Pre-Analysis Plan: Default Effects in Web Search

Hunt Allcott, Juan Camilo Castillo, Matthew Gentzkow, Leon Musolff, and Tobias Salz

February 27, 2024

1 Introduction

In 2023, Google had a 92 percent market share in global web search (Statcounter 2024). Antitrust authorities on both sides of the Atlantic have argued that this dominance is the result of actions such as being the default search engine on Chrome and essentially all mobile devices. It is further argued these defaults are reinforced by economies of scale in data that allow better search quality. Market participants interviewed by the UK Competition and Markets Authority (2020) argue that "we are currently in a catch 22 situation, whereby demand-side remedies would not be sufficiently effective until search engines have access to the level of search data needed to improve their results." Google counterargues that "competition is one click away," economies of scale in data are limited, and Google's success is mainly driven by its high quality (e.g., Varian 2015).

This is a pre-analysis plan (PAP) for a randomized experiment designed to help answer two questions at the core of this debate. First, why is Google so dominant: default effects, users' lack of experience with alternatives, or quality differences? Second, what would be the effects of antitrust policy interventions—such as active choice screens, randomizing default settings, and requiring Google to share data with competitors—on market shares and economic efficiency?

This version of the PAP supersedes an earlier version that we submitted on January 25th, 2024. The updates reflect design changes that we made after examining results from our final pilot. Since we are submitting this updated PAP before any final data collection begins, readers can entirely ignore the earlier PAP.

Section 2 describes the experimental design. Sections 3 and 4 detail our planned analyses.

This document includes "shells" of tables and figures that we plan to include in our analysis. **All results presented in these table and figure shells are from simulated data. The simulated data are not meant to reflect our hypotheses or expectations about the results. We created the simulated data solely to illustrate the format in which we intend to present the results.**

2 Experimental Design

2.1 Overview

Figure 1 illustrates the experimental design and timeline. There are two surveys. The invitation to Survey 2 is sent the morning of the 15th day after participants complete Survey 1.

Recruitment, screening, and demographics. In late February and early March 2024, we will recruit participants from the Prolific online platform. To qualify for the study on Prolific, participants have to be a US resident at least 18 years old on a desktop computer. We instruct the Prolific system to deliver a gender-balanced sample.

Survey 1 begins with screening questions around the participant's device, web browser, and search

engine use. In one question, participants are asked to search for the term "potato" through their browser's address bar and report the search engine that they are directed to. We call that search engine the *baseline default search engine*. For Google (Bing) users, we call Bing (Google) the *alternative search engine*. The survey then asks, "What search engine do you usually use on this web browser?" We call that search engine the *search engine usually used.*

Participants can continue with Survey 1 only if they pass the following screens:

- They report that when browsing the web on their current device, they are using that web browser 100 percent of the time.
- They report that they do not frequently share that computer with other people.
- They report that their baseline default search engine is either Google or Bing.
- They report that their search engine usually used is either Google or Bing.
- Qualtrics records them as using Edge or Chrome and not using a mobile device.

Participants who pass the screening questions and consent to participate are then asked demographic questions. This is followed by a series of questions eliciting opinions about Google and Bing, including why they use the search engine usually used.

Search engine rating questions. We ask the the following four questions on Survey 1 and for some participants (the S group) on Survey 2. We refer to these four questions as the *search engine rating questions*.

***Overall**, how would you rate the quality of **Bing** relative to **Google**?*

Possible answers are *Google is a lot better*, *Google is a little better*, *they are about the same*, *Bing is a little better*, and *Bing is a lot better*. The answer order is randomly flipped, so that half of participants see "Google is a lot better" on the left and half see it on the right.

*How would you rate the quality of **Bing** relative to **Google** on the following dimensions?*

- *Relevance and ordering of search result links*
- *Features on search result pages (e.g., weather info)*
- *Relevance of ads*
- *AI chat*
- *Privacy*
- *Rewards or loyalty points*

- *Please choose "Bing is a lot better" if you are still paying attention*

Possible answers are again *Google is a lot better*, *Google is a little better*, *they are about the same*, *Bing is a little better*, and *Bing is a lot better*. The final row is an attention check. The rows are presented in random order.

*How do you feel about the number of ads on **Google**?*

*How do you feel about the number of ads on **Bing**?*

Possible answers to these two questions are *way too many*, *too many*, *right amount*, *too few*, and *way too few*.

Search Extension. Participants are then asked to install Search Extension, a browser extension developed for this study. Search Extension records the dates, times, and information identifying the source (i.e., address bar vs. the search engine website) of all searches on all general web search engines (google.com, bing.com, etc.) for all searches after installation and (via the browser's recorded search history) for the 20 days before installation. For each search after installation, it records whether the user clicked on a search result, and if so, the rank of the result. Because we observe the source of the search, we can infer the address bar default if the user searched via the address bar. Search Extension does not record any data from Incognito or InPrivate tabs on Chrome or Edge, nor does it record any data from additional browsers on which it is not installed. When the user clicks the Test Extension button, Search Extension opens a new browser window and performs a test search from the address bar, which allows it to record the address bar default search engine. Participants are asked to click the Test Extension button at the end of both Surveys 1 and 2.

Search Extension includes two intervention functionalities that we can turn on or off for use in the treatments described below. First, the *ranking degradation* functionality reverses the order of organic results on search result pages. Thus, the bottom results are moved to the top, and the top results are moved to the bottom. Second, the *ad blocking* functionality removes all ads that it detects on search result pages. Search Extension does not make users aware of these functionalities or whether they are turned on.

Compensation. Participants are paid a base payment of $25: $5 each for completing Survey 1 and Survey 2, $5 for installing Search Extension, and $10 for keeping Search Extension installed for two months after completing Survey 1.

2.2 Treatment Groups

Users are randomized into four treatment groups, one of which has further sub-treatments. We now describe the user experience in each group for the rest of the study after installing Search Extension on Survey 1.

Participants who baseline default search engine is Google are randomized into all groups, with the proportions in Figure 1 and below. Participants whose baseline default search engine is Bing are randomized into only two groups (A and S10CC), with 50-50 probability. We randomize Bing users only into these two

groups to increase power, because there are relatively few Bing users. We do not include a Control group for Bing users because in our pilots, market shares did not change over time.

Control (group "C," 4 percent of baseline Google users). After installing Search Extension, the Control group is shown information about how to change the bookmarks on their web browser, in a similar format to the treatment information that the other groups receive. In our pilots, this information about how to change bookmarks did not change search engine market share.

Active Choice (group "A," 10 percent of baseline Google users). After installing Search Extension, the Active Choice group is told that we will now show them how to change the default search engine. To avoid experimenter demand effects, the survey clearly states that "whether you change it or not is up to you." The survey then asks, "when we get to the screen where you can set your default search engine, what would you like your default to be?" We call this the person's *desired default*. The survey then shows people how to change the default search engine, asks people to copy and paste the correct settings page URL (to confirm that they were on the correct page), asks people to set the address bar default search engine to their desired default, and asks to confirm that they had done so or explain why not. We call this experience the *active choice intervention.*

Default Change (group "D," 10 percent of baseline Google users). After installing Search Extension, the Default Change group is told that we will pay them $10 to switch their default to the alternative search engine and make at least 90 percent of their searches (and at least 4 searches in total) on the alternative search engine over the next 2 days. Participants are then told that we will now show them how to change the default search engine. The survey then asks, "Would you like to accept the additional $10 to make [alternative search engine] your primary search engine for the next 2 days?" As in the Active Choice condition, the survey then shows people how to change the default search engine and asks people to copy and paste the correct settings page URL (to confirm that they were on the correct page). For participants that said they would like to accept the offer, the survey asks people to change the address bar default search engine and to confirm that they have done so or explain why not.

For these first three groups (C, A, and D), we want a second survey only to measure the default search engine and to make sure that all groups have the same number of surveys. Thus, Survey 2 for those three groups is a "placebo survey" where they are simply asked to click the Test Extension button.

Switch Bonus (group "S," 72 percent of baseline Google users). The Switch Bonus condition has the exact same Survey 1 as the Default Change group, with three exceptions. First, the payment may be different: the Switch Bonus group is divided into three subgroups, offered payment p of $1, $10, or $25, with 6, 64, and 6 percent probability among baseline Google users, respectively. Second, the length of the switch is different: participants are told that we will pay them $\$p$ to switch their default to the alternative search engine and make at least 90 percent of their searches (and at least 20 searches in total) on the alternative search engine (i.e., Bing) over the next 14 days (not just the next 2 days). Third, to avoid considerations of future inertia, participants are told that "on the second survey in 14 days, we will remind you how to switch

your default search engine."

On Survey 2, the 14-Day Switch group is asked the search engine rating questions and then receives the active choice intervention described above.

For baseline Google users, the $10 Switch Bonus group (with 64 percent of the sample) is further factorialized into a two-by-two matrix of two search results interventions implemented by Search Extension:

- **Ranking Degradation** (group "R," 50 percent of S group). Starting immediately, Search Extension turns on the ranking degradation functionality on the alternative search engine.
- **Ad Blocking** (group "A," 50 percent of S group). Starting immediately, Search Extension turns on the ad blocking functionality on the alternative search engine.

We refer to the 25 percent subset of the S group assigned to the Ranking Degradation Control and Ad Blocking Control groups as the "Search Extension Intervention Control" group (group "S10CC"). We make the $10 Switch Bonus group relatively large because we expect limited power to detect effects of these two interventions on market shares.

3 Data

Table 1 shows how we plan to present the sample sizes at each point of the recruitment funnel and experiment. The sample reported in each row is a strict subset of the row above. We plan to have a final sample of about 2,000 people completing Survey 2 and keeping Search Extension installed for two months after Survey 1.

Table 2 shows how we plan to present average covariates for the sample and for American adults.

Appendix Table A1 shows how we plan to present tests for balance on those same covariates across the randomized treatment conditions. Since baseline Google users and baseline Bing users have different treatment assignment probabilities, we present separate balance tests for those two samples. Appendix Table A2 shows how we plan to present completion rates and tests for unbalanced attrition.

4 Model-Free Results

Define B_{it} as a variable measuring participants i's use of Bing (instead of Google) in period t, where t could be a day, a week, or some other period. Define G_{it} as the analogous variable for using Google. Bing and Google market shares are the average of B_{it} and G_{it}, respectively, over users i in period t.

4.1 Survey Ratings of Google and Bing

Figure 2 shows how we plan to present responses to the question on Survey 1 about why people use Google or Bing. The figure shows the share of baseline Google and Bing users that responded that they use that

search engine for the listed reason.

Figure 3 shows how we plan to present responses to the search engine rating questions on Survey 1. The top rows present the average quality ratings on the five-point scale, coded from -2 ("Google is a lot better") to +2 ("Bing is a lot better"). The bottom row presents the average feelings about the number of ads on the five-point scale, coded from -2 ("way too few") to +2 ("way too many").

4.2 Effects on Market Shares

Figure 4 shows how we plan to present daily market shares for the treatment groups. The $10 Switch Bonus group presented here is limited to the Search Extension Intervention Control (S10CC) group. Panel (a) limits to Chrome users who search on Google at baseline, whom we refer to as "baseline Chrome-Google users." Panel (b) limits to Edge users who search on Bing at baseline, whome we refer to as "baseline Edge-Bing users." Panel (b) only includes the groups to which baseline Bing users were assigned.

4.3 Effects of Ranking Degradation and Ad Blocking

Figure 5 shows how we plan to present the effects of the Ranking Degradation and Ad Blocking treatments on Bing market shares. This includes only baseline Chrome-Google users; baseline Bing users were not assigned to the Search Extension interventions. We estimate those effects as follows.

The sample includes only the $10 Switch Bonus group. Define R_i and A_i as indicator variables for the Ranking Degradation and Ad Blocking treatments. We construct B_{it} for five periods and let t index periods for this analysis. The periods are: 0–14 days before Survey 1, 0–7 days after Survey 1, 8 days after Survey 1 to the day before Survey 2, 0–7 days after Survey 2, and >7 days after Survey 2. We estimate the effects for each period t:

$$B_{it} = \tau_t^R R_i + \tau_t^A A_i + \beta_t + \varepsilon_{it}. \tag{1}$$

We let the coefficients τ_t^R and τ_t^A vary by period t. We use robust standard errors, clustered by person. Figure 5 shows how we plan to present the τ_t^R and τ_t^A coefficients for each period.

To provide more detail on what people might learn when they switch from Google to Bing, we estimate how the $10 Switch Bonus group's responses to the search engine quality rating questions changed between Survey 1 and Survey 2. Define Y_{i1} and Y_{i2} as responses on Surveys 1 and 2. We run the following regression using all participants in the $10 Switch Bonus group who passed the attention checks on both Survey 1 and Survey 2:

$$(Y_{i2} - Y_{i1}) = \tau^R R_i + \tau^A A_i + \tau^C + \varepsilon_i. \tag{2}$$

Table 3 shows how we plan to present results. The coefficient τ^C on the constant term captures the average change in responses for the S10CC group—i.e., how responses change without either of the Search Exten-

sion interventions. The coefficients on the τ^R and τ^A capture the effects of the Ranking Degradation and Ad Removal treatments on the response changes.

References

StatCounter Global Stats. 2024. "Search Engine Market Share Worldwide." *StatCounter*. https://gs.statcounter.com/search-engine-market-share. (accessed January 2024).

Varian, Hal. 2015. "Is there a data barrier to entry?" *Google*. (accessed January 2024).

UK Competition and Markets Authority. 2019. "Online Platforms and Digital Advertising Market Study." *United Kingdom Competition and Markets Authority*. https://assets.publishing.service.gov.uk/media/5fe36a18d3bf7f08a02c87f6/Appendix_V__-__assessment_of_pro-competition_interventions_in_general_search_1.7.20.pdf. (accessed January 2024).

Table 1: **Sample Sizes [SIMULATED DATA]**

	Sample size
US Prolific users not in pilots	45,219
Started Survey 1	3,042
Passed screening questions	2,501
Consented	2,402
Finished Survey 1	2,402
Not rejected for multiple responses	2,391
Installed Search Extension	1,958
At least 10 baseline searches	1,899
Consistent baseline search engine	1,842
Finished Survey 2	1,749
Kept Search Extension 2 weeks after Survey 2	1,699
Kept Search Extension 2 months after Survey 1	1,681

Notes: This table presents sample sizes at each stage of the experiment. The sample in each row is a strict subset of the row above.

Table 2: **Sample Demographics [SIMULATED DATA]**

	(1) Analysis sample	(2) US adults
Income ($000s)	70.19	40.86
College	0.34	0.33
Male	0.40	0.49
Age	36.87	48.16
White	0.33	0.32
Baseline Google user	0.64	0.82

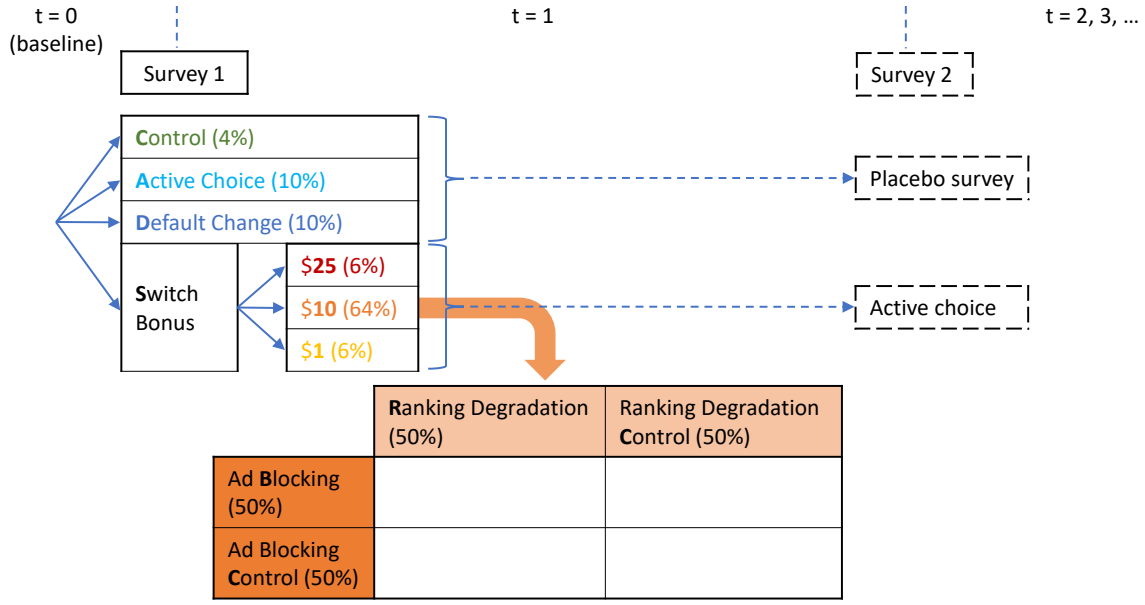
Notes: Column 1 presents average demographics for our analysis sample, and column 2 presents average demographics of American adults. All but the bottom row in column 2 use data from the 2022 American Community Survey 1-year estimates. There is no nationally representative estimate of the share of U.S. adults that use Google vs. Bing, but Statcounter (2024) reports that Google and Bing had 76 and 17 percent desktop search market shares in the U.S. in December 2023, meaning that Google has 82 percent of the Google plus Bing market.

Table 3: **Effects of Learning, Ranking Degradation, and Ad Blocking on Bing Quality Ratings [SIMULATED DATA]**

Dependent variables:	(1) Overall quality	(2) Relevant links	(3) Result page features	(4) Relevant ads	(5) AI chat	(6) Privacy	(7) Rewards	(8) Number of ads
Ranking Degradation	-0.616***	-0.872***	0.175	0.088	0.086	-0.041	0.151	-0.003
	(0.104)	(0.088)	(0.093)	(0.093)	(0.092)	(0.075)	(0.094)	(0.090)
Ad Blocking	0.465***	0.125	-0.123	0.998***	0.085	0.050	-0.075	0.120
	(0.120)	(0.102)	(0.107)	(0.108)	(0.107)	(0.087)	(0.109)	(0.104)
Constant	0.327***	0.567***	0.049	0.105	1.760***	0.982***	1.834***	0.306***
	(0.120)	(0.102)	(0.107)	(0.108)	(0.107)	(0.087)	(0.109)	(0.104)
R^2	0.042	0.076	0.004	0.088	0.001	0.001	0.003	0.001
Observations	903	903	903	903	903	903	903	903

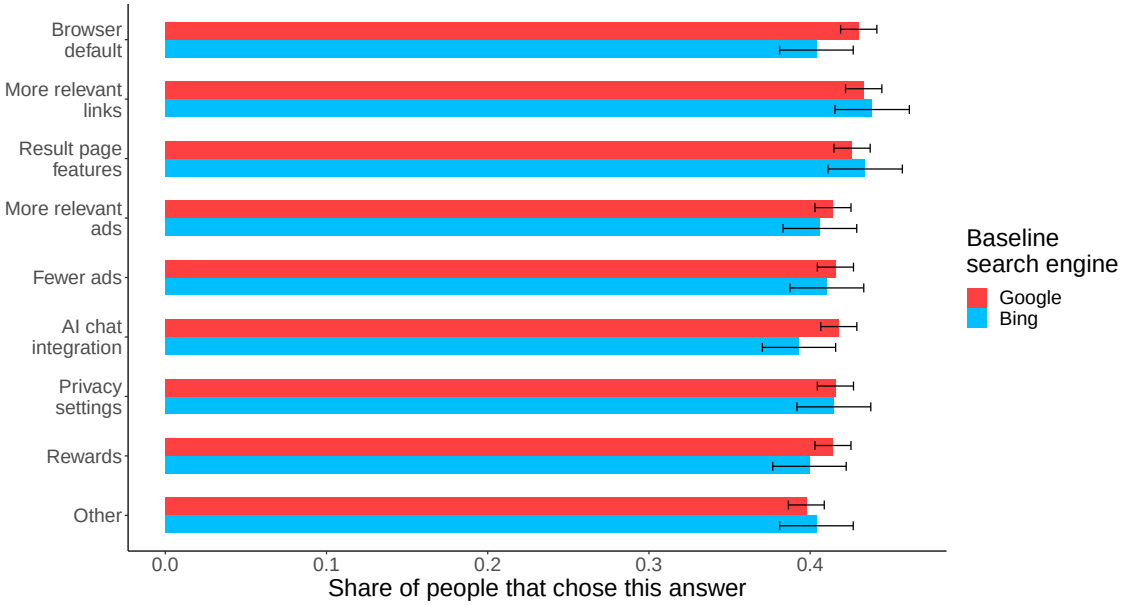
Notes: This table presents the effects of learning, ranking degradation, and ad blocking on Bing quality ratings. The sample includes all baseline Chrome-Google users in the $10 Switch Bonus group who passed the attention checks. These participants were offered $10 to switch to Bing for the 14 days between Survey 1 and Survey 2. Define Y_{i1} and Y_{i2} as responses on Surveys 1 and 2, and define R_i and A_i as indicators for the Ranking Degradation and Ad Blocking groups. The regression is $(Y_{i2} - Y_{i1}) = \tau^R R_i + \tau^A A_i + \tau^C + \varepsilon_i$. The survey questions in column 1 and columns 2–7, respectively, are "Overall, how would you rate the quality of Google relative to Bing?" and "How would you rate the quality of Google relative to Bing on the following dimensions?" Response options were "Bing is a lot better," "Bing is a little better," "They are about the same," "Google is a little better," and "Google is a lot better," coded as 2, 1, 0, -1, and -2, respectively. The survey question in column 8 is "How do you feel about the number of ads on Bing?" Response options were "way too many," "too many," "right amount," "too few," and "way too few," coded as 2, 1, 0, -1, and -2, respectively. *, **, ***: statistically significant with 90, 95, and 99 percent confidence, respectively.

Figure 1: **Experimental Design**



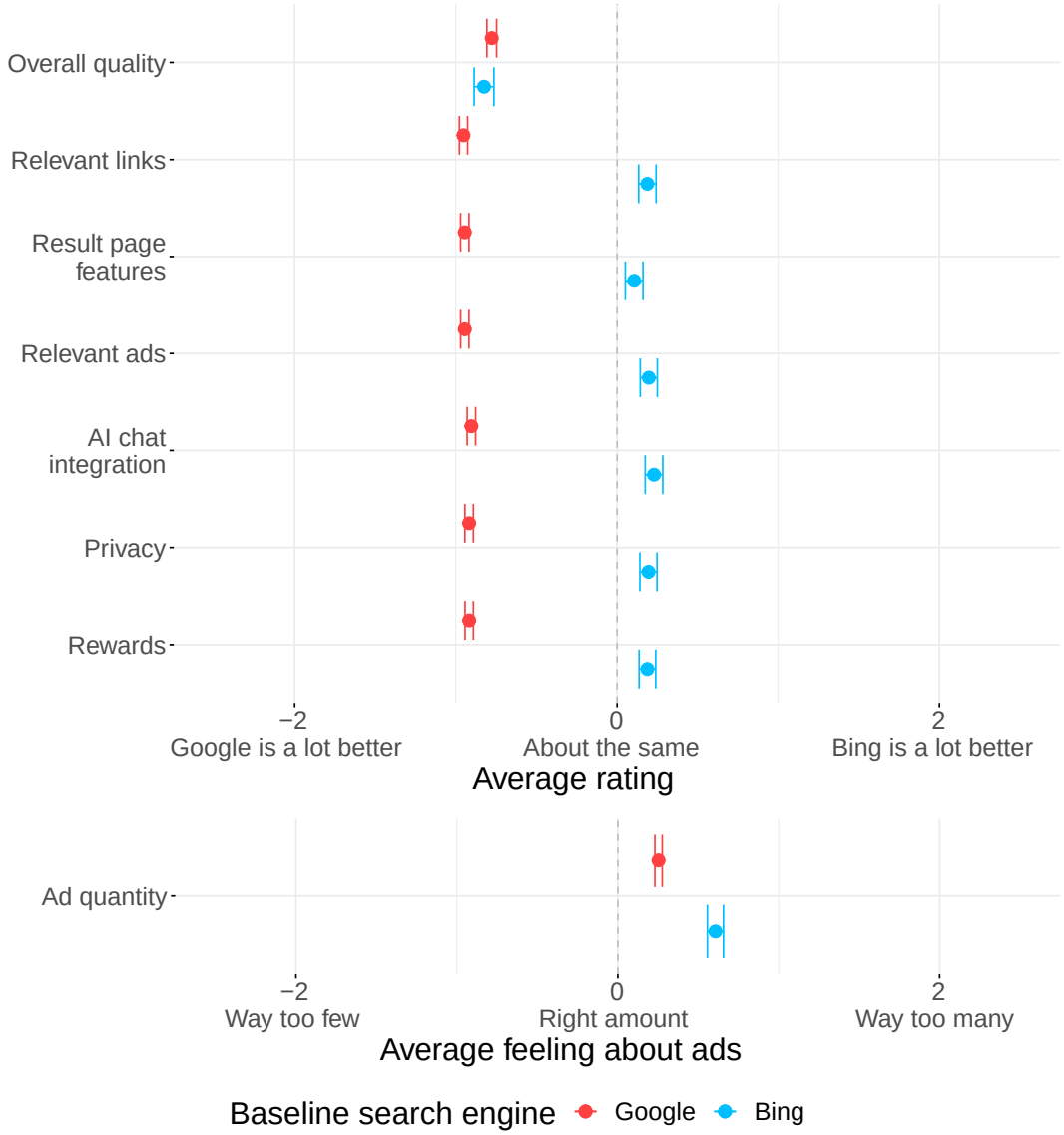
Notes: This figure presents the experimental design. As shown, the $10 Switch Bonus group was further factorialized into the two-by-two matrix of Ranking Degradation and Ad Blocking interventions.

Figure 2: **Why People Use Google or Bing [SIMULATED DATA]**



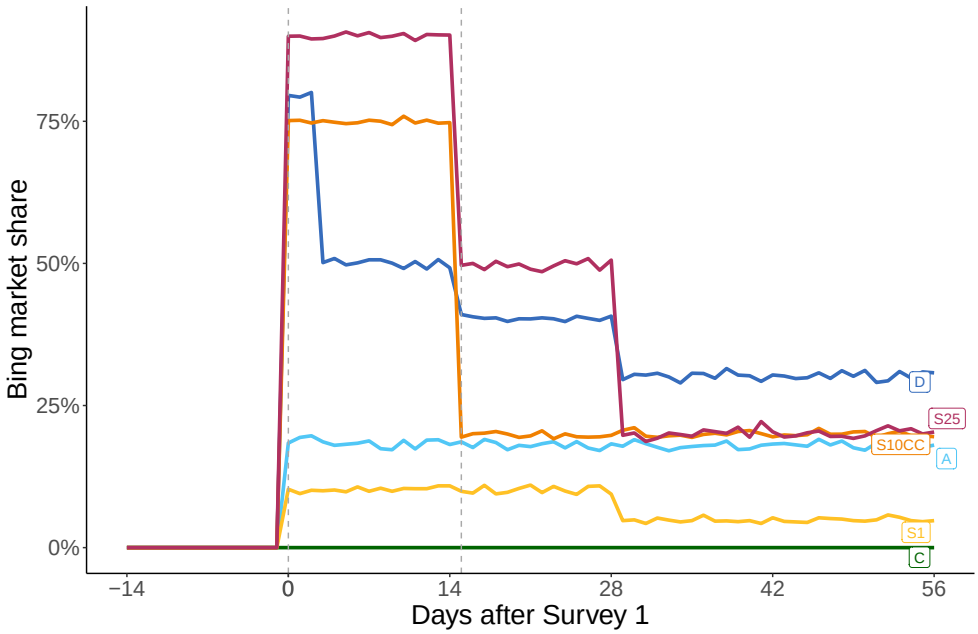
Notes: This figure presents the share of baseline Google and Bing users that chose each answer to the following question: "Why do you use [baseline search engine used] instead of [other search engine] for your searches on this web browser? Choose all that apply." Whiskers indicate 95 percent confidence intervals.

Figure 3: **Ratings of Google and Bing [SIMULATED DATA]**

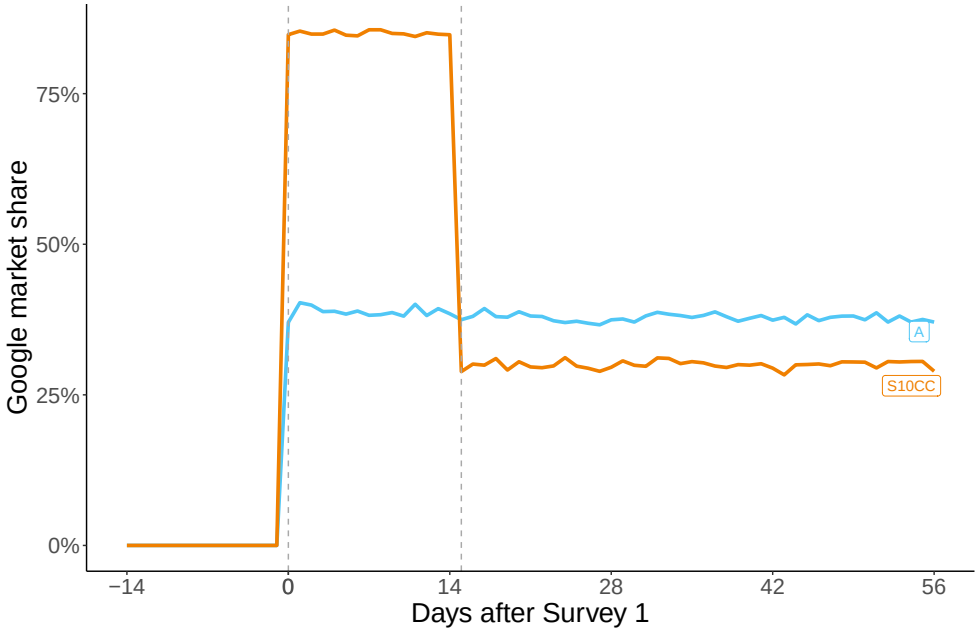


Notes: This figure presents average responses to the search engine rating questions for baseline Google and Bing users. The top rows present the average rating of Google and Bing on each reported dimension, in response to the following questions: "Overall, how would you rate the quality of Google relative to Bing?" and "How would you rate the quality of Google relative to Bing on the following dimensions?" Response options were "Bing is a lot better," "Bing is a little better," "They are about the same," "Google is a little better," and "Google is a lot better," coded as 2, 1, 0, -1, and -2, respectively. The bottom row presents the average response to the following question: "How do you feel about the number of ads on [baseline search engine used]?" Response options were "way too many," "too many," "right amount," "too few," and "way too few," coded as 2, 1, 0, -1, and -2, respectively. Whiskers indicate 95 percent confidence intervals.

Figure 4: **Effects of Active Choice, Default Change, and 14-Day Switch Treatments [SIMULATED DATA]**



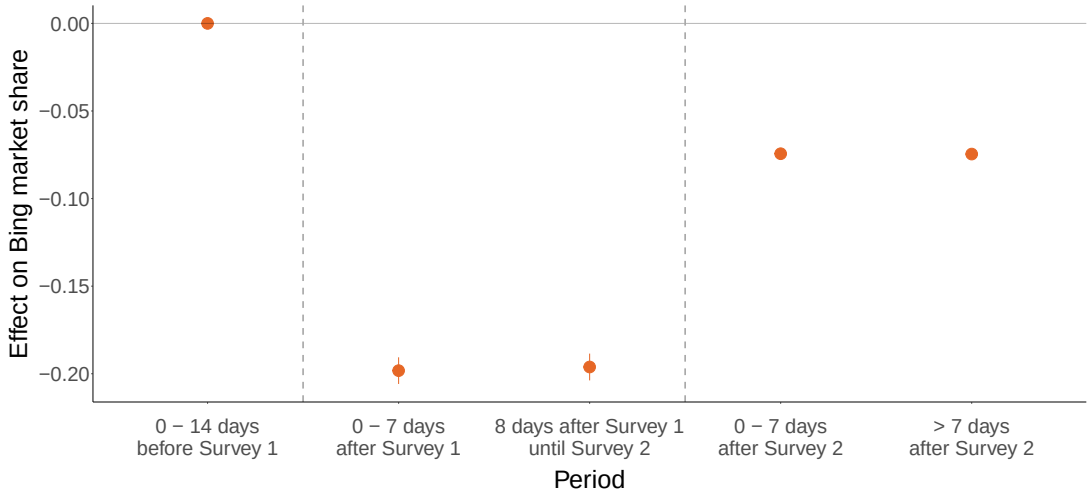
(a) **Baseline Google Users**



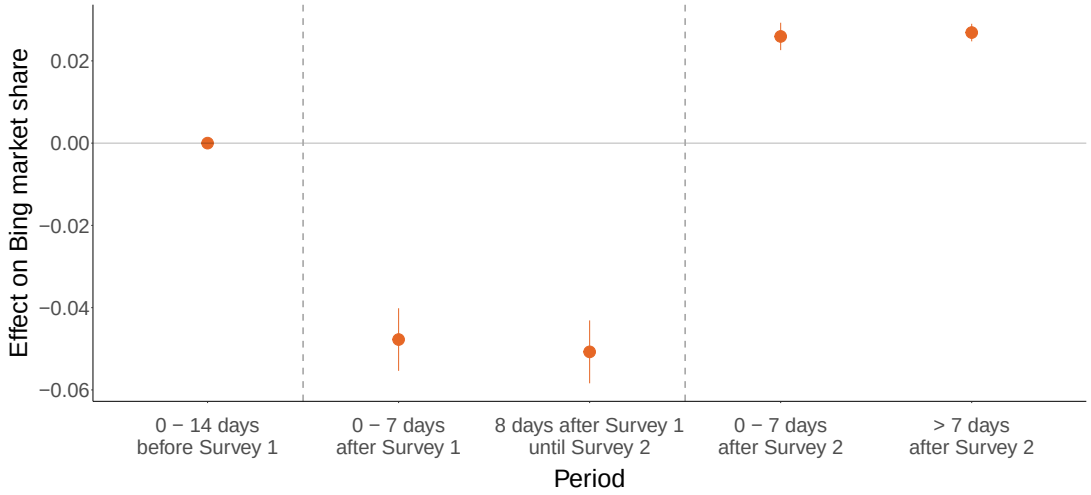
(b) **Baseline Bing Users**

Notes: This figure presents the alternative search engine market share by treatment group for each day of the experiment. Panel (a) presents results for baseline Chrome-Google users, while Panel (b) presents results for baseline Edge-Bing users. The dashed vertical lines mark the dates of the two surveys.

Figure 5: **Effects of Ranking Degradation and Ad Blocking on Market Share [SIMULATED DATA]**



(a) **Effects of Ranking Degradation**



(b) **Effects of Ad Blocking**

Notes: Panels (a) and (b) present the effects of the Ranking Degradation and Ad Blocking treatments on Bing market share by period of the experiment, estimated using equation (1). The sample includes only baseline Chrome-Google users. Whiskers indicate 95 percent confidence intervals. The dashed vertical lines mark the dates of the two surveys.

Online Appendix

Default Effects and Active Choice in the Web Search Market

Hunt Allcott, Juan Camilo Castillo, Matthew Gentzkow, Leon Musolff, and Tobias Salz

Table of Contents

A Data Appendix

Table A1: **Covariate Balance [SIMULATED DATA]**

	(1) Control	(2) Active Choice	(3) Default Change	(4) Switch Bonus ($1)	(5) Switch Bonus ($25)	(6) Switch Bonus ($10 & CC)	(7) Switch Bonus ($10 & BC)	(8) Switch Bonus ($10 & CR)	(9) Switch Bonus ($10 & BR)	(10) F-test p-value
Income ($000s)	68.78	73.48	69.93	71.11	67.75	69.85	67.46	69.86	72.79	0.97
College Degree	0.44	0.45	0.43	0.40	0.36	0.43	0.40	0.45	0.45	0.74
Male	0.43	0.49	0.53	0.44	0.53	0.46	0.52	0.49	0.57	0.10
Age	35.89	37.00	35.55	35.62	37.33	36.40	37.78	37.11	38.04	0.18
White	0.51	0.41	0.39	0.35	0.38	0.38	0.42	0.40	0.40	0.54

(a) **Baseline Google Users**

	(1) Active Choice	(2) Switch Bonus ($10 & CC)	(3) F-test p-value
Income ($000s)	69.45	70.59	0.83
College Degree	0.48	0.44	0.39
Male	0.44	0.53	0.04
Age	36.15	36.72	0.57
White	0.47	0.46	0.84

(b) **Baseline Bing Users**

Notes: Panels (a) and (b) present balance tests within the baseline Google user and baseline Bing user samples, respectively. Columns 1–9 (in Panel (a)) and 1–2 (in Panel (b)) present covariate means for each treatment group. The rightmost column presents the p-value of an F-test of a participant-level regression of that covariate on the treatment group indicators.

Table A2: **Completion Rates [SIMULATED DATA]**

	(1) Control	(2) Active Choice	(3) Default Change	(4) Switch Bonus ($1)	(5) Switch Bonus ($25)	(6) Switch Bonus ($10 & CC)	(7) Switch Bonus ($10 & BC)	(8) Switch Bonus ($10 & CR)	(9) Switch Bonus ($10 & BR)	(10) F-test p-value
Finished Survey 2	0.78	0.74	0.78	0.72	0.67	0.71	0.74	0.71	0.71	0.44
Kept Search Extension 2 weeks after Survey 2	0.76	0.71	0.77	0.72	0.62	0.70	0.72	0.69	0.69	0.25
Kept Search Extension 2 months after Survey 1	0.76	0.71	0.76	0.71	0.62	0.69	0.70	0.68	0.68	0.28

(a) **Baseline Google Users**

	(1) Active Choice	(2) Switch Bonus ($10 & CC)	(3) F-test p-value
Finished Survey 2	0.76	0.72	0.42
Kept Search Extension 2 weeks after Survey 2	0.74	0.70	0.38
Kept Search Extension 2 months after Survey 1	0.72	0.69	0.40

(b) **Baseline Bing Users**

Notes: Panels (a) and (b) present balanced attrition tests within the baseline Google user and baseline Bing user samples, respectively. Columns 1–9 (in Panel (a)) and 1–2 (in Panel (b)) present completion rates for each treatment group. The rightmost column presents the p-value of an F-test of a participant-level regression of completion indicators on the treatment group indicators.