Pre-Registration and Analysis Plan for: Self-Preferencing and Consumer Welfare: Evidence from a Field Experiment

Chiara Farronato, Andrey Fradkin and Alexander MacKay*

June 13, 2023

Contents

1 The Experiment

Below, we explain the constituent parts of the study: the eligibility survey, the experimental conditions, the wish list survey, and the outtake survey.

1.1 Eligibility

We are aiming to recruit 1200 participants with a minimum of 1000. Participants are recruited via Facebook advertisements, and we will continue recruitment until we reach that number as long as recruitment costs per user remain affordable. Participants who

*Farronato: Harvard University, CEPR, NBER, cfarronato@hbs.edu. Fradkin: Boston University, fradkin@bu.edu, MacKay: Harvard University, amackay@hbs.edu. We thank James Dana, Yutao Chen, Yejia Xu, and Hannah Huihan Zhang for outstanding research assistance. We also thank Dean Eckles, Shane Greenstein, Jesse Shapiro, Raluca Ursu, and seminar and conference participants at the NUS Economics of Platforms Workshop, the 2023 ECOP conference, the Online Research Seminar on Digital Businesses, the 2023 ASSA Meeting, and Harvard Business School for valuable comments.

click on the ad are directed to an eligibility survey. The eligibility survey only allows those who meet all the criteria below to participate:

- Live in the US;
- Older than 18;
- English speaker;
- Shop on Amazon at least two to three times a month;
- Primarily use a laptop or desktop for shopping;
- Use Chrome for web browsing;
- Do not work at Amazon;
- Do not share a computer used for online shopping with other members of the household;
- Do not fail to compute 12-4 (an attention check question);
- Their Amazon order history has at least five items purchased since January 1, 2022;
- Their Amazon order history does not belong to another user observed in the sample;
- Formally consent to the study;
- Email provided in the eligibility survey matches email given when signing up on the study web browser extension (see below).

The consent form provides an overview of the study and its compensation schedule. In particular, participants are told that:

- They will receive a $15 gift card and will have a 1 in 3 chance of winning an extra $50 if they complete milestone 1, which requires them to:
  - Install the study browser extension;
  - Share their Amazon order history;
  - Complete an intake survey where they will be asked to create an Amazon wish list and add items to it.
- 8 weeks later, they will receive an additional $15 gift card and will have a 1 in 100 chance of winning another $100 if they complete milestone 2, which requires them to:
  - Not avoid being tracked through the study browser extension;
  - Update their Amazon order history;
  - Complete an outtake survey;
  - Uninstall the study browser extension.

If the survey takers consent to the study, they are asked to install the browser extension and sign up with the email previously used.

1.2 Experimental conditions

Upon installing the browser extension, participants are assigned to one of three groups using simple randomization:

- Group 1 is the *control* group, for which the extension simply monitors participants' behavior. The extension tracks browsing behaviors on Amazon.com, ensuring that confidential information is removed before storing the data. Additionally, the extension tracks URLs visited on other major e-commerce websites, but it does not collect more detailed tracking data or intervene with what the participants see on these websites.
- Group 2 is the *Amazon treatment* group. For this group, the extension hides major Amazon-owned brands[1] from the participants' browsing experience on Amazon.com. This mainly includes removing Amazon-branded products from search and product pages.[2]
- Group 3 is the *random treatment* group. This group is similar to group 2, except that instead of hiding the Amazon-branded products from a web page, the extension hides a random set of products, where the number of hidden products is equal to the number of Amazon-branded products on that web page.[3]

1.3 Milestone 1

Milestone 1 consists of two main tasks:

1. Complete intake survey.

   Study participants are asked to fill out an intake survey. The survey includes a questionnaire about their demographics and shopping habits, after which participants are guided through a shopping task.

   Participants are asked to create an Amazon wish list (`https://www.amazon.com/hz/wishlist/intro`) and share it with the researchers. Participants are then asked to pick a desired product category (e.g., batteries, phone chargers) from a high-level category (e.g., electronics). There are 6 high-level categories in

[2]The extension searches for these brands in the HTML of the product that appears as a search result on search pages or as an alternative recommended product on main product pages: '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'. The extension also searches for 'Amazon Brand' or 'Featured from our brands'. Whenever the extension finds a string match, it hides the product. Rather than showing an empty box, the web page will instantaneously re-adjust to the product removal.

[2]We will also remove the "Compare with similar items" and "Frequently bought together" modals from product pages. We remove these modals from product pages because we were unable to remove just the Amazon products from this modal directly without harming the overall look and feel of the page. Our analysis of pilot data showed that these modals were rarely clicked on.

[3]As in the Amazon treatment group, we will also remove the "Compare with similar items" and "Frequently bought together" modals from product pages.

total, selected based on the likelihood that Amazon-branded products will appear in the relevant search results. In particular, one of the high-level categories (personal care) has no Amazon-branded products appearing in the relevant search results. Thus, it serves as a placebo because all no products will be hidden in any of the treatment arms. All other high-level and corresponding lower-level categories are selected to ensure a meaningful number of Amazon-branded products in the relevant search results.

For each of the 6 categories, study participants are asked to shop in each category and add their preferred product to the previously created Amazon wish list. We induce optimal choices by promising participants that they have a 1 in 3 chance of winning $50 consisting of: 1) purchase of one of the six products in their wish list; and 2) send them an additional gift card for the difference between $50 and the price paid for the product (taxes + shipping included).

In the remainder of this document, we use the term *wish list task* to denote the task of selecting and adding products to the Amazon wish list created for this study.

2. Share Amazon order history.

   Participants are also asked to share their Amazon order history with the researchers. They do so by clicking a link in the browser extension pop-up. We will check this order history for a minimum number of orders and for potential duplication, per the eligibility criteria.

1.4 8-Week Tracking Period

After completing Milestone 1, study participants' organic browsing activity will be tracked, and their shopping options will continue to be tweaked as per the experimental groups described in Section 1.2.

1.5 Milestone 2

At 8 weeks since enrollment, study participants will be asked to complete Milestone 2, which consists of the following:

1. Update their Amazon order history

   Participants are asked to update their Amazon order history, which ensures that the researchers have access to the complete list of purchases on Amazon.com during the study period.

2. Complete Outtake Survey.

   Study participants are asked to fill out an outtake survey. The survey includes questions about their shopping experience during the experimental period, after which participants are guided to answer questions about products purchased during the experiment:

- If the participant was selected to receive a product from their wish list, we will ask them to rate the product;
- If the participant purchased items from Amazon.com during the study period, we will pick at most 3 items from their purchases and ask them to rate these products.

2 Analysis of the experiment

Below, we describe our methods for measuring the reduced-form effects of our treatments on outcomes of interest.

Our analysis will consist of several parts. The first is a reduced-form part, which is aimed at providing precise estimates of average treatment effects and heterogeneous treatment effects on outcomes of interest. This is the pre-registered part of our analysis, which will focus on analyzing the data from wish list task (see intake survey in Section 1.3).

In addition to the pre-registered analyses, we will conduct other analyses. We plan to use structural modeling to estimate demand, which would allow us to identify alternative rankings of products in search results that optimize a variety of objectives (e.g., consumer-welfare, third-party seller surplus). We also plan to engage in exploratory descriptive evidence on the search and purchasing behavior using the participants' organic browsing activity during the 8-week study period. We will also analyze the survey responses and how they vary across treatment groups and over time. The nature of these analyses and the project timeline did not permit fully specifying these analyses in advance.

Our procedure for dealing with attrition is as follows.

1. We will include all users who successfully added products from each of the six categories, finished the intake survey, and successfully installed the browser extension. Any data from individuals who did not complete these steps will not be used for our primary analysis.

2. For outcomes related to search measured through the browser extension, we will exclude users if we do not observe (through the browser) their shopping activity for at least four out of the six wish list products.

We will test for differential attrition across treatment arms based on both criteria. Substantive differences in attrition across treatment groups may cause us to deviate from the pre-analysis plan.

2.1 Reduced form analyses of the data from the wish list task.

Our **primary** outcomes of interest are the following:

- Price outcome: Price of chosen product.
- Review outcome: Log number of ratings of the chosen product + 1.
- Seller outcome: Whether the chosen product is sold by Amazon;
- Branding outcomes: Whether the product is an Amazon-branded product; Whether the product is from a major brand.
- Search outcomes: Number of products clicked on prior to choice. Whether the product selected was accessed through a sponsored search spot

In addition to these primary outcomes, we will also analyze other *secondary* outcomes of interest, including other price outcomes (per-unit price, price including taxes and shipping, logarithm of price variables), count of units in selected product (e.g., 12 paper towels), review outcomes (average star rating), shipping and fulfillment characteristics (whether the product is Prime eligible, whether the chosen product is fulfilled by Amazon,[4], whether the product qualifies for same-day or one-day shipping, whether the product qualifies for free delivery), search outcomes (number of searches and product pages seen prior to choice, time spent searching, ranking of product chosen on the first search results page it was shown), characteristics of search results (log number of reviews, star rating, share carrying a major brand, share sponsored), outtake survey responses (rating of item purchased from the wish list, value of Amazon-branded products versus other products), and purchase behaviors over the 8-week tracking period (whether the participant bought a product placed on the wish list task, number of purchases, share of purchases with an Amazon brand).

We will use three specifications (described below), and plan to present the regression with covariates as our baseline specification. All standard errors will be clustered at the participant level. We will use the method of Anderson (2008) to adjust our false discovery rate for our primary outcomes.

The specifications are described below, from simplest to most complex:

- **Specification without covariates:** The unit of observation is an individual i and category c:

$$Y_{ic} = \beta_0 + \beta_1 * Treatment_i + \gamma_c + \varepsilon_{ic} \quad (1)$$

  In the specification above, Y_{ic} represent outcomes, $Treatment_i$ is a binary variable that takes the value 1 if the individual, i, is in the Amazon treatment group (see group 2 in Section 1.2), and γ_c represent category fixed effects. We will estimate Equations 1 separately for two samples. The first sample will include individuals who are either in the Amazon treatment or in the control groups. The second sample will include individuals who are either in the Amazon treatment or in the random treatment groups. We will also estimate these specifications separately by category.

[4] `https://sell.amazon.com/fulfillment-by-amazon`

- **Specification with covariates:** We plan to augment Equation 1 as follows,

$$Y_{ic} = \beta_0 + \beta_1 * Treatment_i + \beta_2 * (X_i - \bar{X}) + \beta_{1x} * Treatment_i * (X_i - \bar{X}) + \varepsilon_{ic}, \quad (2)$$

  where X_i is a vector of individual-level covariates, and $\bar{X}$ is a vector of covariate means. The vector X_i includes the following covariates:

  - Median split on income
  - Median split on household size
  - Log number of Amazon purchases made in the last year
  - Average log price of Amazon purchases made in the last year
  - Category fixed effects[5]
  - Price of item chosen in the placebo category (personal care)[6]

- **Specification with machine learning for variance reduction** We will also use ML techniques for variance reduction. We plan to use the MLRATE estimator in Guo et al. (2021) with the entire set of covariates in the eligibility and intake surveys (prior to the choices), Amazon order history, and placebo category (see Sections 1.1 and 1.3).

Primary heterogeneous treatment effects of interest We are pre-registering the analysis of heterogeneous effects, which will use Equation 2. These analyses will assess heterogeneity in our outcomes of interest across four dimensions:

- Share of Amazon products that show up in search for the category.
- Heterogeneity based on income. Does the removal of Amazon products affect lower versus higher income household differently?
- Heterogeneity based on Amazon shopping frequency. Does the removal of Amazon products affect less versus frequent Amazon shoppers differently?
- Heterogeneity based on Amazon average price of purchases completed prior to the experiment. Does the removal of Amazon products affect participants with higher versus lower price sensitivity differently?

Generic exploration of treatment effect heterogeneity We will also use the machine learning techniques described in Chernozhukov et al. (2023) to measure the extent of treatment effect heterogeneity. This procedure will use raw values of covariates, rather than median splits.

[5]Note that, by design, all respondents will have one observation per category.

[6]The first category displayed to all users for the wish list task will be personal care. There are no Amazon owned brands in this category. Because of this, we do not expect there to be differences in outcomes across the treatment arms. We plan to use this category as a 'placebo category' that we will use to diagnose any abnormalities in the experiment. If there are no differences in treatment arms for this category, then we can use the choices made in this category as 'pre-treatment' covariates.

References

Anderson, Michael L. 2008. "Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American statistical Association*, 103(484): 1481–1495.

Chernozhukov, Victor, Mert Demirer, Esther Duflo, and Ivan Fernandez-Val. 2023. "Generic machine learning inference on heterogeneous treatment effects in randomized experiments, with an application to immunization in India." National Bureau of Economic Research.

Guo, Yongyi, Dominic Coey, Mikael Konutgan, Wenting Li, Chris Schoener, and Matt Goldman. 2021. "Machine learning for variance reduction in online experiments." *Advances in Neural Information Processing Systems*, 34: 8637–8648.