NBER WORKING PAPER SERIES

VERTICAL INTEGRATION AND CONSUMER CHOICE: EVIDENCE FROM A FIELD EXPERIMENT

Chiara Farronato
Andrey Fradkin
Alexander MacKay

Working Paper 34135
http://www.nber.org/papers/w34135

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2025

We thank James Dana, Connor Haskin, Debbie Liang, Hayden Schrauff, Yejia Xu, and Hannah Zhang for outstanding research assistance. We thank Chris Karr at Audacious Software for developing the software. We thank Dean Eckles, Avi Goldfarb, Ginger Jin, Tai Lam, Ilya Morozov, Devesh Raval, Anna Tuchman, Catherine Tucker, Senthil Veeraraghavan, Giorgos Zervas, and numerous seminar participants for excellent suggestions. We are grateful to Boston University and Harvard Business School for funding this research. Starting August 11, 2025, Fradkin will be an employee at Amazon.com. The research and analysis in this paper were conceived and completed prior to this appointment, and a draft was first submitted as an NBER Working Paper on August 5, 2025. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Vertical Integration and Consumer Choice: Evidence from a Field Experiment
Chiara Farronato, Andrey Fradkin, and Alexander MacKay
NBER Working Paper No. 34135
August 2025
JEL No. C93, D12, K21, L13, L40, L81

ABSTRACT

Platforms, retailers, and other firms often offer their own products alongside products sold by competitors. We study the effects of this practice by combining a field experiment that hides brands owned by Amazon (i.e., private labels) from shoppers on Amazon.com with model-based counterfactuals and welfare analysis. In the absence of private labels, consumers substitute toward products that are similar along most observable dimensions. Removing Amazon brands does not change consumers' search effort or their propensity to shop at other retail websites. Despite the ample availability of observably similar alternatives, our welfare estimates imply that, for the categories we study, removing Amazon brands would reduce consumer surplus by 5.5 percent in the short run, with approximately 10 percent of the impact due to equilibrium price increases by other sellers. The effects are heterogeneous, with consumer surplus reductions exceeding 10 percent in some categories, while other categories realize much smaller decreases when Amazon brands are removed. Demoting private labels in search results to counteract potential self-preferencing does not lead to gains in consumer surplus. This outcome arises because a subset of consumers derive greater utility from private labels and benefit from their high placement in search results.

Chiara Farronato
Harvard University
Harvard Business School
Technology and Operations Management Unit
and NBER
cfarronato@hbs.edu

Andrey Fradkin
Boston University
fradkin@bu.edu

Alexander MacKay
University of Virginia
Department of Economics
mackay@virginia.edu

An online appendix is available at http://www.nber.org/data-appendix/w34135
A randomized controlled trials registry entry is available at https://doi.org/10.1257/rct.11370-1.0

1 Introduction

Digital platforms such as Amazon and Google are so large and influential that many of their actions are under regulatory scrutiny. In Europe, regulators have recently passed the Digital Markets Act[1] and the Digital Services Act[2] to constrain and monitor the behavior of large platforms. In the US, the Department of Justice (DOJ) is involved in legal proceedings against Google[3] and the Federal Trade Commission (FTC) against Amazon,[4] both of which accuse the platforms of abuse of their respective dominant positions.

One issue of regulatory concern is the practice by major technology platforms of vertically integrating and featuring their own products alongside those of third parties. This practice is common among the large firms designated as “gatekeepers” by the Digital Markets Act. For example, Google owns Google Maps and Google Shopping, which directly compete with third-party maps and e-commerce alternatives. Similarly, Amazon sells private-label products—under brands such as Amazon Basics, Solimo, and Mama Bear—on its retail website alongside competing products offered by independent manufacturers and brands. This practice has raised concerns of reduced competition and harm to consumers, especially if the gatekeeper treats its own products more favorably, i.e., engages in self-preferencing.

We study the effects of Amazon-owned brands, or private labels, on consumer decisions and welfare. To answer this question, we develop a browser extension that can manipulate and track browsing behavior and recruit participants to install it. We use the extension to introduce random variation in the set of products available to consumers. For our main treatment, we prevent consumers from seeing Amazon brands when they browse Amazon’s website. This allows us to measure the (short-run) effects on search behavior and product choices in the absence of Amazon’s private labels.

We focus on four demand-side channels by which the presence of private label products can affect consumer welfare. First, their presence increases the number of options that consumers can choose from—the variety effect. Second, consumers may adjust their effort when they search for products on Amazon. Third, private labels may influence whether consumers shop on Amazon instead of Walmart.com, for example. Fourth, competitive pressure can affect equilibrium prices of other products.

We use our experimental variation to provide reduced-form evidence on the extent of the first three mechanisms. We find that when Amazon brands are not available, consumers substitute to broadly similar products, except that those substitutes have fewer reviews. We find no

[1] digital-markets-act.ec.europa.eu.

[2] https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act_en.

[3] https://www.justice.gov/opa/pr/justice-department-sues-google-monopolizing-digital-advertising-technologies

[4] https://www.ftc.gov/news-events/news/press-releases/2023/09/ftc-sues-amazon-illegally-maintaining-monopoly-power.

evidence of changes to search behavior on Amazon, nor evidence that consumers spend more time on other online retail sites. To assess the fourth mechanism, we estimate a structural model of demand and use the model to simulate counterfactual equilibrium prices in the absence of Amazon brands. The model allows us to quantify changes in consumer welfare. The presence of Amazon brands increases welfare, but mainly through the effects of the brands on consumer choices rather than inducing competitors to lower prices.

For this research, we developed a browser extension called Webmunk, recruited US residents to install it, and compensated them for their participation.[5] The extension randomly allocates study participants into three groups: a control group, whose behavior on Amazon.com is tracked for the period of the study; a Hide Amazon group, for whom the extension removes Amazon brands from search results and other pages on Amazon.com; and a Hide Random group, for whom the extension removes a random set of products.

Upon enrollment, we ask participants to complete a set of incentivized shopping tasks and then track their organic browsing behavior on Amazon for the following 8 weeks. In the incentivized tasks, participants are asked to shop for products from pre-defined categories to add to an Amazon wishlist especially created for our study. We pre-selected 23 product categories (e.g., allergy medications, paper towels, socks, batteries) from a set of six meta-categories—health, paper products, household items, apparel, electronics, and personal care. The first five meta-categories are characterized by a high share of Amazon branded options. The sixth, personal care, did not have Amazon brands at the time of our study, and we selected it as a placebo.

The randomization allows us to compare the characteristics of the chosen products across the three treatment groups.[6] In the absence of Amazon brands, consumers choose products with similar price, delivery speed, average star rating, and Prime eligibility, showing no statistically significant differences from those selected by the control group. Though the average star rating is similar, the number of accumulated reviews of the chosen product is much lower when Amazon brands are hidden compared to when Amazon brands are available. The Hide Random group allows us to confirm that these differences are not simply due to a decrease in product variety, but directly linked to the characteristics of Amazon branded products.

Notably, this substitution toward similar products does not appear to require additional effort from consumers. The number of searches performed and the number of products viewed remain comparable across treatment groups, suggesting that consumers are able to find alternative products without engaging in a more extensive search process. Furthermore, traffic to other retail websites outside of Amazon does not increase when Amazon brands are hidden, indicating that consumers remain within the platform and simply shift to comparable alternatives rather than seeking options elsewhere.

[5]Webmunk is open-sourced and available for use by other researchers. Please see Farronato, Fradkin and Karr (2024) or visit `www.webmunk.org` for details.

[6]We pre-registered the reduced form analysis plan under AEARCTR-0011370. `https://doi.org/10.1257/rct.11370-1.0`.

These substitution patterns are consistent with stated consumer preferences in our surveys. Indeed, consumers say they care more about prices, delivery, and quality (as proxied by online ratings) than brand or seller reputation. When asked to rate the chosen products they received from the incentivized shopping task, consumers scored them similarly, regardless of whether they had access to private labels. On average, participants indicated that they valued Amazon private labels $1.75 less than comparable non-Amazon products priced at $25. When the reference product is priced at $10, Amazon brands are valued $0.20 less. The average difference masks substantial heterogeneity across respondents. Stated preferences indicate that the relative preference for Amazon brands compared to other products can be offset by other variables such as average ratings and delivery speed.

To further understand consumer preferences for Amazon private labels and to measure their effects on welfare, we estimate a structural model of demand and supply. We use a discrete choice demand model with heterogeneous preferences across individuals for features like price and Amazon brands. Using individual product selections, we estimate the model via maximum likelihood, while calibrating the mean own-price elasticity to a value of -5 to match reported data on margins for Amazon sellers. Our demand estimates imply that Amazon brands are valued less than other major brands but slightly more than other non-Amazon products. Further, there is meaningful variation in preferences for Amazon products across consumers. Our estimates imply mean marginal costs of roughly $15 and own-price elasticities that range from -8.71 to -2.14 at the 10th and 90th percentiles.

We use the structural model to measure the effects of removing Amazon products from consumers' consideration sets. We remove Amazon brands, shift the remaining products upward in search rankings, and re-compute equilibrium prices. For the product categories in our sample where Amazon brands are present, we estimate that removing Amazon products reduces consumer surplus by 5.5 percent due to reductions in product variety and pricing pressure on competing products. Just over 10 percent of the loss in consumer surplus is due to higher equilibrium prices for the remaining products.

We find sizable heterogeneity in the effect on consumer surplus across categories. In principle, if Amazon excessively prioritizes its own private-label products, their removal can lead to the promotion of alternatives that consumers perceive as superior. As these products move up in the search rankings and are chosen more frequently, the resulting quality improvement can offset—or even outweigh—the effects of higher prices, ultimately increasing consumer surplus. Despite this theoretical possibility, we find that removing Amazon products reduces consumer surplus across all categories. The effect is largest for acid reducers, batteries, monitor cables, and trash bags, whereas it is smallest for laundry detergent, umbrellas, and moisturizers.

Motivated by recent regulatory actions prohibiting self-preferencing, we also examine a counterfactual scenario in which Amazon's private-label products are demoted in search rankings to positions consistent with their observable characteristics or positions consistent with their average utility across consumers. Under both protocols, we find that the demotion of products slightly reduces consumer surplus. Contrary to what regulators might expect, the changes in consumer welfare from these corrections to self-preferencing are not positive. This is because a substantial segment of consumers derive greater utility from Amazon's private labels and benefit from their prominence in search results.

Our study is subject to limitations. Our design ensures that the participants have been active Amazon shoppers. Further, we require that participants use a computer browser when they shop online (rather than a smartphone or a tablet). Because of these margins of selection, as well as the fact that participants elect to join an online study, our results may not completely generalize to the full population of online shoppers. Another limitation is that our study was conducted over a limited time frame: an incentivized shopping task followed by a 8-week observational period. Consumer behavior along various margins, such as search and cross-platform behavior, may take longer than 8 weeks to adjust to the removal of Amazon brands. Thus, the results from our experimental variation are limited to short-run effects. Similarly, our counterfactual simulations only measure price adjustments on the supply side, and depend in part on correctly specifying demand and supply. We abstract away from longer-run effects such as changes to non-price characteristics and the entry of new products.

Related literature. Recent regulatory scrutiny over the market power of digital platforms has given rise to a new literature on vertical integration and biased intermediation (Hagiu, Teh and Wright, 2022; De Corniere and Taylor, 2019; Teng, 2022), specifically on Amazon (Lee and Musolff, 2022; Gutierrez, 2022; Lam, 2022; Crawford et al., 2022; Chen and Tsai, 2021; Raval, 2022; Reimers and Waldfogel, 2023; Waldfogel, 2024). Relative to the above papers, our work has several key advantages. First, our data contain searches and product selections from a variety of real consumers. Second, our field experiment allows us to draw causal links between ranking and consumer choices, and between the availability of Amazon brands and substitution patterns. Third, we link the shopping behavior on Amazon to surveys, order histories, and visits to non-Amazon retailers, shedding light on the generalizability of our results based on incentivized shopping tasks (Morozov and Tuchman, 2024) for more organic search and shopping behavior (Ursu, 2018; Santos, Hortaçsu and Wildenbeest, 2012; Dinerstein et al., 2018).

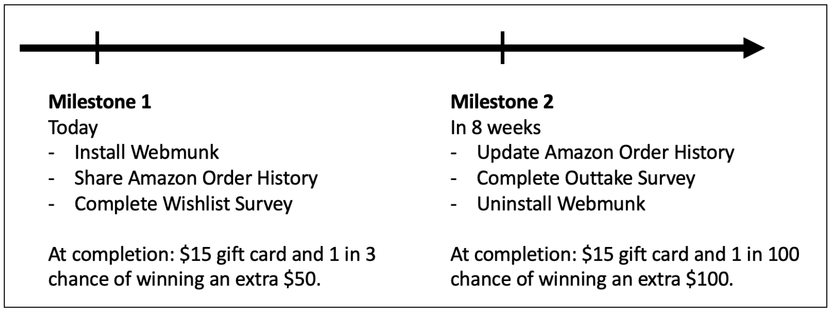
Our approach to collecting data and studying consumer behavior contributes to recent and growing research that uses software to track consumers and run online experiments. Allcott, Gentzkow and Song (2022) study the addiction properties of social media use. Aridor (2022) observes participants' substitution patterns when he experimentally shuts off access to Instagram or Youtube. Levy (2021) differentially exposes study participants to news outlets on social media to study its effects on political polarization. Beknazar-Yuzbashev et al. (2022) study the effects of removing toxic content on social media consumption, highlighting the trade-off between consumption and content toxicity. More recently, Allcott et al. (2024) use a similar approach to identify the reasons behind Google's dominant position in online search.

Our paper relates to a large literature on private labels, mostly focused on offline retail. Private labels are standard practice offline, accounting for almost 20% of products sold (Dubé, 2022). In comparison, we find that Amazon brands are only 2.5% of products sold in the Amazon order histories of our study participants, who are particularly active Amazon shoppers. An older literature has found positive benefits from the introduction of private labels in physical retail, by offering consumers cheaper alternatives of similar quality (Newmark, 1988), without negatively affecting competition, and instead improving it (Adelman, 1949). Research has shown that there are a variety of reasons why retailers may offer private labels (Dhar and Hoch, 1997), from imitating national brands at lower prices (Scott Morton and Zettelmeyer, 2004), especially the most successful brands (ter Braak and Deleersnyder, 2018; Zhu and Liu, 2018), to ensuring quality (Hoch and Banerji, 1993) and offering a variety of premium and value options (Ter Braak, Dekimpe and Geyskens, 2013). Relatedly, Ailawadi, Pauwels and Steenkamp (2008) find that private labels increase store loyalty.

Previous work has also demonstrated how traditional retailers often preferentially treat their private labels (Kumar et al., 2007), by physically placing them prominently (Kotler and Keller, 2016), sometimes side by side with national brands, using similar packaging, discounts, free samples, and comparative messaging (Bronnenberg et al., 2015; Bronnenberg, Dubé and Sanders, 2020; Bronnenberg, Dubé and Joo, 2022). Despite the prevalence and pro-competitive nature of these practices offline, regulators have taken a different approach towards Amazon and its private labels given the dominant position that Amazon and other similarly large platforms have in their respective markets (Dubé, 2022). In our research, we take these concerns over market power seriously and empirically assess whether and to what extent Amazon's private labels distort consumer choice or harm competition. By leveraging randomized experimental variation, we provide direct evidence on the actual impact of private labels on consumer behavior and market outcomes. A key feature of our online setting is that we can directly measure and manipulate product placement, which we find to be meaningful for consumer choice.

The rest of the paper is structured as follows. Section 2 describes our data collection methodology and presents summary statistics about our study population. In Section 3, we present reduced-form evidence on demand effects, including substitution between Amazon brands and non Amazon brands, search effort, and cross-platform effects. The section also discusses our reduced form results in light of perceptions of Amazon brands that consumers report through survey responses. Section 4 presents our structural demand model and counterfactual estimates of impacts to equilibrium prices and welfare. Section 5 concludes.

Figure 1: Study Timeline



2 Data Collection

Our study uses a custom browser extension called Webmunk. Webmunk is an extension similar to an ad blocker and can be installed on the Chrome browser of any computer. The extension has three crucial functionalities. First, it prompts participants to perform specific tasks. Second, it tracks participants' browsing behaviors on pre-determined websites. Third, it allows us to manipulate participants' browsing experience to create different treatment conditions across users and estimate treatment effects of interest. We discuss each of the three functionalities as part of the study design, and then present our sample of study participants.[7]

2.1 Study Design

Recruitment and Study Timeline. We recruited American adults through Facebook advertisements between mid-June and beginning of October 2023.[8] Participants filled out an initial Qualtrics survey, which determined eligibility for the study and collected explicit participant consent to be part of the study. The survey is available in Appendix D. Three eligibility criteria are worth highlighting. First, participants must shop online primarily on a computer, given that Webmunk cannot be installed on a mobile phone or tablet. In our case, 52% of the participants satisfy this condition. Second, participants must use Chrome for their regular browsing, because Webmunk only works on Chrome. This is not a big constraint, since 76% of the respondents who shop on a computer use Chrome. Third, participants needed to be frequent Amazon shoppers (shop at least 2 to 3 times a month on Amazon).

Upon eligibility and consent to participate in the study, participants install Webmunk through the Chrome Web Store and register with their email address. The email address serves two purposes. First, we use it to check that the participant gave their explicit consent to participate in the study, by matching the email address to answers to the initial survey via the Qualtrics API. Only participants who are eligible, consented, and gave matching emails in the

[7]For additional technical details, see Farronato, Fradkin and Karr (2024).
[8]The study was approved under Harvard IRB21-1677.

initial survey and on the browser extension are enrolled into the study. Second, we use the email address to send participants gift cards as compensation for participation.[9]

Figure 1 presents the experiment timeline from the participants' perspective. Upon installing Webmunk, participants are asked to fill out an intake survey, engage in six incentivized shopping tasks (denoted *incentivized shopping tasks* and described in more detail below), and share their past Amazon order history from January 2022 to the day of enrollment.[10] The tasks to complete appear on the extension pop-up window, as shown in Appendix Figure C.1. At completion of these tasks, which we denote as *milestone 1*, participants receive a $15 gift card and a one in three chance of winning an extra $50 (this additional compensation is designed to make the shopping tasks incentive compatible, and is described below). Participants are then asked to keep the extension installed for eight weeks. At the end of the period, we request an update to their Amazon order history, a final survey designed to quantify their satisfaction for products purchased on Amazon, and their preferences towards Amazon brands. The outtake survey is also available in Appendix D. When they complete these final tasks (*milestone 2*), participants receive an additional $15 gift card and a 1/100 chance to win an extra $100.

Incentivized shopping tasks. Upon enrollment, participants are asked to fill out an initial survey, available in Appendix D. In this survey, we ask for basic demographic characteristics and shopping behavior. In addition, we ask participants to engage in a set of incentivized shopping tasks. These shopping tasks are designed to allow us to easily compare choices and behavior across all participants holding constant the product categories. This gives us much more statistical power than just looking at organic browsing behavior, which exhibits substantial heterogeneity, as we confirmed in pilot studies.

The shopping tasks worked as follows. First, we asked participants to select preferred categories from pre-defined lists. They were then instructed to search on Amazon for products within those categories to add to an Amazon wishlist specifically created for our study. Informed by pilot studies, we selected categories (Figure C.4) within health, paper products, household items, apparel, and electronics that contain a sizable share of Amazon brands. The categories within personal care were instead included as a placebo, because we observed no Amazon-branded products in those categories. One concern with shopping tasks is that they are artificial and participants may have no need for the items they are purchasing. For this reason, we picked categories of products that are likely needed in a household. The most popular categories ended up being moisturizer in health; toilet paper in paper products; laundry detergent in household items; t-shirt in apparel; charger in electronics; and deodorant in personal care.

[9]Webmunk employs industry-standard encryption protocols so that personally identifiable information about participants and their actions is not stored in plain-text or observable in transit over the Internet. Additionally, to obfuscate identities further, the mapping between the email address and the participant's anonymous identifier is stored separately from the participant data collected by Webmunk, which are associated to the anonymous identifier.

[10]The order history is automatically crawled by Webmunk after participants click on "Upload your Amazon order history" on Webmunk's pop-up window (see Appendix Figure C.1).

Table 1: Product Meta-Categories and Categories in the Incentivized Shopping Tasks

Meta-Category	Category
Health:	Pain reliever, acid reducer, allergy medication, moisturizer
Paper products:	Paper towels, envelopes, notepads, toilet paper
Household items:	Trash bags, hand soap, umbrella, laundry detergent
Apparel:	Socks, t-shirt, shorts
Electronics:	Extension cord, monitor cable, batteries, charger
Personal care:	Nail clippers, deodorant, toothpaste, comb/brush

In order to make the product choices incentive compatible, we randomly selected one out of every three study participants who successfully completed the incentivized shopping tasks. For them, we randomly picked one product in their wishlist, we purchased it and shipped it to the address listed in the wishlist (not visible to us). The difference between $50 and the price we paid for the product was sent to the study participant in the form of an additional gift card. In practice, the average price of the products added to the wishlist is around $21. Using data from the participants' prior order histories on Amazon, we confirm that this price is similar to the price of their average past order (at $23 from Appendix Table A.2).

Randomization and data tracking. When participants install the extension, Webmunk randomizes them into one of three treatment groups. In the control group, the extension does not modify anything of the participant's browsing experience. In the Hide Amazon condition, the extension identifies and removes products associated with an Amazon brand. To do this, the extension checks the HTML for pre-determined strings related to Amazon Brands (e.g., 'Amazon Basics' and 'Goodthreads').[11] The extension also checks if an item was flagged by Amazon as an Amazon brand.[12] When the extension identifies a product as an Amazon brand, it hides the HTML block corresponding to that product, and the rest of the webpage is automatically adjusted not to show any blank spaces. In the Hide Random treatment, the extension counts the Amazon-branded products appearing on the page, and randomly selects an equal number of products to be removed.

Appendix Figure C.2 displays the same search results page under the three different treatments. Panel (a) shows the search results for the control group when searching for *batteries*. Panel (b) shows that the four products identified as Amazon brands—the 5^{th}, 6^{th}, 7^{th}, and 9^{th} products in Figure C.2a—are removed from the search results and automatically replaced with

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

[12]Amazon started including the Amazon brand badge to search results carrying an Amazon brand. Whenever an Amazon brand is advertised, Amazon shows the Featured from our brands flag below the product image rather than the Sponsored flag. The browser extension identifies both phrases 'Amazon Brand' and 'Featured from Our Brands' as denoting Amazon brands.

the products that immediately follow them in search results. Panel c shows how four products (corresponding to the four Amazon-branded results) are removed at random: the 2^{nd}, 4^{th}, 8^{th}, and 9^{th} products in Figure C.2a. Note that when products appear in sequence—for example, in search results or in carousels of product recommendations contained at the bottom of product pages—the removal of each product is seamless.

Products can also appear in non-search locations on Amazon.com. We employ a few strategies in these cases, since we are unable to precisely remove Amazon items in certain cases without negatively affecting the rendering of the page. In both the Hide Amazon and Hide Random treatments, we completely remove product comparison tables and recommendations of frequently bought together products that appear on the product page (Figure C.3), because a mix of private labels and non-private labels often appear together. In other cases, such as when Amazon brands are featured on the landing page (Amazon.com), or are listed in wishlists or past orders, we do not remove them from either treatment. As a result, even though we remove most Amazon branded products, there is still a small chance that a participant sees and purchases an Amazon brand even in the Hide Amazon treatment. We later show that the Hide Amazon treatment is effective at reducing the availability and purchase rates of Amazon brands.

While installed, the extension tracks a selection of the URLs that participants visit. In particular, when browsing Amazon.com, the extension tracks search, product, cart, checkout, and wishlist pages, removing all personally identifiable information (such as credit card information or shipping addresses) before even storing the data. For pages within Amazon.com, the extension collects information on the products appearing on the pages, and a subset of the clicks that the participant performs. This data collection effort allows us to, for example, identify that a participant searched for a *coffee mug* on Amazon.com, saw a list of search results, each with specific characteristics and position on the page, visited the product pages of a few of those products, and eventually added to cart and purchased one of them (or added one to the wishlist in the case of the incentivized shopping task). Webmunk also tracks visits to other major e-commerce sites, such as Walmart.com and Target.com, but does not record any information except for the page URL.[13] This allows us to identify the extent to which participants shop across competing e-commerce sites.

2.2 Study Population

This section presents descriptive statistics about our study population. We start with Table 2, which presents the number of participants across the various steps of the experiment. Over 14,000 participants started the eligibility survey. Of these, 2,779 qualified and formally consented to the study, but only 74% of them successfully installed Webmunk. Out of the successful installs, 75% completed the incentivized shopping tasks, meaning that they shared with us an

[13]A full list of domains tracked is available in Appendix B.

Table 2: Number of Participants across the Experiment Funnel

Stage	N	Percent	% Female	Avg. Age
IRB Consent	2779	100.0%	76%	44
Webmunk Install	2063	74.2%	78%	44
Wishlist Choice Sample	1549	55.7%	78%	43
Order History	1433	51.6%	81%	45
Milestone 1	1255	45.2%	81%	44
Milestone 2	903	32.5%	81%	44

Notes: This table displays the number of participants across the various steps of the experiment. To be eligible for the study, individuals needed to be US residents, shop online primarily on a computer that is not shared with other household members, use Chrome for their regular browsing, shop at least 2-3 times a month on Amazon, and not work at Amazon. Milestone 1 involves completing the incentivized shopping tasks and uploading the Amazon order history (each of the two tasks separately are displayed as indented rows). Milestone 2 denotes the completion of the final survey and the uploading of the Amazon order history at the end of the study.

Amazon wishlist with 6 items across the categories listed in Table 1 (*wishlist choice sample*). This wishlist choice sample is the main dataset we use in this paper.

The study experienced further participant attrition. Even for those who completed their wishlist task, we were not always able to collect their order history. As a result, the total number of individuals who completed milestone 1 was 1,255. This is close to our planned sample size of 1,000—1,200 people who would complete milestone 1.[14] Finally, not all participants kept the extension installed for the following 8 weeks. 903 participants successfully completed milestone 2 and thus finished the full study.

To understand the patterns of attrition, we compare age and gender along the experiment funnel. The share of female participants increases along the study funnel (76% to 81%), but the average age remains approximately constant around 44 years old. The large proportion of women is clearly not representative of the US population and likely related to the fact that the research was advertised as a study of online shopping behavior. Yet, the large share of women is more likely to be representative of who is responsible for household-related shopping, which is what we are interested in studying here. Indeed, according to the consumer research company Numerator, 75% of Amazon shoppers are female.[15]

For participants who completed the incentivized shopping tasks, we have more demographic characteristics. Participants reside across the US, with the largest states being California (9.1% of participants), Massachusetts (7.4%), New York (6.6%), Texas (6.1%), Pennsylvania (6%), and Florida (5.7%). Their income ranges from less than $25,000 (13% of them) to over $200,000 (9%). The vast majority of the participants are white (74%), followed by Black and Asian participants (both at 10%). A large share, 33%, have a graduate degree, and another 24% have a bachelor degree. 27% live alone, whereas 19% live in a household of at least four

[14] See our pre-registration here: AEARCTR-0011370. `https://doi.org/10.1257/rct.11370-1.0`.

[15] `https://www.homepagenews.com/retail-articles/numerator-average-amazon-shopper-spent-2662-on-site-in-2023/`, accessed June 2024.

people. 30% of the participants have children. Perhaps the most surprising fact is the share of online spending. Based on self-reported metrics, 57% of the respondents spend at least half of their monthly spending online.

Appendix Figure C.5 compares the study population to the US population across four demographic characteristics: geographic location across states (top left), household size (top right), income (bottom left), and race (bottom right). The plots show a remarkably similar distribution between our study participants and the US population, with some minor exceptions.The North-East is overrepresented among the top states (Massachusetts, New York, Pennsylvania), whereas California, Texas, and Florida are underrepresented. For income, the tails (less than $25,000 and over $200,000) are slightly underrepresented in our study. Lastly, white and Asian populations are overrepresented, whereas Black and Hispanic populations are underrepresented.

We perform three sets of checks to verify that: 1) demographics are balanced across the treatment groups; 2) there is no differential attrition across the treatment groups; and 3) since we rely on the extension collecting participant data, there is no differential tracking across treatment groups. We find no differences in participants across these margins. We report these checks in Appendix A.1.

3 Reduced-Form Evidence of Demand Effects

In this section, we study the causal effects of removing Amazon brands on the shopping behavior and choices of participants. We focus on the incentivized shopping tasks, for which we have comparable data across conditions and categories. We provide evidence on how observable characteristics of selected products change when Amazon brands are not available (Subsection 3.1); how search effort is impacted (Subsection 3.2), whether consumers' propensity to shop on Amazon decreases (Subsection 3.3); and finally, how customer satisfaction for the selected products varies with the availability of Amazon brands (Section 3.4). The last subsection offers additional survey-based results that shed light on how consumers make choices on Amazon.

3.1 Substitution Patterns

We first consider the choices individuals make. Individuals for whom Amazon brands are exogenously not available must substitute to other products. Of key interest is which types of products they substitute towards. If consumers substitute towards very different products (in terms of prices and other observable characteristics), then Amazon brands are likely offering alternatives that are distinctly positioned and potentially of high value to consumers. Alternatively, if consumers substitute towards very similar products, then Amazon branded products are likely increasing competition but do not constitute a fundamentally different offering. Lastly, to the extent that Amazon self-preferences, the offering of these products could be harmful to consumers.

We study these effects with simple linear regressions. We pre-registered specifications of the following type:

$$y_{ic} = \beta Hide\_Amazon_i + \gamma_c + \epsilon_{ic}, \tag{1}$$

where y_{ic} denotes characteristics of the product chosen by participant i in category c.[16] The fixed effects γ_c are included to control for category differences (where the categories are defined as in Table 1). $Hide\_Amazon_i$ is a dummy equal to 1 if the participant is in the Hide Amazon treatment group. We compare the choices of participants in the Hide Amazon treatment group with the control and Hide Random groups separately. We cluster standard errors at the participant level.

Before turning to the results, we discuss some measurement issues when conducting this analysis. We would like to measure the characteristics of the products chosen by participants. We observe product characteristics from various sources: search results, product pages, and wishlist pages (we track the latter through both Webmunk and URLs that participants share directly with the study team). The observable characteristics can vary across sources, since search and wishlist pages only have a subset of product information. Furthermore, some characteristics of the product such as the price and reviews may change over time, and can result in measurement error issues when tracked through the wishlist pages.[17] We leverage the repeated observations across pages to fill in values when missing and disambiguate when we have multiple values for a given variable.

Table 3 presents summary statistics of the selected products, separately for meta-categories where Amazon brands are present and for personal care, where Amazon branded products did not yet exist at the time of the study. The first row shows that our treatment is effective at reducing the availability of Amazon brands. In the control and Hide Random groups, between 9 and 10% of products selected carry an Amazon brand. In the Hide Amazon group, that share drops to 2%.[18] On the left-hand side of the table, we highlight the key difference between the Hide Amazon group and the other experimental conditions. When Amazon brands are not available, participants select products that have accumulated a much lower number of reviews (about 19,000 compared to 26,000-27,000 accumulated reviews).

Table 4 presents the treatment effect results. Panel (a) compares the Hide Amazon treatment group with the control group. Column (1) confirms that the treatment was effective at reducing the availability of Amazon brands in the treatment group. Yet, many characteristics of the selected products by treated and control participants are not distinguishable from one

[16]These characteristics are tracked through Webmunk. As discussed in Section 2.2, tracking is missing for a subset of the products. However, Appendix Table A.4 shows no differential tracking rates across treatment groups.

[17]We have HTML screenshots of the wishlist pages as part of our verification process of task completion. These screenshots can happen with a delay of a few days from when the participant completed the task.

[18]Note that people could find Amazon branded products, for example, in their prior orders. Once they land on a product page for an Amazon brand, our extension does not forbid participants in the treatment group from selecting the item and adding it to the wishlist (or purchasing it during the organic phase).

Table 3: Summary Stats

	Meta-Categories with Amazon Brands			Personal Care categories without Amazon Brands		
	Control (1)	Hide Amazon (2)	Hide Random (3)	Control (4)	Hide Amazon (5)	Hide Random (6)
Amazon Brand	0.09	0.02	0.10	0.00	0.00	0.00
Average Star Rating	4.60	4.60	4.60	4.52	4.54	4.53
Fast Delivery	0.45	0.40	0.43	0.49	0.44	0.47
Free Delivery	0.95	0.93	0.93	0.95	0.94	0.96
Nr. Ratings	27,718	19,085	26,024	12,787	14,425	14,010
Price ($)	21.16	20.58	20.67	16.71	15.94	16.65
Prime Eligible	0.74	0.72	0.76	0.75	0.75	0.78

Notes: This table presents summary statistics for the incentivized shopping tasks. The first row shows the share of products selected by participants that are Amazon brands. The other rows show the average price, average star rating, number of reviews, share of products that are Prime eligible, share of products that have free delivery, share of products that have fast delivery, and share of products that are sold by Amazon.

another: prices, average star rating, and Prime eligibility are all comparable. On the other hand, the number of reviews is significantly lower in the Hide Amazon treatment. Note that the increase in major brands is larger than what would be predicted by a simple random substitution pattern (diversion by share). Major brands have a 43% market share, and if 43% of the individuals who would have bought Amazon products (7.7% from column 1) bought major brands instead, the increase in major brands would have been lower, at 3.3 percentage points. The removal of Amazon Brands also increases the share of sponsored products selected compared to the control group, although this effect in not robust to various specifications.[19]

Because our extension removes sections of the product pages—such as product comparisons and frequently bought together recommendations—that may affect consumer choice beyond the simple absence of Amazon brands, Panel (b) of Table 4 compares the Hide Amazon treatment with the Hide Random treatment, both of which experience the same type of page modifications and similarly sized reductions in the number of products returned in search result. Results are similar to those in Panel (a), although there is no difference in the share of selected products that are sponsored.

Overall, the results suggest that, when Amazon brands are not available, consumers choose fairly similar alternatives, except that these alternatives have fewer reviews. Appendix Table C.3 conducts this analysis for the personal care meta-category, where there were no Amazon brands. Except for an increase in the share of sponsored products, we find null effects. Additionally, Appendix Table C.4 combines the observations from the personal care meta-category and all other meta-categories to estimate difference-in-differences coefficients, where the first difference is given by participants being randomized in multiple treatments, and the second difference comes from comparing categories with and without Amazon brands. We again find

[19]We also pre-registered specifications in which covariates are used in an attempt to increase precision. The results from these specifications are similar.

Table 4: Treatment Effect Regressions

(a) Amazon vs Control

	Amazon Brand (1)	Price (2)	Ratings (3)	Stars (4)	Major Brand (5)	Prime Eligible (6)	Sponsored (7)
Hide Amazon	-0.077***	-0.628	-8,981.483***	-0.001	0.039**	-0.024	0.027***
	(0.007)	(0.395)	(1,646.103)	(0.006)	(0.012)	(0.024)	(0.004)
R^2	0.086	0.109	0.114	0.201	0.287	0.011	0.017
Observations	5,350	5,280	5,238	5,238	5,350	5,350	5,350
Mean of Y	0.094	21.106	27718.064	4.604	0.434	0.743	0.000761
Category fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes

(b) Amazon vs Random

	Amazon Brand (1)	Price (2)	Ratings (3)	Stars (4)	Major Brand (5)	Prime Eligible (6)	Sponsored (7)
Hide Amazon	-0.084***	-0.088	-7,066.594***	-0.003	0.052***	-0.037	0.001
	(0.007)	(0.393)	(1,526.730)	(0.006)	(0.013)	(0.024)	(0.005)
R^2	0.085	0.108	0.101	0.211	0.275	0.014	0.007
Observations	5,259	5,174	5,144	5,145	5,259	5,259	5,259
Mean of Y	0.099	20.744	26023.93	4.605	0.425	0.756	0.027
Category fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The pre-registered primary outcomes of interest are the following: whether the chosen product carries an Amazon brand, price, number of reviews, major brand, whether the search result is sponsored, whether the chosen product is sold by Amazon (note, we omitted this variable since we could not reliably measure it). The other outcomes are secondary outcomes.
***$p < 0.001$; **$p < 0.01$; *$p < 0.05$, .$p < 0.10$.

broadly similar results.

The substitution towards products with fewer reviews is the largest and most robust effect of removing Amazon brands, and is confirmed across many specifications. Its interpretation is non-trivial, however. To the extent that the number of reviews signals popularity and underlying quality, the fact that substitute products have fewer reviews means that they are worse. However, Amazon does control which products participants see when searching on the platform, thereby potentially biasing purchases—and thus number of reviews—towards its own products. It can also solicit reviews for its own brands at a higher rate than the reviews of other brands, thereby letting its products accumulate reviews faster than comparable alternatives.[20]

3.2 Search Behavior

Even if participants find close substitutes to Amazon brands, the absence of these Amazon brands may still result in changes in search behavior. For example, if participants need to search for longer or click on more products to find a suitable substitute, this could be a sign that the

[20]Using data from Keepa, we have analyzed the review accumulation pattern of Amazon branded products compared to other similar products, and we find that Amazon brands show a steeper adoption rate in every period, from first review to the most recent time.

Table 5: Effects on Search

	Num. of Search URLs (1)	Num. of Search ASINs (2)	Num. of Product Page URLs (3)	Num. of Visited ASINs (4)
Hide Amazon	-0.088	-7.521**	-0.094	-0.073
	(0.090)	(2.517)	(0.122)	(0.096)
R^2	0.065	0.232	0.128	0.135
Observations	5,350	4,717	5,350	5,350
Mean of Y	1.874	88.957	2.377	1.868
Category fixed effects	Yes	Yes	Yes	Yes

Notes: This table presents regressions of search outcomes on treatment status (Hide Amazon or control). The number of search pages refers to the number of distinct search pages visited (URLs), number of ASINs visited is the unique number of products visited, and number of product pages is the number of product pages visited, including duplicates. $p***< 0.001$; $**p < 0.01$; $*p < .05$.

absence of Amazon brands is harmful to the participant experience. One of the advantages of our setup is that we can observe all search behavior.

To measure the effect of the treatment on search behavior, we consider several outcomes. First, we look at the number of searches participants perform for each product category of the incentivized shopping task. We expect that if consumers are not able to find what they are looking for with an initial search, then they may refine the search in a variety of ways (e.g., they conduct more searches or use different filters), leading to an increase in the number of search URLs. Second, we look at the number of unique products returned in search results, as a measure of the consideration set. Third, we look at the total number of product pages participants visit (including duplicate visits to a specific product page). Finally, we consider the number of unique products inspected through product pages.

In order to map search terms and product pages to our categories in the incentivized shopping tasks, we need to map browsing behavior on Amazon to activity related to the incentivized tasks. We first restrict attention to website activity between the start and end times of the Qualtrics survey that participants fill out as they complete the incentivized shopping tasks. Then, we use OpenAI's gpt4 model to classify search terms (for searches) and product titles (for product pages) to the 23 categories from Table 1 (or to an "Other" category). With this mapping, for every product category we can compute the number of search URLs, which vary because of search terms or the use of filters, the number of product pages visited, and the number of unique ASINs (Amazon product identifiers) inspected. We test whether these outcomes are different between Hide Amazon and the control group.

Table 5 displays the results. We find no statistically significant differences in the number of searches, total product pages visited, or unique products inspected between the treatment and the control groups. The point estimates are all small in size and statistically indistinguishable from zero, implying no change in search behavior. Consistent with our experimental design, we find that participants in the Hide Amazon group are exposed to fewer products in search results

Table 6: Effects on Amazon Traffic Share

	Amazon vs. All (1)	Amazon vs. Target and Walmart (2)	Amazon vs. eBay (3)
Constant	0.529***	0.783***	0.919***
	(0.014)	(0.013)	(0.008)
Hide Amazon	0.002	0.000	-0.004
	(0.019)	(0.019)	(0.012)
Hide Random	0.004	-0.003	0.002
	(0.020)	(0.019)	(0.012)
R^2	0.000	0.000	0.000
Observations	1,237	1,237	1,237

Notes: This regression presents treatment effect regression about participants' satisfaction with price, product quality, and overall (on a scale from 1 to 5). $p < 0.001$; **$p < 0.01$; *$p < 0.05$

than those in the control group, as shown in column (2). (This reduction is similar to the share of Amazon brands found in search results.)

This analysis confirms that when Amazon products are not available, people search similarly, indicating that they find suitable substitutes without much additional effort.

3.3 Cross-Platform Effects

The presence of private labels may attract consumers to Amazon instead of other websites, even when a consumer ultimately purchases another product. In this way, private label brands can play an important role in increasing overall demand for all products on a platform, and this could drive cross-platform competitive effects.

To assess whether Amazon brands affect consumer substitution across platforms, we use the URLs visited by participants during the 8-week organic shopping period, after the incentivized shopping task. During this window, participants are not asked to complete any task; the Webmunk extension collects data on shopping behavior passively. However, the randomization into the three conditions (control, Hide Amazon, Hide Random) persists during this period.

For each participant, we calculate the share of URL visits that belong to the Amazon.com domain out of all retail websites.[21] The full list of tracked domains is in Appendix B. We then regress these shares on treatment indicators to determine whether removing Amazon brands has an influence on cross-platform activity.

Table 6 reports the results. The Constant coefficient in column (1) indicates that Amazon.com accounts for 53 percent of all retail URLs visits tracked by our extension. The near-zero coefficients on Hide Amazon and Hide Random (which should be interpreted as marginal effects

[21] Since we observe participant activity in sequence, we determine new website visits based on whether it is a different web domain from the one that was previously visited. Accordingly, a specific URL—such as the Amazon homepage—may be recorded as multiple distinct webpage visits for a single participant on a given day if the participant returns to it following visits to other domains.

Table 7: Effects on Shopping Again on Amazon – Survey Evidence

	Personal Care (1)	Electronics (2)	Apparel (3)	Household Items (4)	Paper Products (5)	Health (6)
Constant	4.192***	4.302***	3.801***	4.178***	3.900***	4.050***
	(0.044)	(0.038)	(0.049)	(0.043)	(0.052)	(0.046)
Hide Amazon	-0.062	-0.009	-0.063	-0.053	-0.041	-0.012
	(0.062)	(0.054)	(0.069)	(0.060)	(0.073)	(0.065)
R^2	0.001	0.000	0.001	0.001	0.000	0.000
Observations	1,037	1,037	1,037	1,037	1,037	1,037

Notes: This table presents regressions of the answer to the following question: "if you had to buy products in these categories again, would you shop for them again on Amazon.com?" There were five possible answers, from "Definitely not on Amazon.com" to "Definitely yes." We convert the 5 categories to a Likert scale where 5 is definitely yes. $p***<0.001$; $**p<0.01$; $*p<.05$.

compared to the Control group) indicate that neither treatment group experienced meaningful changes in the share of website visits to Amazon. The standard errors allow us to reject the hypothesis that the presence of Amazon brands have a moderate effect on Amazon's website traffic, relative to other retailers.

We consider specific platform substitution channels in columns (2) and (3). For column (2), we construct the share of Amazon URL visits out of a narrower set of domains: Amazon.com, Target.com, and Walmart.com. For column (3), we construct the share of Amazon URL visits out of visits to Amazon and eBay. Amazon has a 78 percent share of webpage visits out of the group that includes Target and Walmart, and it has a 92 percent share of out Amazon and eBay. As in column (1), we do not find evidence that either treatment affected Amazon's website traffic relative to these other retailers, and the standard errors are precise enough to reject moderate decreases in traffic.

Additionally, we provide corroborating evidence using survey responses about whether participants would shop again on Amazon. Table 7 presents the results for each of the 6 product categories. The outcome is the participant's answer to the following question: "If you had to buy products in these categories again, would you shop for them again on Amazon.com?" There were five possible answers, from "Definitely not on Amazon.com" to "Definitely yes." We convert the 5 categories to a Likert scale where 5 is definitely yes. While the propensity to shop on Amazon for these products varies by categories, with the lowest for apparel and the highest for electronics, there is no differential effect of removing Amazon brands—as indicated by the Hide Amazon coefficients.

Our results suggest that Amazon private labels do not play a large role in steering consumers to the Amazon platform in the short run. It may well be the case that, over a horizon longer than the 8 weeks in our study, the absence of Amazon brands may have a larger effect.

Table 8: Effects of Treatment on Participant Satisfaction

	Price (1)	Product Quality (2)	Overall Rating (3)
Constant	4.039***	4.416***	4.364***
	(0.095)	(0.078)	(0.074)
Hide Amazon	-0.002	0.152	0.093
	(0.132)	(0.109)	(0.103)
R^2	0.000	0.012	0.005
Observations	158	158	158

Notes: This regression presents treatment effect regression about participants' satisfaction with price, product quality, and overall (on a scale from 1 to 5). $p < 0.001$; $**p < 0.01$; $*p < 0.05$

3.4 Consumer Satisfaction and Additional Survey Evidence

Although we find no differences in search behavior across treatment groups, it is possible that participants are less satisfied with the products they choose when Amazon brands are not available. To study this, in the final survey, we asked a variety of questions about participant satisfaction with the products they purchased on Amazon (including, when applicable, the product we purchased from their wishlist), their preferences towards Amazon brands, and their general shopping behavior.

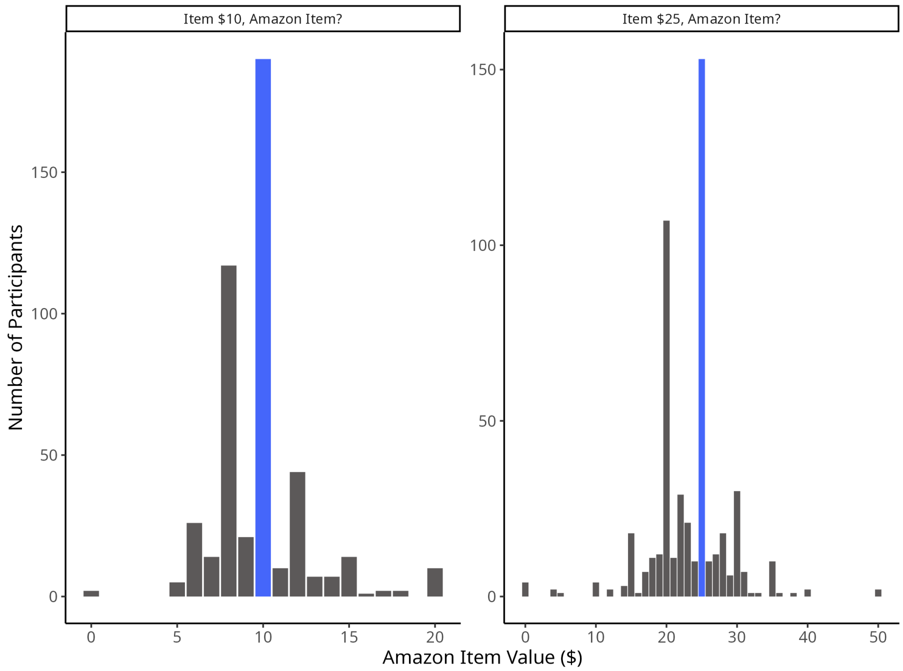
First, for those who received an item from the incentivized shopping tasks, we asked how they would rate the product, overall and separately for price and quality.[22] Table 8 displays the results of regressions comparing satisfaction (on a scale from 1 to 5) across treatment conditions. Note that the number of observations is substantially lower since a) only a third of the participants received an item from their wishlist, and b) there was substantial attrition between milestone 1 and milestone 2 (see Table 2). We find no differences in participant satisfaction, overall nor for price or quality separately. Even if the statistical power of this analysis is limited due to small sample size, the estimates exclude large changes in satisfaction.

Next, we analyze a set of survey questions designed to assess participants' preferences for Amazon brands. Half of the participants were asked how much they would be willing to pay for an Amazon-branded product with the same characteristics as a product they desired. To account for reference price effects, the price of the desired product was randomly set at either \$10 or \$25 with equal probability. The other half of the participants were asked the reverse question–how much they would be willing to pay for a non-Amazon product that matched all the characteristics of an Amazon-branded item they wanted.

Figure 2 displays the distributions of these responses. The left panel plots willingness to pay for the Amazon product when the reference item costs \$10. The right panel is analogous for a \$25 reference item. A large share of respondents value an Amazon branded product exactly the

[22]Delivery speed is an important aspect of product quality that consumers value. However, it is not directly relevant in the context of the incentivized shopping tasks, as the products were purchased on their behalf with a delay after participants made their selection.

Figure 2: Willingness to Pay for Amazon Branded Products



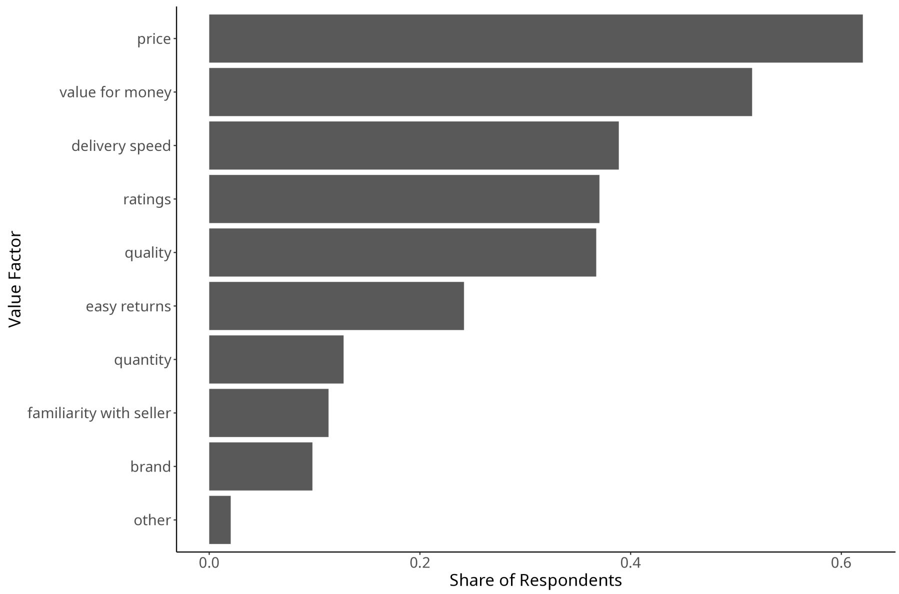
Notes: This plot presents the distribution of participants' responses to hypothetical questions in which they are asked their willingness to pay for an Amazon branded product, if a similar non-Amazon branded product they want costs either $10 (left) or $25 (right).

same as another product. Yet, there is large heterogeneity above and below the reference value. On average, participants are willing to pay less for an Amazon brand than a comparable alternative. When the alternative costs $10, participants value Amazon items -$0.20 less. When the alternative costs $25, participants value Amazon brands -$1.75 less. We obtain similar results for the alternative question wording, in which people are asked their willingness to pay for a non-Amazon item when an Amazon product is the one they want (Appendix Figure C.6).

These preferences are greatly affected by slight changes in ratings and delivery characteristics. To test this, we asked participants how they may trade off Amazon brand with ratings and delivery speed. We first consider the trade off between Amazon brand and ratings. We asked the willingness to pay for an Amazon brand if it had 0.5 lower star rating than the alternative product (4 versus 4.5 stars). We find that consumers are highly sensitive to the rating difference. In particular, for a $10 reference item, their value for the Amazon branded product drops from $9.80 when star ratings are the same (Figure 2) to $8.25 when ratings are 0.5 stars lower. For a $25 reference product, the difference decreases from -$1.75 to -$5.49.

We find that participants also care about delivery speed, but less so than ratings. We asked how much they would be willing to pay for an Amazon product with a faster delivery speed (1

Figure 3: Share of Respondents Valuing Each Factor



Notes: This plot presents the share of respondents indicating that a given factor was one of the ones they valued the most when making purchasing decisions.

day vs 3 days for the alternative product). Participants were willing to pay \$0.59 more for the Amazon product in the \$10 condition (up from -\$0.20), and -\$0.49 less in the \$25 condition(up from -\$1.75).

To finish, we asked individuals to list the three most important factors when shopping online. Figure 3 displays the distribution of responses. Price and value for money were the most important, with more than 50% of respondents listing them as a top factor. Delivery speed, quality, and ratings were chosen by almost 40% of respondents, which corroborates the results on willingness to pay between Amazon and non-Amazon brands described above. Easy returns was also important for more than 20% of respondents. Lastly, quantity, familiarity with seller, and brand were the least important product features.

The survey results provide additional evidence that Amazon branded products are not much worse than alternative products on average, but that there is large heterogeneity around that mean. A sizable group of people is willing to pay more for Amazon brands than similar alternatives. We also find that relative preferences for Amazon brands versus other products can be offset by differences in delivery speed and ratings, which is consistent with our findings on substitution patterns described previously. The next section translates these reduced-form findings into measures of consumer surplus.

4 Model and Counterfactuals

In this section, we develop a model of demand and supply that will allow us to estimate potential welfare effects of private labels. In our model, consumers have heterogeneous preferences and choose among differentiated goods. Suppliers set prices to maximize profits conditional on the set of competing products observed by the consumer and how the products appear in search (i.e., the search rankings). We use the estimated demand model along with the supply-side assumptions to calculate counterfactual equilibrium prices and consumer choice probabilities, allowing us to assess the welfare effects of private labels and separate these effects into changes in product assortment and changes in pricing pressure.

4.1 Demand and Supply

We assume that consumers make a discrete choice over the options they observe while searching for a product in a category. These products could be observed either on search results pages or on product pages. For a given product category s from the incentivized shopping task, consumer i's consideration set J_{is} includes all ASINs appearing on search results pages, product pages (including product variants and recommended alternatives below the main product), and the chosen product. We define the sequence of pages visited in a category before selecting a product as a single *search*, even if, for example, the consumer enters multiple distinct search terms in this sequence of pages.

For a given search s, consumer i's indirect utility for product j is given by:

$$u_{ijs} = \alpha_{mi} p_{js} + \mathbf{x}_{js}\beta_i + \ \zeta_i A_j + \gamma r_{js} + \xi_j + \epsilon_{ijs}$$

where p_{js} is price, $\mathbf{x}_{js}$ is a vector of observed product characteristics, A_j is an indicator for whether or not the product is an Amazon brand, and r_{js} is the log search rank observed by the consumer when selecting options, which accounts for the fact that consumers are more likely to buy products near the top of search results. ξ_{js} a product-specifc demand shock. We allow preference parameters $(\alpha_{mi}, \beta_i, \zeta_i)$ to vary across individuals by demographic characteristics. We also allow the price coefficient α_{mi} to vary across meta-categories, to allow for varying price sensitivities in different contexts. ϵ_{ijs} is distributed Type I extreme value, and independent of product characteristics, rank, and price. We also assume that consumers can choose an outside option with mean utility of 0. (We further discuss this assumption within the context of the incentivized shopping task below.) This formulation yields the standard mixed logit choice probabilities.

Our demand model abstracts away from the details of the search process. For the purposes of the counterfactuals we consider (removing products), we think this assumption is reasonable since our experimental results yielded precise null treatment effects on measures of search behavior (i.e., clicks on product pages, number of searches).

On the supply side, we assume that multi-product brands (e.g., Amazon, Duracell) set prices to maximize profits separately for each search, while treating the consumer demographics as unobserved. This assumption reflects the fact that sellers cannot generate search-specific prices on Amazon, and instead sellers treat each possible search as drawn at random.[23] We use this assumption to back out marginal costs for each product when simulating counterfactual equilibrium prices.

For search ranks, we assume that when products are removed, remaining products slide up to fill in the missing search result slots, mirroring the behavior of our web extension. A key consideration that we do not account for in our counterfactuals is the role of the platform in ordering products for consumers to see. In theory, the platform's ranking policies can have equilibrium effects on pricing decisions by the suppliers, and prices and ranks would both be endogenously determined in equilibrium. Our current specification does not explicitly account for this, in part because reverse-engineering Amazon's ranking algorithm is challenging with our data. However, holding Amazon's existing ranking fixed, we can determine whether removing Amazon brands is beneficial—which could arise if they receive preferential placement over better alternatives—or harmful, if they exert beneficial pricing pressure and their ranking aligns with consumer preferences.

4.2 Empirical Specification and Estimation

In addition to price, the Amazon brand indicator, and log search rank, we include the following product characteristics in estimation: the average star rating, log number of reviews, and indicators for major brand (distinct from Amazon brands), Prime eligibility, fast delivery (within one day), and whether the listing is sponsored.[24] We also include a constant, an indicator for whether we do not observe the star ratings, and several page location indicators.[25] We construct each of these variables by identifying the window in which a consumer was searching for a given product category and the eligible products in the category, and then aggregating over instances in which an ASIN is observed in this search to obtain a unique consumer-search-ASIN combination.

To account for the fact that products that show up higher in search rankings may be of higher

[23]This assumption is equivalent to one in which a population of consumers all make identical searches and search results are not personalized. With our estimates, we are able to test whether search results are indeed personalized according to user demographics, and we do not find any evidence of this personalization, in line with public statements by Amazon.

[24]For products not on search pages, we assign a rank of 0 and we construct r_{js} as $\ln(\text{search rank} + 1)$. If a product appears with multiple search ranks within a search, we use the smallest value (highest rank).

[25]The page location variables indicate whether the product appears: on a search page, in the banner at the top of a search page (above the results), in a product swatch but not on search page, in a similar products bar but not on search page, and on neither a search nor a product page. Participants could have navigated to other pages, such as their previous orders, and avoided search and product pages.

quality in ways that we do not observe, we exploit variation introduced by our experiment. We include in $\mathbf{x}_{js}$ the search rank that a product was initially assigned, before the browser extension intervened to hide products. After controlling for the initial assigned rank, the coefficient on the search rank shown to the consumer r_{js} (realized rank), provides an estimate of the causal effect of search rank on utility, γ. We interpret γ as capturing the search cost of considering products lower on the page. Thus, γr_{js} is included in consumer utility for our welfare calculations.[26] The difference between assigned and realized rank is due to products being hidden for our treatment groups, and the coefficient on the assigned rank allows us to control for the correlation between rank and unobserved quality.

We include the following consumer characteristics to construct individual preference parameters: log household income, presence of children at home, Amazon Prime membership, whether the participant had purchased an Amazon brand prior to the experiment, and an indicator for whether we did not obtain an order history for the participant. We allow price, Amazon brand, and star rating coefficients to vary with all of these demographics. We also allow utility for Prime eligible products to vary with Prime subscriber status. We also include two "unobserved" demographics that are drawn independently from a standard normal distribution. The first influences the price coefficient, and the second jointly affects the Amazon brand and major brand coefficients. We restrict the rest of the demographic interactions to zero.

Conditional on the vector of 41 demand parameters $(\alpha_{mi}, \beta_i, \zeta_i, \gamma)$, we construct choice probabilities for each individual for each product in each realized search, and we use these choice probabilities to form an empirical likelihood. We solve for the parameters that maximize the sum of the log likelihood, subject to two additional restrictions that we introduce to address limitations in our experimental data.

First, even after controlling for a rich set of variables, it is possible that price is correlated with unobserved quality. We do not have an instrument for price. To address this, we model endogeneity directly in estimation. Note that, without loss of generality, we can express the unobserved demand shock as $\xi_{js} = \lambda p_{js} + \eta_{js}$, where λ is the coefficient obtained from projecting ξ on p and η is the residual. This implies that we can express the linear utility components $\alpha p_{js} + \xi_{js}$ as $(\alpha + \lambda)p_{js} + \eta_{js}$, where α represents the mean price parameter, and estimate $(\alpha + \lambda)$ jointly under a distributional assumption for η. We assume that $\eta_{js} = \eta_j \sim N(0, \sigma)$, representing a ASIN-level random effect. We then choose α to calibrate the mean own-price elasticity to a value of -5, which we set from survey-based estimates of seller margins obtained by Scout (2022) and similarly used by Yu (2024). The implied value of λ captures the correlation between price and quality.

A second limitation of our data is that, due to the nature of the incentivized shopping task, all experiment participants had to make a choice among the products from Amazon, so we cannot directly measure substitution to the outside option. Instead, we use our eight-week

[26] For comparison, we also report rank-independent utility measures where the impact of r_{js} on utility is omitted.

observational period to measure website traffic to Amazon and other retailers, as shown in Table 6. We use the share of URL visits to other retailers as a proxy for the outside option share (i.e., 0.47). We constrain our estimates to hit this target by imposing a penalty in the estimation objective function if the implied outside option share deviates from this value. We include a constant in $\mathbf{x}_{js}$, which allows our demand estimates to hit this calibrated value exactly.[27]

We make some additional data cleaning steps to generate our estimation data. Due to our experimental design, we drop products with prices less than $2 and greater than $50. Finally, we keep searches for which the consumer observes at least 10 valid product options and the consumer sees products from multiple brands. After these steps, we obtain a mean of 5.5 searches per consumer (given a maximum of 6), with a median of 57 and a mean of 81 products in the consideration set per search. For our estimation sample, we have 1,504 consumers who were assigned to one of our three experimental treatment groups.[28]

Estimated Demand Parameters

Table 9 reports the estimated point estimates and standard errors. Column (1) shows the mean effects of the covariates as well as the parameters governing the distribution of ξ. The coefficient on price is negative, as imposed by the calibration. We quantify substantial positive correlation between price and unobserved quality. The coefficient on realized search rank is negative, as expected. The coefficient on assigned search rank is equally negative, indicating that products that are initially assigned higher search rankings yield greater utility (separately from the realized rank on a page). Our decomposition implies that about half of the relationship between rank and utility reflects unobserved quality. (We plot the average mean utility (Delta) by search rank in Appendix Figure C.7, including and excluding realized rank.)

Average star rating, log number of reviews, Prime eligibility, and fast delivery have positive effects on utility. Our estimates yield a small, statistically insignificant positive effect of Amazon brand. This is in contrast to a positive effect of major brand, equivalent to about $1 in value. Thus, our point estimates indicate that consumers view Amazon brands as lower in quality than major brands, but slightly better than other non-major brands on Amazon. One note is that these values are conditional on other characteristics. Amazon brands may still yield higher mean utility than major brands if they have better reviews or faster shipping.

Columns (2) through (5) report interactions between observed consumer characteristics and covariates. We see limited variation in price and star rating coefficients across consumers, with

[27]For implementation, we use a single draw of unobserved demographics per customer. Based on additional checks, the other demand coefficients appear to be unaffected when we use many (i.e., 50) draws per consumer or whether we omit the unobserved demographics entirely. For the product-level random effects (η_j), our reported estimates use 100 draws. Increasing the number of draws does not appear to affect the other estimated parameters.

[28]We observe product selections for 1,579 participants, but 75 participants do have a sufficient number of valid products in any category. For some participants, our extension did not track any data, either because they turned it off or because they used a different electronic device from the one where they installed Webmunk. Appendix Table A.4 shows that the rate at which participants are tracked is not different across treatment conditions.

Table 9: Demand Estimates

Variable	Mean	Interactions with Demographics						
		ln(Income)	**Children**	**Prime Subs.**	**Prior AB**	**No History**	**Unobs. 1**	**Unobs. 2**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Price	-0.2674 (0.0028)	0.0000 (0.0016)	0.0038 (0.0029)	0.0148 (0.0038)	0.0039 (0.0032)	-0.0135 (0.0050)	0.0011 (0.0009)	-
Stars	0.3925 (0.0654)	0.0008 (0.0092)	-0.0161 (0.0167)	-0.2028 (0.0230)	-0.0343 (0.0180)	0.1113 (0.0277)	-	-
Amazon Brand	0.0600 (0.0629)	-0.0458 (0.0684)	0.3367 (0.1239)	-0.3068 (0.1481)	0.3730 (0.1523)	0.4656 (0.2043)	-	0.0749 (0.0551)
Major Brand	0.2312 (0.0333)	-	-	-	-	-	-	0.0192 (0.0256)
ln(Reviews)	0.0268 (0.0062)	-	-	-	-	-	-	-
Prime	0.6602 (0.0499)	-	-	0.7903 (0.1184)	-	-	-	-
Fast Delivery	0.1385 (0.0294)	-	-	-	-	-	-	-
Sponsored	3.8340 (0.1604)	-	-	-	-	-	-	-
ln(Rank Realized)	-0.5107 (0.1752)	-	-	-	-	-	-	-
ln(Rank Assigned)	-0.5091 (0.1753)	-	-	-	-	-	-	-
Price×Apparel	-0.0232 (0.0039)	-	-	-	-	-	-	-
Price×Electronics	0.0112 (0.0035)	-	-	-	-	-	-	-
Price×Health	0.0035 (0.0034)	-	-	-	-	-	-	-
Price×Household	0.0006 (0.0034)	-	-	-	-	-	-	-
Price×Personal Care	-0.0024 (0.0031)	-	-	-	-	-	-	-
σ	0.1819 (0.1422)	-	-	-	-	-	-	-
λ	0.2803	-	-	-	-	-	-	-

Notes: Table presents parameter estimates for our demand model. Column (1) presents the mean coefficients, while columns (2) through (7) present interactions of participant demographics with product characteristics. Prior AB indicates that the consumer had previously purchased an Amazon Brand. No History indicates that we do not observe Amazon order histories for that consumer. The final two columns represent unobserved demographics, which are drawn independently from a standard normal distribution. Constant, Missing Stars, and the following page location indicators are not displayed: In Search Results, In Swatch (Not Search), In Similar (Not Search), In Top Banner, and Not in Search or Product Page. The last two rows give the parameters governing the distribution of ξ. No standard error is reported for λ because it is calibrated jointly with the price coefficient.

the largest coefficients on the indicator for Prime subscriber. The coefficients on Amazon brand indicate meaningful heterogeneity across individuals, with higher valuations for consumers with children, those who previously bought an Amazon brand, and those for whom we do not have an order history. Prime subscribers have a lower valuation for Amazon brands. We find small coefficients on the unobserved demographics interacted with price, Amazon brand, and major brand in columns (6) and (7), which suggests that our observed demographics likely capture a reasonable amount of the variation in preferences across consumers.

Elasticities and Marginal Costs

Because the realized ranks in our treatment groups diverged from the ranks normally seen by consumers on the platform, we use only data from the 499 consumers in our control group for subsequent results on elasticities and marginal costs. Integrating across consumers, we obtain a mean elasticity of -5.11, which is similar to the calibration value of -5. Our estimates provide variation in own price elasticities across products: the 10th and 90th percentiles of the own price elasticity are -8.71 and -2.14, respectively.

We estimate marginal costs by inverting the profit maximizing first-order conditions for each (multi-product) brand. For Amazon, we assign all of Amazon's brands to a single firm. For other products, we obtain brand information from Keepa for about 95 percent of the observations and treat their pricing decisions similarly to what we do for Amazon brands. We assume that the remaining 5% of products for which we do not have brand information are owned by single-product firms. We do not model the commission rate structure or other features of the Amazon platform.

Summary statistics for the estimated marginal costs and own-price elasticities are displayed in Table 10. The mean marginal cost is \$14.66, as compared with a mean price of \$18.75. This implies a mean margin of \$4.09, or Lerner index of 0.22. Within a search, Amazon brands have marginal costs that are \$3.68 lower on average.

We observe roughly a quarter of the ASINs across multiple searches, whereas we back out marginal costs separately for each search. As a sensibility check, we measure the standard deviation in recovered marginal costs across these observations. The median standard deviation is 0.53 cents, representing less than 4 percent of the mean estimate of marginal costs.

4.3 Counterfactuals and Welfare

We use the model to consider counterfactuals in which products are no longer available or are demoted in search results. In our main counterfactual, we remove all Amazon brands. We adjust search rankings for the remaining products by moving them up into vacated slots. Given the adjusted consideration set, the new search rankings, and the estimated marginal costs, we re-compute equilibrium prices without Amazon brands.

Table 10: Estimated Marginal Costs and Own-Price Elasticities

Variable	**Mean**	**Percentiles**				
		10^{th}	25^{th}	50^{th}	75^{th}	90^{th}
Price ($)	18.75	7.99	11.88	16.99	23.99	32.00
Marginal Cost ($)	14.66	3.90	7.49	13.17	20.18	28.13
Own-Price Elasticity	-5.11	-8.71	-6.66	-4.65	-3.16	-2.14

Notes: Table displays the estimated marginal costs and own-price elasticities for products in the 2,742 searches from the 499 consumers in our control group. We integrate over consumer demographics when calculating elasticities and marginal costs.

We construct the average consumer surplus by integrating over the demographics of all 499 consumers in the control group across all searches. We start with the 2,742 consumer-category searches from the control group, and narrow our focus to the 2,283 searches where Amazon brands are present. In the placebo meta-category of personal care, where Amazon had not yet widely introduced its own brands, effects are near-zero by design. Within the non-placebo categories in this sample, 5.3 percent of products are Amazon brands. Of these, 19 searches with an unusually large number of products (mean of 629) did not converge, leaving us with 2,264 searches in our analysis. In each case, we compute consumer surplus in dollar terms following Small and Rosen (1981).

Table 11 presents estimates of consumer surplus across scenarios. In the baseline, across the 19 product categories with Amazon brands, we estimate that consumers receive an average of $3.22 in consumer surplus per completed search. The second row shows that removing Amazon brands reduces consumer surplus by 5.5 percent (or $0.18 per search), while increasing prices by 0.2 percent. This implies that Amazon's presence contributes to slightly lower equilibrium prices. For comparison, we present an alternative scenario where Amazon brands are removed and search ranks adjust, but prices are unchanged (third row). In this case, consumer surplus would decline by 4.9 percent. Thus, price effects account for 10 percent of the total benefits that Amazon products provide to consumers in our sample. Finally, the removal of Amazon-branded products results in a small shift toward the outside option, which is within the 95% confidence interval of our reduced form results.[29]

As a benchmark, we also conduct a counterfactual in which we remove products at random (fourth row in Table 11). Specifically, we identify the number of Amazon brands observed in each search, and we remove an equal number of products at random from that search (which may include some Amazon brands). We then adjust search ranks and compute new equilibrium prices in the same manner as above. On average, the removal of random products reduces consumer surplus by 3.2 percent, or about 60 percent of the impact of removing Amazon products.

[29]For comparison, we also consider welfare calculations using rank-independent utility. Specifically, we consider consumer surplus where realized rank affects choice probabilities but does not enter the consumer's utility function, and we use the approach of Allcott (2013) and Train (2015). These results are similar, and are reported in Table C.5.

Table 11: Counterfactual Welfare Effects

Scenario	Consumer Surplus ($)	% Change in CS	% Change in Prices	Inside Share
Baseline	3.22	—	—	0.524
Remove Amazon	3.04	-5.50	0.20	0.502
Remove Amazon (No Price Adj.)	3.06	-4.87	0.00	0.504
Remove Random	3.11	-3.20	-0.05	0.511

Notes: This table reports the effects of removing products on consumer welfare, excluding placebo personal care categories (nail clippers, deodorant, toothpaste, and comb/brush). The second row reports the impacts of removing Amazon products, adjusting search ranks, and computing equilibrium prices. Impacts without equilibrium price adjustments are reported in the third row. For comparison, the impacts of removing products at random and re-computing equilibrium prices are reported in the fourth row. Consumer surplus is calculated following Small and Rosen (1981). Percent changes are reported relative to the baseline scenario. Price changes are weighted by choice probabilities (shares). Inside share represents the total probability of purchasing any product within the choice set.

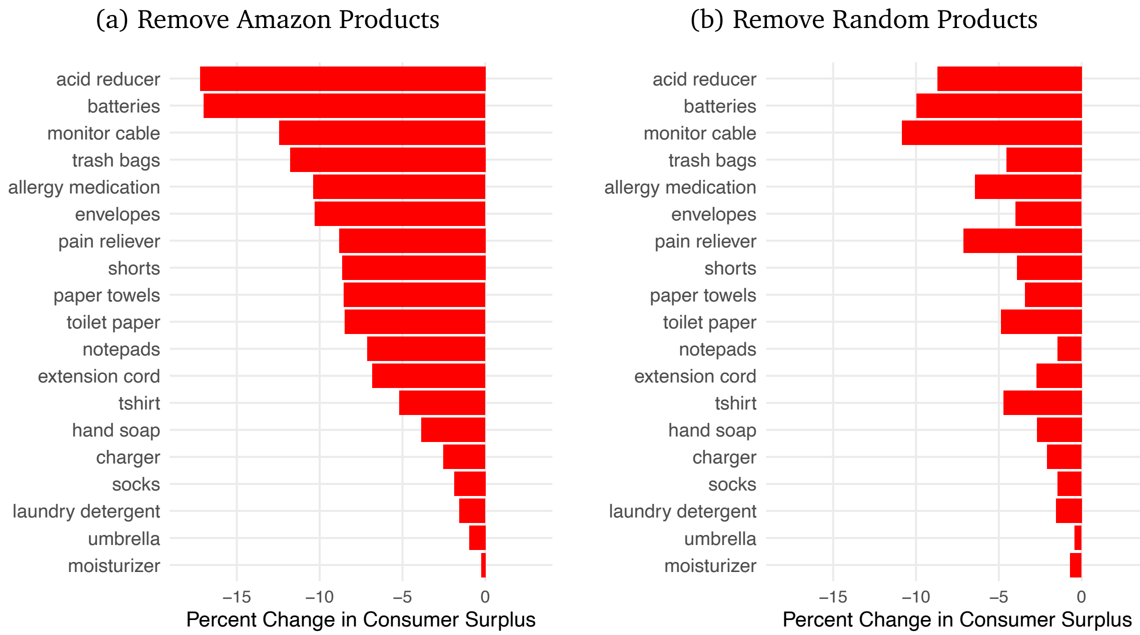
This implies that Amazon branded products provide greater consumer surplus than a random set of products.

The aggregate results mask sizable heterogeneity across product categories. Figure 4 reports the category-level changes in consumer surplus from removing Amazon products (panel (a)) and removing random products (panel (b)). Removing products at random results in smaller welfare reductions overall, though the order of impacts is not perfectly correlated with the impacts of removing Amazon products. For thirteen categories, the counterfactuals imply that Amazon brands increase consumer welfare by over 5 percent. By contrast, removing an equal number of products at random has a 5 percent impact in only five categories. We estimate that in only one category (moisturizers), removing random products makes consumers worse off than removing Amazon brands. This is possible because Amazon products may displace more preferred products in rankings, causing consumers to buy worse alternatives.

These results imply that the value of Amazon products are heterogeneous across categories. In categories such as acid reducers and batteries, Amazon products are valued more by consumers than the average third-party product, and their removal yields consumer harms substantially greater than removing random products. We compare the effects explicitly in Appendix Figure C.8. For categories where removing Amazon brands has a greater impact than random removal, Amazon products tend to yield greater mean utility than other products, but they are not perfectly correlated. This implies that a large part of the effect is due to heterogeneity in preferences—including heterogeneity in preferences for Amazon brands.

An additional counterfactual we consider is one in which Amazon brands are shifted down in the search rankings. We motivate this counterfactual by recent policies that prohibit self-preferencing on digital platforms. In earlier work, we found that Amazon brands appear higher in search rankings than observably similar products (Farronato, Fradkin and MacKay, 2023). We confirm that finding with our new data (see Appendix Table C.6), where products are ranked roughly 4 slots higher up in search rankings when conditioning on variables that the consumers

Figure 4: Category-Level Effects on Consumer Surplus



Notes: Figure reports the percent change in consumer surplus from removing Amazon products (panel (a)) and random products (panel (b)). Consumer surplus is calculated following Small and Rosen (1981) after search ranks are adjusted and new equilibrium prices are computed. Categories are sorted in ascending order based on the impact from the removal of Amazon products.

can see on the search page. We thus estimate the welfare effects of demoting Amazon brands down 4 positions, which would eliminate this difference. This counterfactual yields a small (-0.33 percent) decline in consumer surplus on average and slight decreases across all categories (left panel in Appendix Figure C.9).

While this may seem surprising, the apparent prioritization of Amazon brands in search results—-only conditional on product characteristics observable to consumers—may simply reflect the influence of unobserved attributes that favor Amazon's private-label products, as we discussed in (Farronato, Fradkin and MacKay, 2023). Using our demand estimates, we can explore the extent to which unobserved utility components account for the higher rankings of Amazon brands. We find that unobservables partially explain the higher rankings, but they do not fully account for the extent of the observed prioritization. Column (3) of Appendix Table C.6 indicates that Amazon products are still ranked about 2 slots higher in search rankings even when controlling for unobserved components of utility that we recover from our demand estimates. If we shift all Amazon brands down by 2 positions in search rankings to correct for this average difference, we estimate similarly small declines in consumer surplus across all categories (right panel of Appendix Figure C.9).

The finding that demoting Amazon's private-label products by the estimated extent of self-preferencing does not lead to improvements in consumer welfare runs counter to what regulators envisioned when prohibiting self-preferencing practices. Our evidence shows the important role of heterogeneity in consumer preferences: ranking products solely based on mean utility overlooks the fact that some consumers have strong preferences for Amazon brands.

These results suggest that it is worth exercising caution when assessing the role of vertically integrated products on consumer welfare. Our results imply that consumer surplus may not easily be improved by enforcing broad prohibitions on self-preferencing. Enforcing a neutral ranking on average slightly decreases consumer surplus in our calculations, and removing products—even when similar products are available—may substantially reduce consumer surplus. One reason for this is that consumer heterogeneity in preferences, including for vertically-integrated brands, may drive a meaningful component of welfare. We are able to provide an assessment of the benefits by estimating underlying utility parameters and simulating counterfactual prices and choices.

Our welfare analysis is subject to caveats. We consider short-run effects where only prices can adjust. We do not address how non-price characteristics might change in response to these policies, and we do not address how new products might enter in response to the removal of private labels. Other supply-side features that we abstract away from and may influence short-run welfare are the determination of sponsored products, adjustments to the order of search results in response to changing consideration sets and prices, and strategies that Amazon might use to incentivize sellers to adjust prices in ways that fall outside of the Bertrand-Nash assumption we maintain here.

5 Conclusion

In this paper, we explore the effects of removing Amazon brands from the choice set of Amazon consumers using a field experiment. We find that participants have ample choices in categories where Amazon brands exist. As a result, removing Amazon branded products leads participants to select substitutes with fairly similar characteristics, from price to ratings to delivery speed. A notable exception is that in the absence of Amazon brands, the alternative selection has a lower number of accumulated reviews. Survey evidence from our experiment supports our empirical findings. Consumers typically care most about price, quality, and delivery speed, rather than brands per se.

We also find that removing Amazon brands does not lead to significant changes in consumer search behavior. The number of searches performed, the total number of product pages visited, and the number of unique products viewed remain largely unchanged, suggesting that participants can find alternative products without additional search effort. Furthermore, we observe no increase in shopping at other retailers, indicating that consumers remain within Amazon's ecosystem even when its private labels are unavailable, at least within a couple of months since intervention. Our survey evidence also confirms that consumer satisfaction with purchased products does not differ meaningfully between treatment groups, reinforcing the conclusion that substitutes are perceived as largely comparable.

Despite the availability of observably similar substitutes, the presence of Amazon brands may have non-negligible welfare effects. We estimate a structural demand model to quantify the effects of Amazon brands on consumer welfare. Our counterfactual simulations indicate that removing Amazon products results in a 5.5 percent decrease in consumer surplus ($0.18 per search). About 10 percent of this effect stems from the price pressure Amazon brands exert on their competitors. We observe significant heterogeneity across product categories, with acid reducers and batteries benefiting the most from private labels and umbrella and moisturizers benefiting the least.

In response to recent regulatory scrutiny of self-preferencing, we evaluate counterfactual scenarios in which Amazon's private labels are ranked according to their mean utility across consumers. Contrary to regulators' intended objectives, we find that demoting these products in the search rankings to correct for apparent self-preferencing leads to modest reductions in consumer surplus. This outcome is driven by substantial heterogeneity in consumer preferences for private labels, which average-based rankings fail to capture.

These findings highlight the complex trade-offs involved in platforms' vertical integration strategies. While Amazon's private labels generally enhance competition and provide price benefits, their impact varies across categories. Our study underscores the importance of empirical analysis in evaluating policies toward vertically integrated platforms, as blanket restrictions on private labels or ranking practices may have unintended consequences for consumer welfare.

Our analysis is subject to limitations. We evaluate only the short-term effects of removing Amazon brands, meaning we do not capture how vertical integration influences third-party sellers' long-term pricing strategies, advertising decisions, innovation efforts, or market entry and exit. The effects we observe are also shaped by the types of categories Amazon chooses to enter with its private labels. These tend to be well-established product categories, such as toilet paper or batteries, where numerous alternatives already exist. In these cases, the benefits of entry stem less from product innovation and more from economies of scale and fast delivery—advantages that Amazon leverages to offer lower prices and faster shipping, which benefit consumers. In contrast, for categories such as designer handbags or specialty cameras, product differentiation and innovation may play a more critical role than cost reductions or delivery speed, making the competitive dynamics fundamentally different.

References

Adelman, M. A. 1949. "The A & P Case: A Study in Applied Economic Theory." *The Quarterly Journal of Economics*, 63(2): 238–257.

Ailawadi, Kusum L, Koen Pauwels, and Jan-Benedict EM Steenkamp. 2008. "Private-label use and store loyalty." *Journal of marketing*, 72(6): 19–30.

Allcott, Hunt. 2013. "The welfare effects of misperceived product costs: Data and calibrations from the automobile market." *American Economic Journal: Economic Policy*, 5(3): 30–66.

Allcott, Hunt, Juan Camilo Castillo, Matthew Gentzkow, Leon Musolff, and Tobias Salz. 2024. "Sources of Market Power in Web Search: Evidence from a Field Experiment."

Allcott, Hunt, Matthew Gentzkow, and Lena Song. 2022. "Digital addiction." *American Economic Review*, 112(7): 2424–2463.

Aridor, Guy. 2022. "Market definition in the attention economy: An experimental approach."

Beknazar-Yuzbashev, George, Rafael Jiménez-Durán, Jesse McCrosky, and Mateusz Stalinski. 2022. "Toxic Content and User Engagement on Social Media: Evidence from a Field Experiment." *Working Paper*.

Bronnenberg, Bart J, Jean-Pierre Dubé, and Robert E Sanders. 2020. "Consumer misinformation and the brand premium: A private label blind taste test." *Marketing Science*, 39(2): 382–406.

Bronnenberg, Bart J, Jean-Pierre Dubé, Matthew Gentzkow, and Jesse M Shapiro. 2015. "Do pharmacists buy Bayer? Informed shoppers and the brand premium." *The Quarterly Journal of Economics*, 130(4): 1669–1726.

Bronnenberg, B, JP Dubé, and J Joo. 2022. "Millennials and the Take-Off of Craft Brands: Preference Formation in the US Beer Industry. forthcoming Marketing Science."

Chen, Nan, and Hsin-Tien Tsai. 2021. "Steering via algorithmic recommendations." *Available at SSRN 3500407*.

Crawford, Gregory S, Matteo Courthoud, Regina Seibel, and Simon Zuzek. 2022. "Amazon entry on Amazon marketplace." *CEPR Discussion Papers*, , (17531).

De Corniere, Alexandre, and Greg Taylor. 2019. "A model of biased intermediation." *The RAND Journal of Economics*, 50(4): 854–882.

Dhar, Sanjay K, and Stephen J Hoch. 1997. "Why store brand penetration varies by retailer." *Marketing science*, 16(3): 208–227.

Dinerstein, Michael, Liran Einav, Jonathan Levin, and Neel Sundaresan. 2018. "Consumer price search and platform design in internet commerce." *American Economic Review*, 108(7): 1820–1859.

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

Farronato, Chiara, Andrey Fradkin, and Alexander MacKay. 2023. "Self-Preferencing at Amazon: Evidence from Search Rankings." *AEA Papers and Proceedings*, 113: 239–43.

Farronato, Chiara, Andrey Fradkin, and Chris Karr. 2024. "Webmunk: A New Tool for Studying Online Behavior and Digital Platforms." *Working Paper*.

Gutierrez, German. 2022. "The Welfare Consequences of Regulating Amazon." Working paper.

Hagiu, Andrei, Tat-How Teh, and Julian Wright. 2022. "Should platforms be allowed to sell on their own marketplaces?" *The RAND Journal of Economics*, 53(2): 297–327.

Hoch, Stephen J, and Shumeet Banerji. 1993. "When do private labels succeed?" *MIT Sloan Management Review*, 34(4): 57.

Kotler, Philip, and Kevin Lane Keller. 2016. *Marketing Management.* Pearson Education.

Kumar, Nirmalya, et al. 2007. *Private label strategy: How to meet the store brand challenge.* Harvard Business Review Press.

Lam, H Tai. 2022. "Platform Search Design and Market Power." Working paper.

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

Levy, Ro'ee. 2021. "Social media, news consumption, and polarization: Evidence from a field experiment." *American economic review*, 111(3): 831–870.

Morozov, Ilya, and Anna Tuchman. 2024. "Where does advertising content lead you? We created a bookstore to find out." *Marketing Science*.

Newmark, Craig M. 1988. "Is Antitrust Enforcement Effective?" *Journal of Political economy*, 96(6): 1315–1328.

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

Reimers, Imke, and Joel Waldfogel. 2023. "A Framework for Detection, Measurement, and Welfare Analysis of Platform Bias." *NBER Working Paper No. 31766*.

Santos, Babur De los, Ali Hortaçsu, and Matthijs R Wildenbeest. 2012. "Testing models of consumer search using data on web browsing and purchasing behavior." *American economic review*, 102(6): 2955–2980.

Scott Morton, Fiona, and Florian Zettelmeyer. 2004. "The strategic positioning of store brands in retailer–manufacturer negotiations." *Review of industrial organization*, 24: 161–194.

Scout, Jungle. 2022. "State of the Amazon Seller 2022." Accessed: 2025-03-20.

Small, Kenneth A, and Harvey S Rosen. 1981. "Applied Welfare Economics with Discrete Choice Models." *Econometrica*, 105–130.

Teng, Xuan. 2022. "Self-preferencing, Quality Provision, and Welfare in Mobile Application Markets."

ter Braak, Anne, and Barbara Deleersnyder. 2018. "Innovation cloning: The introduction and performance of private label innovation copycats." *Journal of Retailing*, 94(3): 312–327.

Ter Braak, Anne, Marnik G Dekimpe, and Inge Geyskens. 2013. "Retailer private-label margins: The role of supplier and quality-tier differentiation." *Journal of Marketing*, 77(4): 86–103.

Train, Kenneth. 2015. "Welfare calculations in discrete choice models when anticipated and experienced attributes differ: A guide with examples." *Journal of choice modelling*, 16: 15–22.

Ursu, Raluca M. 2018. "The power of rankings: Quantifying the effect of rankings on online consumer search and purchase decisions." *Marketing Science*, 37(4): 530–552.

Waldfogel, Joel. 2024. "Amazon Self-preferencing in the Shadow of the Digital Markets Act." National Bureau of Economic Research.

Yu, Chuan. 2024. "The Welfare Effects of Sponsored Product Advertising." *Available at SSRN 4817542*.

Zhu, Feng, and Qihong Liu. 2018. "Competing with complementors: An empirical look at Amazon. com." *Strategic management journal*, 39(10): 2618–2642.