. All panels show binned scatter plots using 50 bins.

I further examine market-level characteristics that could predict the accuracy of organic rankings in a market. Intuitively, organic rankings could prove less accurate in markets with many new products lacking established sales records on the platform. Panel B of Figure 8 illustrates a positive correlation between the welfare impact and the log of total reviews averaged across all products in a given market. When the average number of reviews is small, the platform may lack the necessary information to generate accurate organic rankings, and sponsored product advertising can help bridge this information gap. Therefore, removing advertising in such situations may have a more negative effect on consumers.

Organic rankings may also be less accurate in markets with significant product differentiation. In contrast, markets with relatively homogeneous or standardized products likely have limited information not reflected in their organic rankings. To investigate this hypothesis, I calculate the standard deviation of estimated product utility, $\hat{\delta}_{jt}$, across all products within a given market, which serves as a proxy for product differentiation. Panel C of Figure 8 demonstrates a negative correlation between the impact of removing sponsored product advertising and the degree of product differentiation, providing support for this hypothesis.

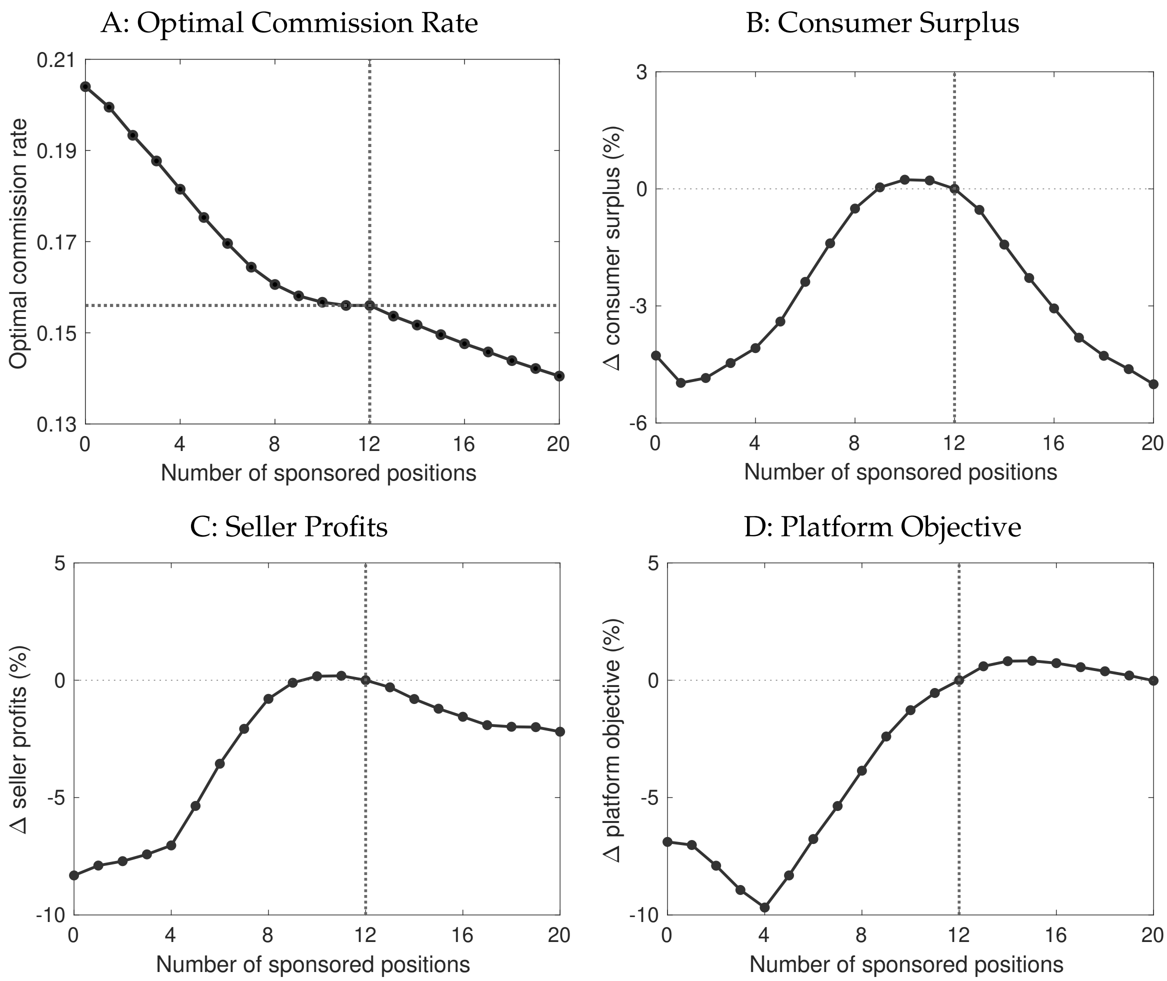
This analysis emphasizes that the effect of sponsored product advertising on consumer surplus varies with market conditions. Advertising tends to generate greater gains in markets characterized by fewer established products and higher levels of product differentiation.

7.4 Alternative Numbers of Sponsored Positions

For policymakers concerned about the rise in sponsored product advertising, completely removing sponsored products may not be a practical or desirable policy choice. Instead, a more pragmatic approach might involve an intermediate solution. In this final analysis, I examine the effects of altering the number of sponsored positions in search results.

Starting from the baseline scenario with 12 sponsored positions, I explore two directions. First, I gradually replace sponsored positions with organic ones, starting from the lowest one and progressing to the highest. This approach represents an intermediate step between maintaining the status quo and eliminating all sponsored positions. Second, I incrementally introduce more sponsored positions in the middle of search results, starting by substituting the 5th organic position with a sponsored one and continuing down to the 12th one. This exercise allows me to speculate on potential future strategies for Amazon. Panel A of Figure 9 shows that as the number of sponsored positions increases, the optimal commission rate consistently decreases.

Figure 9: Welfare Effects Under Varying Numbers of Sponsored Positions



Notes: This figure examines the welfare effects of altering the number of sponsored positions in search results. I change the number of sponsored positions in two directions: (i) a gradual removal of sponsored positions, starting from the lowest one and progressing to the highest; and (ii) a gradual addition of more sponsored positions in the middle, starting by substituting the 5th organic position with a sponsored one and continuing down to the 12th one. Panel A illustrates the optimal commission rate under each number of sponsored positions. Panels B to D depict the aggregate changes in consumer surplus, seller profits, and platform revenues relative to the status quo with 12 sponsored positions. Appendix Figure A.14 presents the results under a fixed commission rate.

Panel B of Figure 9 reveals a nonmonotonic relationship between consumer surplus and the number of sponsored positions in search results. Starting from the status quo, as the number of sponsored positions decreases, consumer surplus initially sees a slight increase. However, as more sponsored positions are removed, consumer surplus begins to decline. In the other direction, when additional sponsored positions are introduced, consumer surplus gradually decreases. A similar trend exists for seller profits, as illustrated in Panel C.[33]

[33] Appendix Figure A.14 presents the results under a fixed commission rate. As more sponsored positions are added,

The nonmonotonic relationship underscores that sponsored product advertising does not always benefit consumers and sellers. When comparing the status quo to a scenario with no sponsored positions at all, the findings in Section 7.1 indicate that both consumers and sellers generally benefit from the presence of advertising. However, introducing more sponsored positions, while boosting the platform's revenues, can hurt consumers and sellers. Notably, both the consumer- and seller-optimal numbers of sponsored positions are slightly lower than the current number. In contrast, as illustrated in Panel D of Figure 9, the number of sponsored positions that maximizes the platform's objective is slightly higher than the current number, suggesting that the platform may have an incentive to introduce several more sponsored positions.

Hence, for policymakers primarily concerned with the welfare of consumers and sellers, a potential solution could involve implementing a limit on the number of sponsored positions in search results instead of completely banning sponsored product advertising. This approach aims to strike a balance that preserves the benefits of sponsored product advertising while safeguarding the interests of consumers and sellers.

8 Conclusion

In recent years, there has been a significant expansion in sponsored product advertising across numerous digital retail platforms offering a wide array of products and services. These platforms, which traditionally relied primarily on sales commissions for revenue, have now embraced advertising as an essential component of their revenue streams. This shift has brought about a transformation in their business model, with potentially profound implications for all market participants.

This paper introduces a comprehensive framework for analyzing the welfare effects of sponsored product advertising. It begins with a stylized example that highlights the theoretical ambiguity regarding the welfare impact. The empirical analysis leverages extensive datasets on Amazon searches, purchases, and advertising auction bids. Using estimates derived from an equilibrium model that integrates consumers, sellers, and the platform, I find that, under a fixed commission rate, removing sponsored product advertising generally benefits both consumers and sellers on Amazon. These effects stem from the removal of low-utility sponsored results and the elimination of the pass-through from advertising expenses to prices.

However, if sponsored product advertising were to be prohibited, the platform would likely

consumer surplus and seller profits consistently decrease, and platform revenues initially increase and then level off.

respond by increasing its commission rate. The distortion caused by a higher commission rate would outweigh the benefits of removing advertising and result in a negative overall effect for consumers and sellers. These findings highlight the importance of considering various factors, including the platform's strategic response, when evaluating the implications of sponsored product advertising. For policymakers concerned with the welfare of consumers and sellers, implementing a limit on the number of sponsored positions emerges as a more suitable policy target than instituting an outright ban on sponsored product advertising.

I want to highlight several limitations in my analysis, which could motivate further research. First, my main dataset covers a relatively short period, which limits my ability to investigate longer-term effects like seller entry and exit dynamics. In Appendix F.5, I present suggestive evidence using a small dataset collected in March 2024, roughly 20 months after the initial sample period. I find that while advertising may help products with lower visibility ascend organic rankings, it could also lead to the exits of high-ranking products. Future studies with extended time coverage could provide a more comprehensive insight into these effects.

Second, my analysis relies on observational data at an aggregate level, necessitating specific assumptions in modeling and estimation. For example, my model assumes perfect consumer awareness of product utility. In Appendix F.1, I explore some extensions that relax this assumption and find similar results. Researchers with access to microdata on consumers' search behavior or experimental variations on product arrangement could estimate a more flexible model. Additionally, I observe only a subset of keywords for which products appear. For keywords that are not observed, I have to make simplifying assumptions regarding consumer demand and seller competition in my model estimation.

Finally, my analysis abstracts from several practices employed by Amazon aimed at enhancing the quality of sponsored results, such as relevance scores used to rank sellers in auctions and personalized sponsored results. Not incorporating these aspects in my analysis may lead to underestimating the potential benefits of sponsored product advertising.

References

Amazon (2017). 2016 Amazon.com Annual Report. https://s2.q4cdn.com/299287126/files/doc_financials/annual/2016-Annual-Report.pdf. Accessed April 4, 2024.

Amazon (2024). 2023 Amazon.com, Inc. Form 10-K. https://s2.q4cdn.com/299287126/files/doc_financials/2023/q4/c7c14359-36fa-40c3-b3ca-5bf7f3fa0b96.pdf. Accessed April 4, 2024.

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

Athey, S. and Nekipelov, D. (2011). A Structural Model of Sponsored Search Advertising Auctions. Working Paper.

Baker, L. (2018). Amazon's Search Engine Ranking Algorithm: What Marketers Need to Know. *Search Engine Journal*. https://www.searchenginejournal.com/amazon-search-engine-ranking-algorithm-explained/265173. Accessed April 4, 2024.

Blake, T., Nosko, C., and Tadelis, S. (2015). Consumer Heterogeneity and Paid Search Effectiveness: A Large-Scale Field Experiment. *Econometrica*, 83(1):155–174.

Castillo, J. C. (2023). Who Benefits From Surge Pricing? *Econometrica*, Forthcoming.

Chan, T. Y. and Park, Y.-H. (2015). Consumer Search Activities and the Value of Ad Positions in Sponsored Search Advertising. *Marketing Science*, 34(4):606–623.

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

Choi, H. and Mela, C. F. (2019). Monetizing Online Marketplaces. *Marketing Science*, 38(6):948–972.

Edelman, B., Ostrovsky, M., and Schwarz, M. (2007). Internet Advertising and the Generalized Second-Price Auction: Selling Billions of Dollars Worth of Keywords. *American Economic Review*, 97(1):242–259.

Errico, M. and Lashkari, D. (2022). The Quality of US Imports and the Consumption Gains From Globalization. Working Paper.

Farronato, C., Fradkin, A., and MacKay, A. (2023). Self-Preferencing at Amazon: Evidence From Search Rankings. Working Paper 30894, National Bureau of Economic Research.

Federal Trade Commission (2023). FTC Sues Amazon for Illegally Maintaining Monopoly Power. https://www.ftc.gov/news-events/news/press-releases/2023/09/ftc-sues-amazon-illegally-maintaining-monopoly-power. Accessed April 4, 2024.

Fowler, G. A. (2022). It's Not Your Imagination: Shopping on Amazon Has Gotten Worse. *Washington Post*. https://www.washingtonpost.com/technology/interactive/2022/amazon-shopping-ads. Accessed April 4, 2024.

Grennan, M. (2013). Price Discrimination and Bargaining: Empirical Evidence From Medical Devices. *American Economic Review*, 103(1):145–177.

Gutierrez, G. (2022). The Welfare Consequences of Regulating Amazon. Working Paper.

Hui, X. and Liu, M. (2022). Quality Certificates Alleviate Consumer Aversion to Sponsored Search Advertising. CESifo Working Paper.

Ilango, A. (2022). Sponsored Search: Theory and Evidence on How Platforms Exacerbate Product Market Concentration. Working Paper.

Instacart (2024). Letter to Shareholders, Q4 2023. https://investors.instacart.com/static-files/c858fabf-2254-4376-8018-2a3d60801dd4. Accessed April 4, 2024.

Jungle Scout (2023). State of the Amazon Seller 2022. https://www.junglescout.com/amazon-seller-report/2022-results. Access April 4, 2024.

Kaziukėnas, J. (2021). Amazon Tops Six Million Third-Party Sellers. *Marketplace Pulse*. https://www.marketplacepulse.com/articles/amazon-reaches-six-million-third-party-sellers. Accessed April 4, 2024.

Kaziukėnas, J. (2024a). Amazon Sales Top $700 Billion in 2023. *Marketplace Pulse*. https://www.marketplacepulse.com/articles/amazon-sales-top-700-billion-in-2023. Accessed April 4, 2024.

Kaziukėnas, J. (2024b). E-commerce Needs Walmart. *Marketplace Pulse*. https://www.marketplacepulse.com/articles/e-commerce-needs-walmart. Accessed April 4, 2024.

Lam, H. T. (2023). Platform Search Design and Market Power. Working Paper.

Lee, K. H. and Musolff, L. (2023). Entry Into Two-Sided Markets Shaped by Platform-Guided Search. Working Paper.

Long, F., Jerath, K., and Sarvary, M. (2022). Designing An Online Retail Marketplace: Leveraging Information From Sponsored Advertising. *Marketing Science*, 41(1):115–138.

Marx, L. M. and Shaffer, G. (2010). Slotting Allowances and Scarce Shelf Space. *Journal of Economics & Management Strategy*, 19(3):575–603.

Matsakis, L. (2022). Small Businesses Say They Are Hurt by Rising Costs to Advertise on Amazon. *NBC News*. https://www.nbcnews.com/tech/small-businesses-say-are-hurt-rising-costs-advertise-amazon-rcna16685. Accessed April 4, 2024.

Motta, M. and Penta, A. (2022). Market Effects of Sponsored Search Auctions. CEPR Discussion Paper No. DP17401.

Ostrovsky, M. and Skrzypacz, A. (2022). Pure-Strategy Equilibrium in the Generalized First-Price Auction. Working Paper.

Pacvue (2022). Q2 2022 CPC Report. https://pacvue.com/guides-reports/q2-2022-cpc-report/.

Sahni, N. S. and Nair, H. S. (2020). Does Advertising Serve as a Signal? Evidence From a Field Experiment in Mobile Search. *Review of Economic Studies*, 87(3):1529–1564.

Schein, A. I., Popescul, A., Ungar, L. H., and Pennock, D. M. (2002). Methods and Metrics for Cold-Start Recommendations. In *Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 253–260.

Sudhir, K. and Rao, V. R. (2006). Do Slotting Allowances Enhance Efficiency or Hinder Competition? *Journal of Marketing Research*, 43(2):137–155.

Sullivan, M. W. (1997). Slotting Allowances and the Market for New Products. *Journal of Law and Economics*, 40(2):461–494.

Sweeting, A. (2013). Dynamic Product Positioning in Differentiated Product Markets: The Effect of Fees for Musical Performance Rights on the Commercial Radio Industry. *Econometrica*, 81(5):1763–1803.

Tanner, M. A. and Wong, W. H. (1987). The Calculation of Posterior Distributions by Data Augmentation. *Journal of the American statistical Association*, 82(398):528–540.

Ursu, R. M. (2018). The Power of Rankings: Quantifying the Effect of Rankings on Online Consumer Search and Purchase Decisions. *Marketing Science*, 37(4):530–552.

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

Wang, Z. (2022). I'm an Amazon Advertiser. Here's How I Used My Ad Budget To Pull In $100 Million in Sales in 2021. https://www.businessinsider.com/amazon-advertiser-hero-cosmetics-breaks-down-how-it-spends-its-ad-budget-2022-5. Accessed April 4, 2024.

Weise, K. (2019). Amazon Knows What You Buy. And It's Building a Big Ad Business From It. *New York Times*. https://www.nytimes.com/2019/01/20/technology/amazon-ads-advertising.html. Accessed April 4, 2024.

Yao, S. and Mela, C. F. (2011). A Dynamic Model of Sponsored Search Advertising. *Marketing Science*, 30(3):447–468.

Appendix

A Details on Stylized Example

The stylized example presented in Section 3 illustrates the economic factors that influence the welfare implications of sponsored product advertising. It incorporates the essential elements in this environment: consumer search frictions, seller competition, and a profit-maximizing platform. In this section, I provide formal definitions of consumer demand and the auction rule and derives the equilibria.

Consider a market with two sellers, labeled $j = 1, 2$, each offering a single product with quality δ_j, a marginal cost c_j, and a price p_j. The platform's search results consist of two positions, displaying either organic or sponsored results. Without loss of generality, I assume that product 1 is ranked higher in the organic ranking, even though this ranking may not necessarily reflect the actual ranking of δ_j. This discrepancy can arise when the organic ranking fails to fully capture product quality. The platform collects a commission for each unit sold, calculated as a percentage of the product's price, denoted by $\tau \in [0, 1)$.

Consumer i's utility of purchasing product j is expressed as follows:

$$u_{ij} = \delta_j - \alpha p_j + \epsilon_{ij}, \tag{A.1}$$

where ϵ_{ij} is an idiosyncratic preference shock following a type I extreme value distribution. A fraction $\lambda \in (0, 1)$ of consumers consider only the product in the first position, while the remaining consumers consider products in both positions. This assumption approximates real-world search behavior, where consumers may face search costs and evaluate only a subset of products in search results. Such behavior motivates sellers to pursue a higher location in search results. Consumers choose either a product they consider or an outside option of not buying any product in the search results, whose mean utility is normalized to be 0.

Let $s_j(p_j)$ denote the market share of product j among consumers who consider only product j, and $s_j(p_1, p_2)$ denote the market share of product j among consumers considering both products. Specifically,

$$s_j(p_j) = \frac{\exp(\delta_j - \alpha p_j)}{1 + \exp(\delta_j - \alpha p_j)}, \quad s_j(p_1, p_2) = \frac{\exp(\delta_j - \alpha p_j)}{1 + \sum\limits_{j'=1,2} \exp(\delta_{j'} - \alpha p_{j'})}.$$

Organic Only In the absence of sponsored product advertising, both positions display organic results. By assumption, product 1 appears in the first position and is considered by all consumers, while product 2 appears in the second position and is considered only by $(1-\lambda)$ consumers who consider both options. Let $m_j = (1-\tau)p_j - c_j$ denote seller j's margin and $\boldsymbol{p} = (p_1, p_2)$. Each seller sets a price p_j to maximize profits given its rank in search results:

$$\max_{p_1} \ \pi_1^N(\boldsymbol{p};\tau) = \big(\lambda s_1(p_1) + (1-\lambda)s_1(p_1,p_2)\big)m_1,$$
$$\max_{p_2} \ \pi_2^N(\boldsymbol{p};\tau) = (1-\lambda)s_2(p_1,p_2)m_2. \tag{A.2}$$

Let $p_j^N(\tau)$ and $s_j^N(\tau)$ denote the equilibrium prices and market shares under a commission rate τ. The platform chooses a commission rate τ^N to maximize total commissions:

$$\tau^N = \arg\max_{\tau} \ \tau\big(s_1^N(\tau)p_1^N(\tau) + s_2^N(\tau)p_2^N(\tau)\big). \tag{A.3}$$

With Sponsored Product Advertising Next, consider a scenario where the first position is allocated through a first-price auction on a pay-per-click basis.[34] When seller j submits a bid b_j, the realized bid $\tilde{b}_j$ is stochastic and follows $\log(\tilde{b}_j) \overset{\text{i.i.d.}}{\sim} N(\log(b_j), \sigma^2)$. The seller with a higher realized bid wins the auction and is placed in the first position. The second position displays the top-ranked organic product (product 1). Given both sellers' bids $\boldsymbol{b} = (b_1, b_2)$, the probability that each seller wins the auction is given by:

$$q_1(\boldsymbol{b}) = \Pr(b_1\omega_1 > b_2\omega_2) = \Phi\Big(\frac{\log(b_1/b_2)}{\sqrt{2}\sigma}\Big), \quad q_2(\boldsymbol{b}) = 1 - q_1(\boldsymbol{b}). \tag{A.4}$$

Conditional on winning the auction, the seller pays its realized bid for each click on the listing. Let $F(\cdot)$ denote the distribution of $\boldsymbol{\omega} = (\omega_1, \omega_2)$. The expected payment per click is given by:

$$\ell_1(\boldsymbol{b}) = \frac{1}{q_1(\boldsymbol{b})}\int_{b_1\omega_1 > b_2\omega_2} b_1\omega_1 \mathrm{d}F(\boldsymbol{\omega}) = \frac{b_1}{q_1(\boldsymbol{b})}\int \Phi\Big(\frac{\log(b_1/b_2)}{\sigma} + u\Big)\exp(\sigma u)\phi(u)\mathrm{d}u,$$
$$\ell_2(\boldsymbol{b}) = \frac{1}{q_2(\boldsymbol{b})}\int_{b_1\omega_1 < b_2\omega_2} b_2\omega_2 \mathrm{d}F(\boldsymbol{\omega}) = \frac{b_2}{q_2(\boldsymbol{b})}\int \Phi\Big(\frac{\log(b_2/b_1)}{\sigma} + u\Big)\exp(\sigma u)\phi(u)\mathrm{d}u. \tag{A.5}$$

I assume that the number of clicks is equal to a linear combination of the number of sponsored units sold and the number of impressions, with weights $\gamma_1 \geq 0$ and $\gamma_2 \geq 0$, respectively. Each seller sets a price p_j and submits a bid b_j to maximize expected profits:

[34] The stylized example assumes a first-price auction to generate reasonable results with two bidders. The model in Section 5 considers a generalized second-price auction following Amazon's practice.

$$
\begin{aligned}
\max_{p_1,b_1}\ \pi_1^S(\boldsymbol{b},\boldsymbol{p};\tau) =& \big(q_1(\boldsymbol{b})s_1(p_1) + q_2(\boldsymbol{b})(1-\lambda)s_1(p_1,p_2)\big)m_1 - q_1(\boldsymbol{b})\ell_1(\boldsymbol{b})\big(\gamma_1 s_1(p_1)+\gamma_2\big),\\
\max_{p_2,b_2}\ \pi_2^S(\boldsymbol{b},\boldsymbol{p};\tau) =& q_2(\boldsymbol{b})\big(\lambda s_2(p_2) + (1-\lambda)s_2(p_1,p_2)\big)\big(m_2 - \gamma_1\ell_2(\boldsymbol{b})\big) - \gamma_2 q_2(\boldsymbol{b})\ell_2(\boldsymbol{b}).
\end{aligned} \tag{A.6}
$$

Let $p_j^S(\tau)$, $b_j^S(\tau)$, and $s_j^S(\tau)$ denote the equilibrium prices, bids, and market shares under a commission rate τ. Define $q_j^S(\tau) = q_j\big(\boldsymbol{b}^S(\tau)\big)$ and $\ell_j^S(\tau) = \ell_j\big(\boldsymbol{b}^S(\tau)\big)$. The platform chooses a commission rate τ^S to maximize the sum of total commissions and advertising revenues:

$$
\begin{aligned}
\tau^S = \arg\max_{\tau}\ & \tau\big(s_1^S(\tau)p_1^S(\tau) + s_2^S(\tau)p_2^S(\tau)\big) + q_1^S(\tau)\ell_1^S(\tau)\Big(\gamma_1 s_1\big(p_1^S(\tau)\big)+\gamma_2\Big)\\
&+ q_2^S(\tau)\ell_2^S(\tau)\Big(\gamma_1\lambda s_2\big(p_2^S(\tau)\big) + \gamma_1(1-\lambda)s_2\big(p_1^S(\tau),p_2^S(\tau)\big)+\gamma_2\Big).
\end{aligned} \tag{A.7}
$$

Simulations Throughout all simulations, I set $\lambda = 0.4$, $\alpha = 0.8$, $\sigma = 0.5$, $\gamma_1 = 1$, and $\gamma_2 = 0.1$. In the main results discussed in Section 3, I assume both products have the same marginal cost $c_1 = c_2 = 4$. I hold the quality of product 1 constant at $\delta_1 = 4$ while varying the quality of product 2, δ_2, within a range from 2 to 6, centered around the value of δ_1. Figure 1 and Appendix Figure A.5 depict the simulation results. In the supplementary analysis presented in Appendix Figures A.15 and A.16, I assume both products have the same quality $\delta_1 = \delta_2 = 4$. I hold the marginal cost of product 1 constant at $c_1 = 5$ while varying the marginal cost of product 2, c_2, within a range from 3 to 7, centered around the value of c_1.

Given any parameter, I first solve the equilibrium without sponsored product advertising, characterized by the platform's commission rate τ^N and sellers' prices $\boldsymbol{p}^N(\tau^N)$. Sponsored product advertising can impact welfare through various channels. To examine the contribution of each channel, I explore three assumptions regarding how sellers and the platform can respond to the introduction of a sponsored position. In the most restrictive scenario, sellers cannot adjust prices, and the platform cannot modify the commission rate. So the commission rate remains at τ^N and sellers determine their bids taking prices $\boldsymbol{p}^N(\tau^N)$ as given. In the second assumption, sellers have the flexibility to adjust their prices, while the commission rate remains unchanged at τ^N. In the most flexible assumption, I allow both sellers to adjust their prices and the platform to select a different commission rate. For each assumption, I calculate the difference in market outcomes and welfare between the equilibrium with solely organic results and that with a sponsored result.

B Details on Data Collection

B.1 Selection of Keywords and Construction of Markets

The main dataset introduced in Section 4.1 comprises a large set of scraped search results from Amazon for 3,237 keywords. In this section, I describe the procedure of constructing the set of keywords, with the objective of identifying high-search-volume keywords on Amazon and categorizing them into different markets using a data-driven method.

I started from the frequently searched keywords reported by Yin and Jeffries (2021). The authors collected the 300 most popular keywords from the first to the third quarters of 2020 reported by Amazon across nine categories, including "Automotive," "Baby," "Beauty," "Electronics," "Grocery," "Office Products," "Softlines," "Toys," and "Amazon.com." This compilation yielded a total of 2,700 unique keywords.[35]

To further enrich this dataset, I searched these keywords on Jungle Scout, a prominent e-commerce intelligence service provider for Amazon sellers. For each keyword, Jungle Scout furnished an extensive list of related keywords, ranging from a few hundred to more than ten thousand, alongside their estimated 30-day search volumes for the prior month.[36] Within this expanded dataset of keywords, I kept keywords with a 30-day search volume of at least 100,000 and then manually eliminated keywords falling into four categories: (i) general terms, like "my orders" or "prime"; (ii) keywords that are too broad, like "car accessories" or "home decor"; (iii) keywords specific to a brand, like "Nintendo switch" or "iPhone"; and (iv) keywords related to a specific time, like "Christmas lights," or "2021 monthly planner."

To identify a comprehensive set of high-search-volume keywords, I executed an iterative procedure, repeatedly following the aforementioned steps: identify relevant keywords, keep those with high search volumes, and exclude keywords falling into certain categories. This iterative procedure ended when no new relevant keywords could be identified. The process took place in April 2022 and yielded an extensive list of 904 keywords that had a 30-day search volume of at least 100,000 and satisfied the exclusion criteria outlined earlier, which I referred to as *focal keywords*.

Within the compiled list of focal keywords, some keywords could be close substitutes, poten-

[35] The original data can be found at https://github.com/the-markup/investigation-amazon-brands/blob/master/data/input/combined_queries_with_source.csv.

[36] I focused on the *exact* search volume provided by Jungle Scout, which accounted for searches with identical phrases, excluding plurals, misspellings, additional or omitted words, or synonyms.

tially encompassing products within the same market. I implemented a data-driven method to define markets by grouping keywords that shared a significant number of overlapping products in their search results within the same market. Specifically, I conducted searches for each focal keyword on Amazon over three separate days in April 2022 and kept all products appearing on the first page of the search results. Let $\mathcal{J}_k$ denote the set of products that appeared in the search results of keyword k. I defined the similarity of two keywords k_1 and k_2 as the fraction of overlapping products in their search results:

$$s_{k_1,k_2} = \frac{|\mathcal{J}_{k_1} \cap \mathcal{J}_{k_2}|}{\sqrt{|\mathcal{J}_{k_1}||\mathcal{J}_{k_2}|}}. \tag{B.8}$$

This metric reflected the extent of product overlap between the two keywords.

I represented the set of focal keywords using a graph, where each focal keyword was depicted as a node. To establish connections between nodes, I introduced edges between two nodes if they shared a minimum of 10% of their respective products, i.e., $s_{k_1,k_2} \geq 0.1$. I defined a *market* as a connected component within this graph, which partitioned the focal keywords into distinct markets. An average market encompassed 1.7 focal keywords. Two keywords not directly linked by an edge in the graph could still belong to the same market if they were connected through other intermediary keywords in the graph.

As mentioned earlier, for each focal keyword, Jungle Scout provided a list of *relevant keywords* along with a relevance score, rated on a scale of 1-100, which indicated the degree of relevance each keyword had to the focal keyword. I retained all relevant keywords that met two criteria: they had a relevance score of at least 30, and their 30-day search volume had to be at least 10,000. When a single market encompassed multiple focal keywords, I merged their lists of relevant keywords. In the rare instances where a keyword appeared in multiple markets, I only retained that keyword in the market with the highest relevance score.

This process produced a final list of 3,237 keywords categorized into 546 markets. These keywords collectively accounted for a total daily search volume of 8.8 million. Appendix Table A.1 presents examples of several markets along with associated keywords. The descriptive results in Section 4 use all keywords, while the estimation in Section 5 focuses on the highest-search-volume keyword in each market.

B.2 Conversion of Best Sellers Ranks to Sales Quantities

A crucial element for demand estimation is the daily sales quantities of each product. However, Amazon does not publicly disclose sales quantities for products sold on Amazon. Instead, it assigns a Best Sellers Rank (BSR) to each product in a category, which is updated several times a day. Amazon uses the BSR system to report a product's sales performance relative to that of other products within the same category over a recent period of time. In this section, I outline how I utilize a well-established tool to convert a product's daily BSRs to an estimate of its daily sales quantity. Additionally, I validate the accuracy of this conversion method using actual sales data.

The BSR is defined within one of approximately 25 product categories on Amazon, such as "Baby Products," "Beauty & Personal Care," "Clothing, Shoes & Jewelry," "Health & Household," "Home & Kitchen," "Office Products," and "Pet Supplies," among others. While Amazon maintains confidentiality regarding the exact algorithm employed to calculate the BSR, it is generally believed that the BSR takes into account both recent and historical sales, with more recent sales carrying significantly higher weights.

Prior studies have documented that when the number of products is sufficiently large, there exists a relatively stable and predictable relationship between a product's sales rank and its sales quantities (e.g., Chevalier and Goolsbee 2003; Brynjolfsson et al. 2011). This relationship is often modeled using functional forms such as power functions or splines.

I utilized a tool developed by Jungle Scout for this conversion. Jungle Scout provided marketing services to many Amazon sellers of varying sizes selling products in different categories, thus has access to a substantial amount of actual sales data from Amazon sellers. This tool used these actual sales data to fit a mapping between the BSR and daily sales quantities in each category. I obtained these mappings estimated by Jungle Scout on May 16, 2022. Appendix Figure A.17 presents mappings for six common categories. Using these mappings, I converted the BSR to the sales quantity for each product on each day in my sample.

I validated the conversion from daily BSRs to daily sales quantities using actual daily sales data obtained in collaboration with an Amazon seller, Forum Brands. This dataset encompassed several hundred products spanning several product categories on Amazon and covered a two-year period. For each product, I compared its estimated daily sales quantities derived from the BSR-to-quantity mappings with its actual daily sales quantities.

Appendix Figure A.6 depicts these two series for several randomly chosen products with a wide range of sales volumes. For most products, the two series align remarkably well, both in the cross-section and for high-frequency variation over time. I also computed the correlation between estimated and actual daily quantities for each product. Appendix Figure A.18 displays the distribution of the correlation across products sold by the company. The median correlation is 0.81, suggesting that the conversion method is reasonably accurate.

C Details on Demand Estimation

C.1 Calibration of Fraction of Sales From Focal Keywords

The demand model introduced in Section 5.1.1 calculates the market share of each product in one keyword. However, what I observe in the data is a product's total sales quantity on the platform. In reality, a product can appear in the search results of multiple keywords. Disregarding these additional keywords in the estimation can lead to biased estimates of search frictions.

To illustrate this bias, consider a product that ranks poorly in the keyword I observe. Search frictions are so high that few consumers have this product in their consideration sets. However, this same product might rank higher in other keywords that are a better match, resulting in sales from those keywords. Attributing all of its sales solely to the observed keyword would lead to an incorrect conclusion that search frictions are low.

In this section, I describe the method to bridge the gap. As in Appendix B.1, I refer to the keyword with the highest search volume in each market as the "focal keyword." I calibrate the fraction of a product's total sales quantity attributed to the focal keyword, represented by $\kappa_{jt} \in (0,1)$. Let q_{jt} denote the total sales quantity of product j on day t, and M_t the market size of the focal keyword, calibrated in Appendix C.2. For each keyword, I conducted a total of $F = 6$ searches each day, indexed by f. These searches are assumed to represent what consumers typically encounter on that day, and I calculate the average market share across these searches. Therefore, equation (4) can be reformulated as follows:

$$s_{jt}(\boldsymbol{p}) = \frac{\kappa_{jt} q_{jt}}{M_t} = \frac{1}{F} \sum_{f=1}^{F} \sum_{n=1}^{N} \mathbf{1}(j \in \mathcal{J}_{nt}^{f})(\lambda_n - \lambda_{n+1}) \frac{\exp(\delta_{jt})}{1 + \sum_{j' \in \mathcal{J}_{nt}^{f}} \exp(\delta_{j't})}. \tag{C.9}$$

Here, $\mathcal{J}_{nt}^{f}$ represents the set of products in the top n positions within search f on day t.

The calibration of κ_{jt} relies on external data. I randomly selected 480 products across various product categories, which appeared in the search results of the focal keywords with a wide range of ranks. I refer to this sample of products as the "prediction sample," and the full set of products in the main sample as the "estimation sample." I leveraged a tool developed by Jungle Scout. For each product j, this tool provided a comprehensive list of keywords for which the product had recently appeared on the first page of search results, indexed by $k \in \mathcal{K}_j$. By construction, the focal keyword, denoted by k_j^0, belongs to $\mathcal{K}_j$. In addition, the tool also reported the rank of this product in each keyword, represented by n_{jk}, and the search volume of the keyword, represented by V_k.

Similar to Section 5.1.2, I parametrize the fraction of consumers whose consideration sets include the product in the n-th position of the search results as $\lambda_n(\beta) = 1/\exp(\beta(n-1))$, where $\beta \geq 0$ is a parameter determining the magnitude of search frictions. I assume that the mean utility of a product in the n-th position is given by $\delta_n(\gamma) = -\gamma \log(n)$. Consequently, product j's total sales quantity from any arbitrary set of keywords $\mathcal{K}$ can be calculated as:

$$q_j(\mathcal{K};\beta,\gamma) = \sum_{k\in\mathcal{K}} V_k \sum_{n=n_{jk}}^{N} \left(\lambda_n(\beta) - \lambda_{n+1}(\beta)\right) \frac{\exp\left(\delta_{n_{jk}}(\gamma)\right)}{1 + \sum\limits_{n'=1}^{n} \exp\left(\delta_{n'}(\gamma)\right)}. \tag{C.10}$$

Then, given the parameters β and γ, I can determine the fraction of product j' total sales quantity attributed to the focal keyword with the following equation:

$$\kappa_j(\beta,\gamma) = \frac{q_j(\{k_j^0\};\beta,\gamma)}{q_j(\mathcal{K}_j;\beta,\gamma)}. \tag{C.11}$$

The next step involves predicting κ_{jt} for all products in the estimation sample by leveraging the relationships observed in the prediction sample. To achieve this, I utilize several predictors that can be constructed in both the prediction sample and the estimation sample. As discussed in Section 4.1 and Appendix B.1, in addition to the 546 focal keywords, I also keep track of search results for several thousand other keywords, with an average of about five keywords in each market. Let $\overline{K}$ represent the set of all tracked keywords (so $k_j^0 \in \overline{K}, \forall j$). For each product j, I construct the following predictors:

- Log of the product's rank in the focal keyword, $\log(n_{jk_j^0})$.
- Indicator for the product appearing in other tracked keywords, $\mathbf{1}(\overline{K} \cap (\mathcal{K}_j \setminus \{k_j^0\}) \neq \emptyset)$.
- The search volume of the focal keyword as a fraction of the search volume of all tracked keywords that this product appears in, $V_{k_j^0} / \sum_{k\in\overline{K}\cap\mathcal{K}_j} V_k \in (0,1]$.

- Log of the average product's rank in other tracked keywords, $\log(\frac{\sum_{k \in \overline{K} \cap (\mathcal{K}_j \setminus \{k_j^0\})} n_{jk}}{\sum_{k \in \overline{K} \cap (\mathcal{K}_j \setminus \{k_j^0\})} 1})$.[37]

I regress $\kappa_j(\beta, \gamma)$ on these four variables. Appendix Table A.2 presents the regression results under different values of β. The R^2 ranges from 0.35 to 0.38, indicating a good model fit. The signs of the coefficients are reasonable. For instance, the fraction of sales attributed to the focal keyword tends to be lower when a product ranks low in the focal keyword but high in other tracked keywords. Panel A of Appendix Figure A.19 compares the predicted and actual fractions, confirming that the predictions perform reasonably well and capture the observed patterns.

The final step involves applying the prediction model to the estimation sample to calibrate κ_{jt} for each product on each day. This calibration process depends on two parameters, β and γ. β is also a parameter to be estimated and γ depends on the estimated product utility. I iterate between the calibration exercise and the demand estimation until the values of β and γ used in the calibration match the estimated β and implied γ averaged across all markets. Panel B of Appendix Figure A.19 plots the distribution of κ_{jt} under the final calibration in the estimation sample, where I restrict κ_{jt} within the range of $[0.001, 1]$.

C.2 Calibration of Market Size

The market size is another input to a product's market share in equation (C.9). In this section, I describe the calibration of the market size. Estimating the market size for each keyword on each day is challenging, as consumers may search for multiple keywords during the same visit to Amazon or the same keyword across different visits when looking for a product. While industry reports commonly suggest that Amazon holds around a 40% market share of online purchases (Lam 2023), this share can be significantly different at the *search* level.

To address this challenge, I calibrate the market size using aggregate data from industry reports. The calibration depends on two essential statistics: (i) the total number of units sold on Amazon, and (ii) the total number of searches on Amazon.

First, Amazon's earnings release reveals that total sales on Amazon's marketplaces amounted to approximately \$118.4 billion in the second quarter of 2022,[38] and North America accounted for

[37] If $\overline{K} \cap (\mathcal{K}_j \setminus \{k_j^0\} = \emptyset$, this variable is set to a constant. The specific value of this constant does not affect the analysis, as there is already an indicator variable $\mathbf{1}(\overline{K} \cap (\mathcal{K}_j \setminus \{k_j^0\}) \neq \emptyset)$ in the regression.

[38] According to Amazon's 2022 Q2 earnings release, available at https://s2.q4cdn.com/299287126/files/doc_financials/2022/q2/Q2-2022-Amazon-Earnings-Release.pdf (page 17, accessed in May 2022), Amazon's online stores (i.e., first-party products) generated revenues of \$50.9 billion in 2022 Q2, and 57% of units were sold

62%. My analysis focuses on Amazon.com, a domain primarily used in the U.S. market. Across all markets in my sample, the sales-weighted average price is \$35.5. Therefore, I estimate the total quarterly units sold in North America to be roughly 2.1 billion units (\$118.4 $\times$ 62% $\div$ \$35.5).

Second, there were approximately 2.4 billion monthly visits to Amazon.com in May 2022.[39] I assume that, on average, consumers perform two keyword searches during each visit. Hence, the total number of quarterly searches is 14.4 billion (2.4 billion visits $\times$2 searches per visit $\times$3 months). Combining these two statistics, I estimate the average inside market share as 2.1 billion units sold divided by 14.4 billion searches, or 14.6%. As a result, I calibrate the total market size of each market, denoted as M_t, as follows:

$$M_t = \frac{1}{0.146} \sum_{j \in \mathcal{J}} \kappa_{jt} q_{jt}. \tag{C.12}$$

C.3 Estimation and Empirical Bayes Procedure

In the demand estimation, I impose several restrictions at the product level. First, I exclude products with missing data for sales quantities or prices. Second, I drop products that appear in the search results only once out of six times on any given day. Third, I exclude products with prices exceeding five times or falling below 20% of the median price in each market. Fourthly, I remove products with calibrated fractions of sales attributed to the focal keyword less than 5%. Last, I drop products that appear in the bottom one-third of search results but have market shares exceeding the 90th percentile of all products in each market, or appear in the bottom two-thirds of search results but have market shares exceeding the 95th percentile of all products. This last restriction is based on spot checks and aims to eliminate products that may be popular in other markets but less relevant in the focal market.

I also make several restrictions at the market level. I drop markets where (i) the search results do not consistently contain 60 positions, (ii) over 40% of products are excluded based on the previous criteria, or (iii) where the median bid exceeds 15% of the median price, which likely reflects miscoded data.

I estimate the parameters (β, ρ) using the Method of Moments separately in each market with the moment conditons in equation (8). The estimation in each market takes the following steps.

by third-party sellers. Since the prices of products sold by third-party sellers and by Amazon were similar, total sales amounted to \$118.4 billion.

[39]Source https://www.statista.com/statistics/623566/web-visits-to-amazoncom/.

1. Given β, invert market shares s_{jt} to get δ_{jt} using equation (C.9).

2. Regress δ_{jt} on product dummies, time dummies, product-specific linear time trends, and p_{jt}. Calculate the residuals as the unobserved demand shock ξ_{jt}.

3. Given ρ, calculate the innovation $\eta_{jt} = \xi_{jt} - \rho\xi_{j,t-1}$.

4. Construct the moment in equation (8) and search for parameters (β, ρ) that minimize $|m|$.

To adjust for sampling errors, I employ an empirical Bayes procedure to shrink the market-specific parameter $\hat{\beta}$ in the demand estimation towards its mean value across markets. Let m index markets. Let $\hat{\beta}^m$ denote the estimated parameter in market m and $se(\hat{\beta}^m)$ its standard error. I compute the shrunken parameter using the following formula:

$$\tilde{\beta}^m = \frac{\sigma_\beta^2}{\sigma^2 + se(\hat{\beta}^m)^2}\hat{\beta}^m + \frac{se(\hat{\beta}^m)^2}{\sigma_\beta^2 + se(\hat{\beta}^m)^2}\bar{\beta}, \tag{C.13}$$

where $\bar{\beta}$ represents the mean of $\hat{\beta}^m$ across markets, and σ_β denotes the standard deviation of $\hat{\beta}^m$. This procedure effectively adjusts market-specific parameter estimates to account for variability and ensures more robust estimates.

D Details on Supply Estimation

D.1 Construction of Profits From Other Keywords

A seller on the platform typically makes multiple keyword-specific bidding decisions and a single pricing decision that applies to the entire platform. However, estimating such a model is not feasible due to data and computational limitations. On the data side, I do not observe search results from all possible keywords. Even if I did observe search results from a comprehensive set of keywords, estimating a model where each seller simultaneously makes many interdependent decisions is computationally challenging.

To address this problem, I employ a partial-equilibrium approach, focusing on sellers' decisions within the focal keyword, which is the keyword with the highest search volume in a market (for more details, see Appendix B.1). Specifically, when seller j submits a bid b_j for this keyword and sets a price p_j, it considers its bidding decisions in other keywords as fixed but acknowledges that changing its price can impact its sales across all keywords.

As introduced in Section 5.2.1, seller j's expected total profits can be expressed as:

$$\pi_j(\boldsymbol{p},\boldsymbol{b}) = \left(s_j^e(\boldsymbol{p},\boldsymbol{b}) + M_j^{other} s_j^{other}(p_j)\right)\left((1-\tau)\,p_j - c_j\right) - ad_j^e(\boldsymbol{p},\boldsymbol{b}) - ad_j^{other}(p_j), \tag{D.14}$$

In this equation, $s_j^e(\boldsymbol{p},\boldsymbol{b})$ and $ad_j^e(\boldsymbol{p},\boldsymbol{b})$ represent the seller's expected market share and advertising payment within the focal keyword, as defined in equation (10). $s_j^{other}(p_j)$ denotes its aggregate market share in other keywords, M_j^{other} is the market size of other keywords relative to the focal keyword, and $ad_j^{other}(p_j)$ represents the seller's advertising payment in other keywords. This section derives these three elements: $s_j^{other}(p_j)$, M_j^{other}, and $ad_j^{other}(p_j)$.

Let s_j^0 and $s_j^{0,other}$ denote seller j's market shares in the status quo within the focal keyword and in other keywords, respectively. While I can calculate s_j^0 based on the demand model and realized search results of the focal keyword, I cannot directly calculate $s_j^{0,other}$ due to a lack of search results in other keywords. Therefore, I assume

$$s_j^{0,other} = s_j^0. \tag{D.15}$$

This assumption is grounded in the idea that the competitive environment and strategic interactions among sellers in other keywords closely resemble those in the focal keyword. After all, these sellers are all part of the same market. Hence, I consider the market outcome observed in the focal keyword as a reasonable approximation for the market outcomes in other keywords.

In Appendix C.1, I calibrate the fraction of product j's total sales quantity from the focal keyword on day t, denoted as $\kappa_{jt} \in [0,1]$. When a product does not appear in the search results on a given day, I set $\kappa_{jt} = 0$. I take the average of this fraction across all days within a week, denoted as κ_j. The market size of other keywords relative to the focal keyword can then be derived as

$$\kappa_j = \frac{s_j^0}{s_j^0 + M_j^{other} s_j^{0,other}} = \frac{1}{1+M_j^{other}} \quad \Rightarrow \quad M_j^{other} = \frac{1-\kappa_j}{\kappa_j}. \tag{D.16}$$

A product can generate both sponsored and organic sales within the focal keyword or in other keywords. Let $s_{j,O}^0$ and $s_{j,S}^0$ represent seller j's market shares for organic and sponsored sales within the focal keyword in the status quo, respectively. Let $s_{j,O}^{0,other}$ and $s_{j,S}^{0,other}$ denote the corresponding market shares in other keywords. By definition, these quantities satisfy the relationships $s_{j,O}^0 + s_{j,O}^0 = s_j^0$ for the focal keyword and $s_{j,O}^{0,other} + s_{j,S}^{0,other} = s_j^{0,other}$ for other keywords. I can derive $s_{j,O}^0$ and $s_{j,S}^0$ within the focal keyword using the demand model. To extend this division to other keywords, I assume that the division between sponsored and organic sales in other keywords mirrors the observed split in the focal keyword, with one modification described below.

To approximate the division between sponsored and organic sales in other keywords, I rely on the ratio of seller j's sponsored sales to total sales within the focal keyword in the status quo. This ratio is calculated as $s^0_{j,S}/s^0_j$ and falls within the range [0,1]. It quantifies the relative importance of sponsored sales for seller j within the focal keyword. I extend this ratio to other keywords using the following formula:

$$s^{0,other}_{j,S} = \min\{0.5, \frac{s^0_{j,S}}{s^0_j}\} \times s^{0,other}_j, \;\; s^{0,other}_{j,O} = s^{0,other}_j - s^{0,other}_{j,S}. \tag{D.17}$$

It is worth highlighting that in equation (D.17), I cap the fraction of sponsored sales at 0.5. This constraint ensures that, in the status quo, no product can have more than 50% of its sales originating from sponsored positions in other keywords. The reason for imposing this limitation is to prevent the model from implying an unreasonably high proportion of sponsored sales on Amazon, which would not align with observed data. Specifically, I can calculate the ratio between sponsored and organic sales aggregated across all keywords as

$$r = \frac{\sum_j p_j (s^0_{j,S} + M^{other}_j s^{0,other}_{j,S})}{\sum_j p_j (s^0_{j,O} + M^{other}_j s^{0,other}_{j,O})}. \tag{D.18}$$

Without the cap imposed in equation (D.17), $s^{0,other}_{j,S} = s^0_{j,S}$ and $s^{0,other}_{j,O} = s^0_{j,S}$, resulting in a ratio between sponsored and organic sales in all keywords of 34%, significantly higher than the actual ratio of around 20% in 2022. By imposing a cap of 0.5 in equation (D.17), the ratio in equation (D.18) has a more reasonable value of 21%.

Having derived seller j's organic and sponsored market shares in other keywords in the status quo, I now need to specify how these shares respond to changes in the product's price. I define $s^{other}_{j,O}(p_j)$ and $s^{other}_{j,S}(p_j)$ as seller j's organic and sponsored market shares in other keywords under any given price p_j. I employ a logit model to approximate both functions. For either organic ($G = O$) or sponsored ($G = S$) market share and for any price p_j, I assume the following relationship:

$$\log\left(s^{other}_{j,G}(p_j)\right) = \log(s^{0,other}_{j,G}) - \alpha(p_j - p^0_j), \;\; G \in \{O, S\}. \tag{D.19}$$

Here, p_j^0 represents product j's price in the status quo, and α characterizes consumers' price sensitivity and is the same parameter used in equation (3).

The final step is to construct seller j's advertising payment in other keywords, $ad_j^{other}(p_j)$. Let $\ell_j(\boldsymbol{p}, \boldsymbol{b})$ represent the advertising payment per sponsored unit for product j in the focal keyword, which is defined as $\ell_j(\boldsymbol{p}, \boldsymbol{b}) = ad_j^e(\boldsymbol{p}, \boldsymbol{b})/s_{j,S}^e(\boldsymbol{p}, \boldsymbol{b})$. Similarly, I define $\ell_j^{other}(p_j)$ as product j's expected advertising payment per sponsored unit in other keywords. I can then express its advertising payment in other keywords as

$$ad_j^{other}(p_j) = M_j^{other} \ell_j^{other} s_{j,S}^{other}(p_j), \tag{D.20}$$

where M_j^{other} is determined using equation (D.16) and $s_{j,S}^{other}(p_j)$ is defined by equation (D.19).

Auctions in different keywords are independent, so a seller's bid in the focal keyword affects only its sales and advertising payment within that keyword without impacting other keywords. In the partial equilibrium I consider, each seller takes its bidding decisions in other keywords as given. Therefore, for a given price, the expected advertising payment in other keywords does not affect the seller's bidding decision in the focal keyword. In equilibrium, I assume that the expected advertising payment per sponsored unit in other keywords is the same as that in the focal keyword:

$$\ell_j^{other}(p_j) = \ell_j(\boldsymbol{p}, \boldsymbol{b}). \tag{D.21}$$

It is worth emphasizing that equation (D.21) holds only in equilibrium. When a seller makes its bidding decision in the focal keyword, it does *not* consider its sponsored sales or advertising payment in other keywords.

D.2 Gibbs Sampler and Estimation of Bids and Organic Ranks

When sellers make bidding and pricing decisions every week, they take into account the uncertainty in the realized search results on Amazon, which arises from two sources. First, organic ranks are determined in real time by Amazon's ranking algorithm, which is influenced by a variety of factors and subject to variations across different searches. Second, auction outcomes can undergo frequent changes due to various real-world elements, such as sellers' budget constraints, personalized advertising, and experimentation, as discussed in Section 2. This section describes the method to estimate the uncertainty associated with realized search results. Specifically, this involves separately estimating sellers' bids and the distribution of organic ranks.

Instead of observing individual sellers' bids, I only have access to realized auction outcomes from numerous searches, along with aggregate statistics of the winning bids for each keyword, including the median, lowest, and highest. The goal is to infer sellers' bidding decisions and the associated uncertainty most likely to generate the observed auction outcomes.

As discussed in Section 5.2.1, I assume that when seller j submits a bid b_j, the realized bid in search f is subject to randomness and follows a log-normal distribution:

$$\log(\tilde{b}_j^f) \overset{\text{i.i.d.}}{\sim} N(\log(b_j), \sigma_b^2). \tag{D.22}$$

Sellers are ranked based on their realized bids $\tilde{b}_j^f$. Let r_j^{f*} represent the underlying rank of product j in the ad auction associated with search f. However, only the top $R = 12$ sellers are displayed in sponsored positions, and thus, I can only observe the rank when $r_j^{f*} \leq R$. The observed rank, denoted by r_j^f, is determined as follows:

$$r_j^f = \begin{cases} r_j^{f*}, & \text{if } r_j^{f*} \leq R \\ \emptyset, & \text{if } r_j^{f*} > R \end{cases}. \tag{D.23}$$

The parameters to be estimated are $\boldsymbol{\theta} = \{\{b_j\}_{j=1}^J, \sigma_b\}$. The data consists of observed auction ranks $\{r_j^f\}_{j,f}$ and two aggregate statistics: the median winning bid denoted as $\overline{M}$ and the range of winning bids denoted as $\overline{R}$, which is calculated as the highest winning bid minus the lowest winning bid.

In the model, the relative ranking of sellers is preserved if I replace b_j with $a_1(b_j)^{a_2}$ and σ_b with $a_2\sigma_b$ for any $a_1, a_2 > 0$. However, this transformation will impact the median and range of winning bids. The estimation follows two steps. First, I fix two parameters in $\boldsymbol{\theta}$ (e.g., b_1 and σ_b) and estimate the remaining parameters using the observed auction ranks. Then I estimate the two parameters to match the two aggregate statistics.

In the first step of estimation, a conceptually straightforward approach involves using the maximum likelihood estimator (MLE). Let $(R_1, \ldots, R_J)$ be a vector of random variables representing the observed rank of each product. The likelihood function is given by:

$$f(\boldsymbol{\theta}) = \prod_{f=1}^{F} \Pr\left((R_1, \ldots, R_J) = (r_1^f, \ldots, r_J^f)\right) = \prod_{f=1}^{F} \Pr(\tilde{b}_{j_{1f}}^f > \tilde{b}_{j_{2f}}^f > \cdots > \tilde{b}_{j_{Rf}}^f > \max_{j: r_j^f = \emptyset} \tilde{b}_j^f), \tag{D.24}$$

where j_{rf} refers to the product that ranks the r-th in the auction of search f (i.e., $r^f_{j_{rf}} = r$). However, the likelihood function does not have an analytical expression. For a large number of bidders (ranging from 30-150 in my data), numerical integration or approximation methods may not be feasible and could introduce significant errors.

I estimate the model parameters using the Gibbs sampler. The Gibbs sampler is a Markov chain Monte Carlo (MCMC) algorithm to generate a sequence of observations approximated from a specified multivariate probability distribution. In my model, the Gibbs sampler is tractable because the econometric model is fully parameterized. A notable advantage of the Gibbs sampler is that it allows for data augmentation of latent variables (Tanner and Wong 1987). I augment the realized bids in each search, $\tilde{b}^f_j$, as pseudo parameters of the model. This augmentation approach allows me to break down simulations from the posterior distribution into smaller steps. Specifically, conditional on draws of b_j and the observed data, the posterior distribution of $\log(\tilde{b}^f_j)$ is a truncated normal distribution. Conditional on draws of $\tilde{b}^f_j$, the posterior distribution of $\log(b_j)$ is a normal distribution.

In the first step of estimation, I fix the parameters $b_1 = 1$ and $\sigma_b = 0.1$, resulting in a set of parameters to be estimated denoted as $\tilde{\boldsymbol{\theta}} = \{\{b_j\}^J_{j=2}, \{\tilde{b}^f_j\}_{j,f}\}$. For any parameter $\delta \in \tilde{\boldsymbol{\theta}}$, I use $\delta|\tilde{\boldsymbol{\theta}}^-$ to denote the parameter δ conditional on all other parameters in $\tilde{\boldsymbol{\theta}}$ and all the observed data. I employ a conventional diffuse prior distribution for each parameter in $\tilde{\boldsymbol{\theta}}$: $\log(b_j) \sim N(0, \Sigma_0)$ and $\log(\tilde{b}^f_j) \sim N(0, \Sigma_0)$ with $\Sigma_0^{-1} = 0$ for all j and f. I can then express the posterior distribution of each parameter, given all other parameters and the observed data, as follows:

$$\log(b_j)|\tilde{\boldsymbol{\theta}}^- = N(\frac{1}{F}\sum_{f=1}^{F}\log(\tilde{b}^f_j), \frac{\sigma_b^2}{F}), \;\; j = 2, \ldots, J,$$

$$\log(\tilde{b}^f_j)|\tilde{\boldsymbol{\theta}}^- \sim TN\left(\log(b_j), \sigma_b^2, \log(\underline{b}^f_j), \log(\overline{b}^f_j)\right), \;\; j = 1, \ldots, J,$$

$$\underline{b}^f_j = \begin{cases} 0, & r^f_j = \emptyset \\ \max\{\tilde{b}^f_{j'}\}_{r^f_{j'} = \emptyset}, & r^f_j = R \\ \tilde{b}^f_{j_{r^f_j+1,f}}, & r^f_j < R \end{cases}, \quad \overline{b}^f_j = \begin{cases} \min\{\tilde{b}^f_{j'}\}_{r^f_{j'} \neq \emptyset}, & r^f_j = \emptyset \\ \tilde{b}^f_{j_{r^f_j-1,f}}, & 1 < r^f_j \leq R \\ \infty, & r^f_j = 1 \end{cases}. \tag{D.25}$$

Here, $TN(\mu, \sigma^2, a, b)$ represents a truncated normal distribution with support $[a, b]$. I ran the algorithm for 5,000 iterations. A common practice in Markov chain Monte Carlo (MCMC) algorithms is to run a sufficient number of iterations to ensure that the posterior distributions have converged and stabilized. In my setting, most posterior distributions became stable after several hundred iterations, indicating convergence to the target distributions.

Following a standard procedure, I discarded the initial half of the iterations, often referred to as the "burn-in" period, and calculated parameter estimates by averaging the values across the latter half of the iterations. This approach helps minimize potential bias introduced during the initial iterations, ensuring that the parameter estimates are based on samples drawn from the actual posterior distributions.

In the second step, I calculate the aggregate statistics for the winning bids $\{\tilde{b}_j^f\}_{j:r_j^f \neq \emptyset}$, including the median M_f, the second lowest bid L_f, and the second highest bid H_f.[40] The relative ranking of sellers is preserved if I replace b_j with $a_1(b_j^f)^{a_2}$ and σ_b with $a_2\sigma_b$, in which case $\tilde{b}_j^f$ is replaced with $a_1(\tilde{b}_j^f)^{a_2}$, as well as the statistics M_f, L_f, and H_f. I search for the parameters (a_1, a_2) that yield aggregate statistics $(\overline{M}, \overline{R})$ matching the observed values:

$$\min_{a_1,a_2}\left(\frac{1}{F}\sum_{f=1}^{F} a_1 M_f^{a_2} - \overline{M}\right)^2 + \left(\frac{1}{F}\sum_{f=1}^{F}(a_1 H_f^{a_2} - a_1 L_f^{a_2}) - \overline{R}\right)^2. \tag{D.26}$$

After obtaining the estimated values $\hat{a}_1$ and $\hat{a}_2$, I transform all estimated parameters using these values. Specifically, $\hat{b}_j$ is rescaled to $\hat{a}_1(\hat{b}_j)^{\hat{a}_2}$ and σ_b is rescaled to $\hat{a}_2\sigma_b$.

Appendix Figure A.20 summarizes the estimates of the auction model. Panel A depicts the distribution of the winning bids, pooling data from all market–week pairs. Notably, the average winning bid is \$1.24, closely mirroring industry estimates.[41] In Panel B, I plot the parameter that measures uncertainty, σ_b, against the number of unique auction winners. A higher σ indicates greater uncertainty in realized bids, resulting in more unique auction winners. Panel C illustrates the strong alignment between predicted frequencies of products winning the auction and observed winning frequencies, with a correlation coefficient of 0.966.

I apply a similar methodology to estimate the distribution of organic ranks. I do not have access to Amazon's proprietary ranking algorithm, so I need to approximate this process. I assume that when a consumer enters a keyword, he platform's algorithm assigns each product a stochastic score drawn from a product-specific distribution:

$$\log(\tilde{o}_j) \overset{\text{i.i.d.}}{\sim} N(\log(o_j), \sigma_o^2). \tag{D.27}$$

[40] To avoid potential distortions from extreme values, I exclude the minimum and maximum bids from this analysis.
[41] For instance, Pacvue, the average PPC on Amazon in 2022 Q2 was estimated to be \$1.24 (Pacvue 2022).

Products are then ranked based on their realized scores. This assumption parallels equation (D.22) with the submitted bid b_j replaced with o_j, making the process similar in nature to the determination of auction outcomes. A product with a higher o_j is more likely to be ranked higher in the organic search results. However, there exists a distinction in my model. For auction outcomes, sellers can strategically influence the distribution of their realized bids by selecting their targeted bids. In contrast, in my model, the distribution of organic ranks is assumed to be fixed, thus sellers are unable to exert influence over it.

Likewise, I employ the Gibbs sampler to estimate the parameters $\left\{\{o_j\}_{j=1}^{J}, \sigma_o\right\}$ in each market using realizations of organic ranks within a given week. I normalize the parameters with $o_1 = 1$ and $\sigma_o = 0.1$. Unlike the bids, the absolute level and dispersion of these scores do not possess economic significance. Appendix Figure A.21 indicates that the predicted frequencies of products appearing in organic positions closely aligns with the observed frequencies, with a correlation coefficient of 0.972.

E Details on Estimation of Platform's Objective

E.1 Definition of Platform's Objective Function

This section provides a formal definition of the platform's objective function introduced in Section 5.3.1. The platform sets a commission rate τ to maximize a combination of short-term revenues, consumer surplus, and seller profits. Consistent with prior research on multi-sided platforms (e.g., Gutierrez 2022; Castillo 2023), Amazon can invest in user acquisition and retention to maximize its long-run profitability. Incorporating the welfare of consumers and sellers into the platform's objective approximates these long-term considerations that are not explicitly modeled.

Let $\boldsymbol{p}(\tau)$ and $\boldsymbol{b}(\tau)$ denote the equilibrium prices and bids under commission rate τ, respectively. Consumer surplus under product arrangement Γ is defined as follows

$$CS(\Gamma;\tau) = \frac{1}{\alpha}\sum_{n=1}^{N}(\lambda_n - \lambda_{n+1})\log\left(1 + \sum_{j\in\mathcal{J}_n(\Gamma)}\exp(\delta_j)\right) \tag{E.28}$$

The expected consumer surplus is given by

$$CS(\tau) = \int CS\Big(\Gamma\big(\boldsymbol{b}(\tau), \boldsymbol{\omega}, \Gamma_0\big);\tau\Big)\mathrm{d}F(\boldsymbol{\omega})\mathrm{d}G(\Gamma_0). \tag{E.29}$$

Expected seller profits are given by

$$PS(\tau) = \sum_{j \in \mathcal{J}} s_j^e\big(\boldsymbol{p}(\tau), \boldsymbol{b}(\tau)\big)\big((1-\tau)p_j(\tau) - c_j\big) - ad_j^e\big(\boldsymbol{p}(\tau), \boldsymbol{b}(\tau)\big), \tag{E.30}$$

where $s_j^e(\boldsymbol{p}, \boldsymbol{b})$ and $ad_j^e(\boldsymbol{p}, \boldsymbol{b})$ represent seller j's expected market share and advertising payment defined in equation (10). The platform's short-term revenues consist of commissions and advertising revenues given by

$$\begin{aligned} COM(\tau) =& \tau \sum_{j \in \mathcal{J}} s_j^e\big(\boldsymbol{p}(\tau), \boldsymbol{b}(\tau)\big) p_j(\tau), \\ AD(\tau) =& \sum_{j \in \mathcal{J}} ad_j^e\big(\boldsymbol{p}(\tau), \boldsymbol{b}(\tau)\big). \end{aligned} \tag{E.31}$$

The platform chooses the commission rate τ to maximize the following objective function:

$$\max_{\tau} \Pi(\tau) = COM(\tau) + AD(\tau) + \mu\big(CS(\tau) + PS(\tau)\big), \tag{E.32}$$

where $\mu \geq 0$ is a nonnegative weight that reflects the platform's consideration of the welfare of other market participants relative to its short-term revenues. All these quantities are aggregated across markets using the market sizes calibrated in Appendix C.2.

E.2 Alternative Set of Sponsored Positions

In the model, I do not assume that the platform optimally selects the set of sponsored positions. The number of possible configurations is enormous, making it implausible for the observed choice to be entirely optimal. With 60 positions in search results, there are 2^{60} possible sets of sponsored positions, leading to approximately 10^{18}. In reality, the arrangement of sponsored positions on Amazon has been constantly evolving in recent years, likely due to ongoing experimentation.

To offer additional evidence, I explore small deviations from the existing set of sponsored positions. Specifically, I replace a single sponsored position with an organic one and evaluate the resulting change in consumer surplus, seller profits, and platform revenues. As depicted in Appendix Figure A.22, replacing the first or the second sponsored position with an organic one can increase consumer surplus, seller profits, and platform revenues simultaneously. Therefore, regardless of the weight placed on consumers and sellers, the platform would be better off from these deviations, which implies that the current layout is not optimal.

F Model Extensions

F.1 Aversion to Sponsored Products

In the baseline model, I estimate product utility based on consumers' revealed preferences, assuming they can accurately discern product utility. However, in reality, this assumption may not hold. Consumers might exhibit aversions to sponsored products, choosing to skip them entirely when making purchase decisions. Such behaviors could result in an underestimation of the utility of sponsored products. Despite sponsored products often being intermingled with organic search results on digital retail platforms, making them more challenging for consumers to disregard, platforms are legally obligated to differentiate between sponsored and organic results. For instance, on Amazon, sponsored products are labeled as "Sponsored" on the screen. In this section, I employ two methods to investigate how such behaviors might impact my findings.

F.1.1 A Fixed Fraction of Consumers Skip Sponsored Positions

Let λ_n represent the fraction of consumers considering the product in the n-th *organic* position. Due to consumers' aversion to sponsored products, a fraction $\Omega \in [0, 1]$ of consumers consistently skip sponsored products entirely, resulting in only a fraction $(1 - \Omega)\lambda_n$ of consumers considering the product in the n-th *sponsored* position. Unfortunately, since I do not have microdata on consumers' search behavior or experimental variation in product arrangement or labeling, I cannot directly determine the value of Ω from aggregate demand data. Instead, I explore how different values of Ω might impact the results. Researchers with access to more detailed data could employ a more flexible model to estimate this parameter.

I explore various values of Ω ranging from 0 to 0.6 with increments of 0.1. For each Ω, I re-estimate the model, including consumer demand, product utility, seller primitives, and the platform's weight, and then conduct counterfactual simulations. Appendix Figure A.23 presents the utility estimates for products in each position under different values of Ω. The first panel, with $\Omega = 0$, corresponds to the baseline result and replicates Panel B of Figure 3. As Ω increases, the estimated utility of sponsored products gradually improves. Notably, when Ω exceeds 0.4, sponsored products have higher utility than their organic counterparts in comparable positions, except for products located at the very top.

Appendix Figure A.24 summarizes the welfare implications of removing sponsored positions under different values of Ω. As in Section 7.1, I consider two scenarios: one with a fixed commission rate and one with endogenous commission rates. There are two types of consumers in this model: "skippers," who consistently skip all sponsored products, and "non-skippers," who consider all products. I calculate consumer surplus separately for these two types.

Panel A of Appendix Figure A.24 illustrates that eliminating sponsored positions benefits skippers across all Ω values. The effect shrinks with endogenous commission rates due to increased prices resulting from higher commission rates. For non-skippers, the impact under a fixed commission rate varies depending on the value of Ω. When Ω surpasses 50%, the initially positive effect reverses. This reversal occurs because, at higher values of Ω, sponsored products are estimated to exhibit higher utility, and their removal negatively affects non-skippers. With endogenous commission rates, the negative effect on non-skippers becomes more pronounced as Ω increases, for the same reason.

When considering the aggregate effect on consumers, which weighs the impacts on both skippers and non-skippers based on their respective shares, the overall change in consumer surplus remains positive under a fixed commission rate and negative under endogenous commission rates. In the latter case, the losses experienced by non-skippers outweigh the benefits enjoyed by skippers, resulting in a negative overall impact on consumer surplus.

Panel B of Appendix Figure A.24 depicts the effects on seller profits and platform revenues. Under endogenous commission rates, as Ω increases, the negative impact of removing sponsored positions on seller profits diminishes significantly. This change happens because, at higher values of Ω, the estimate of the platform's weight on consumers and sellers is higher, leading to a smaller increase in the commission rate following the removal of sponsored positions.

Moreover, as Ω increases, the negative impact of removing sponsored positions on the platform's total revenues becomes less pronounced. Under endogenous commission rates, when Ω is high, removing sponsored product advertising does not notably reduce the platform's revenues. Therefore, if a substantial portion of consumers consistently ignore sponsored products, it is plausible that the platform could benefit from eliminating such products.

In summary, within the reasonable range of values considered for Ω, the findings remain qualitatively consistent with those in the baseline model.

F.1.2 Predicted Utility Based on Observables

To mitigate potential biases in estimating the utility of sponsored products due to consumers' aversion to them, I construct an alternative measure of product utility that relies on objective observables. This measure aims to offer a more robust assessment of utility for sponsored products.

To construct this measure, I first select products that never appear in a sponsored position within a week and pool their estimated product utility, $\hat{\delta}_{jt}$, across all market–week pairs. Next, I regress $\hat{\delta}_{jt}$ on market fixed effects and four observables: the log of the number of reviews, average consumer rating, an indicator for Prime eligibility, and an indicator for Amazon's Choice. The coefficients on these observables have reasonable signs: products with more reviews and higher average consumer ratings generally have higher utility, and products eligible for Prime and chosen as Amazon's Choice also tend to have higher utility. Then I extend this relationship to all products, including those appearing in sponsored positions, and the measure represents the linear prediction from the regression. The resulting predicted product utility is denoted as $\tilde{\delta}_{jt}$.

In Panel A of Appendix Figure A.25, I compare product utility based on revealed preferences, $\hat{\delta}_{jt}$, with product utility based on observables, $\tilde{\delta}_{jt}$. While these two measures display a moderately high correlation of 0.7, the variation in the latter measure is considerably smaller. Panel B illustrates the average observable-based utility for all products in each position. Compared to the baseline result in Panel B of Figure 3, the utility difference between organic and sponsored products in similar locations shrinks significantly and is even reversed in the bottom positions. This change suggests that the observable-based utility measure could alleviate potential biases in utility estimates based on consumers' revealed preferences.

When consumers' preferences are biased, it becomes crucial to distinguish between consumers' perceived utility and actual utility (Allcott 2013; Train 2015). The former determines consumer choices, while the latter determines consumer welfare. In this context, consumers' perceived utility is represented by $\hat{u}_{ijt} = \hat{\delta}_{jt} + \epsilon_{ijt}$, while I assume the actual utility follows $\tilde{u}_{ijt} = \tilde{\delta}_{jt} + \epsilon_{ijt}$. The discrepancy between the two measures may stem from consumers' aversion to sponsored products.

When maximizing profits, sellers are only concerned with consumers' decision utility. Similarly, I assume that when the platform selects the commission rate, it only cares about consumers' decision utility. Therefore, all estimation results remain valid. The only distinction arises in how I calculate consumer surplus. Specifically, let s_{jt} denote product j's market share. Consumer surplus is given by

$$CS = \frac{1}{\alpha}\mathbb{E}(\max_j \tilde{u}_{ijt}) = \frac{1}{\alpha}\mathbb{E}(\max_j \hat{u}_{ijt}) + \frac{1}{\alpha}\sum s_{jt} \times (\tilde{\delta}_{jt} - \hat{\delta}_{jt}), \tag{F.33}$$

where the last term captures the welfare losses arising from biased preferences.

Appendix Figure A.26 presents the welfare results. Panel A illustrates consumer surplus in the status quo and counterfactual scenarios, calculated using both the baseline and the observable-based utility measures. Consumer surplus experiences a decline under the observable-based utility measure due to the discrepancy between perceived and actual utility. Panel B depicts the change in consumer surplus following the removal of sponsored positions. Notably, since sponsored products exhibit higher observable-based utility compared to the baseline utility, eliminating sponsored positions leads to a less pronounced positive effect under a fixed commission rate and a more substantial negative impact under endogenous commission rates. All findings remain qualitatively robust under this alternative utility measure.

F.2 Endogenous Consideration Sets

In the baseline model, consumers have heterogeneous consideration sets in the form of the top n products in search results, where $\lambda_n \in [0, 1]$ represent the fraction of consumers whose consideration sets contain the product in the n-th position. However, the population shares of different consumer types are exogenous and do not adapt to changes in prices or product arrangement. In this section, I consider an extension where consideration sets are determined endogenously in a sequential search model. This approach allows λ_n to change in a counterfactual scenario where all sponsored positions are eliminated.

F.2.1 Model

The sequential search model is built upon the framework developed by Lam (2023). In this model, consumers navigate the search results in a fixed order from the top to the bottom, gradually revealing the product arrangement and incorporating new products into their consideration sets. This model focuses on a stage of the search process different from the optimal search model developed by Weitzman (1979). In the latter, consumers start with a known utility component for a predefined set of options. They explore the products in an *unrestricted* order to disclose the remaining utility. However, in my model, consumers start with neither a utility component nor a predefined set of options. Prior to reaching a product, a consumer must traverse all the products located above it, which is consistent with the typical search process on a platform.

Before reaching a position, consumers do not observe the utility of the product, denoted as δ_j, in that position. Rather, they form expectations about a product's utility based on its location and whether it is a sponsored result. Let $\tilde{\delta}_n$ represent the expected utility of the product in the n-th position and j_n denote the actual product appearing in the n-th position. Consumers behave as though they have perfect knowledge of product utility, such that $\delta_{j_n} = \tilde{\delta}_n$.

Let $\mathcal{J}_n = \{j_1, ..., j_n\}$ denote the set of products in the top n positions. Consumer i with a consideration set $\mathcal{J}_{n_1}$ will add new products to her consideration set to form a new consideration set $\mathcal{J}_{n_2} \supset \mathcal{J}_{n_1}$ (i.e., $n_2 > n_1$), incurring a search cost of $(n_2 - n_1)s_i$, if and only if the following condition holds:

$$\mathbb{E}_\epsilon[\max\{\delta_{j_n} + \epsilon_{ij_n}\}_{n \le n_1} \cup \{\tilde{\delta}_n + \epsilon_{ij_n}\}_{n_1 < n \le n_2}] - \mathbb{E}_\epsilon[\max\{\delta_{j_n} + \epsilon_{ij_n}\}_{n \le n_1}] > (n_2 - n_1)s_i. \tag{F.34}$$

In this equation, $(n_2 - n_1)$ is the number of steps between the two consideration sets, and s_i denotes the *per-step* search cost for consumer i, drawn from a distribution F. The left-hand side of equation (F.34) represents the incremental expected utility gained from expanding the consideration set, while the right-hand side represents the total search cost incurred when transitioning from $\mathcal{J}_{n_1}$ to $\mathcal{J}_{n_2}$. The consumer will continue searching if there exists $n_2 > n_1$ for which equation (F.34) holds, meaning the incremental expected utility exceeds the cost of searching. Otherwise, she will stop her search and maintain her current consideration set.

The search process in this model is dynamic because of the uncertainty in the realizations of δ_{j_n}. A consumer may choose to stop searching if she encounters a high realization of δ_{j_n}, while she might continue if she encounters a low realization. The idiosyncratic preference shock, denoted as ϵ_{ij}, follows a type I extreme value distribution and is only realized after the search process.

To determine whether a consumer stops searching, I define a threshold function as follows:

$$s(n_1) = \max_{n_2 > n_1} \frac{\mathbb{E}_\epsilon[\max\{\delta_{j_n} + \epsilon_{ij_n}\}_{n \le n_1} \cup \{\tilde{\delta}_n + \epsilon_{ij_n}\}_{n_1 < n \le n_2}] - \mathbb{E}_\epsilon[\max\{\delta_{j_n} + \epsilon_{ij_n}\}_{n \le n_1}]}{n_2 - n_1}. \tag{F.35}$$

Given a consideration set $\mathcal{J}_n$ formed after reaching the n-th product, consumer i decides to stop searching if and only if her per-step search cost s_i exceeds $s(n)$. Thus, consumers with a per-step search cost s_i falling within the range $\big(s(n), \min_{n' < n} s(n')\big]$ have a consideration set $\mathcal{J}_n$.[42] The fraction of consumers who consider the n-th product, λ_n, is determined by the following expression:

$$\lambda_n = \sum_{n'=n}^{N} \int \mathbb{1}\Big(s_i \in \big(s(n), \min_{n' < n} s(n')\big]\Big) \mathrm{d}F(s_i). \tag{F.36}$$

[42] The set $\big(s(n), \min_{n' < n} s(n')\big]$ can be empty, indicating that no consumers have a consideration set $\mathcal{J}_n$. This scenario may occur, for example, if δ_{j_n} has a very low realization.

Equations (F.35) and (F.36), combined with consumers' expectations regarding product utility $\{\tilde{\delta}_n\}_{n=1}^N$, the realized product utility $\{\delta_{j_n}\}_{n=1}^N$, and the distribution of consumers' per-step search cost $F(\cdot)$, collectively determine the population shares of different consideration sets.

The model has a limitation that sellers and the platform do not respond to changes in consumers' search behaviors in counterfactual scenarios. In reality, sellers and the platform are likely to adjust their strategies based on consumers' behaviors, leading to a feedback loop where consumers may further change their search behaviors in response. This dynamic interplay creates a complex environment where each player's decision is influenced by the choices of the others. The equilibrium, thus, represents a "fixed point" where each player's choice is optimal given the choices of the others. However, estimating such a model is challenging and is beyond the scope of this paper. Despite its simplifications, the current model offers valuable insights into the formation of consideration sets and suggests potential avenues for future research.

F.2.2 Estimation

The estimation involves two steps: (1) constructing consumers' expectations regarding product utility in each position $\{\tilde{\delta}_n\}_{n=1}^N$, and (2) estimating the distribution of consumers' per-step search cost $F(\cdot)$ to match the observed population shares of different consideration sets.

I assume that consumers hold the following expectations regarding product utility:

$$\begin{aligned}\tilde{\delta}_n^{O,m} &= \bar{\delta}^{O,m} - \beta^O \log(n),\\ \tilde{\delta}_n^{R,m} &= \bar{\delta}^{R,m} - \beta^R \log(n).\end{aligned} \tag{F.37}$$

In this equation, m indexes markets, O refers to organic positions, and R stands for sponsored positions. The parameters $\bar{\delta}^{O,m}$ and $\bar{\delta}^{R,m}$ indicate the expected utility of the product in the first organic or sponsored position in market m. $\beta^O > 0$ and $\beta^R > 0$ measure the decrease in expected product utility as the rank increases, and I assume these parameters are common across markets. I estimate these relationships and construct consumers' expectations.

I assume that consumers' per-step search cost s_i follow a log-normal distribution, i.e., $\log(s_i) \sim N(\mu, \sigma^2)$. Given parameters (μ, σ), I can predict $\tilde{\lambda}_{nf}$ using equations (F.35) and (F.36) for each search, indexed by f, and calculate $\tilde{\lambda}_n$ as the average across all searches within a market. I estimate parameters (μ, σ) to minimize $\sum_{n=1}^N (\tilde{\lambda}_n - \hat{\lambda}_n)^2$, where $\hat{\lambda}_n = 1/\exp(\hat{\beta}(n-1))$ represents the fraction of consumers considering each position implied by the estimates in Section 5.1.1.

F.2.3 Results

Panel A of Appendix Figure A.27 plots the distribution of the average per-step search cost across markets. In each market, I calculate the average search cost in dollars as $\exp(\mu + \sigma^2/2)/\alpha$, where α represents the price sensitivity in consumers' utility, as defined in equation (3). The average per-step search cost is around a few cents in most markets, suggesting that moving from the first to the 20th position in search results incurs an estimated search cost of approximately half a dollar.

I use the estimated distribution of consumers' per-step search cost to predict the proportion of consumers considering each position in the following three scenarios:

1. $\tilde{\lambda}_n^0$: the status quo with sponsored positions;

2. $\tilde{\lambda}_n^1$: the scenario with sponsored positions removed under a fixed commission rate;

3. $\tilde{\lambda}_n^2$: the scenario with sponsored positions removed under endogenous commission rates.

For each set of λ_n, I calculate the average number of products in consumers' consideration sets as:

$$\sum_{n=1}^{N} n(\lambda_n - \lambda_{n+1}). \tag{F.38}$$

Panel B of Appendix Figure A.27 evaluates the model fit by comparing the average numbers of products in consumers' consideration sets under $\hat{\lambda}_n$, the estimates in Section 5.1.2, and under $\tilde{\lambda}_n^0$, the values predicted above. The two sets of numbers closely align across all markets, indicating a good fit. In Panel C. I compare the average numbers of products under $\tilde{\lambda}_n^0$ and $\tilde{\lambda}_n^1$. When sponsored positions are removed, consumers, on average, search for 2.6 products less (30.9 vs. 28.3). The removal of sponsored positions leads consumers to encounter high-utility products more quickly, prompting them to stop searching sooner. Panel D compares the average numbers of products under $\tilde{\lambda}_n^1$ and $\tilde{\lambda}_n^2$. With an increase in the commission rate, consumers, on average, search for 1.2 products more (28.3 vs. 29.5). The higher prices may push consumers to explore more options before making a purchase decision.

In Appendix Figure A.28, I assess the overall impact on consumer surplus with consideration sets endogenized and search costs incorporated. The first block illustrates consumer surplus without accounting for search costs, the second block quantifies total search costs incurred, and the third block presents net consumer surplus, calculated as the difference between consumer surplus and search costs. The results align qualitatively with the baseline findings: while removing sponsored positions increases net consumer surplus when holding the commission rate fixed, it decreases consumer surplus under endogenous commission rates.

F.3 Endogenous Organic Rankings

In the baseline model, I assume that the distribution of organic rankings remains fixed in the counterfactual exercises. However, in practice, the platform's ranking algorithm considers factors such as a product's historical sales and prices, which would vary if sponsored product advertising were removed. In this section, I estimate the ranking algorithm and allow for changes in the distribution of organic rankings in counterfactual scenarios.[43]

As introduced in Section 5.2.2, for each search f in week t, the platform assigns a score $\tilde{o}_{jt}^{f}$ to product j drawn from a product-specific distribution, $\log(\tilde{o}_{jt}^{f}) \overset{\text{i.i.d.}}{\sim} N(\log(o_{jt}), \sigma_{o,t}^2)$, where o_{jt} denotes the average score of product j in week t. Products are ranked based on their realized scores. I employ a Gibbs sampler to estimate $\boldsymbol{o}_t = (o_{1t}, \ldots, o_{Jt})$ in each week to match the realized distribution of organic rankings.

I assume that the average score o_{jt} is determined by the following ranking algorithm:

$$o_{jt} = \gamma_j + \beta_1 \log(p_{j,t-1}) + \beta_2 \log(s_{j,t-1}) + \epsilon_{jt}, \tag{F.39}$$

where $p_{j,t-1}$ and $s_{j,t-1}$ represent the price and sales of product j in the previous week, respectively. The equation incorporates product fixed effects γ_j, which remain constant throughout the sample period. Therefore, the estimation of equation (F.39) leverages variations within a product over time. The estimated coefficients β_1 and β_2 are reasonable: a higher price is associated with a lower average score, while a larger sales volume is associated with a higher average score.

With this ranking algorithm established, I can now allow the distribution of organic rankings to change in the counterfactual exercises. For the first week in my sample, I assume that the average scores are fixed to determine the new equilibrium after the removal of advertising. However, for $t \geq 2$, I utilize the new equilibrium prices and sales from the preceding week, $\tilde{p}_{j,t-1}$ and $\tilde{s}_{j,t-1}$, to compute the updated average scores as follows:

$$\tilde{o}_{jt} = \gamma_j + \beta_1 \log(\tilde{p}_{j,t-1}) + \beta_2 \log(\tilde{s}_{j,t-1}) + \epsilon_{jt}. \tag{F.40}$$

[43] Sellers and the platform may have other strategic responses. For example, sellers can invest in improving their organic rankings by setting lower prices. The platform may manipulate organic rankings to intensify the competition in auctions (Long and Liu 2023). Fully incorporating these strategic responses is beyond the scope of this paper.

This algorithm retains product fixed effects and unobserved residuals ϵ_{jt}. Subsequently, I use the new average scores to simulate organic rankings in week t and solve for the new equilibrium. The welfare results under endogenous organic rankings, as shown in Appendix Table A.3, are all qualitatively similar.

F.4 Market-Specific Commission Rates

On Amazon, commission rates exhibit limited variation across different markets. For most markets in my sample, the commission rate is either 15% or 17%. In this section, I explore the potential impact of Amazon implementing market-specific commission rates.

For each market, I determine the optimal commission rate that maximizes the platform's objective function, both with and without sponsored positions. As depicted in Panel A of Appendix Figure A.29, the optimal commission rate without sponsored positions consistently surpasses the rate when sponsored positions are present, confirming the intuition from the baseline results. In Panel B, I calculate the averaged predicted commission rate separately for markets with observed commission rates of 15% and 17%. Notably, the average predicted commission rate is higher in the second group of markets and closely aligns with the actual 17%. In Panel C, the welfare effects of removing sponsored positions under market-specific commission rates are qualitatively the same as the baseline results.

F.5 Long-Term Effects

The main dataset utilized in this paper covers a relatively short period, which limits my ability to investigate some longer-term impacts. For instance, sponsored product advertising can enhance the visibility of new entrants, thereby incentivizing more sellers to enter the market. On the other hand, the increase in advertising expenses may diminish the profitability of selling on Amazon, potentially prompting some sellers to leave the platform.

To provide suggestive evidence on these longer-term effects, I collected supplementary data on March 25, 2024, approximately 20 months after the initial sample period. During this data collection, I gathered search results for the same set of keywords five times. For each product in the main sample, I can identify whether it appears in the organic search results of the same keyword in the new sample. Below, I consider two types of products: those present in organic positions in the main sample and those solely present in sponsored positions.

First, I document the survival rate for products present in organic positions in the main sample. As depicted in Panel A of Appendix Figure A.30, the turnover rate is notably high. Among products consistently occupying an organic position across all searches in the main sample, only about 50% remain in such a position 20 months later. This percentage declines sharply for products that appear in organic positions less frequently. Moreover, the survival rate is higher for products that invested in advertising and also appeared in sponsored positions. This effect is consistent across the board. These observations underscore the intense competition on Amazon and suggest that even high-ranking products, if they do not invest in advertising, have a higher chance of exiting the market in the long run, potentially due to decreased profitability.

Second, I investigate whether advertising can assist sellers with low visibility in climbing organic rankings. I focus on products that initially had no presence in organic search results but reached consumers only through sponsored product advertising. Panel B of Appendix Figure A.30 illustrates that the majority of these products have not transitioned to organic positions 20 months later. However, a higher frequency of appearances in sponsored positions is indeed associated with a moderate increase in the likelihood of appearing in an organic position. While this correlation is not definitive, it implies that advertising could facilitate the successful breakthrough of a small subset of lesser-known products.

This exercise suggests that while advertising may help products with lower visibility ascend organic rankings, it could also contribute to the exits of high-ranking products. Future studies with extended time coverage could offer a more comprehensive insight into these effects.

G Robustness Checks

G.1 Alternative Autoregressive Models

In Section 5.1.2, the identification of consumers' search friction hinges on the assumption that the unobserved demand shock ξ_{jt} follows an AR(1) process, as defined in equation (5). To explore the robustness of the estimates, I consider a more generalized autoregressive model, which allows for a more flexible characterization of the dynamics of the unobserved demand shock over time. Specifically, I assume that ξ_{jt} follows an AR(p) process given by:

$$\xi_{jt} = \sum_{\tau=1}^{p} \rho_\tau \xi_{j,t-\tau} + \eta_{jt}, \quad \eta_{jt} \sim \text{i.i.d.}, \quad \mathbb{E}(\eta_{jt}) = 0, \quad \eta_{jt} \perp \xi_{j,t-\tau}, \forall \tau \geq 1, \tag{G.41}$$

where p represents the order of the autoregressive process. In the baseline analysis, $p = 1$, leading to equation (5). As shown in Panel A of Appendix Figure A.31, the estimates of β under alternative values of $p = 2, 3, 4, 5$ closely resemble the estimates derived under the specification with $p = 1$. Panel B presents the estimates of ρ_τ when $p = 5$. Notably, while ρ_1 averages around 0.4, ρ_2 diminishes to less than 0.1 in most markets, and ρ_3 approaches 0. This empirical pattern suggests that the AR(1) model is a reasonable approximation, reinforcing the validity of the baseline findings.

G.2 Alternative Extent of Search Frictions

An important factor in estimating seller primitives and the platform's weight, as well as in conducting counterfactual simulations, is the extent of consumers' search frictions, which determines how a product's position in search results impacts its sales. Given the lack of experimental variation, establishing the causal effect of positions on sales is challenging.

In Section 5.1.2, I parametrize the fraction of consumers whose consideration sets contain the n-th product as $\lambda_n = 1/\exp(\beta(n-1))$. To identify the parameter β, I examine how changes in a product's daily positions affect its sales while accommodating demand persistence. This identification approach relies on a few assumptions regarding the evolution of organic ranks and demand shocks on Amazon. To address concerns regarding the sensitivity of the welfare results to this parameter estimate, I conduct a robustness check to evaluate the impact of varying values of β on the welfare outcomes.

In this robustness exercise, I consider six values of $\beta = 0.015, 0.02, 0.025, 0.03, 0.035, 0.04$,[44] and I assume all markets share the same β. These values offer a range that encompasses various alternative estimates for β. First, Ursu (2018) rely on experimental variation to identify the impact of ranks on consumer search. The results depicted in Figure 2 of Ursu (2018) suggest a β value of 0.021. Second, Lam (2023) adopts a similar identification strategy as in my paper to estimate consumers' search behavior in the "Home & Kitchen" category on Amazon. The results illustrated in Figure 11 of Lam (2023) imply a β value of 0.04. Third, Column 6 of Table 2 in my paper presents a reduced-form estimate that corresponds to a β value of 0.016.

For each value of β, I re-estimate the model, including product utility, seller primitives, and the platform's weight, and then conduct counterfactual simulations. Appendix Figure A.32 presents the utility estimates for products in each position under different values of β. When β is small, the

[44] Appendix Figure A.11 provides examples of this function under different values of β. In the demand estimation, the average $\hat{\beta}$ across markets is 0.027.

slope between product utility and position is notably steep, possibly reflecting the mechanical correlation between positions and product utility when search frictions are not adequately accounted for. As β increases, the slope becomes flatter.

Appendix Figure A.33 illustrates the welfare effects of eliminating sponsored product advertising. Under a fixed commission rate, the impact on consumers and sellers is more sensitive to the degree of search frictions. Specifically, when search frictions are higher (i.e., when β is larger), removing advertising proves more beneficial for consumers and sellers. This is because higher search frictions exacerbate the welfare losses incurred when low-utility sponsored products displace high-utility organic alternatives. Thus, the elimination of sponsored product advertising would yield more substantial welfare gains.

With endogenous commission rates, the welfare effects remain consistent across different values of β. The intuition behind this consistency lies in the fact that when search frictions are high, advertising revenues play a more significant role in the platform's total revenues. Consequently, I estimate that the platform assigns a lower weight to consumers and sellers in such cases. Upon removing sponsored positions, the platform will raise its commission rate by a slightly larger magnitude. This adjustment offsets the increased benefits brought by high-utility organic options, resulting in relatively stable overall effects.

This analysis emphasize that while the magnitude of welfare effects may vary with different values of β, the qualitative conclusions remain robust across reasonable parameter ranges.

G.3 Assumption on Duplicated Listings

In the baseline analysis, I assume that the second appearance of the same product in search results does not receive an independent idiosyncratic preference shock when calculating consumer surplus. Under this assumption, any duplicated listings resulting from a product appearing in both sponsored and organic positions reduce product variety, a source of inefficiency associated with sponsored product advertising. To test the robustness of my findings, I explore an alternative assumption where the second appearance also receives its independent preference shock. Appendix Table A.4 and Appendix Figure A.34 present the main welfare results under this alternative assumption, confirming the qualitative robustness of my findings. With this duplication no longer considered a source of inefficiency, the impact of removing sponsored product advertising on consumer surplus differs in magnitude, resulting in a smaller positive effect under a fixed commission rate and a larger negative impact under endogenous commission rates.

References

Allcott, H. (2013). The Welfare Effects of Misperceived Product Costs: Data and Calibrations from the Automobile Market. *American Economic Journal: Economic Policy*, 5(3):30–66.

Brynjolfsson, E., Hu, Y., and Simester, D. (2011). Goodbye Pareto Principle, Hello Long Tail: The Effect of Search Costs on the Concentration of Product Sales. *Management Science*, 57(8):1373–1386.

Chevalier, J. and Goolsbee, A. (2003). Measuring Prices and Price Competition Online: Amazon.com and BarnesandNoble.com. *Quantitative Marketing and Economics*, 1:203–222.

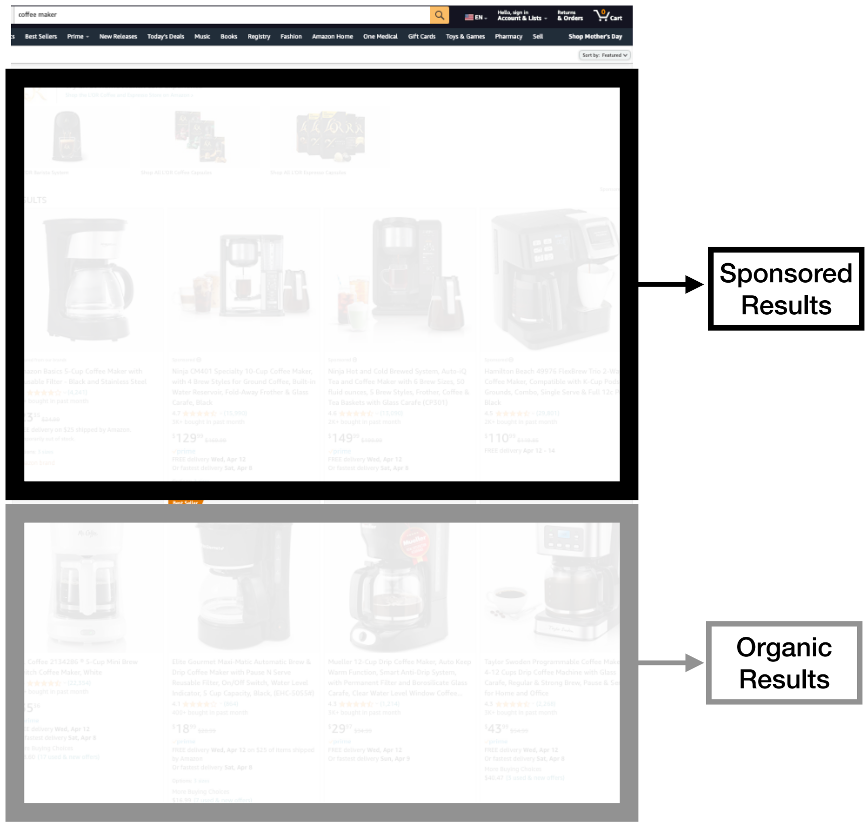
Long, F. and Liu, Y. (2023). Platform Manipulation in Online Retail Marketplace with Sponsored Advertising. *Marketing Science*.

Train, K. (2015). Welfare Calculations in Discrete Choice Models When Anticipated and Experienced Attributes Differ: A Guide with Examples. *Journal of Choice Modelling*, 16:15–22.

Weitzman, M. L. (1979). Optimal Search for the Best Alternative. *Econometrica*, 47(3):641–654.

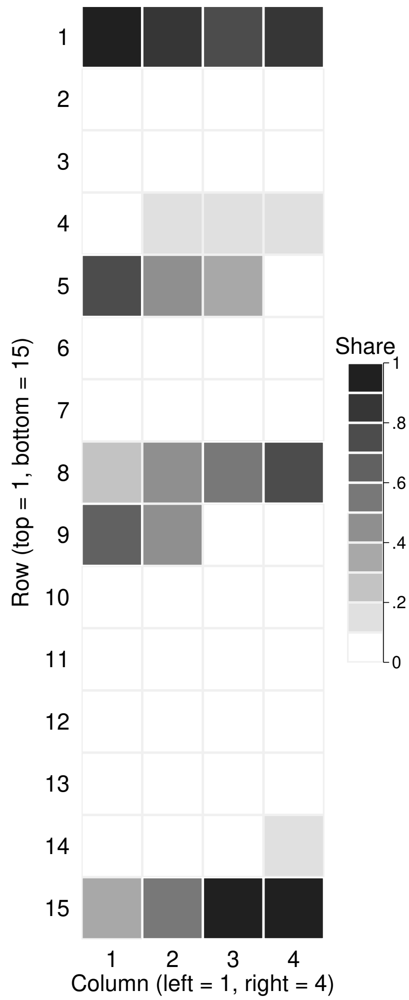
Yin, L. and Jeffries, A. (2021). How We Analyzed Amazon's Treatment of Its "Brands" in Search Results. https://themarkup.org/amazons-advantage/2021/10/14/how-we-analyzed-amazons-treatment-of-its-brands-in-search-results. Accessed in April 2023.

Figure A.1: Example of Search Results on Amazon



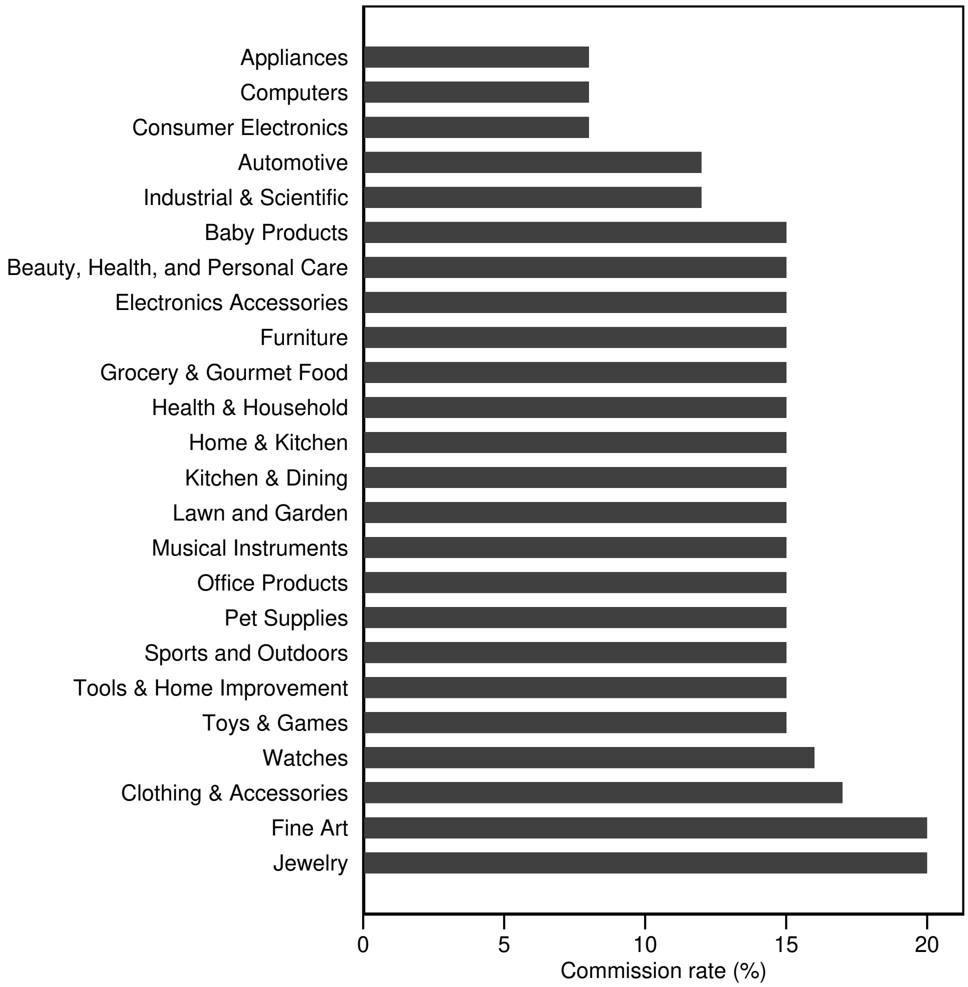
Notes: This figure provides an example of the search results on Amazon. The majority of search results are organic, determined by Amazon's ranking algorithm based on factors such as relevance to the searched keyword, sales performance, and consumer reviews. The remaining results are sponsored, where sellers pay the platform to display their products. Sponsored products are determined through ad auctions, with sellers ranked according to their bids. They are labeled onscreen as "Sponsored" to distinguish them from organic results.

Figure A.2: Distribution of Sponsored Positions in Search Results



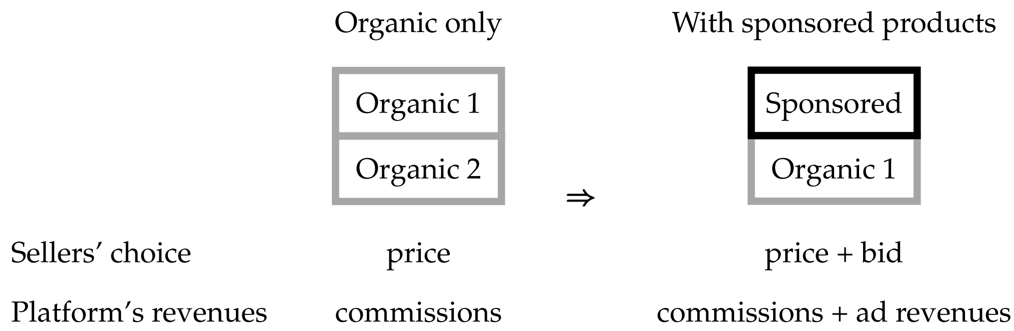
Notes: This figure visually represents the distribution of sponsored positions in search results, which consist of a total of 60 positions on the first page. These results are typically organized in 15 rows and 4 columns when viewed on a desktop or laptop browser. Each block in the figure corresponds to one position in the search results. For each of these positions, I calculate the fraction of searches where a sponsored product appears in that position. A fraction of 1 indicates that the position always displays a sponsored product, while a fraction of 0 indicates that the position always displays an organic product.

Figure A.3: Amazon's Commission Rates Across Product Categories



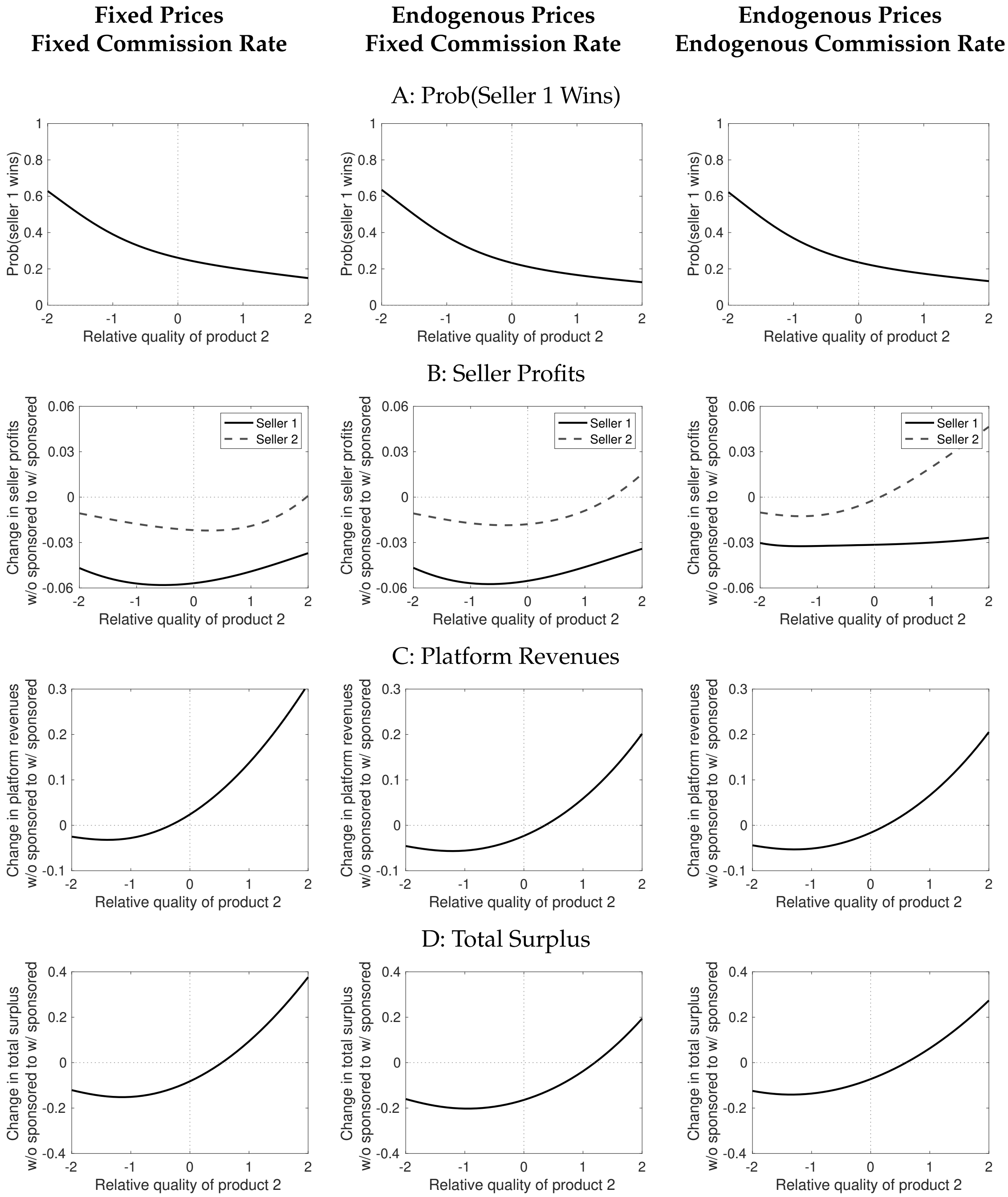
Notes: This figure presents the commission rates set by Amazon across various product categories. The data was collected from Amazon's disclosure, available at https://sell.amazon.com/pricing#referral-fees, accessed in May 2022. For each unit sold on Amazon's marketplace, Amazon collects a commission fee from the seller equal to a percentage of the selling price.

Figure A.4: Setup of the Stylized Example



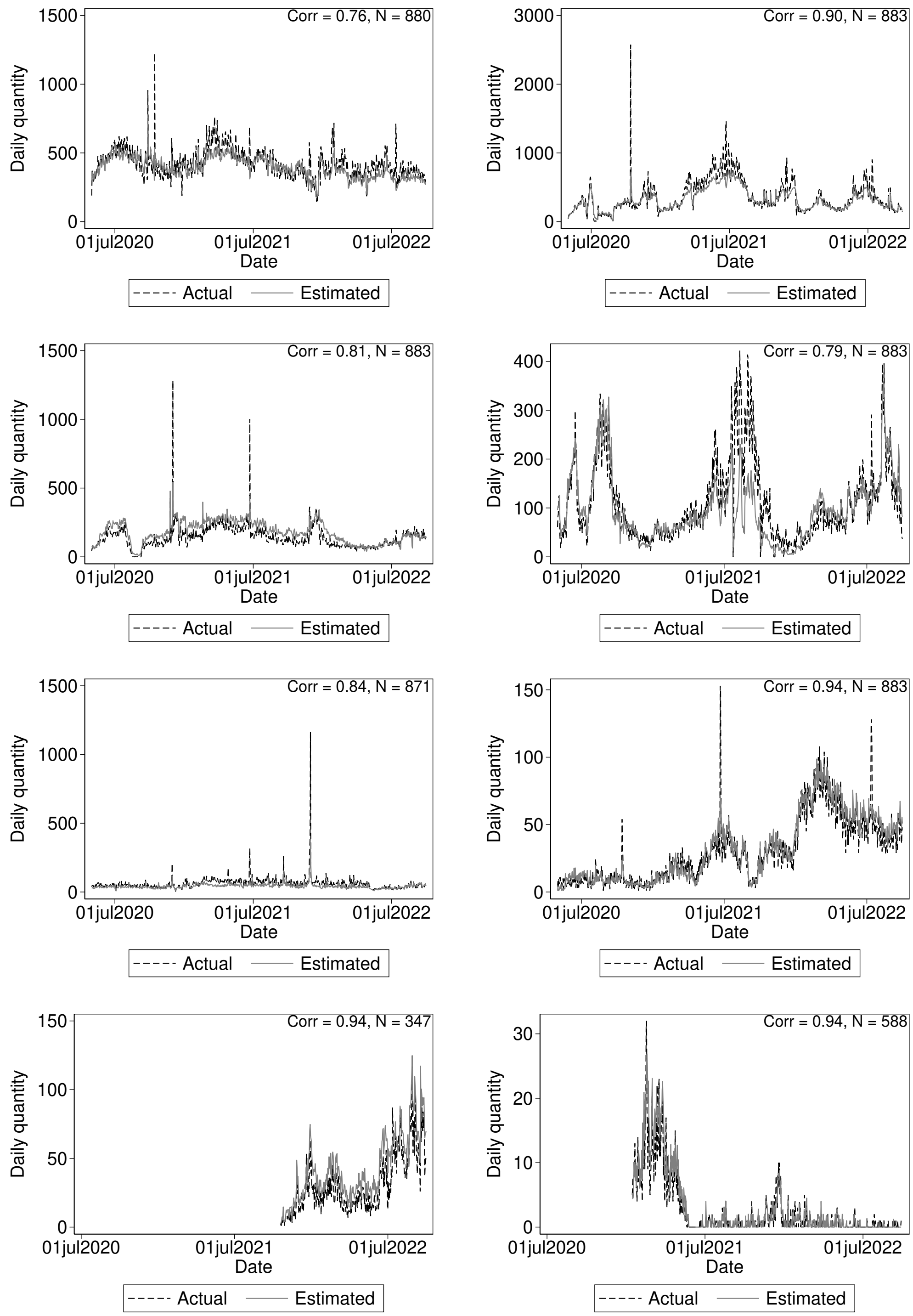
Notes: This figure provides a visual representation of the stylized example's setup in Section 3.1. In the scenario with only organic results, both positions in search results are determined based on organic ranks. Each seller sets a price to maximize its profits. The platform chooses a commission rate to maximize its total commissions. In the scenario with sponsored products, the first position is allocated through an ad auction, while the second position displays the top-ranked organic product. Each seller sets a price and submits a bid to maximize its profits, and the platform chooses a commission rate to maximize the sum of total commissions and advertising revenues.

Figure A.5: Additional Simulation Results of the Stylized Example



Notes: This figure presents additional simulation results of the stylized example. For each panel, the y-axis (except for Panel A) shows the change in outcomes from the equilibrium with two ranked organic positions and no sponsored positions to the equilibrium with a sponsored position at the top and an organic position at the bottom. The x-axis represents the relative quality of product 2 compared to product 1. Each column corresponds to a different assumption regarding how sellers and the platform can respond to the introduction of a sponsored position. See the notes of Figure 1 for these assumptions. Panel A displays the probability that seller 2 wins the sponsored position. Panels B to D display the changes in the profits of both sellers, the platform's total revenues, and total surplus, respectively. See Section 3 for the setup of the stylized example.

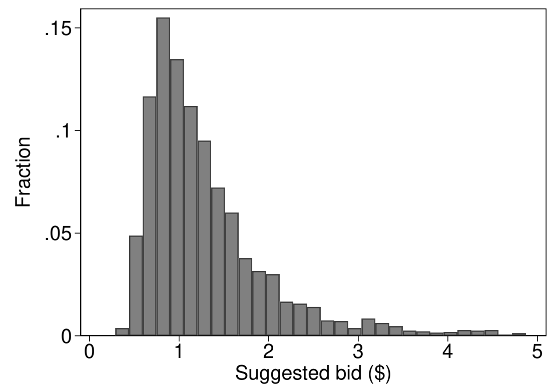
Figure A.6: Comparison of Estimated and Actual Daily Sales Quantity



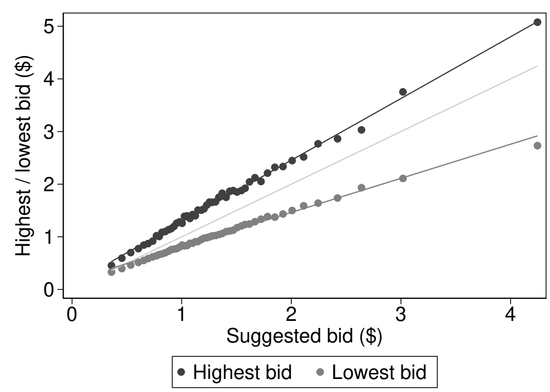
Notes: This figure compares estimated and actual daily sales quantities for eight products sold on Amazon. I obtained actual daily sales quantities over a two-year period for a few hundred products sold on Amazon in collaboration with an Amazon seller. The figure displays the actual and estimated daily sales quantities for eight randomly chosen products with varying sales volumes. The estimation is based on the Best Seller Ranks using the method detailed in Appendix B.2. Each panel includes the correlation coefficient between actual and estimated daily sales quantities.

Figure A.7: Aggregate Statistics of Winning Bids on Amazon

A: Distribution of the Median Winning Bids

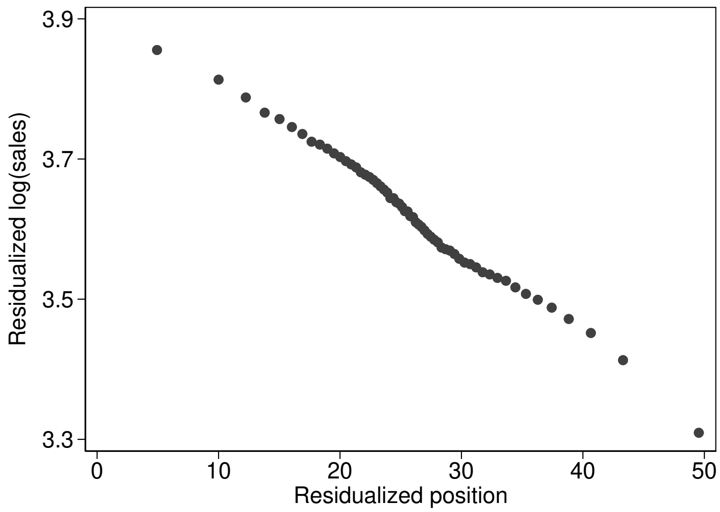


B: Relationships Between Median, Lowest, and Highest Winning Bids



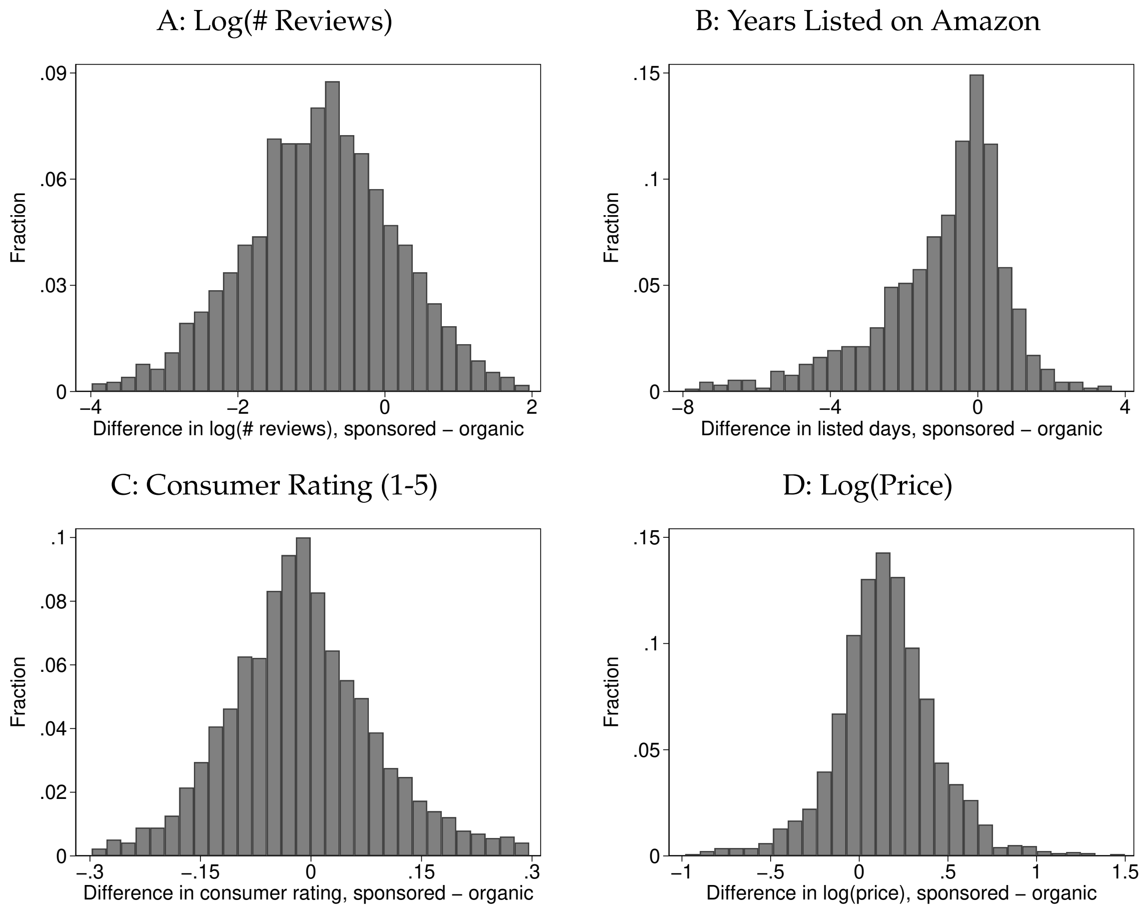
Notes: This figure summarizes the median, lowest, and highest bids of recent auction winners for each keyword reported by Amazon. Panel A plots the distribution of median winning bids across all keywords in the sample. Panel B presents binned scatter plots between the highest or lowest winning bids and the median winning bid.

Figure A.8: Relationship Between Positions and Sales



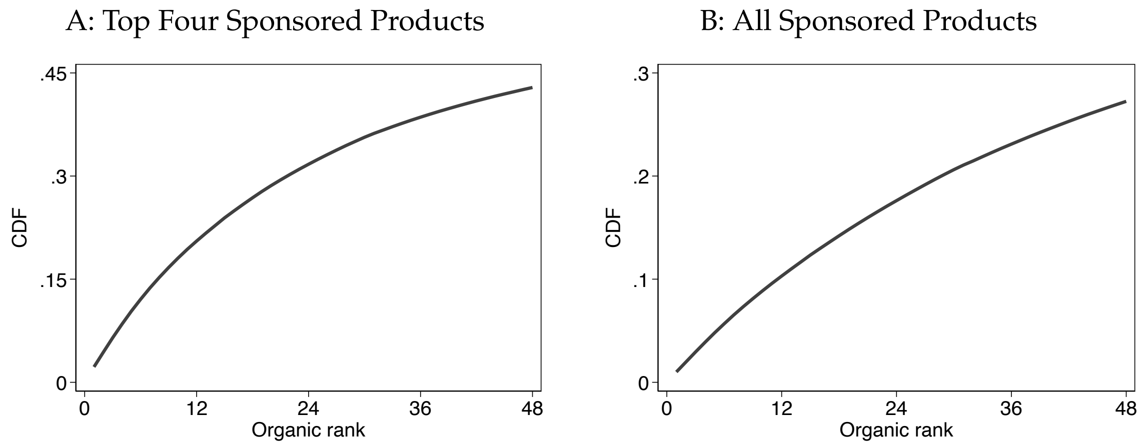
Notes: This figure illustrates the relationship between a product's position in search results and its sales quantity. Each observation is a product on a given day. I first calculate the product's average position in the search results for a keyword, with lower values indicating higher positions. When a product appears in the search results for multiple keywords, I retain only the keyword with the highest frequency of appearance. The sample includes products that appear in the search results for a keyword in more than 50% of the searches. I then residualize the product's average position and the log of its sales quantity by product fixed effects and keyword–day fixed effects, recenter them at their respective sample means, and plot a binned scatter plot of the two residualized variables.

Figure A.9: Distribution of Differences Between Top Sponsored and Organic Products



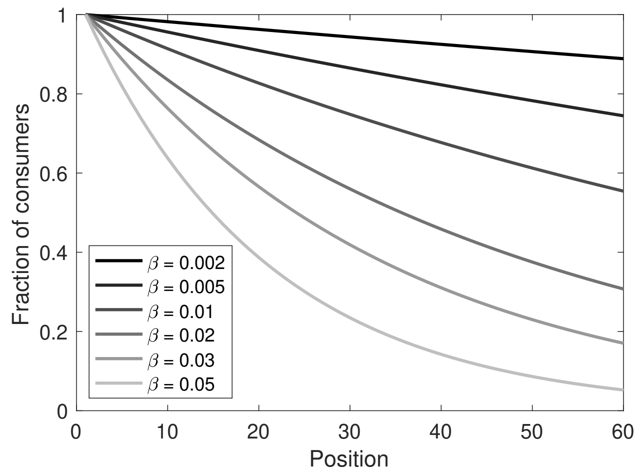
Notes: This figure displays the distribution of the differences between the top four sponsored and four organic products across all the keywords in the sample. I estimate equation (2) separately for each keyword in my sample following the same procedure used to construct Figure 2. In Panels A to D, the variables are the log of the number of reviews, years listed on Amazon, average consumer rating, and the log of price, respectively.

Figure A.10: Overlap Between Sponsored and Organic Products



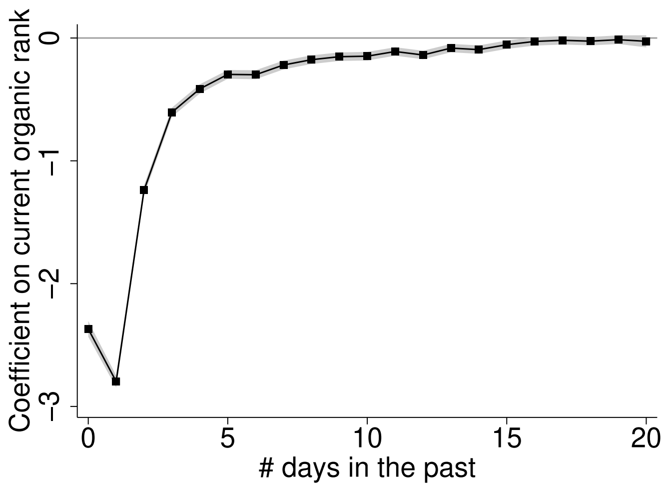
Notes: This figure examines the degree of overlap between sponsored and organic products in Amazon's search results. For each sponsored product, I identify whether it appears in an organic position on the same result page and, if so, its organic rank. This figure plots the cumulative distribution function of organic ranks for sponsored products. Panel A focuses on the top four sponsored products in each search. Panel B includes all sponsored products on the first page.

Figure A.11: Examples of Parametrization in Consumers' Consideration Sets



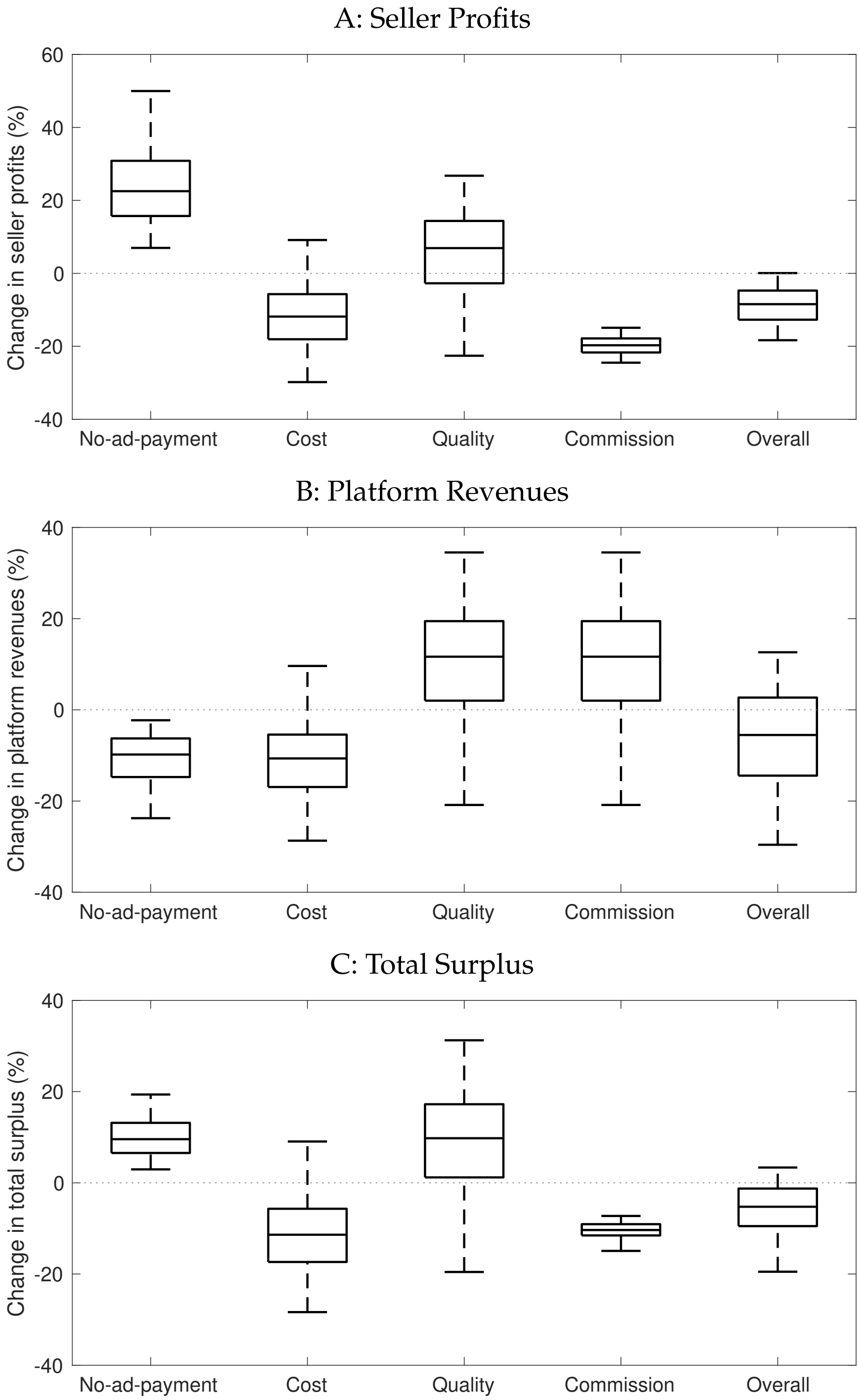
Notes: This figure illustrates the parametrization used to define the fraction of consumers whose consideration sets contain each position in the search results, as introduced in Section 5.1.2. This parametrization is denoted as $\lambda_n = 1/\exp(\beta(n-1))$ for $1 \leq n \leq N$. Different lines represent this function under different values of the parameter β.

Figure A.12: Effects of Past Sales Performance on Current Organic Rank



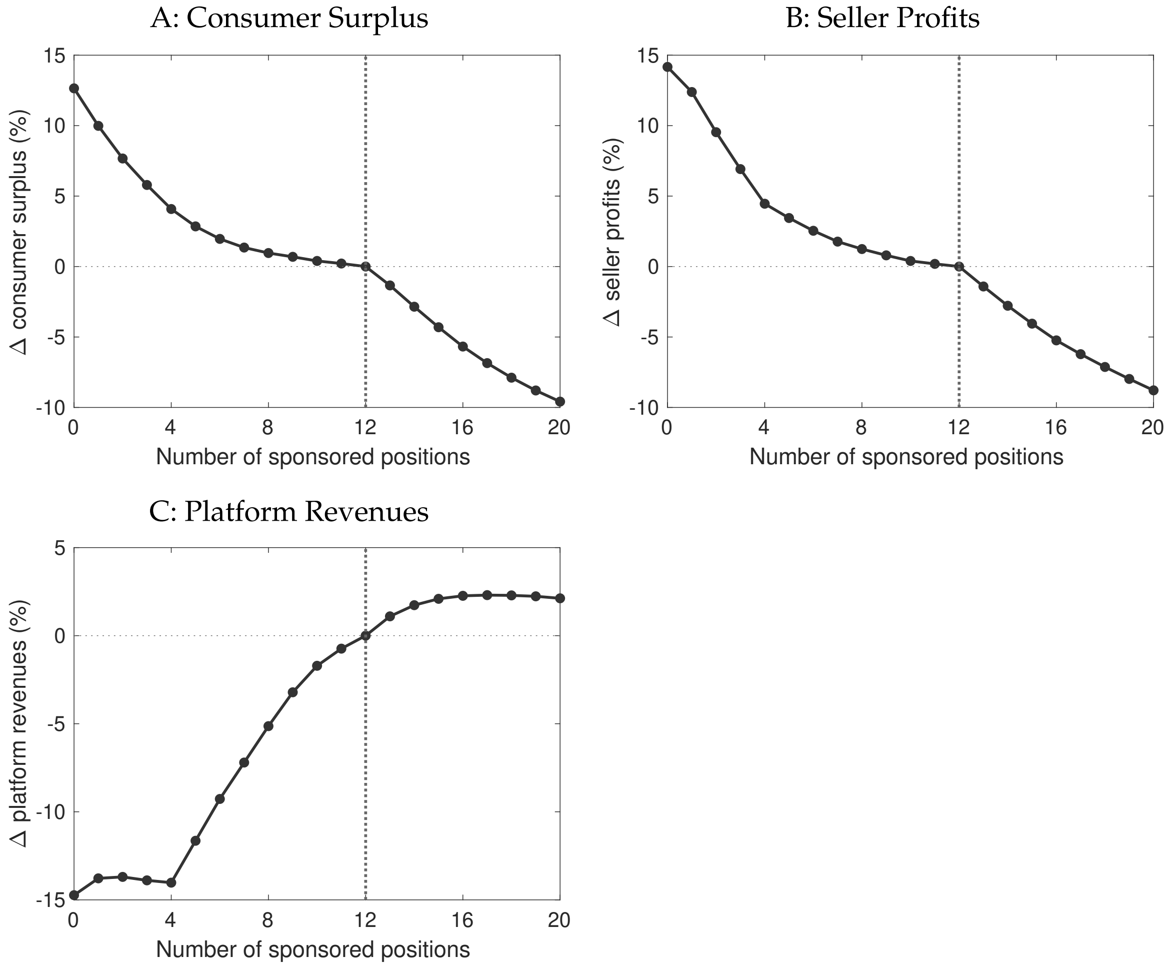
Notes: This figure illustrates the effects of a product's current and past sales performance on its organic rank. Each observation is a product on a given day. I calculate the product's average organic rank in the search results for a keyword, with lower values indicating higher positions. When a product appears in the search results for multiple keywords, I retain only the keyword with the highest frequency of appearance. The sample includes products that appear in the search results for a keyword in more than 50% of the searches. I regress the product's average organic rank on the current and lagged log of its sales quantities while controlling for product fixed effects and keyword–day fixed effects. This figure plots the coefficients on the current and lagged log of sales quantities along with 95% confidence intervals in shaded areas. Standard errors are clustered at the product level.

Figure A.13: Decomposition of Effects of Removing Sponsored Product Advertising



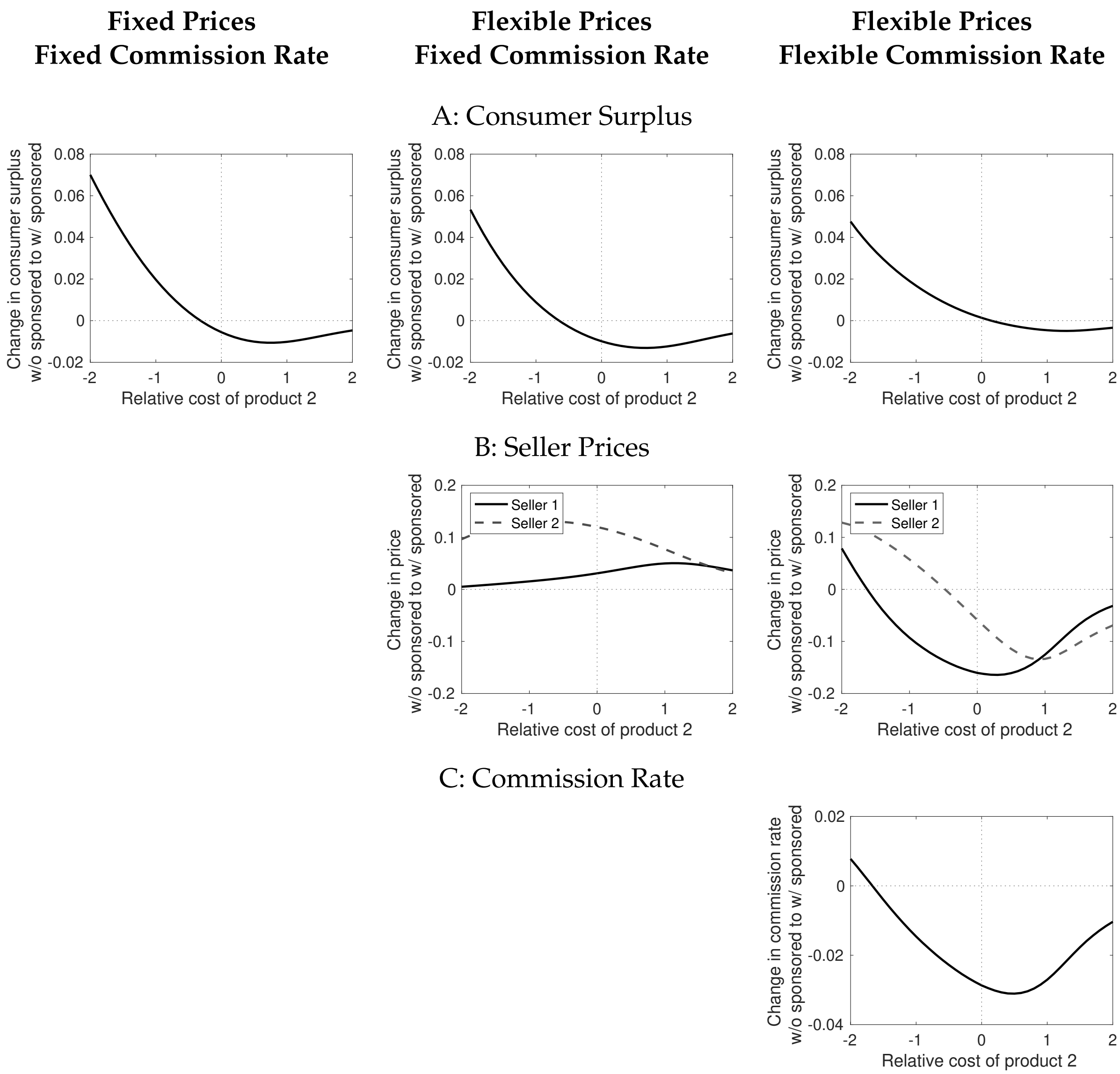
Notes: This figure decomposes the aggregate effect of removing sponsored product advertising on seller profits (Panel A), platform revenues (Panel B), and total surplus (Panel C). See Section 7.2 and the notes of Figure 7 for the definitions of these components. Each box presents the distribution of one of the four effects and the overall effect across market–weeks, with the middle line indicating the median, the edges of a box indicating the 25th and 75th percentiles, and the outer lines indicating the 5th and 95th percentiles. Figure 7 plots the decomposition of the aggregate changes in consumer surplus.

Figure A.14: Welfare Effects Under Varying # of Sponsored Positions, Fixed Commission



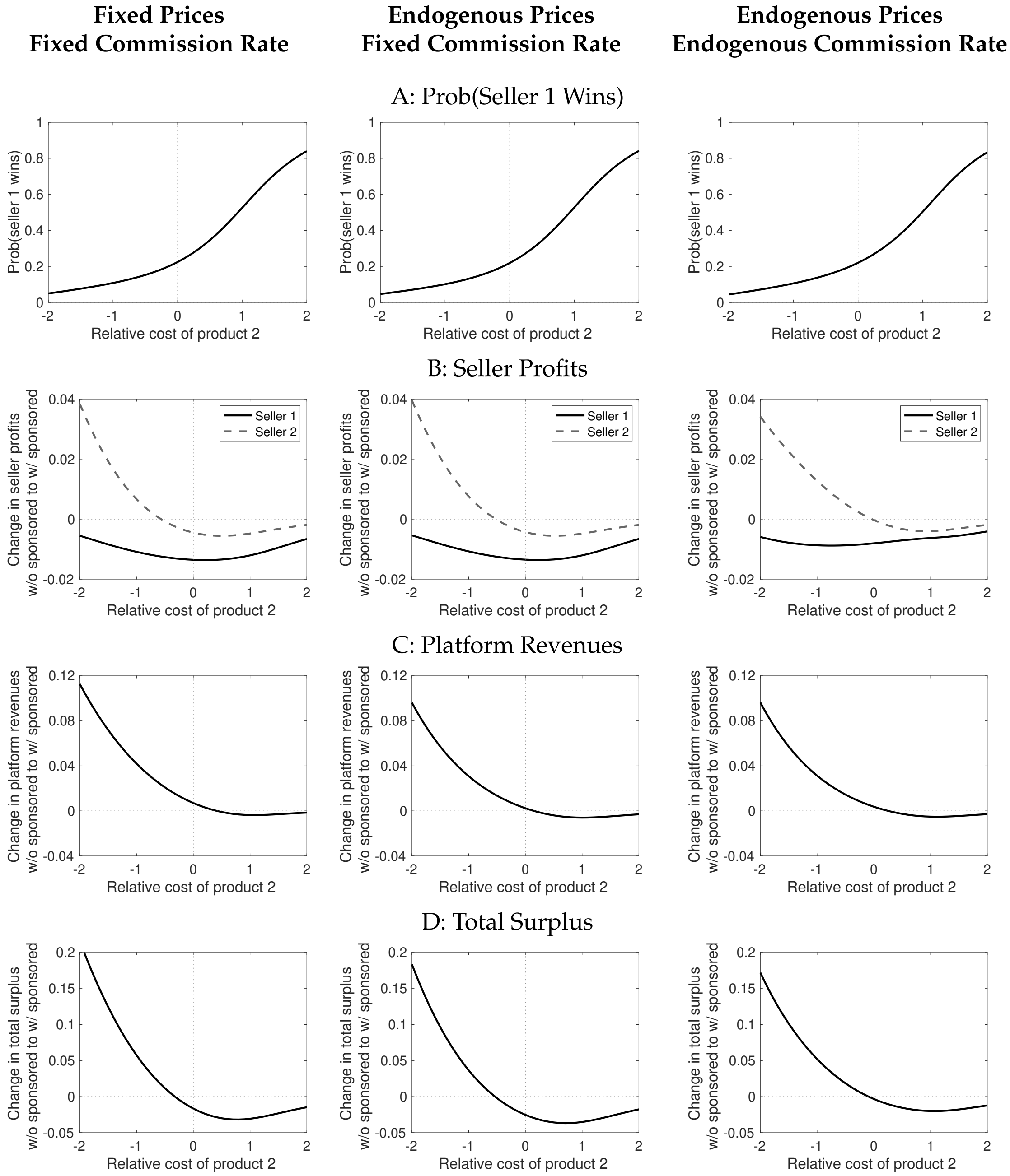
Notes: This figure examines the welfare effects of altering the number of sponsored positions in search results under a fixed commission rate of 15.6%. I change the number of sponsored positions in two directions: (i) a gradual removal of sponsored positions, starting from the lowest one and progressing to the highest; and (ii) a gradual addition of more sponsored positions in the middle, starting by substituting the 5th organic position with a sponsored one and continuing down to the 12th one. Panels A to C depict the aggregate changes in consumer surplus, seller profits, and platform revenues relative to the status quo with 12 sponsored positions. Figure 9 presents the results under endogenous commission rates.

Figure A.15: Simulation Results of the Stylized Example, Varying Cost



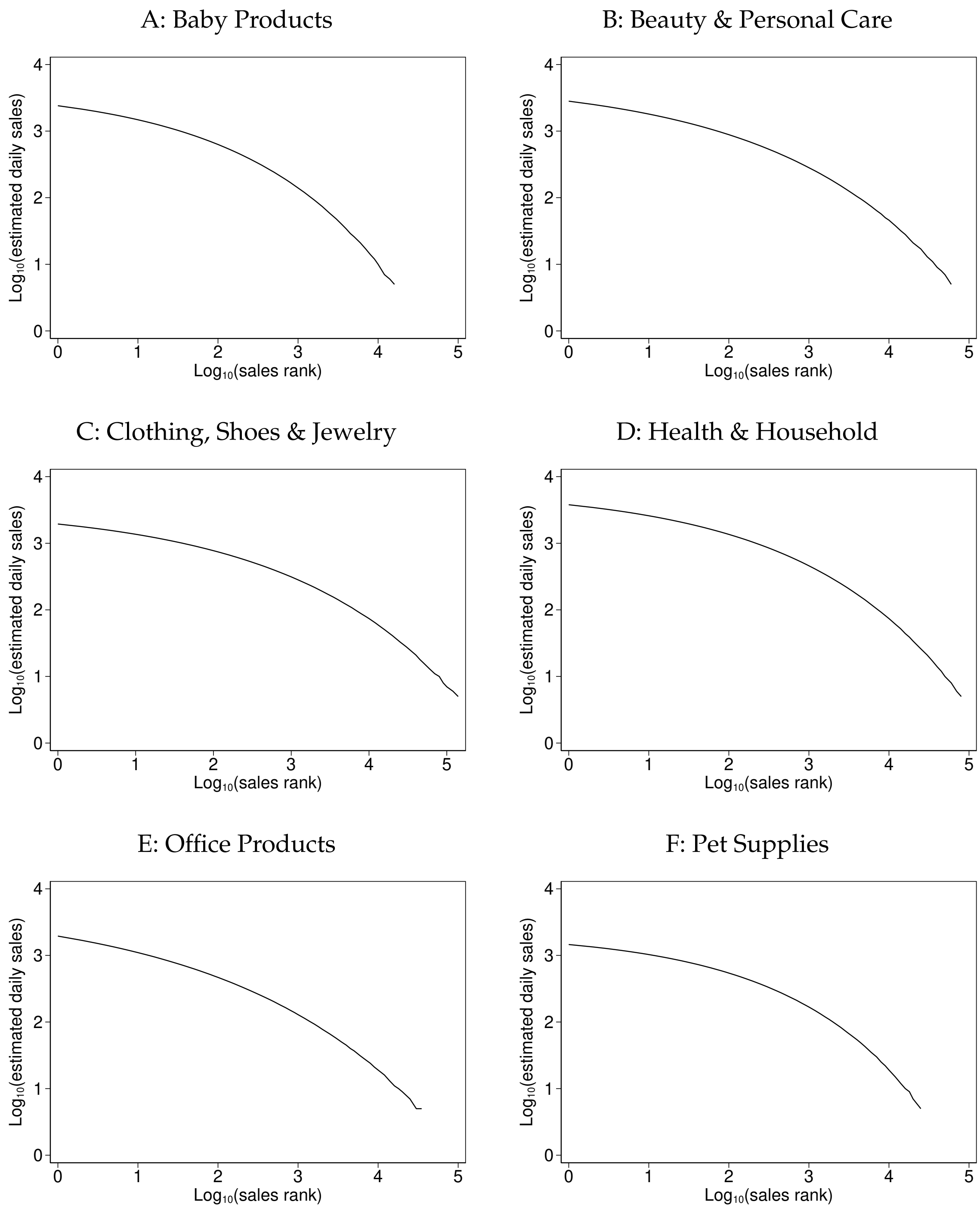
Notes: This figure presents the simulation results of the stylized example. In contrast to the baseline simulations, both products have the same quality, and I hold the marginal cost of product 1 constant while varying the marginal cost of product 2. For each panel, the y-axis shows the change in outcomes from the equilibrium with two ranked organic positions and no sponsored positions to the equilibrium with a sponsored position at the top and an organic position at the bottom. The x-axis represents the relative marginal cost of product 2 compared to product 1. Each column corresponds to a different assumption regarding how sellers and the platform can respond to the introduction of a sponsored position. See the notes of Figure 1 for these assumptions. Panels A to C display the changes in consumer surplus, the prices of both products, and the commission rate, respectively. See Section 3 for the setup of the stylized example and Appendix Figure A.16 for additional simulation results.

Figure A.16: Additional Simulation Results of the Stylized Example, Varying Cost



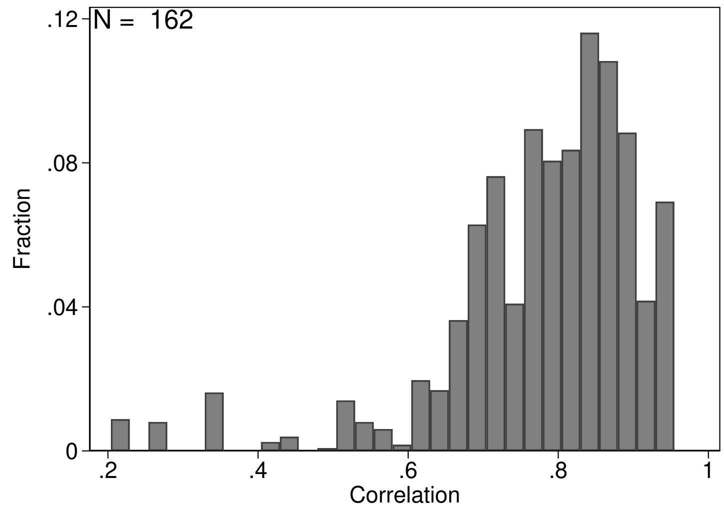
Notes: This figure presents additional simulation results of the stylized example. In contrast to the baseline simulations, both products have the same quality, and I hold the marginal cost of product 1 constant while varying the marginal cost of product 2. For each panel, the *y*-axis (except for Panel A) shows the change in outcomes from the equilibrium with two ranked organic positions and no sponsored positions to the equilibrium with a sponsored position at the top and an organic position at the bottom. The *x*-axis represents the relative marginal cost of product 2 compared to product 1. Each column corresponds to a different assumption regarding how sellers and the platform can respond to the introduction of a sponsored position. See the notes of Figure 1 for these assumptions. Panel A displays the probability that seller 2 wins the sponsored position. Panels B to D display the changes in the profits of both sellers, the platform's total revenues, and total surplus, respectively. See Section 3 for the setup of the stylized example.

Figure A.17: Mapping Between Best Sellers Rank and Sales Quantity



Notes: This figure depicts the mappings from the Best Sellers Ranks to sales quantities for six product categories on Amazon. These mappings were constructed by Jungle Scout for each product category using actual sales data obtained from Amazon sellers. See Appendix B.2 for additional details.

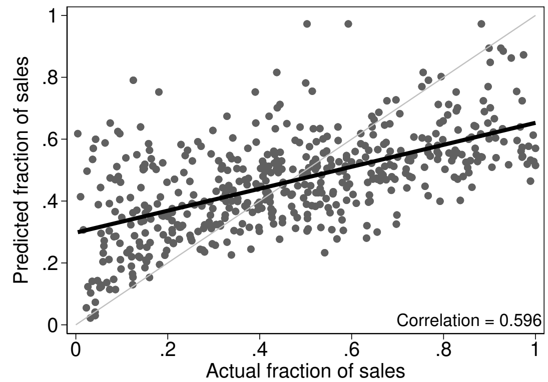
Figure A.18: Correlation Between Estimated and Actual Daily Sales Quantity



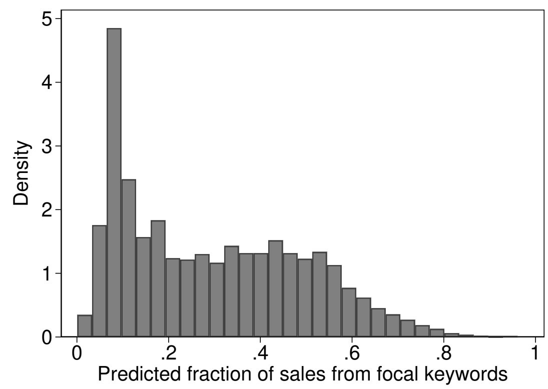
Notes: This figure displays the distribution of correlation coefficients between actual and estimated daily sales quantities for a few hundred products sold on Amazon. I obtained actual daily sales quantities over a two-year period for a few hundred products in collaboration with an Amazon seller. The estimation is based on the Best Seller Ranks using the method detailed in Appendix B.2.

Figure A.19: Calibrated Fraction of Sales from Focal Keywords

A: Actual and Predicted Fractions in the Prediction Sample

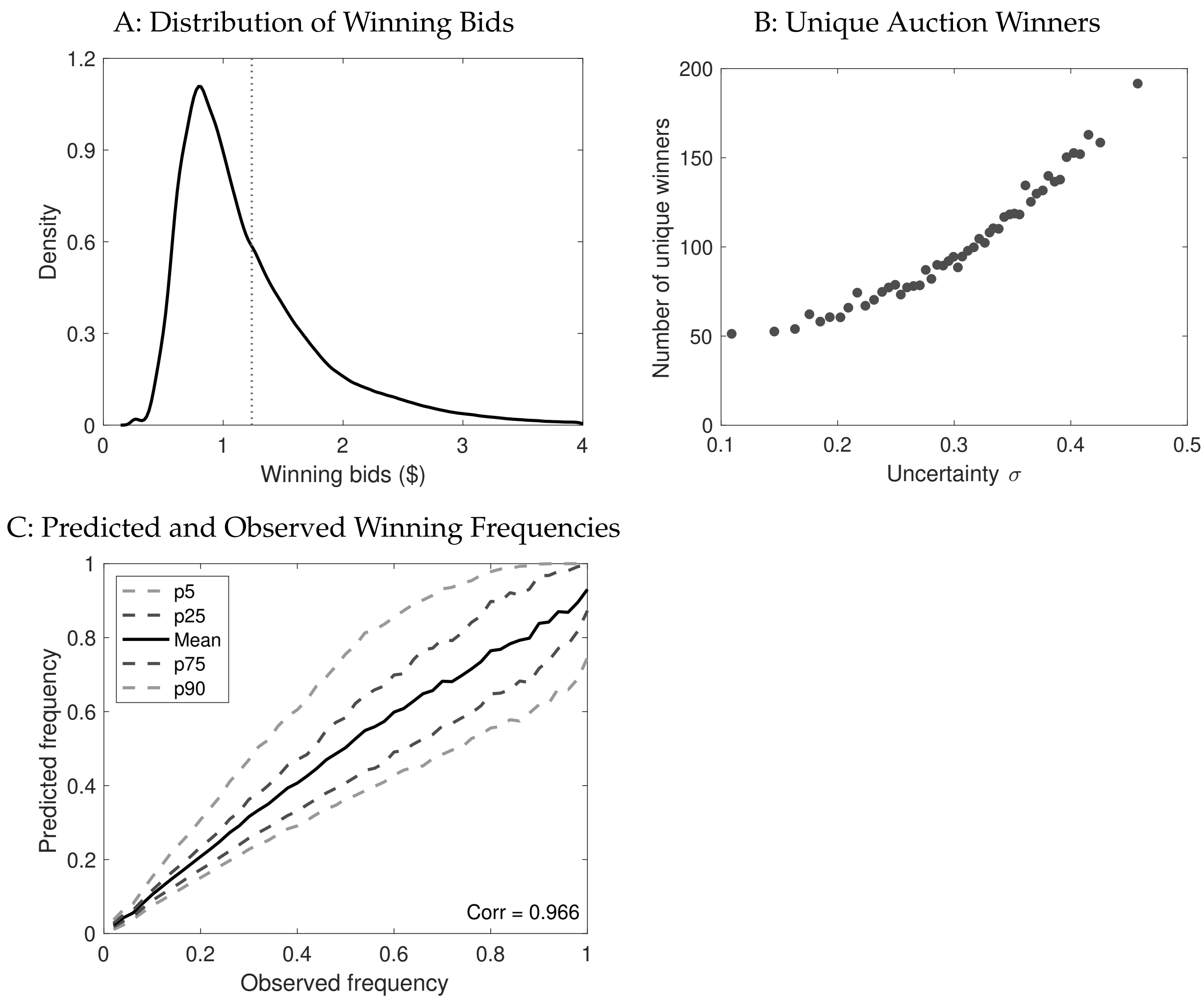


B: Distribution of Predicted Fractions in the Estimation Sample



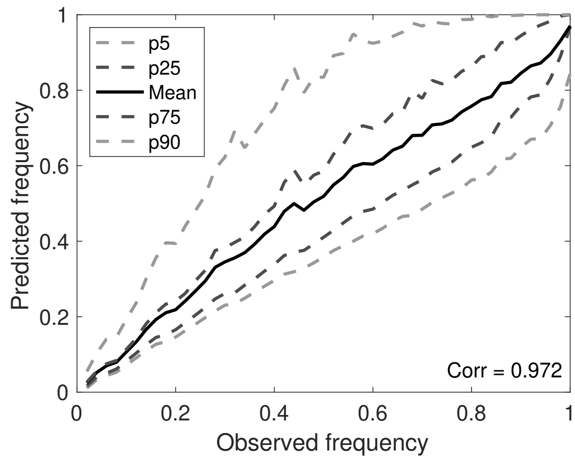
Notes: This figure illustrates the calibrated fraction of a product's total sales quantity attributed to the focal keyword. Panel A compares the actual and predicted fractions in the prediction sample of 480 products, where I can observe a comprehensive list of keywords whose search results contain each product. The actual fraction is constructed using equation (C.11). The predicted fraction is constructed by regressing the actual fraction on four predictors defined in Appendix C.1 and forming a linear prediction. Panel B plots the distribution of the calibrated fractions for all products in the estimation sample, which are bounded within $[0.001, 1]$.

Figure A.20: Estimates of the Auction Model



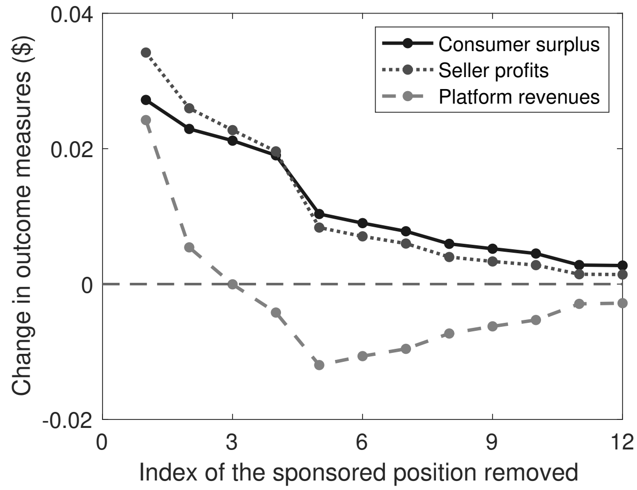
Notes: This figure summarizes the estimates of the auction model. Panel A depicts the distribution of the winning bids, pooling data from all market–week pairs. The dashed line indicates the mean. Panel B illustrates the relationship between the parameter that measures uncertainty, σ_b, and the number of unique auction winners in each market. Panel C compares the predicted frequencies of products winning the auction and observed winning frequencies. For each percentile value in the observed winning frequency, the figure displays the corresponding average predicted winning frequency, along with the 5th, 25th, 75th, and 90th percentiles.

Figure A.21: Estimates of Organic Ranks



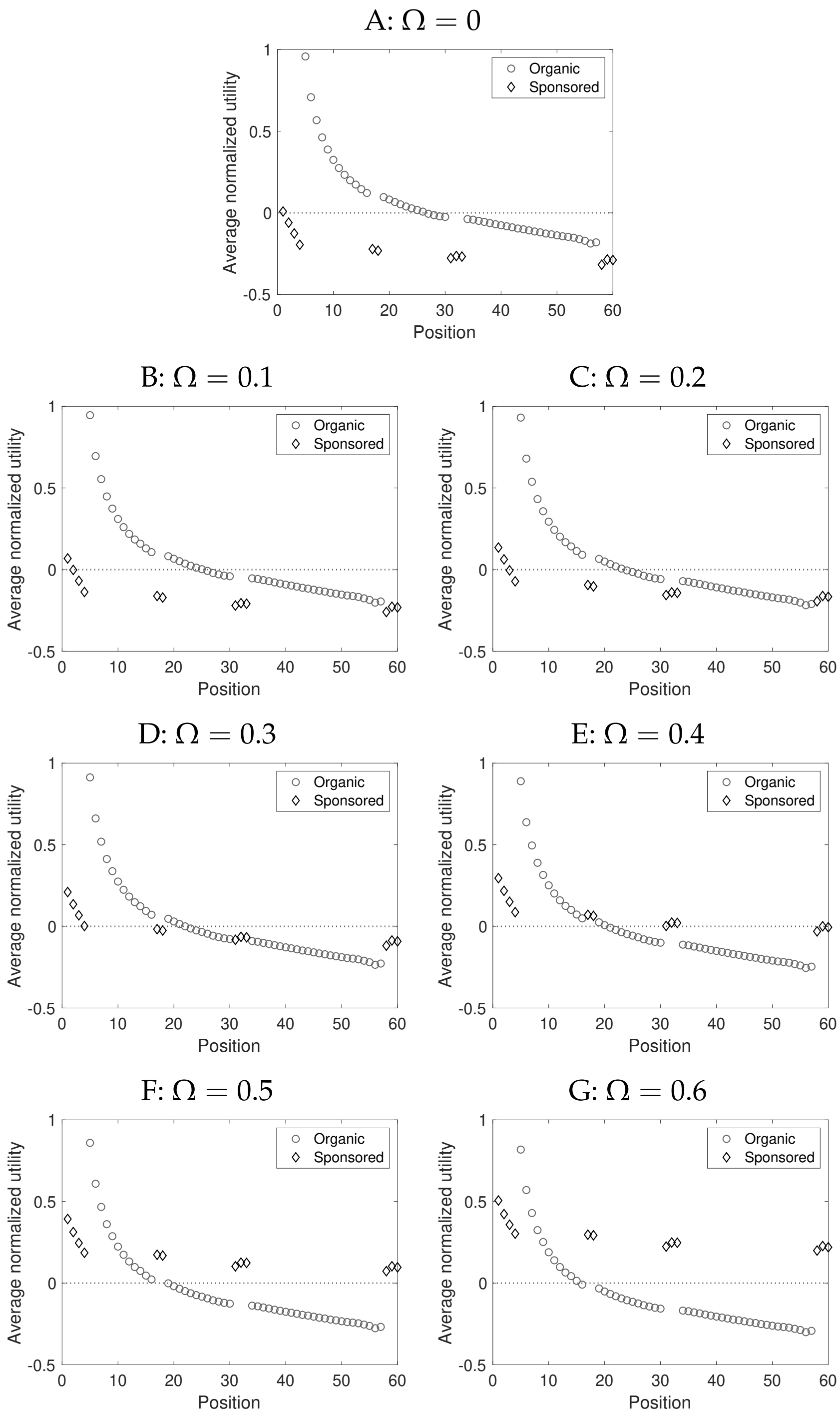
Notes: This figure summarizes the estimates of the distribution of organic ranks. It compares the predicted frequencies of products appearing in organic positions with the observed frequencies. For each percentile value in the observed frequency, the figure displays the corresponding average predicted frequency, along with the 5th, 25th, 75th, and 90th percentiles.

Figure A.22: Impact of Replacing One Sponsored Position with An Organic Position



Notes: This figure depicts the effects of replacing one sponsored position with an organic one. The *x*-axis represents the index of the sponsored position being replaced. The *y*-axis shows the change in consumer surplus, seller profits, and platform revenues.

Figure A.23: Product Utility in Each Position When Consumers Skip Sponsored Products

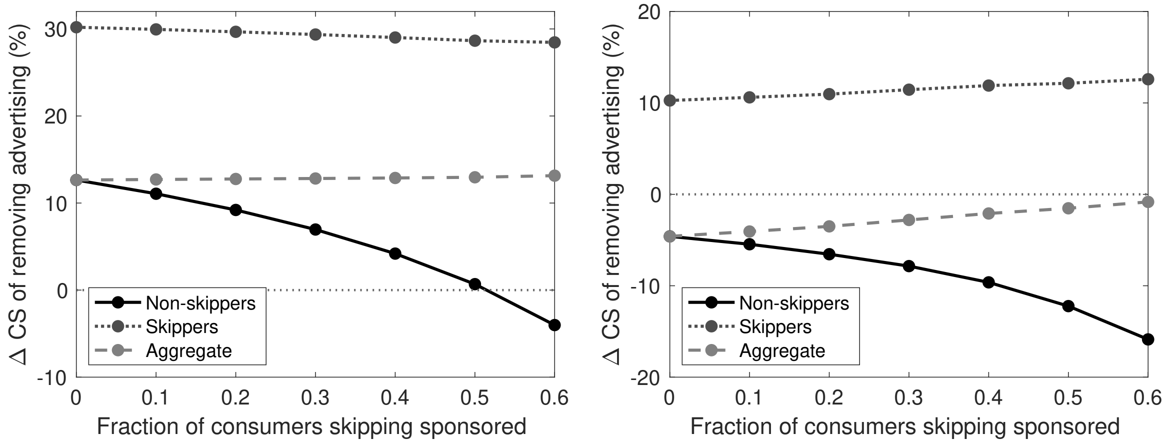


Notes: This figure presents the average normalized utility of products appearing in each position under different assumptions regarding consumer behaviors. Each panel corresponds to a different fraction of consumers who consistently skip sponsored products, denoted by Ω. When a fraction λ_n of consumers consider a product in the n-th *organic* position, only a fraction $(1-\Omega)\lambda_n$ of consumers consider a product in the n-th *sponsored* position. The first panel, with $\Omega = 0$, corresponds to the baseline results and replicates Panel B of Figure 3. See Appendix F.1.1 for additional details of the model and Figure 3 for the construction of the utility estimates.

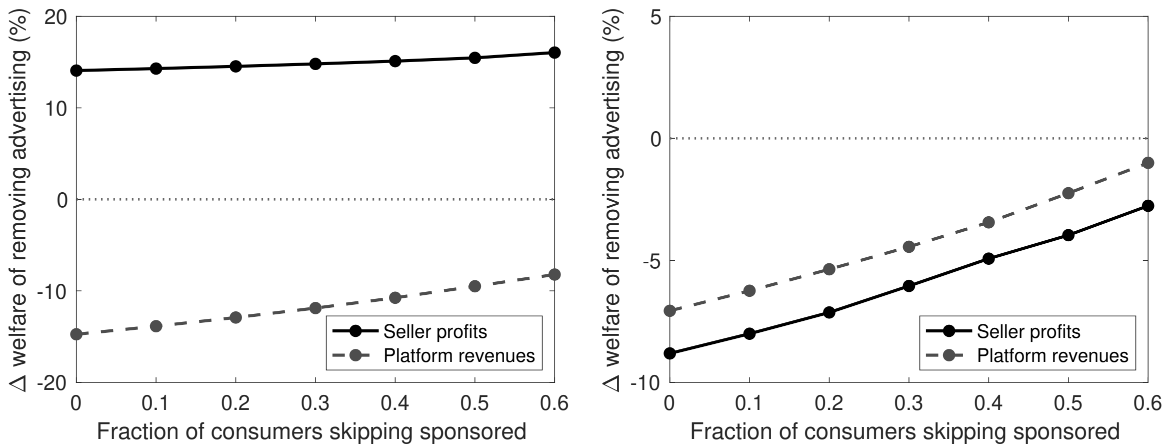
Figure A.24: Welfare Effects When Consumers Skip Sponsored Products

Fixed Commission Rate **Endogenous Commission Rate**

A: Consumer Surplus

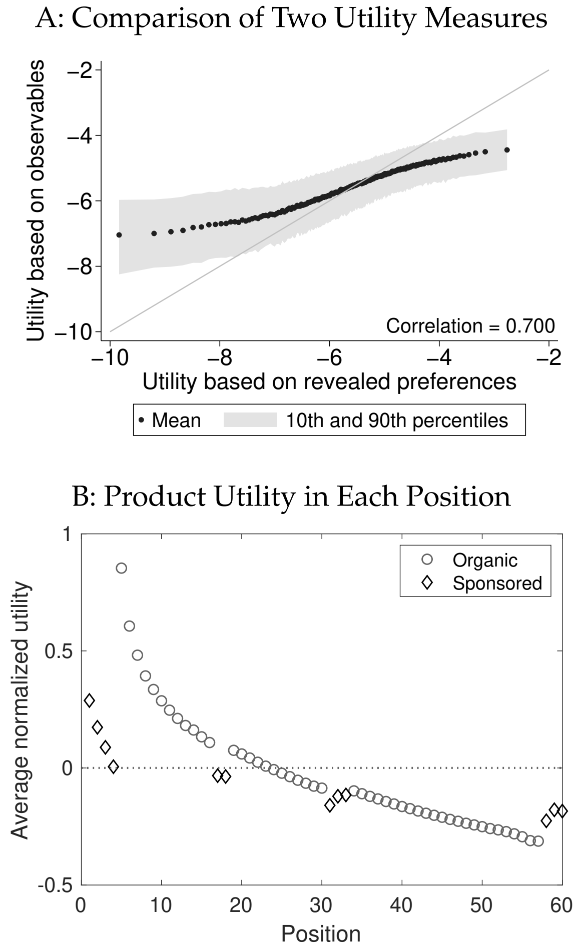


B: Seller Profits and Platform Revenues



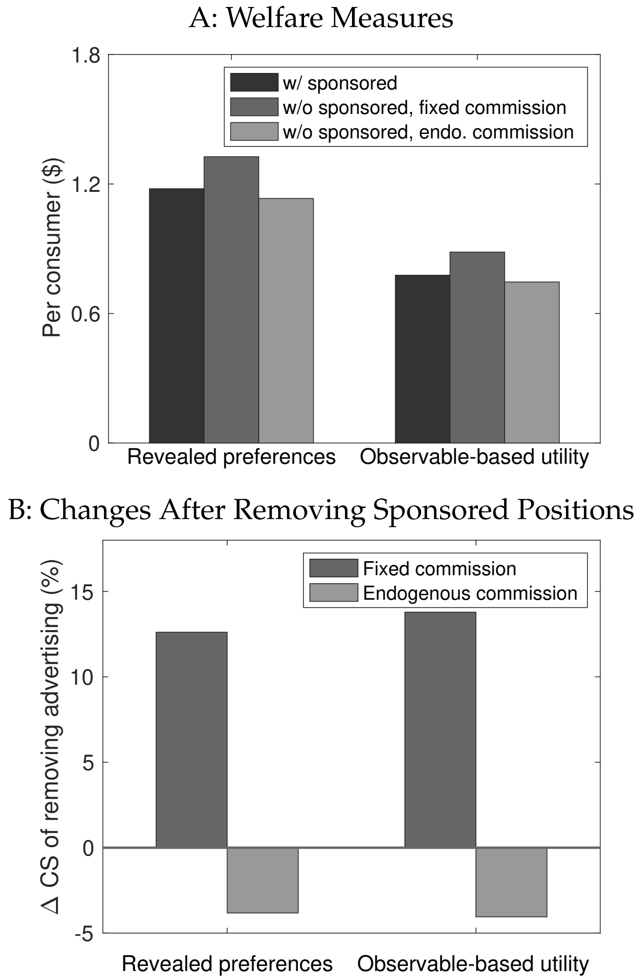
Notes: This figure depicts the welfare effects of eliminating sponsored positions under different assumptions regarding consumer behaviors. In each panel, the x-axis represents the fraction of consumers who consistently skip sponsored positions, ranging from 0 to 0.6 in increments of 0.1. Panel A illustrates the changes in consumer surplus for skippers (those who always skip sponsored products), non-skippers (those who consider all products), and all consumers. Panel B shows the changes in seller profits and platform revenues. The left column considers scenarios with a fixed commission rate, and the right column considers scenarios with endogenous commission rates. The very left point in each panel corresponds to the baseline results reported in Table 3. See Appendix F.1.1 for additional details of the model.

Figure A.25: Utility Measure Based on Observables



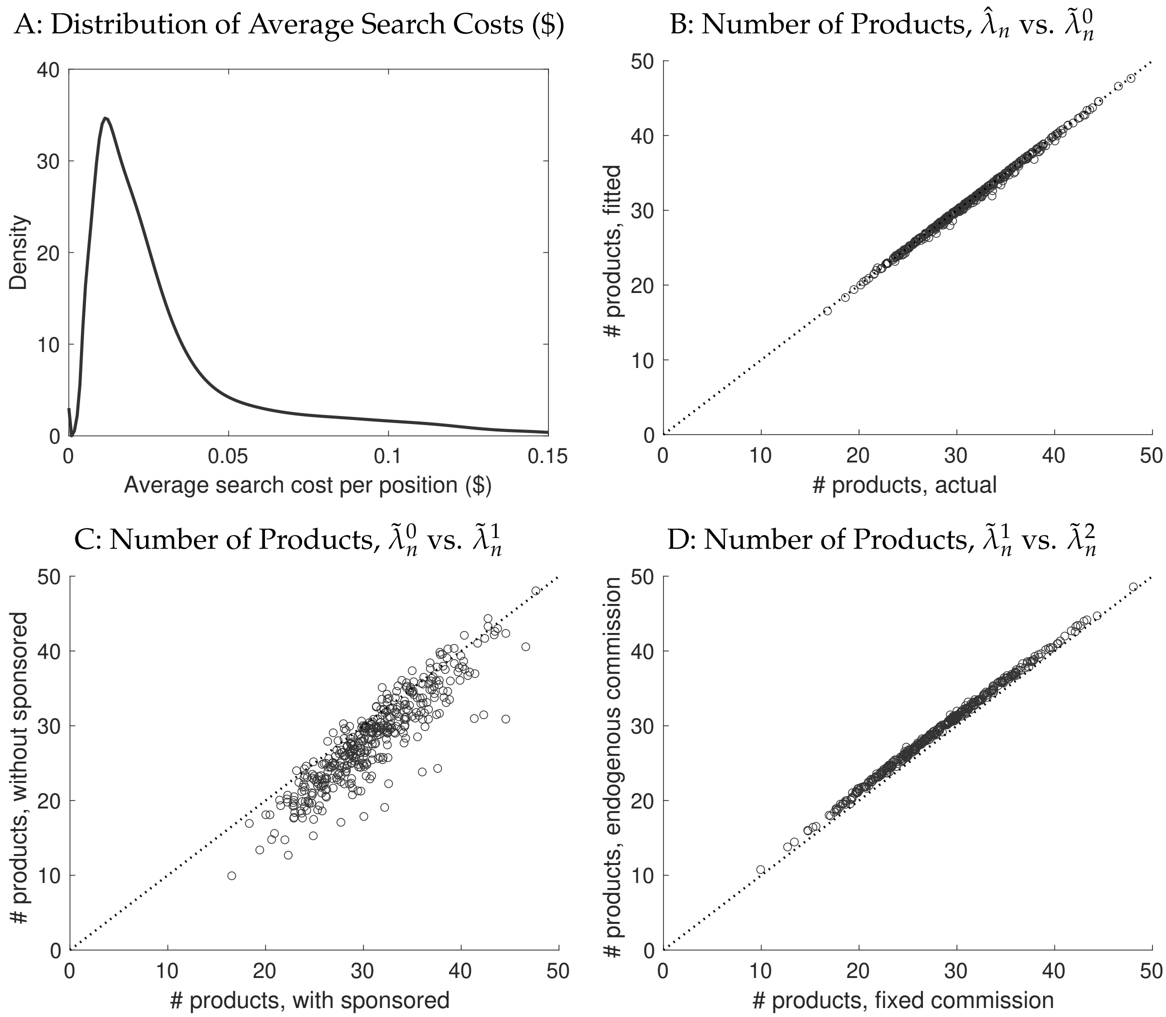
Notes: This figure presents a summary of a utility measure based on observable product characteristics. To construct this measure, I first select products that never appear in a sponsored position within a week. Then I regress estimated product utility, derived from consumers' revealed preferences obtained in Section 6.1, on market fixed effects and four observables. These observables include the log of the number of reviews, average consumer rating, an indicator for Prime eligibility, and an indicator for Amazon's Choice. The measure represents the linear prediction from the regression. Panel A compares product utility based on observables (y-axis) with product utility based on consumers' revealed preferences (x-axis). For each value on the x-axis, the dot represents the mean of the utility measure on the y-axis, and the shaded area depicts the 10th and 90th percentiles of the utility distribution. Panel B displays the normalized average utility for products appearing in each position using the observable-based utility measure. See Appendix F.1.2 for further details on the construction of this utility measure.

Figure A.26: Welfare Effects Using Observable-Based Utility Measure



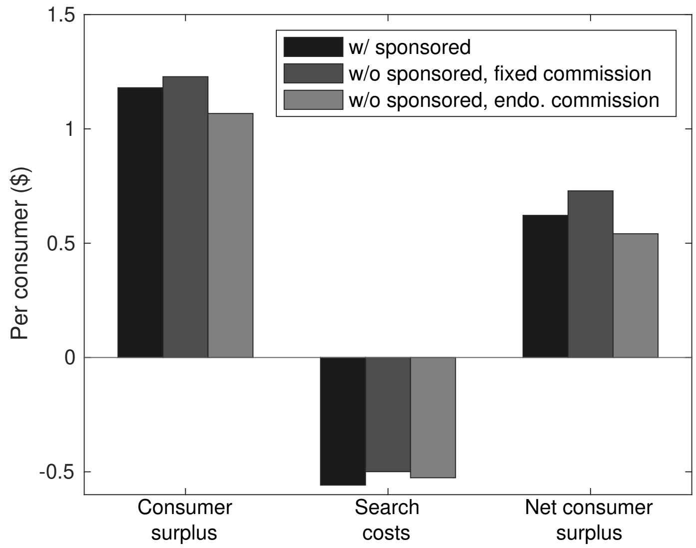
Notes: This figure illustrates the welfare effects of removing sponsored positions using the observable-based utility measure. Panel A presents consumer surplus under two utility measures: one based on consumers' revealed preferences (left block) and one based on observables (right block). Each block presents three scenarios: the status quo with sponsored positions, a scenario removing sponsored positions under a fixed commission rate, and a scenario removing sponsored positions under endogenous commission rates. Panel B displays the change in consumer surplus under the two utility measures, considering both fixed and endogenous commission rates. In both panels, the left block corresponds to the baseline results reported in Table 3. See Appendix F.1.2 for further details on the construction of this utility measure and the calculation of consumer surplus.

Figure A.27: Distribution of Search Costs and Number of Products Considered



Notes: This figure illustrates the distribution of the average search cost and the average number of products in consumers' consideration sets. Appendix F.2 outlines a sequential search model where consumers' per-step search cost follows a log-normal distribution. The model can predict the fraction of consumers considering each position in search results. Panel A displays the distribution of the average per-step search cost across markets, converted into dollars. I use the estimated distribution of consumers' per-step search cost to predict the proportion of consumers considering each position in four scenarios and calculate the average number of products in consumers' consideration sets. These four scenarios are: (i) the actual estimates in the status quo; (ii) the search process in the status quo implied by the estimated distribution of search costs; (iii) the scenario with sponsored positions removed under a fixed commission rate; and (iv) the scenario with sponsored positions removed under endogenous commission rates. Panels B to D compare the average number of products in consumers' consideration sets for (i) vs. (ii) (as an evaluation of model fit), (ii) vs. (iii), and (iii) vs. (iv), respectively. See Appendix F.2 for further details of the model.

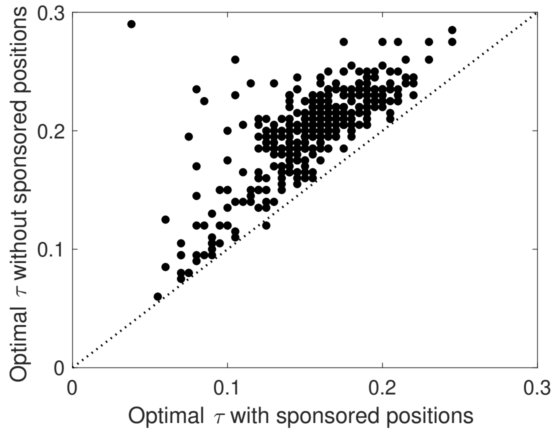
Figure A.28: Welfare Effects on Consumer Surplus with Endogenous Consideration Sets



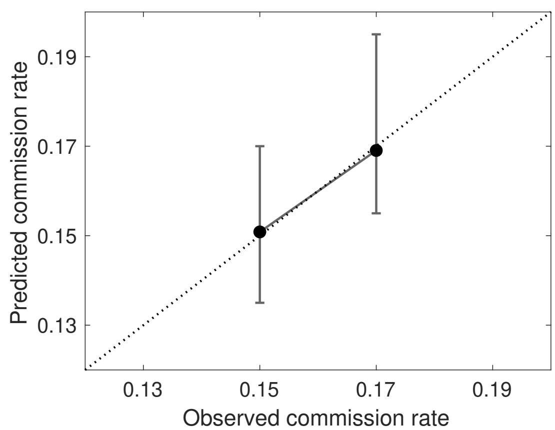
Notes: This figure illustrates the overall effect on consumer surplus with consideration sets endogenized and search costs incorporated. The first block illustrates consumer surplus without accounting for search costs, the second block quantifies total search costs incurred, and the third block presents net consumer surplus, calculated as the difference between consumer surplus and search costs. Each block includes three scenarios: the status quo with sponsored positions, a scenario with sponsored positions removed under a fixed commission rate, and a scenario with sponsored positions removed under endogenous commission rates. See Appendix F.2 for further details of the model.

Figure A.29: Market-Specific Commission Rates

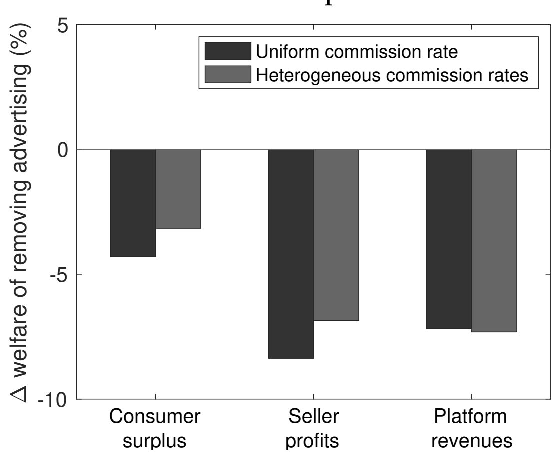
A: Optimal Commission Rates With and Without Sponsored Positions



B: Comparison of Predicted and Observed Commission Rates



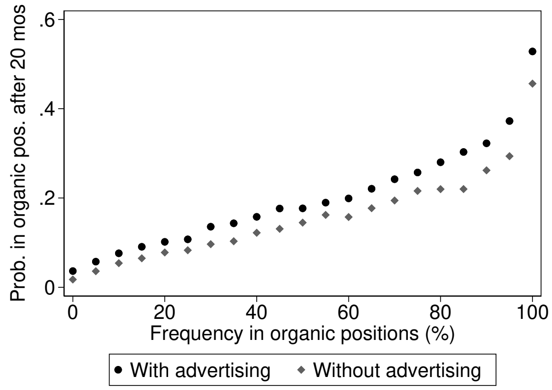
C: Welfare Effects with Market-Specific Commission Rates



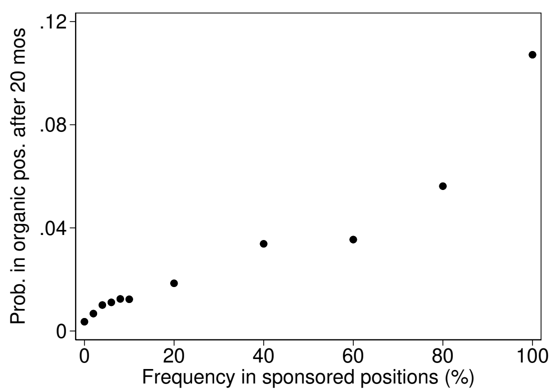
Notes: This figure examines the implication of Amazon implementing market-specific commission rates. Panel A compares the optimal commission rates with and without advertising in each market that maximize the platform's objective. Panel B compares the averaged predicted commission rate for markets with an observed commission rate of 15% and 17%. The dot represents the mean, and the error bar represents the 25th and 75th percentiles. Panel C displays the welfare effects of removing sponsored product advertising under uniform or market-specific commission rates.

Figure A.30: Longer-term Effects

A: Survival Rate of Organic Products

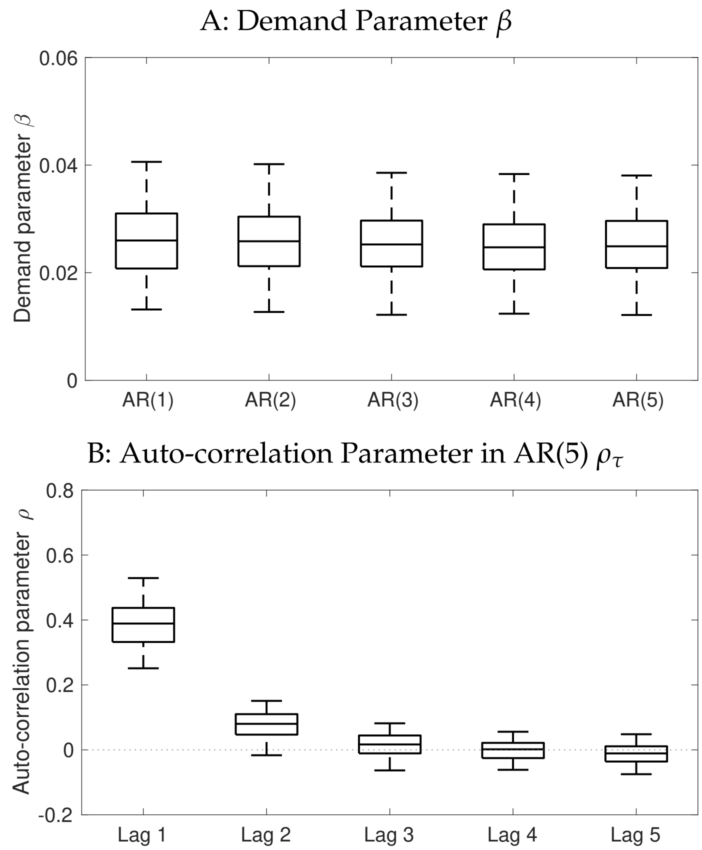


B: Effects of Advertising on Survival Rate of Sponsored Products



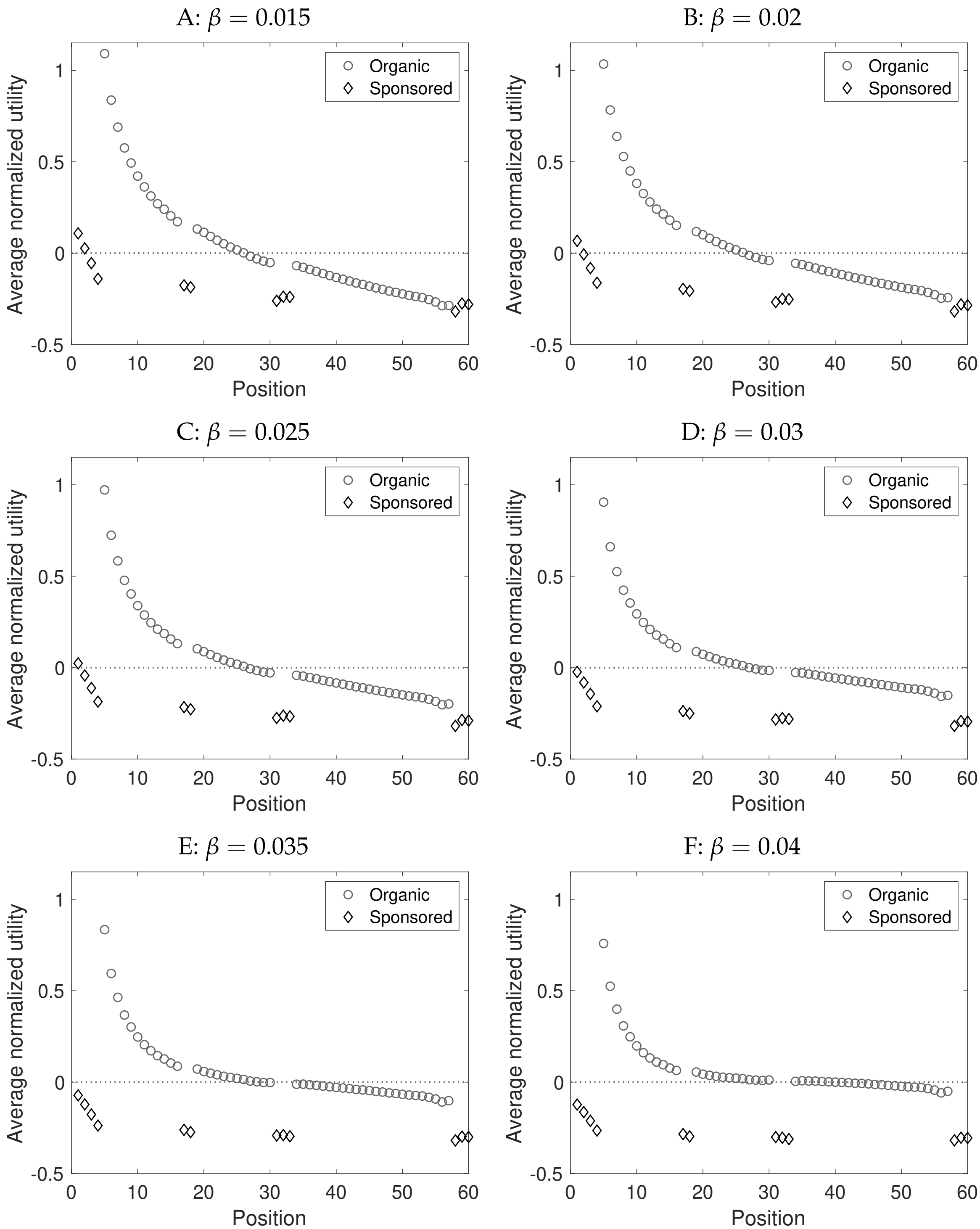
Notes: This figure presents suggestive evidence regarding the longer-term impacts of sponsored product advertising. I collected supplementary data on March 25, 2024, approximately 20 months after the initial sample period. During this data collection, I gathered search results for the same set of keywords five times. For each product in the main sample, I identify whether it appears in the organic search results of the same keyword in the new sample. Panel A focuses on products present in organic positions in the main sample, separately for products also present in sponsored positions and products not. Panel B focuses on products solely present in sponsored positions. The x-axis represents the frequency of appearances in organic positions (Panel A) or sponsored positions (Panel B). The y-axis represents the fraction of products that appear in the organic search results of the same keyword in the new sample.

Figure A.31: Demand Estimates Under Alternative Autoregressive Models



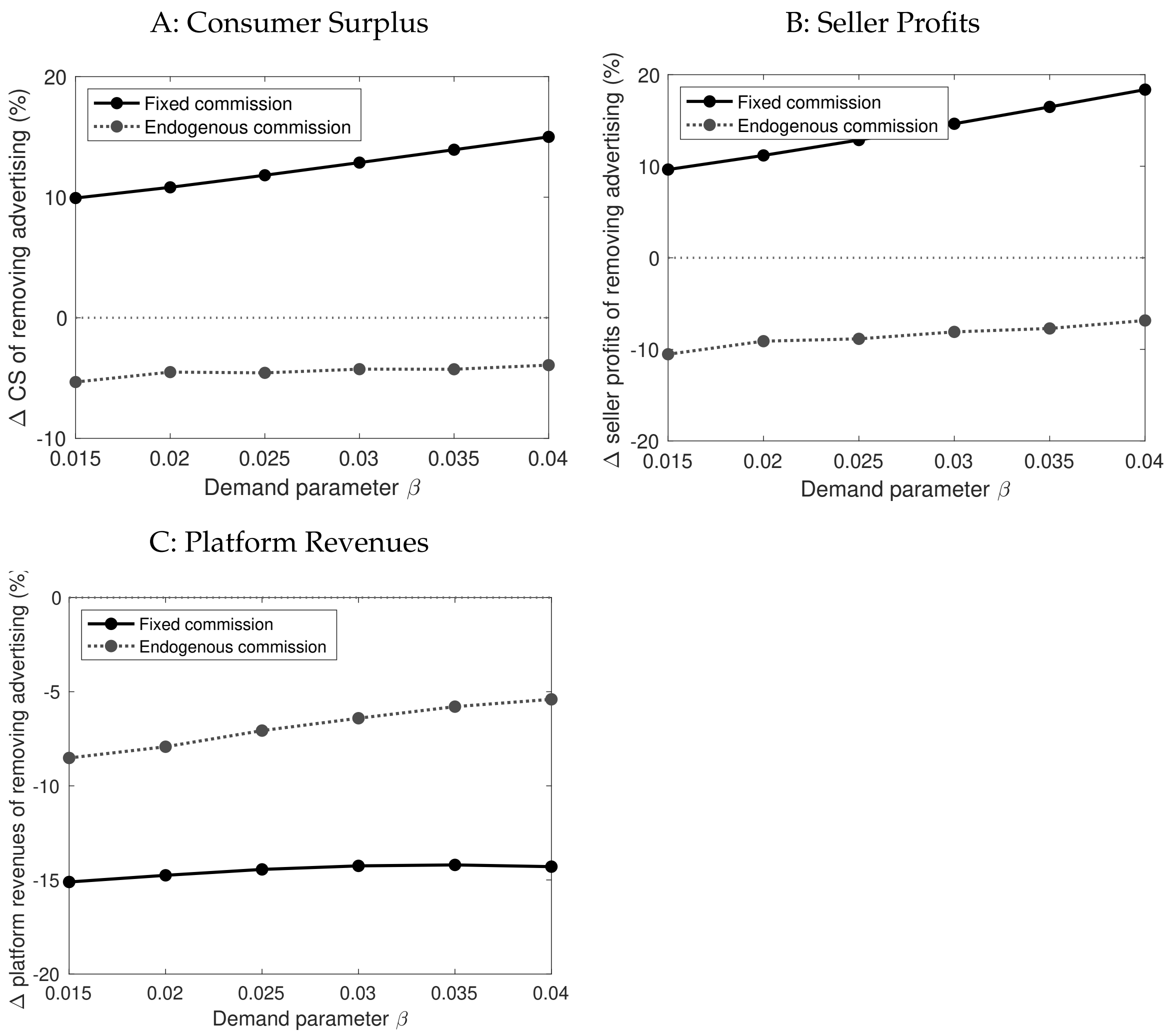
Notes: This figure presents demand estimates under a generalized autoregressive model where the unobserved demand shock ξ_{jt} follows an AR(p) process given by $\xi_{jt} = \sum_{\tau=1}^{p} \rho_\tau \xi_{j,t-\tau} + \eta_{jt}$, with $\eta_{jt} \perp \xi_{j,t-\tau}, \forall \tau \geq 1$. The baseline analysis in Section 5.1.2 sets $p = 1$. Panel A displays the estimates of the demand parameter β under different assumptions for p. Panel B shows the estimates of the autoregressive parameters ρ_τ for $\tau = 1, ..., 5$ when $p = 5$.

Figure A.32: Product Utility in Each Position Under Different Consumer Search Frictions



Notes: This figure presents the average normalized utility of products appearing in each position under different assumptions regarding the extent of consumers' search frictions. The fraction of consumers whose consideration sets contain the n-th product in search results is parametrized as $\lambda_n = 1/\exp(\beta(n-1))$. Each panel corresponds to a different value of the parameter β.

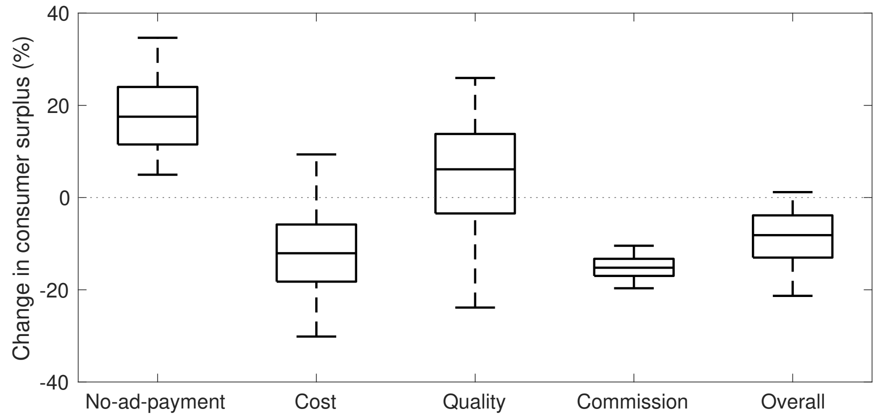
Figure A.33: Welfare Effects Under Different Consumer Search Frictions



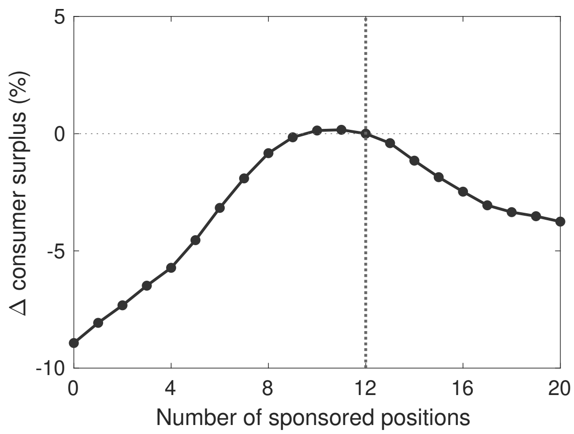
Notes: This figure depicts the welfare effects of eliminating sponsored positions under different assumptions regarding the extent of consumers' search frictions. The fraction of consumers whose consideration sets contain the n-th product in search results is parametrized as $\lambda_n = 1/\exp(\beta(n-1))$. In each panel, the x-axis represents the parameter β, ranging from 0.015 to 0.04 in increments of 0.005. Panels A to C illustrate the change in consumer surplus, seller profits, and platform revenues. In each panel, I consider both fixed and endogenous commission rates.

Figure A.34: Consumer Surplus Under Alternative Assumption on Duplicated Listings

A: Decomposition of the Change in Consumer Surplus



B: Change in Consumer Surplus Under Varying Numbers of Sponsored Positions



Notes: This figure presents the welfare results under an alternative assumption regarding duplicated listings. The baseline analysis assumes that the second appearance of the same product in search results does not receive an independent idiosyncratic preference shock when calculating consumer surplus. Appendix G.3 considers an alternative assumption where the second appearance also receives its independent preference shock. Panel A decomposes the aggregate change in consumer surplus into four components, corresponding to Figure 7. Panel B presents the change in consumer surplus when varying the number of sponsored positions in the search results relative to the status quo with 12 sponsored positions, corresponding to Panel B of Figure 9.

Table A.1: Examples of Markets and Keywords

Market	Keyword	30-Day Search Volume
baby walker	baby walker	137,093
	baby walkers for boys	25,303
	walker for baby boy	24,834
	baby walkers for girls	17,125
	walker for baby girl	15,509
pedicure kit	pedicure kit	118,896
	manicure kit	26,475
	pedicure tools	23,148
	pedicure	12,891
	pedicure supplies	12,396
paper towels	paper towels	513,567
	bounty paper towels	52,401
	paper towel	49,573
	paper towels bulk	32,874
	viva paper towels	14,100
coffee maker	coffee maker	520,363
	coffee pot	53,636
	coffee machine	45,263
	coffee makers	30,407
	drip coffee maker	12,390
sticky notes	sticky notes	113,604
	post it notes	106,387
	notes	15,766
	post it	11,244
	post its	10,918
dog bed	dog bed	488,901
	dog beds	75,156
	large dog bed	59,034
wedding dress	wedding dress	253,725
	wedding dresses for bride	103,461
	wedding dresses	40,288
outdoor toys	outdoor toys	104,029
	kids outdoor toys	45,879
	toddler outdoor toys	43,277
	outside toys	27,196
	outdoor toys for kids	10,840
stapler	stapler	113,008
	heavy duty stapler	10,246

Notes: This table provides several examples of markets and keywords in the sample. Keywords whose search results contain a substantial overlap of products are grouped into the same market using the method described in Appendix B.1. The sample contains 3,237 keywords in 546 markets. Data on search volumes are collected from Jungle Scout.

Table A.2: Prediction of the Fraction of Sales From Focal Keyword

	Dep. Var. = Frac. of Sales From Focal Keyword				
	(1)	(2)	(3)	(4)	(5)
	$\beta = 0.01$	$\beta = 0.02$	$\beta = 0.03$	$\beta = 0.04$	$\beta = 0.05$
Log(Rank)	-0.063	-0.091	-0.113	-0.132	-0.147
	(0.010)	(0.011)	(0.011)	(0.012)	(0.012)
Log(# Other Keywords + 1)	-0.018	-0.025	-0.031	-0.036	-0.039
	(0.024)	(0.025)	(0.026)	(0.027)	(0.028)
Fraction of Search Volume	0.608	0.566	0.525	0.486	0.449
	(0.058)	(0.060)	(0.062)	(0.064)	(0.066)
Log(Avg. Rank in Other Keywords)	0.072	0.080	0.085	0.088	0.089
	(0.013)	(0.014)	(0.014)	(0.015)	(0.015)
Appear in Other Keywords	-0.224	-0.260	-0.284	-0.299	-0.308
	(0.053)	(0.055)	(0.057)	(0.059)	(0.060)
Constant	0.256	0.361	0.447	0.519	0.580
	(0.066)	(0.069)	(0.072)	(0.074)	(0.076)
N	496	496	496	496	496
R^2	0.382	0.364	0.355	0.352	0.353

Notes: This table presents the regression results of predicting the fraction of a product's total sales attributed to focal keywords, as outlined in Appendix C.1. Each observation corresponds to a product. The dependent variable represents the fraction of the product's total sales from the focal keyword, as constructed in equation (C.11). I include five predictors in the regressions, as defined in Appendix C.1. Each column in the table corresponds to different values of the demand parameter β, where the fraction of consumers considering each position in search results is parametrized as $\lambda_n = 1/\exp(\beta(n-1))$ for $1 \leq n \leq N$. Standard errors are reported in parentheses.

Table A.3: Welfare Effects Under Endogenous Organic Rankings

	(1)	(2)	(3)	(4)	(5)
	Status Quo	Counterfactual: Without Sponsored Product Advertising			
		Fixed Commission		Endogenous Commission	
	Level ($)	Level ($)	Change (%)	Level ($)	Change (%)
Consumer Surplus	1.18	1.33	13.0%	1.13	-4.0%
Seller Profits	0.81	0.93	14.3%	0.74	-8.7%
Platform Revenues	0.89	0.76	-14.9%	0.82	-7.4%
Commission Revenues	0.71	0.76	7.3%	0.82	16.7%
Advertising Revenues	0.18	0.00	-100.0%	0.00	-100.0%
Total Surplus	2.88	3.02	4.8%	2.70	-6.4%

Notes: This table presents the welfare effects under endogenous organic rankings. Instead of assuming that the distribution of organic rankings remains fixed in the counterfactual scenarios, I estimate the ranking algorithm and allow for changes in the distribution of organic rankings in counterfactual scenarios. See Appendix F.3 for more details. This table replicates Table 3 under this alternative assumption.

Table A.4: Welfare Effects Under Alternative Assumption on Duplicated Listings

	(1)	(2)	(3)	(4)	(5)
	Status Quo	Counterfactual: Without Sponsored Product Advertising			
		Fixed Commission		Endogenous Commission	
	Level ($)	Level ($)	Change (%)	Level ($)	Change (%)
Consumer Surplus	1.24	1.33	7.1%	1.13	-8.9%
Seller Profits	0.81	0.93	14.1%	0.75	-8.4%
Platform Revenues	0.89	0.76	-14.8%	0.83	-7.2%
Commission Revenues	0.71	0.76	7.4%	0.83	16.9%
Advertising Revenues	0.18	0.00	-100.0%	0.00	-100.0%
Total Surplus	2.94	3.01	2.4%	2.70	-8.3%

Notes: This table presents the welfare effects under an alternative assumption regarding duplicated listings. The baseline analysis assumes that the second appearance of the same product in search results does not receive an independent idiosyncratic preference shock when calculating consumer surplus. Appendix G.3 considers an alternative assumption where the second appearance also receives its independent preference shock. This table replicates Table 3 under this alternative assumption.