Sources of Market Power in Web Search: Evidence from a Field Experiment

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

September 9, 2024

PRELIMINARY – DO NOT CITE

Abstract

We evaluate the economic forces that contribute to Google's large market share in web search. We develop a model of search demand in which consumer choices are influenced by switching costs, quality beliefs, and inattention, and estimate it using data from a field experiment with US desktop internet users. We find that (i) requiring Google users to make an active choice among search engines does not change market shares, implying that switching costs play a limited role; (ii) Google users who are paid to try Bing for two weeks update positively about its relative quality, with a significant share preferring to continue using it; (iii) switching defaults increases Bing market share by more than correcting quality beliefs, consistent with persistent inattention. Correcting beliefs and removing choice frictions would increase Bing's market share by 17.5 percentage points and increase consumer surplus by $2.83 per consumer-year. Policies that affect defaults lower Google's market share more than those requiring active choice. In the final part of the paper, we use Microsoft search logs to assess the impact of additional data on search result relevance. The results suggest that sharing Google's click-and-query data with Microsoft may have a limited effect on market shares.

*We are grateful to Marc Cartright, Anlei Dong, Glenn Ellison, Jingyi Guo, Minha Hwang, Aadharsh Kannan, Widad Machmouchi, Markus Mobius, Aviv Nevo, Michael Schwartz, Fiona Scott Morton, and Catherine Tucker for helpful comments. We especially thank Mike Whinston for many helpful discussions and Chiara Farronato and Andrey Fradkin for their conceptual feedback and for allowing Search Extension to use code developed for their work. We thank Chris Karr and Audacious Software for dedicated work on the Search Extension browser extension. For exceptional research assistance we thank Shotaro Beppu, Jack Cenatempo, Juan Carlos Cisneros, Grace Coogan, Sameer Nair-Desai, Shiqi Yang, and especially Wonjoon Choi. We are grateful to the Sloan Foundation for generous support. The experiment was approved by the MIT Committee on the Use of Human Subjects (Protocol # 2308001088) and was registered in the American Economic Association Registry for randomized control trials under trial AEARCTR-0012884; the preanalysis plan is available from https://www.socialscienceregistry.org/trials/12884. Disclosures: Allcott, Castillo, and Musolff have previously worked at Microsoft Research. Gentzkow has done litigation consulting for Google and has been a member of the Toulouse Network for Information Technology, a research group funded by Microsoft.
Allcott: Stanford University and NBER, email: allcott@stanford.edu. Castillo: University of Pennsylvania and NBER, email: jccast@sas.upenn.edu. Gentzkow: Stanford University and NBER, email: gentzkow@stanford.edu. Musolff: University of Pennsylvania, email: lmusolff@wharton.upenn.edu. Salz: Massachusetts Institute of Technology and NBER, email: tsalz@mit.edu.

1 Introduction

Search engines are the gateway to the internet, the starting point for 69 percent of all online activities and 44 percent of online purchases.[1] According to a bipartisan report by The U.S. House Judiciary Subcommittee on Antitrust (2020), search engines are part of "the infrastructure of the digital age" and have the potential to "pick winners and losers throughout [the] economy."[2] Due to this key position in the online ecosystem, better search engines can unlock substantial benefits for consumers and firms. Strengthening competition and improving efficiency in the web search market is therefore an important goal for policy.

Google currently holds approximately 90 percent of the global web search market. Antitrust authorities on both sides of the Atlantic allege that Google has cemented a dominant market position through anticompetitive practices, such as contracts that make Google the default search engine for many consumers, and that Google's advantages are reinforced by economies of scale in data.[3] Google counters that its success is driven by its high quality (e.g., Varian 2015), that competition is "one click away" given the ease of switching, and that increasing returns to data are small over the relevant range.

We study two questions at the core of this debate. First, why is Google's market share so large? It may be due simply to higher quality. It might stem from users' lack of exposure to alternative search engines, which leads to misperceptions about their quality. It might also result from default effects, which could strengthen Google's position directly through switching costs and inattention, and indirectly by limiting exposure to alternatives. Economies of scale could reinforce many of these advantages. Second, how would widely discussed policy interventions such as active choice screens, alternative defaults, and mandatory data sharing impact market efficiency? We answer these questions using a field experiment, observational data analysis, and counterfactual simulations based on model estimates, focusing on the competition between Google and Bing in the U.S. desktop search market.[4]

In our model, internet users make a binary choice between two search engines (Google and Bing) in each period. In an initial period, their web browser determines the initial default search engine. There are two sources of inertia in switching away from the default: a switching cost that reduces switching by marginal users, and inattention that prevents some users from making an attentive choice. Users may also

[1] See SEO Statistics (2024) and Ecommercedb (2024).

[2] Companies that operate essential online services have recently faced more regulatory scrutiny. The EU's Digital Markets Act (DMA) regards services that are "integral to the operation of digital businesses" as core platform services (CPS) and companies that operate such CPS as "Gatekeepers". Gatekeepers fall under the more stringent regulatory oversight of the DMA. Google Search is classified as a CPS and Google as a Gatekeeper.

[3] The UK Competition and Markets Authority (2020) has expressed concerns 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." Based on similar arguments, the U.S. Department of Justice sued Google in 2020 (United States v. Google LLC), alleging that Google has unlawfully exploited its monopoly power to dominate the search engine industry (Department of Justice, 2020). The EU's competition authority brought a case against Google in 2003, the outcome of which required Google to pay a $4.1 billion fine for requiring handheld manufacturers for Android devices to make Google the default search engine (European Commission, 2018). Following this decision, Google implemented a choice screen where users are prompted to select their preferred search provider when setting up a new device.

[4] We focus on Google and Bing because they had 79 and 14 percent shares of the desktop general search market in 2023 (StatCounter 2023), and no other competitor had more than 4 percent.

begin with incorrect beliefs about the quality of the other search engine. Users switch away from the default if they are paying attention and if the perceived quality improvement of the other search engine outweighs the switching cost.

We designed our experiment to provide both model-free evidence on the key forces and information to identify the model's parameters. We recruited a sample of 2,354 desktop internet users from Prolific and asked them to take a baseline survey and a follow-up survey two weeks later. Participants first answered questions about their web search preferences. Then they installed a web browser extension that records search engine queries and clicks for a period starting two weeks before the first survey and ending two months after.[5] We also conducted a qualitative exit survey at the end of the study. We limit our analysis to the 73 percent of users whose self-reported search engine usage is consistent across several measures before the first study.[6]

We randomized participants into a control group and three treatment groups. The Active Choice group was asked what search engine they would like to be their default and then was guided step-by-step through implementing their choice. The Default Change group was offered $10 to change their default search engine for two days; users that accepted were similarly guided through implementing their choice and then received no further instructions or incentives on the follow-up survey. The Switch Bonus group was offered a payment to change their default search engine for 14 days, and then were asked to make an active choice on the follow-up survey. The majority of this group was offered a $10 payment; smaller groups were offered either $1 or $25. The $10 Switch Bonus group was further subject to a factorial design based on two interventions implemented by our browser extension: (i) the Ranking Degradation condition, which decreased the relevance of Bing's results by reversing the order of organic search results on the first page, and (ii) the Ad Blocking condition, which removed most ads from Bing's search result pages.

During the two weeks before Survey 1, 96 percent of participants used Google for the majority of their searches.[7] They rated Google substantially better than Bing across several dimensions, including overall quality. The main reasons they used Google were that it presents the most relevant links (69 percent), and/or that it is the browser default (56 percent). By contrast, 4 percent of participants used Bing for the majority of their searches. They rated Bing only somewhat better or equal to Google among most dimensions. The only exceptions were rewards and AI chat integration, for which they rated Bing substantially better. They primarily used Bing because of rewards (66 percent)—consistent with the fact that Bing has a rewards program whereas Google does not—and because it was the browser default (53 percent). Most of our results focus on Google users, given their greater prevalence in our sample.

[5] We can record queries based on the browser history before the first survey but only see click data beginning with the first survey. Furthermore, our extension does not record the text of users' search terms or the URLs they click on.

[6] To be exact, we limit our analysis to users who for whom the same search engine (Bing or Google) is (i) reported on Survey 1 as the search engine they "usually use," (ii) reported on Survey 1 as the search engine used for address bar search, (iii) the search engine for more than half of searches recorded by Search Extension before installation, (iv) the address bar default recorded by Search Extension before installation. These users also need to have made at least 10 searches on this search engine in the baseline period.

[7] Relative to the universe of desktop users, Google users are overrepresented in our sample. For our counterfactual computations we will reweight the sample to match the shares in the US population.

The Active Choice intervention had almost no effect on Google users, of whom only 1.84 percent chose to switch to Bing. The small rate of switching suggests that eliminating switching costs would not meaningfully reduce Google's market share. By contrast, among Bing users, 19.7 percent chose Google when required to make an active choice, suggesting that Bing users may have larger switching costs and/or may not have made an attentive choice in the past.

In the $10 Switch Bonus group, 72 percent of users accepted our offer to switch to Bing for two weeks. Assignment to this group substantially increased users' self-reported perceptions of Bing's quality, with an increase in their overall quality rating of 0.452 standard deviations and an increase in their ratings on specific dimensions of up to 0.52 standard deviations. Of those who accepted the offer, 33 percent actively chose to keep using Bing after Survey 2. Our endline survey confirms that this reflected quality updating: 64 percent of participants who actively decided to keep using Bing reported that it was better than expected. Additionally, 59 percent reported that they had gotten accustomed to using Bing. Thus, Google users' lack of experience with Bing is a significant driver of Google's large share at baseline in our sample. For Bing users, the discrepancy between choices before and after exposure to Google is lower, indicating that Bing users are generally well-informed about Google's quality.

The Default Change intervention led to a persistent increase in Bing's market share among baseline Google users, with 81 percent of users in this group accepting the $10 to switch their baseline default search engine to Bing for two days. Among those who complied, we observe a gradual decline in Bing's market share in this group from the initial 100 percent, to 66.4 percent after one week, to 60 percent after two weeks.

Through the lens of our model, these patterns suggest that the Default Change group participants who keep using Bing do so for two reasons. First, like users in the Switch Bonus group, their valuation of Bing increases due to experience. Second, an additional group of users may continue to prefer Google but may not switch back due to persistent inattention. Our exit survey is consistent with these hypotheses: 34.9 percent of Default Change group users who keep using Bing do so because they prefer it, while 44.2 percent forgot to switch back or were too lazy to do so. The takeaway from this result is two-fold. First, defaults create a lasting mismatch between preferences and choices. Second, the friction caused by defaults persists long enough to induce learning about unobserved product quality, leading to a lasting effect by altering users' perceptions.

Our randomized price variation uncovers substantial heterogeneity in participants' willingness to accept switching search engines. Among Google users, a $1 payment to switch to Bing for two weeks raised Bing's market share to 31.5 percent, meaning that many users are close to indifferent. When instead offered $10, Bing's market shared increased to 64.4 percent. At $25, Bing's market share still only increased to 74.1 percent, meaning that many users have strong preferences.

Users noticed our Ad Blocking and Ranking Degradation changes: on Survey 2, Ranking Degradation users reported that Bing results were significantly less relevant, and the Ad Blocking group reported that Bing had fewer ads. However, market share responses were limited. Among baseline Google users who accepted the offer to try Bing, we do not detect significant changes in subsequent market shares: Ranking

Degradation reduced Bing's market share by 3.44 percent (s.e.=2.85), and Ad Blocking decreased Bing's market share by 1.28 percent (s.e.=5.2).

To estimate the parameters of our model, we use the generalized method of moments (GMM) to match predicted and observed market shares for different treatment groups at different time periods of the experiment. We find negligible switching costs, that 33.4 percent of users are never attentive, and that the remaining 66.6 percent of users are attentive with probability 82 percent in each two-week period. We also find that on average people would have to be paid $3.11 to use Bing instead of Google for two weeks. After two weeks of experience with Bing, the average required payment shrinks to $2.84. Although this difference is small when measured in terms of dollars, many users have weak preferences over search engines and hence learning is enough to shift market shares significantly. Parameters capturing the responsiveness to search result quality and ad-loads are statistically indistinguishable from zero.

For our analysis on the scale economies in data, we utilize internal search logs from Bing. To distinguish the impact of increased data on search result relevance from time-invariant confounders, we follow novel search terms that have not previously been used on Bing over time as they gather additional searches.[8] Our results suggest that the relationship between result relevance and volume of data is log-linear. Furthermore, if Bing had access to Google's data, click-through rates (a standard industry measure of result relevance) would increase from 23.50% to 24.79%. In an appendix, we extend this analysis to also account for the possibility of cross-query learning, yielding similar results.

We then consider several counterfactuals. First, we simulate the effect of shutting down all demand-side frictions—switching costs, inattention, and misperceptions about quality. In our sample, Bing's market share increases from 4.73 percent to 22.2 percent, implying that these frictions are a substantial barrier that prevents Bing from increasing its market share. Consumer surplus increases by $2.83 per consumer per year.

We then simulate the effects of active choice screens that appear when web browsers are first installed, as currently required in the European Union. To simulate this scenario, we shut down switching costs and inattention, but we continue to allow users to have misperceptions about quality. Driven by the limited effects of our Active Choice intervention, our model predicts that choice screens would increase Bing's market share by only 0.87 percentage points. This small increase underscores that, although choice screens could eliminate certain forms of frictions, they would only have a limited impact because they do nothing to avoid the larger barrier to competition that exists because of users' misperceptions about quality.

These results suggest that if regulators want to significantly impact market shares, they should focus their efforts on exposing a large fraction of consumers to Bing. We consider some interventions that could achieve this. We simulate the effects of switching search engine defaults to Bing on Chrome and Google on Edge. This switch increases Bing's market share by 42.1 percentage points. However, it decreases consumer surplus by $77.1 per consumer per year. Consumer surplus decreases because a large number of users will now use Bing even though they strongly prefer Google. We find similar results in a counterfactual where all

[8] This strategy addresses (e.g.) the existence of navigational queries ("facebook.com") which are both easy and common.

browsers are mandated to use Bing as their default. [9]

Our previous results highlight a conundrum for competition policy. While choice screens increase consumer surplus by a modest amount, they have almost no effect on market shares. Changing defaults, on the other hand, has a large effect on market shares, but only at the expense of a large decrease in consumer surplus. This raises the question of whether a policy exists that can reduce market concentration without (substantially) lowering consumer surplus. One possible approach is to mandate that Bing be set as the default on all browsers upon installation, followed by a requirement that browsers present a choice screen *after some time*. This would allow users to experience Bing before making a choice. Such policy would reduce Google's market share by 18.6 percentage points, while only decreasing consumer surplus by $0.25 cents per consumer per year—a small fraction of the harm caused by simply setting Bing as the default.

Finally, in a more speculative analysis, we simulate the effects of providing Google's search results and click data ("click-and-query" data) to Bing, using the economies of scale estimated from the Bing search logs. This data sharing has a modest effect on market shares and consumer surplus. Although there is a moderate increase in the quality of Bing, the overall effect on the market is small as our experimental results suggest that consumers respond little to the relevance of search results.

Our results have several important limitations. First, our sample is more educated, has higher income, and is more white than the population of U.S. adults and may, therefore, also not be representative on unobserved factors such as price sensitivity. Second, we focus only on desktop search in browsers. Desktop is important *per se*—representing 55 percent of the total (search + non-search) web traffic in 2023 (StatCounter 2024b)—but switching costs could be higher on mobile and on non-browser search bars that are integrated into Windows, Android, and iOS operating systems. Third, the economy of scale analysis requires stronger identifying assumptions than our experimental estimates.

Our work contributes to several related literatures. First, we contribute to the literature on competition and antitrust concerns in the web search market (Ostrovsky 2021, Duque 2022, Decarolis, Li and Paternollo 2023) and the surrounding policy discussion (Patterson, 2012; Stigler Committee on Digital Platforms, 2019; UK Competition and Markets Authority, 2020; Dinielli et al., 2023; Heidhues et al., 2023).[10] Decarolis et al. (2023) is particularly related and uses observational data to investigate the effect of remedies by European, Russian, and Turkish regulators. The EU and Russia mandated a choice screen that allowed users to pick their preferred search provider. The Turkish Competition Authority prohibited Google from granting itself privileged access on devices through contractual arrangements with device manufacturers. They find large effects from the Turkish and Russian interventions but small effects from introducing a choice screen in the EU. The results from our active choice treatment line up with the small effect size of the EU's choice screen. The results from our switch treatment suggest that an active choice intervention would be a more

[9] Although regulators are unlikely to impose this mandate, they may instead prevent Google from bidding to be the default while allowing its competitors to bid. This would likely result in Bing becoming the default on all browsers.

[10] Existing work on the search engine market has focused on the advertising side: (Varian, 2007; Edelman et al., 2007; Athey and Ellison, 2011; Blake et al., 2015). In addition there are studies assessing the value of digital services that don't charge prices to consumers (Brynjolfsson et al., 2019), of which search engines are an important example.

successful remedy if users were fully informed about product quality and that defaults prevent some users from becoming informed. In addition, our experiment more broadly investigates the sources of Google's large market share and quantifies the equilibrium effects of different remedies.

Second, we extend previous work on the competitive effect of defaults and inattention (DellaVigna and Malmendier, 2006; Carroll et al., 2009; Handel, 2013; Ericson, 2014; Ho et al., 2017; Fowlie et al., 2021; Einav et al., 2023). Our results highlight an important new role of defaults in a market setting. Even when switching costs and inattention are relatively small, defaults can matter by preventing consumers from gaining experience with an alternative whose quality they initially underestimate.

Third, we extend previous empirical work on economies of scale in search (Chiou and Tucker 2017; He et al. 2017; Azevedo et al. 2020; Schaefer and Sapi 2023), and in data more broadly (Bajari et al. 2019; Tucker 2019), by combining our estimates of the returns to scale with experimental estimates to quantify the equilibrium implications of anti-trust remedies and the resulting welfare effects for consumers.

Fourth, and more broadly, we contribute to a literature that experimentally studies digital markets. For example, prior studies have focused on consumer surplus from social media (Allcott et al., 2020), price salience Blake et al. (2021), addiction to digital services (Allcott et al., 2022), substitution pattern across online services (Aridor, 2022), and the welfare consequences of platforms when people experience fear of missing out (Bursztyn et al., 2023). To facilitate such studies, Farronato et al. (2024) introduces an open source browser extension, which has also benefitted the development of the extension for this project.

Sections 2–8, respectively, present the model, experimental design, data, model-free experimental results, structural estimation, economy of scale analyses, counterfactuals, and conclusion.

2 Model

We now present our model of demand for search engines.

2.1 Search engine choices

Consumers indexed by i make a binary choice between two search engines $j \in \{B, G\}$ in periods indexed by t. We think of this choice as determining the search engine for both direct navigation and address bar searches.[11] Consumer i's search engine choice in period t is denoted x_{it}.

The consumer's web browser is exogenous. We define $t = \underline{T} < 0$ as the period when the browser is first installed on the computer, and we define $t = 0$ as the period before the experiment begins.We refer to the original default search engine on the consumer's web browser as o and to the alternative search engine as $-o$. For example, for Chrome users, the original default is Google ($o = G$), and the alternative search engine is Bing ($-o = B$).

[11] In our experimental sample, only 6.64 percent of users search more than 10 percent of the time on a search engine that is not their current browser default.

The price agent i receives for using search engine j in period t is p_{ijt}. Normally, $p_{ijt} = 0$ but our experimental interventions will vary prices. Variables a_j^*, r_j^*, and ξ_j^* refer to j's ad load, search result relevance, and other unobserved characteristics. We define $\zeta_j^* \coloneqq \alpha a_j^* + \rho r_j^* + \xi_j^*$ as j's "quality." The term ε_{ij} is an idiosyncratic preference shifter that does not vary across time periods.

Each period's flow utility from j is

$$u_{ijt}^* = \eta p_{ijt} + \zeta_j^* + \varepsilon_{ij}, \tag{1}$$

with $\eta > 0$.

Users may be imperfectly informed about the quality of search engines. We denote the quality agent i perceives at time t by $\zeta_{ijt} \coloneqq \alpha a_{ijt} + \rho r_{ijt} + \xi_{ijt}$. Although the true quality ζ_j^* is constant, the perceived quality ζ_{ijt} depends on time because users' perceptions may change over time, as we explain below. The perceived flow utility is given by

$$u_{ijt} = \eta p_{ijt} + \zeta_{ijt} + \varepsilon_{ij}. \tag{2}$$

We assume that users have correct beliefs about their original default search engine o, so $\zeta_{iot} = \zeta_o^*$ and $u_{iot} = u_{iot}^*$. Users that have never chosen the alternative search engine $-o$, on the other hand, may be imperfectly informed about its quality. In that case, perceived quality $\zeta_{i,-o,t}$ takes a different value $\tilde{\zeta}_{-ot} \coloneqq \alpha\tilde{a}_{-o} + \rho\tilde{r}_{-o} + \tilde{\xi}_{-o},$. and perceived flow utility is

$$\tilde{u}_{ijt} = \eta p_{ijt} + \tilde{\zeta}_j + \varepsilon_{ij}. \tag{3}$$

After one period of experience with $-o$, consumers become fully informed, so their perceived quality becomes $\zeta_{i,-o,t} = \zeta_{-o}^*$.

There are two sources of inertia. First, there is a switching cost σ of getting to the choice screen and changing the default. Second, there is inattention among consumers. A fraction ϕ of users never pay attention (we say they are "permanently inattentive"), so they always stay with their original default search engine. The remaining fraction $1-\phi$ of users are probabilistically inattentive. In each period, with exogenous probability π they are "attentive" and thus consider the choice between search engines, and with probability $1-\pi$ they are "inattentive" and instead continue with their default.

If inattentive in period t, consumers stick with the search engine they used in the previous period: $x_{it} = x_{i,t-1}$. If attentive in period t, the consumer maximizes utility over an infinite horizon with per-period discount factor δ:

$$x_{it} = \arg\max_{j\in\{G,B\}} \left\{ u_{ijt} - \sigma \mathbf{1}_{j\neq x_{i,t-1}} + \delta V_{i,t+1}(j) \right\}. \tag{4}$$

where $V_{i,t+1}(j)$ is the perceived continuation value after having chosen search engine j.

We now define variables (without j subscripts) for differences between the alternative search engine and the original default search engine: $\Delta v_t \coloneqq v_{-o} - v_o$ for any variable v. Concretely, $\Delta p_{it} \coloneqq p_{i,-o,t} - p_{iot}$,

$\Delta\zeta^* \coloneqq \zeta^*_{-o} - \zeta^*_o$, $\Delta\zeta_{it} \coloneqq \zeta_{i,-o,t} - \zeta_{i,o,t}$, and $\Delta\varepsilon_i \coloneqq \varepsilon_{i,-o} - \varepsilon_{i,o}$. We also define $\Delta\tilde{\zeta} \coloneqq \tilde{\zeta}_{-o} - \zeta^*_o$, since users always perceive the utility of the original default search engine correctly. Define $s_{jt} \in [0,1]$ as j's market share in period t.

We make three assumptions, each of which is empirically realistic, that simplify this dynamic switching problem into an effectively static decision. Firstly, consumers do not perceive any uncertainty about quality ζ_{-o}, so there is no option value to exploration. Secondly, consumers weakly underestimate the quality of the alternative search engine ($\tilde{\zeta}_o \leq \zeta^*_{-o}$), so experience weakly increases $-o$'s market share. Finally, the time between browser installation ($t = \underline{T}$) and the beginning of the experiment ($t = 0$) is long enough that all consumers that are not permanently inattentive have made an attentive choice, so market shares are in steady state at $t = 0$.

We believe that the first assumption is psychologically realistic, and it is also consistent with what we observe in the Active Choice treatment. The second assumption is also consistent with our experimental estimates of experience effects. The final assumption is consistent with the lack of market share trends before our experiment and in our Control group.

Given these assumptions, and since $\Delta\varepsilon_i$ remains constant over time, the consumer's decision is static: if it is optimal to switch in the future, it is optimal to switch immediately. Thus, attentive consumers effectively make a static decision of whether to permanently choose B or G, where they account for the perceived discounted utility $u_{ixt}/(1-\delta)$ of each search engine. The problem 4 thus simplifies to

$$x_{it} = \arg\max_{j \in \{G,B\}} \left\{ \frac{u_{ijt}}{1-\delta} - \sigma \mathbf{1}_{j \neq d_t} \right\}. \tag{5}$$

At baseline—before we treat users—prices are equal to zero. Hence, the perceived discounted utility from permanently continuing with the original default search engine is

$$\frac{\zeta^*_o + \varepsilon_{io}}{1-\delta}. \tag{6}$$

The utility from permanently switching to the alternative search engine is

$$\frac{\tilde{\zeta}_{-o} + \varepsilon_{i,-o}}{1-\delta} - \sigma. \tag{7}$$

Differencing these two equations and re-arranging, an attentive consumer chooses to switch to $-o$ if

$$\Delta\tilde{\zeta} + \Delta\varepsilon_i - \sigma(1-\delta) > 0. \tag{8}$$

Therefore, the probability that a consumer that pays attention chooses $-o$ is $\mathbb{P}(\Delta\tilde{\zeta} - \sigma(1-\delta) > -\Delta\varepsilon_i)$.

A fraction ϕ of users is permanently inattentive, and thus will keep o forever. The remaining fraction $(1-\phi)$ will eventually pay attention at some point and choose the search engine that maximizes perceived utility. Thus, at time $t = 0$, $-o$'s market share is

$$s_{-o,0} = (1-\phi)S\left(\Delta\tilde{\zeta} - \sigma(1-\delta)\right). \tag{9}$$

where $S(x)$ is the cumulative density function of $-\Delta\varepsilon_i$.

Equation 9 illustrates four reasons why Google might have high steady-state market share: (i) consumers correctly perceive that Bing is worse ($\Delta\tilde{\zeta} = \Delta\zeta^* < 0$), (ii) consumers incorrectly perceive Bing is worse ($\Delta\tilde{\zeta} < \Delta\zeta^*$), (iii) Google is many users' initial default and the switching cost σ is large, or (iv) Google is many users' initial default and the fraction ϕ of permanently inattentive users is large. The social optimum is attained in a counterfactual where consumers learn the true ζ and make active choices with zero switching cost. In that scenario, the market share of $-o$ is $s_{-o,0} = S(\Delta\zeta^*)$.

In Section 5.1 we show formally how our experimental treatments identify the parameters (η, σ, π, ϕ,$\Delta\zeta^*$, $\tilde{\zeta}_{-o} - \zeta^*_{-o}$, α, and ρ) of this model and Section 5 how to estimate them.

3 Experimental Design

3.1 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. In addition, we sent an exit survey to some S and D group users.

Recruitment, screening, and demographics. In late February and early March 2024, we recruited 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: (i) They report that when browsing the web on their current device, they are using that web browser 100 percent of the time, (ii) they report that they do not frequently share that computer with other people, (iii) they report that their baseline default search engine is either Google or Bing, (iv) they report that their search engine usually used is either Google or Bing, (v) they access the Survey through a Desktop computer using Edge or Chrome. Our analysis is restricted to participants for whom the same search engine (Bing or Google) is (i) reported on Survey 1 as the search engine they "usually use," (ii) reported on Survey 1 as the search engine used for

address bar search, (iii) the search engine for more than half of searches recorded by our browser extension before installation, (iv) the address bar default recorded by our browser extension before installation. These users also need to have made at least 10 searches on this search engine in the baseline period.

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 four questions on Survey 1 and for some participants on Survey 2. We refer to these four questions as the *search engine rating questions*. We first ask people to rate Google versus Bing in terms of overall quality and then on specific dimensions. 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.

The specific quality dimensions we ask for are: *(i) Relevance and ordering of search result links, (ii) Features on search result pages (e.g., weather info), (iii) Relevance of ads, (iv) AI chat, (v) Privacy, (vi) Rewards or loyalty points, (vii) Please choose "Bing is a lot better" if you are still paying attention.* The possible answers are the same as before. The last question is an attention check. The rows are presented in random order.

Lastly, we ask participants about how they feel about the number of ads, separately for Bing and Google. 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. Figure A1 shows what users see when they first install the extension. As a standard Chrome/Edge browser extension, it is unobstrusive and it is not visible on the browser interface after the installation. 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. The main screen of the extension has a Test Extension button that, when clicked, opens a new browser window and performs a test search from the address bar, allowing us 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. These interventions occur at a split-second when the page loads and are imperceptible to the user.

Compensation. In addition to the treatment's incentive payments, 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.

3.2 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 whose 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, see below), with 50-50 probability. We randomize Bing users only into these two groups to increase power, because there are relatively few Bing users. Moreover, we expect most Bing users to be familiar with the quality of Google and therefore do not expose them to our extension interventions, which could be perceived as artificial. 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 if they switch their default to Bing and make at

least 90 percent of their searches (and at least 4 searches in total) on Bing over the next 2 days. Participants are told that we will show them how to change the default search engine. The survey then asks, "Would you like to accept the additional $10 to make Bing 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 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 Bing.
- **Ad Blocking** (group "A," 50 percent of S group). Starting immediately, Search Extension turns on the ad blocking functionality on Bing.

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

3.3 Exit Survey

We sent our exit survey at the end of the experiment to a random subset of Switch Bonus and Default Change group participants who kept the new default to understand why they did so. Eligible participants

were offered $5 for completing it. The survey experience was similar across the Switch Bonus and Default Change groups, each of which was asked "Why did you decide to keep using [Bing/Google]? Please choose all that apply." However, the available options are different.

For the Switch Bonus group, the possible answers *were (i) Before the experiment, I had always wanted to use [Bing/Google] but hadn't gotten around to it, (ii) [Bing/Google] was better than I thought it would be, (iii) I got accustomed to [NotReportedDefaultSE] or (iv) Other, with forced free form response.*

For the Default Change group, the possible answers were *(i) I wanted to keep using Bing, (ii) I forgot to change back to Google, (iii) Changing back to Google was too much effort, (iv) Other, with forced free form response.*

4 Experimental Results

4.1 Data

Table 1 shows the sample sizes at each point of the recruitment funnel. The sample reported in each row is a strict subset of the row above. Out of 5,280 participants who saw the study advertisement on Prolific, 2,543 either chose not to participate in the survey or did not pass our screening questions. Out of 2,737 who passed the screening questions, 218 were either not comfortable installing the extension or did not want to commit to staying in the study for three months. An additional 128 users were disqualified by us for attempting to complete the survey multiple times. Of the 2,390 remaining users, 664 users did not pass our requirement of (i) having conducted at least ten searches with their stated preferred search engine during the 20 days preceding the experiment and (ii) reporting that they "usually use" the search engine that our extension identifies them as using. After these restrictions, there are 1,660 users who finished Survey 2. Our final analysis sample consists of 1,461 users who stayed in the survey until the end of the experiment (two months past Survey 1).[12]

Table 2 shows demographic covariates for the sample and for U.S. adults. Our sample is nearly balanced on gender but younger, more educated, more white, and higher-earning than the U.S. adult population. It also includes a higher share of desktop Google users.

Appendix Table A1 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. For neither Google nor Bing users, are we able to reject the null hypothesis that our treatment assignment is balanced on these observed demographic variables.

Appendix Table A2 shows our tests for differential attrition across treatments at various stages of the experiment. Among Google users we reject the null of balanced attrition after Survey 2. Surprisingly, this is mostly driven by the S25 group who can earn the largest incentive payment. Among users assigned to

[12] Our results are essentially identical if we use the 1,660 users who stayed two days past Survey 2

the S25 group, 92 percent finish Survey 2 whereas completion rates in the other treatments range from 92 percent to 99.4 percent. However, importantly, we cannot reject the null of balanced deinstallation of the extension (even between S25 and other groups). Depending on the treatment, we collect extension data from 92.7 percent to 97.4 percent of participants two weeks after Survey 2. Starting two weeks after Survey 2 more users deinstall the extension.

We find that our subjects are attentive while taking our survey: our ratings survey includes an attention check that 96.1 percent of Google users and 100 percent of Bing users pass.

Table 3 shows summary statistics. Participants conduct searches on over 60 percent of days, averaging about 11 searches daily. Using pre-experiment browser history data collected by the search extension, we compare search behavior before and during the experiment (i.e. the eight weeks after Survey 1). We find no statistically significant difference in either the frequency of search days or the average number of daily searches between these periods, indicating that participants are not substituting to other browsers or devices. We further investigate search volumes by treatment groups in A3. Lastly, the table confirms that there is only a minimal amount of multi homing in our sample for both Google and Bing users.

4.2 Initial Survey Ratings of Google and Bing

Figure 2 shows the responses to the question on Survey 1 that asks participants about why they use Google or Bing. A significant share of participants attribute their usage to the browsers default option. These shares are comparable across Bing (52.6 percent) and Google users (56 percent). The share of Google users who report that they use their search engine because it serves the most relevant links is 68.7 percent while it is only 26.3 percent for Bing users, a large and statistically significant difference. The starkest difference across Google and Bing users is in terms of rewards: 65.8 percent of all Bing users say they use Bing because of rewards while only 1.94 percent of the Google users mentions rewards. This is consistent with the fact that using Bing can earn users Microsoft Rewards worth up to $10 per month, whereas Google offers no such rewards program.

Figure 3 shows how participants responded to the search engine rating questions on Survey 1. The top rows present the average quality ratings on a 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 a five-point scale, coded from -2 (*"way too few"*) to +2 (*"way too many"*).

Google users strongly prefer their search engine on almost all dimensions, including overall quality, the relevance of search results, result page features, relevance of ads, AI chat integration, and privacy. The only exception are rewards for which they express a slight preference for Bing. Meanwhile, Bing users only slightly prefer Bing on overall quality, result page features, AI chat integration, and privacy. However, Bing users strongly prefer Bing in terms of rewards. Regarding the relevance of links and ads, Bing users' responses are not statistically distinguishable from zero (indicating indifference).

4.3 Effects of Treatment Interventions

Our main focus will be the effect of treatment interventions on market shares over time, which we present in Figure 5. To compute the market share at time t, we first compute the market share for each subject separately and then average shares across subjects.[13] Panels (a) and (b) presents results for baseline Google users. The $10 Switch Bonus group presented here is limited to Search Extension Intervention Control (S10CC) participants who did not experience the Ad Blocking or Ranking Degrade interventions. Panel (c) limits to baseline Bing users. It only includes the two groups to which they were assigned, Active Choice and Switch Bonus with Search Extension Intervention Control (S10CC).

In panels (a) and (b), we see that the market share of Bing among baseline Google users was virtually zero before the experiment. This reflects the fact that most of these users exclusively used Google. We also see in panel (a) that control group users did not change their behavior during the experiment: less than 0.314 percent of overall searches in the fourteen days since Survey 1 were conducted with Bing. In panel (c) we observe that baseline Bing users did use Google to some extent before the experiment, but its market share was only about 5 percent.

Active Choice. Participants in the Active Choice treatment were asked about their preferred default search engine and then guided trough the process of implementing this default. Figure 5 shows that this intervention had almost no effect on the search engine usage of baseline Google users, increasing Bing's market share to a mere 1.86 percent. This result is consistent with the relatively small effect of the choice screen that Google implemented on Android devices in response to the European Commission's antitrust ruling in 2018 (Decarolis et al., 2023; European Commission, 2018). This suggests that most baseline Google users would choose Google even when there are no switching costs, indicating that switching costs play only a minor role in Google's large market share.

Among baseline Bing users, on the other hand, we find that the market share of Google increased to 18.2 percent in the Active Choice group. This suggests that switching costs are more effective in retaining Bing users or that more Bing users have not yet made an active choice. Thus, active choice screens may end up increasing the market share of Google.

Switch Group. We now focus on the market shares of baseline Google users in the $10 Switch Bonus group. Recall that this is the treatment in which we offered participants $10 to switch to the alternative search engine for two weeks at the end of which they made an active choice. In this section, we focus on Switch group participants who experienced Bing without any interventions from our browser extension. A large fraction of these users switched to Bing, and 51.2 percent satisfied the conditions to receive the $10. During the two-week incentive period, the market share of Bing in this treatment group was 64.4 percent, as can be seen in Figure 5.

[13] We can also define market shares by first computing the search engine chosen by every subjects (the one with the most searches) and aggregating across subject. We obtain very similar results.

Many of these users kept on using Bing after the end of the incentive period and after Survey 2, when they were asked to make an active choice (following the exact same steps as Active Choice users in Survey 1). During the week after Survey 2, the market share of Bing among those who accepted our offer and qualified for the incentive payment was 37.8 percent, and it was still 35.1 percent during the last week of the experiment.

This marks a stark departure from the Active Choice group. The resulting marketshare of 22 percent for the Switch Group in the last week of the experiment (which includes participants who declined our offer) is substantially higher than the 2.46 percent in the Active Choice group. The only difference between these two groups is that Switch Group participants were exposed to Bing for two weeks. Our results hence suggest that they corrected their beliefs about the quality of Bing upwards.

This interpretation is reinforced by our results from ratings surveys and our exit survey. Recall that our ratings survey asked participants to rate the relative quality of Bing versus Google on various dimensions. Table 5 shows the averages and standard deviations of the initial ratings in Survey 1 along with the change in the relative ratings from Survey 1 to Survey 2. A higher relative rating indicates a stronger preference for Bing. On average, Google users' rating of Bing significantly improves on all dimensions except the number of shown ads. These changes are sizable if we compare them to the baseline survey ratings and their standard deviation. For instance, the effect on the overall quality rating between Google and Bing corresponds to a third of the initial gap between Google and Bing and more than half a standard deviation. In our exit survey 64.1 percent of former Google users responded that Bing is better than expected, 59 percent that they got accustomed to it, 5.13 percent that they always wanted to use it, and 28.2 percent gave other reasons in a free form text response.[14]

Among baseline Bing users, we find that 38 percent of S10CC users satisfy the conditions to obtain our $10 payment. Among those who satisfied the conditions, the market share of Google is 98.2 percent during the two-week incentivized period. We also find that an important fraction of these users decide to stay with Google: the market share of Bing among S10CC was 36.6 percent during the week after Survey 2, and it was still 35.1 percent during the last week of the experiment. However, this market share is not statistically different from the market share among Active Choice users. Thus, the market share response suggests that Bing users appear to be well informed about the quality of Google. The results from the ratings survey are mostly consistent with this interpretation. Bing users in general update less towards Google with only the update on overall quality and AI chat being significant and the quality update being smaller compared to the standard deviation of baseline ratings than for Google users.

Default Change. Recall that participants in the Default Change treatment, which only includes baseline Google participants, were asked to change their default for two days in exchange for $10. After the initial survey there was no further communication with these subjects until they participated in the placebo survey which was sent out on day fourteen. 57.7 percent satisfied the conditions to receive the $10 payment, and

[14] These responses do not sum to 100 percent because participants could give multiple responses.

among them Bing's market share during the two-day incentivized period was 97.7 percent.

After the end of the incentive period, we see that the use of Bing trends down gradually during the first few weeks. Among participants that received the payment, Bing's market share was 66.4 percent after seven days, and 60 percent after fourteen days.[15] This gradual decline is consistent with our stochastic model of inattention; in the absence of inattention, we would have expected to see a sudden drop in usage after the end of the incentivized period. The market share eventually stabilizes at 53.2 percent after four weeks, significantly higher than the 38.2 percent who still use Bing in the Switch group at this time.

There are different explanations for this persistently high market share of Bing after the default change. First, some users may never be attentive. Indeed, we know that some participants must be persistently inattentive because the market share of Bing in the Default group is significantly higher than the 38.2 percent in the Switch group at the end of the experiment. A second possible explanation is that users have changed their perceptions of Bing during the incentive period and during the subsequent days they used Bing because of inattention. Such learning would be consistent with the significant amount of learning we observe in the Switch Bonus Group after two weeks of exposure to Bing.

The results from our exit survey suggest that both explanations play a role: while 44.2 percent of participants reply that they still use Bing because they forgot to switch back or that switching back was too much effort, 34.9 percent of participants reply that they kept using Bing because they prefer it. The significant share of people who revise their perceptions about Bing after the default change suggests another important effect of Google's default agreements, which is to prevent users from learning about about the search engine that would otherwise be the default.

Price Responsiveness. Panel (b) of Figure 5 shows the market shares of Switch Bonus users in our different price groups. We see evidence of a clear price response during the incentive period. Even a $1 payment leads to a 31.5 percent Bing market share—which is 29.64 percent higher than the Active Choice market share—suggesting that an important fraction of users have very weak preferences. A $10 payment increases Bing's market share to 64.4 percent and a $25 payment results to 74.1 percent. The small additional gains in Bing's marketshare for higher prices implies that the distribution of willingness to accept is fat-tailed. There is a drop in all Bing market shares once the incentive period ends during Survey 2. Post-Survey 2 market shares are higher for higher payments, consistent with our interpretation that users that are exposed to Bing update their beliefs about its quality upwards: in higher payment groups a larger fraction of users update their beliefs, resulting in higher market shares after the incentive period ends.

4.4 Effects of Ranking Degradation and Ad Blocking

We discuss the effect of Ranking Degradation and Ad blocking on click-through rates (CTR), search engine ratings, and market shares. Recall that Bing users were randomized into $10 Switch Bonus Group but not

[15] We do not observe any meaningful change in the time trend when these users received the placebo Survey 2.

into the sub-treatments that experienced Ad Blocking or Ranking Degrade. All results in this section are therefore based on Google users in the $10 Switch Bonus group.

Click-through rates. We first investigate the effect of these interventions on interactions with the result page. Letting w_i^R and w_i^A be dummies for the Rank Degradation and Ad Blocking group respectively, we measure these effects in the cross section of participants in the $10 Switch Bonus group, using the following specification:

$$y_i = \mu + \tau^R w_i^R + \tau^A w_i^A + \varepsilon_i. \tag{10}$$

Here, y_i captures ways in which a participant can interact with the result page, τ^R and τ^A are the coefficients of interest and μ a constant. Results are shown in Table 6.

The first specification is a manipulation check that shows that our Ranking Degradation treatment was successful. The outcome in this specification is the average original rank of clicked links. The original rank refers to the rank that a search result would have had were it not for our interventions (e.g., a result with original rank 1 would have been the top-ranked result.) We find that participants click on worse results as measured by Bing's original ranking. Absent interventions, users click on results at the beginning of the results page with an average original rank of 1.68. With Ranking Degrade, the original rank significantly increases to 5.2, which means that users click on results that they would normally encounter at the end of the result page.

Specifications (3)-(5) measure substantive effects of Ranking Degradation on various click-through rates. These click-through rates measure whether participants find the results useful and will also serve as our main measure of result relevance in the returns-to-scale analysis. Our results suggest that participants find the results after Ranking Degradation less useful, as indicated by a significantly reduced overall click-through-rate on organic links by almost 19 percent (see specification (4)), and especially of the (new) top link whose CTR falls by 63 percent (see specification (5)). When the results are degraded, participants more likely click on ads, which also leads to an overall smaller effect on the CTR if ads are considered.

Specification (2) is another manipulation check, showing that the Ad Blocking intervention works as intended: participants exposed to the intervention click on significantly fewer ads (as almost all have been removed.) Because we are able to block most ads, the effect size is even large when compared to the constant, which measures the number of ad clicks in the control group. Ad Blocking, however, does not significantly affect participants interaction with organic search results nor does it much affect overall click-through rates.

Survey ratings. Table 6 shows how these interventions affect the responses to our ratings survey questions. Define y_{i1} and y_{i2} as responses on Surveys 1 and 2:

$$(y_{i2} - y_{i1}) = \mu + \tau^R w_i^R + \tau^A w_i^A + \varepsilon_i. \tag{11}$$

The coefficient μ on the constant term captures the average change in responses for the S10CC group—i.e., how responses change without either of the Search Extension 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. Table 6 shows the results, where larger values indicate a more favorable view of Bing relative to Google.

Ranking degradation significantly reduces the positive updating towards Bing on overall quality, the relevance of results, the result page features, and the relevance of ads. Interestingly, the effect size of Ranking degradation on the relevance of results is similar to (though a little larger than) the constant. This indicates that Google users in Survey 1 expect the results of Bing to be (almost) as bad as we make them with Ranking Degradation. Ad Blocking significantly reduces the positive update about the relevance of ads and the result page features.

Market shares. Figure 6 shows the effects of the Ranking Degradation and Ad Blocking treatments on Bing market shares. As before, r_i and a_i are indicator variables for the Ranking Degradation and Ad Blocking treatments. Let y_{it} be Bing's market share with t indexing 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:

$$y_{it} = \tau_t^R w_i^R + \tau_t^A w_i^A + \beta_t + \varepsilon_{it}. \tag{12}$$

We let the coefficients τ_t^R and τ_t^A vary by period t. We use robust standard errors, clustered by person. Figure 6 shows the τ_t^R and τ_t^A coefficients for each period. We would expect the interventions to have zero effect before Survey 1 (when they are not yet active), a limited effect between Survey 1 and Survey 2 (when they are active but users are incentivized to stick with their choice) and a more stark effect after Survey 2 (on which all users make an unincentivized active choice after experiencing Bing with the respective modification). Both interventions reduce Bing's market share after the active choice in Survey 2. These effects, however, are not significant despite the fact that the extension intervention changes both participants behavior on the result page and their relative rating of Bing versus Google. A possible explanation for this result is that participants simply do not place enough weight on these attributes relative to other considerations, such as the interface and Bing rewards.

5 Model Estimation

Our main estimation is through generalized method of moments (GMM). To gain intuition for identification, we first derive transparent equations that connect structural parameters to observed market shares.

We let t index (approximately) two-week periods. Period $t = 0$ correspond to the 14 days before Survey 1, $t = 1$ to the 14 days between Survey 1 and when Survey 2 is sent, and $t = 2, 3, \ldots$ are successive 14-day periods after Survey 2.

In the model, consumers make a binary choice between Bing and Google. In the data, a small number of people multi-home or occasionally use other search engines. For estimation purposes, we say those users chose the search engine where they conducted the most searches.

Our model predicts browser-specific market shares, but our experimental treatment varies at the baseline search engine level.For instance, Bing users are paid to switch to Google in the S group, and Google users are paid to switch to Bing. To simplify the computation of market shares, we assume that people who had already switched away from their browser's default search engine (e.g., Chrome users who search on Bing) would also continue to do so if we payed them an incentive payment to stick with their currently chosen search engine. This allows us to compute market shares at the browser level, as in our model.

5.1 Identification

For identification, we derive equations that relate structural parameters to observed market shares. Figure 7 guides our discussion. It highlights which elements of our experimental data identify each parameter.

Error distribution and price response η. By comparing market shares of the Switch Bonus group at different payment offers during the incentivized period, we can identify the error distribution $S(\cdot)$ and the price response η, as shown in Figure 7. Consider a Switch Bonus user who is offered a price p to use search engine n after Survey 1. The utility from declining the Switch Bonus and staying with search engine o is

$$\zeta_o^* + \varepsilon_{io} + \delta V_{i,t=2}(o). \tag{13}$$

The utility from accepting the Switch Bonus and switching the default to $j = -o$ is

$$\eta p + \tilde{\zeta}_{-o} + \varepsilon_{i-o} + \delta V_{i,t=2}(-o). \tag{14}$$

Survey 1 tells participants that, regardless of their $t = 1$ choice, they will be guided through the choice screens on Survey 2 and consumers therefore do not expect any switching costs at time $t = 2$. Furthermore, we have assumed that consumers believe they know $\tilde{\zeta}_{-o}$ with certainty, so there is no perceived value from exploration. Therefore, $V_{i,t=2}(o) = V_{i,t=2}(-o)$ and the continuation values drop out from the comparison. Having this in mind, the consumer chooses $-o$ if

$$\eta p + \Delta\tilde{\zeta} + \Delta\varepsilon_i > 0. \tag{15}$$

The modeled market share of search engine $-o$ at time $t = 1$ is thus

$$s^{Sp}_{-o,t=1} = S\left(\eta p + \Delta\tilde{\zeta}\right). \tag{16}$$

The parameters η and $\Delta\tilde{\zeta}$ simply play the role of a scale factor and a shifter, so we can rewrite this expression as

$$s^{Sp}_{-o,t=1} = H(p), \tag{17}$$

where $H(\cdot)$ is the cumulative density function of willingness to accept,$-1/\eta \cdot (\Delta\varepsilon_i + \Delta\tilde{\zeta})$, a linear transformation of the error. As the left hand side of Equation 17 is data, different price offers directly identify values of $H(\cdot)$ at different points. As is standard in discrete choice models, we normalize the mean and variance of $\Delta\varepsilon_i$. This gives us the distribution $S(\cdot)$ of the normalized error $\Delta\varepsilon_i$ from the shape of $H(p)$.[16]

When $S(\cdot)$ is known, any two price points identify η, for instance:

$$\eta = \frac{S^{-1}(s^{S25}_{-o,t=1}) - S^{-1}(s^{S1}_{-o,t=1})}{25-1}$$

Perceived difference in quality $\Delta\tilde{\zeta}$. The perceived difference in quality is identified by market shares among Active Choice users, as depicted in Figure 7. Since these users made an active choice on Survey 1, we assume that their market shares from that point on are not driven by switching costs or inattention. Instead, they are entirely determined by perceived differences in quality.

Formally, for Active Choice users, the market share of the alternative search engine $-o$ is as in equation (9), with zero price difference Δp and without the switching cost $\sigma(1-\delta)$:

$$s^{A}_{-o,t\geq 1} = S(\Delta\tilde{\zeta}). \tag{18}$$

The relevant difference in quality is $\Delta\tilde{\zeta}$, capturing the fact that these users have not used search engine $-o$ and might thus have wrong perceptions about its quality. We invert this equation to obtain the following expression for the perceived difference in quality $\Delta\tilde{\zeta}$:

$$\Delta\tilde{\zeta} = S^{-1}(s^{A}_{-o,t\geq 1}), \tag{19}$$

where $S^{-1}(x)$ is the inverse of the cumulative density function of $-\Delta\varepsilon_i$.

Learning $\zeta^*_{-o} - \tilde{\zeta}_{-o}$. To identify learning, we compare the active choices made by Switch Bonus users after they had two weeks to learn about search engine $-o$ with the choices made by Active Choice users, who are not familiar with search engine $-o$ (see Figure 7).

When Switch Bonus subjects in the Search Extension Intervention Control group make their Survey 2 active choice, they have had time to learn ζ^*_{-o}, the true quality of search engine $-o$. They thus choose it if

$$\Delta\zeta^* + \Delta\varepsilon_i > 0. \tag{20}$$

[16] Technically, $H(\cdot)$ would be non-parametrically identified if we had a switch treatment for each price point p. In that case, $S(\cdot)$ would be non-parametrically identified up to a scale and location normalization. In practice, we have enough price points to determine that a log-normal fits our data well, according to which some participants are close to indifferent between Google and Bing whereas others require large payments to abandon Google for two weeks.

Assuming $\eta p_1 > \zeta^* - \tilde{\zeta}$, that is, that the Switch Bonus was large enough that all consumers who would choose Bing under perfect information were induced to try Bing, the market share of n is thus

$$s^{S10CC}_{-o,t\geq 2} = S(\Delta\zeta^*). \tag{21}$$

This experienced quality difference ζ^* depends on whether the consumer is in the Ranking Degrade or Ad Block groups.

Comparing this expression with the market share for Active Choice users (equation 18) and rearranging gives the following expression:

$$\zeta^*_{-o} - \tilde{\zeta}_{-o} = S^{-1}\left(s^{S10CC}_{-o,t\geq 2}\right) - S^{-1}\left(s^{A}_{-o,t\geq 1}\right). \tag{22}$$

Attention probability π. To explain how we identify π, we will first explain how the market share in the default group evolves.[17] Let $\tilde{s}^*$ be long-run share of the alternative search engine in the D group (that is the share after everybody who is not permanently attentive has made an attentive choice). Also, lets^D_{-o*} be the share of the alternative search engine directly after treatment. Let $\tilde{\pi}$ be the rate at which participants who are not permanently attentive become attentive in a given week.[18]With these definitions, the share of users who still use the alternative search engine in a given week is given by a weighted average of permanently inattentive users and users that become stochastically attentive; the users who become stochastically attentive converge away from s^D_{-o*} to $\tilde{s}^*$.

$$s^D_{-o,\text{week}=w} = \phi s^D_{-o*} + (1-\phi)\left[(1-\tilde{\pi})^w \cdot s^D_{-o*} + (1-(1-\tilde{\pi})^w)\cdot \tilde{s}^*\right]. \tag{23}$$

Intuitively, among those users who (i) are not permanently inattentive and (ii) would like to switch back, only a fraction $\tilde{\pi}$ actually switch back during a given week (which corresponds to half a period in our model). Therefore, the share of users switching away from the alternative search engine in a given week decays geometrically with a rate $\tilde{\pi}$, and we can identify that rate of decay from the following expression (which follows directly from Equation 23):

$$s^D_{-o,\text{week}=2} - s^D_{-o,\text{week}=1} = (1-\tilde{\pi})\left(s^D_{-o,\text{week}=1} - s^D_{-o,*}\right), \tag{24}$$

where s^D_{-o*} is the initial market share directly after treatment—that is, the fraction of users who switched to obtain our payment—and $s^D_{-o,\text{week}=w}$ represents the market share among the D group *at the end* of week d.Hence, we can derive the following expression for π directly as a function of market shares

[17] Our explanation assumes that learning has not yet occurred, which means that the derived expressions are only correct for week one and two. This suffices for identification of the parameters.

[18] Given that $\pi$is defined for a two-week period, this is given by $\tilde{\pi} = 1-(1-\pi)^{1/2}$.

$$\pi = 1 - \left(\frac{s^D_{-o,\text{week}=2} - s^D_{-o,\text{week}=1}}{s^D_{-o,\text{week}=1} - s^D_{-o,*}} \right)^2 . \tag{25}$$

Switching costs σ and permanent inattention ϕ. As suggested by Figure 7, the switching cost σ and inattention parameter ϕ are jointly identified from the difference between the active choice and control marketshare as well as the long run D group market-share. In particular, switching costs and inattention both create inertia: consumers are less likely to switch from the search engine they previously used, increasing the difference between the active choice and control group and the long-run default marketshare. However, as we will argue next, they are separately identified because σ affects each of these quantities symmetrically while inattention has a stronger effect on the long-run D group marketshare.

First, both types of inertia create a gap between the shares for Control users—who are subject to both forms of inertia—and for Active Choice users—who are not subject to them. The gap between those market shares is given

$$s^A_{-o,t\geq 0} - s^C_{-o,t\geq 0} = S\left(\Delta\tilde{\zeta}\right) - (1-\phi)S\left(\Delta\tilde{\zeta} - \sigma(1-\delta)\right), \tag{26}$$

and is increasing in both σ and ϕ.

Second, both types of inertia lead to higher market shares for the Default Change group after the incentive period. High switching costs and permanent inattention both imply that the geometric decay process that describes the market share over time will settle at a higher level. To state this formally, we now obtain an expression for $s^D_{-o,\infty}$, the value the market share $s^D_{-o,t}$ converges to as $t \to \infty$. Remember that $\tilde{s}^* = S\left(\Delta\tilde{\zeta} + \sigma(1-\delta)\right)$ is the hypothetical long-run marketshare of the alternative search engine among D group users, assuming (i) everybody is attentive, and (ii) nobody has yet learned its true quality ζ^*_{-o}. Similarly, let $s^* = S(\Delta\zeta^* + \sigma(1-\delta))$ be the hypothetical long-run marketshare of the alternative search engine among Dgroup users, assuming (i) everybody is attentive and (ii) users *have* learned its true quality.With these definitions, we can express the actual long-run marketshare of the alternative search engine as

$$s^D_{-o\infty} = \phi s^D_{-o*} + (1-\phi)\left[\tilde{s}^* + (1-\pi)^2(s^* - \tilde{s}^*)\right], \tag{27}$$

which is indeed increasing in ϕ and σ.

To see why $s^D_{-o\infty}$ is the long-run marketshare, suppose first that agents learn the true quality of the alternative search engine instantaneously after switching. Then the geometric decay process described by equation (23) would start at s^D_{-o*} and would converge to s^* (instead of $\tilde{s}^*$), and the market share would converge to $\phi s^D_{-o*} + (1-\phi)s^*$ as time goes to infinity. However, since we assume that it takes two weeks to learn the true quality of $-o$, there is a fraction $\left[1-(1-\tilde{\pi})^2\right](s^* - \tilde{s}^*)$ of attentive uses who switch back too early—they switched back in the first two weeks but they would not want to switch back if they had waited

and learned. After accounting for those users, we obtain Equation 27.[19]

$$s^D_{-o\infty} = \phi s^D_{-o*} + (1-\phi)\left[\tilde{s}^* + (1-\pi)^2 (s^* - \tilde{s}^*)\right]. \tag{28}$$

This market share increases with inattention, as more people will be stuck with the search engine they are paid to use. It also increases with switching costs: σ increases the market shares $\tilde{s}^*$ and s^*, as explained earlier. This expression assumes that learning is a discontinuous process: agents have misperceptions on day 13, but they are suddenly corrected on day 14. In future versions of this paper, we will allow for alternative models of learning.The parameters σ and ϕ are thus jointly identified by two expressions that relate them to market shares, equations (26) and (28). Rather than proving formally that the solution to these two expressions is unique, we will provide an intuitive argument that both parameters affect both expressions quite differently. To make the argument simpler, we will consider what happens if there are no misperceptions, in which case we get $s^D_{-o\infty} = \phi s^D_{-o*} + (1-\phi)s^*$.

First, suppose there is no switching cost. In that case, $s^A_{-o,t\geq 0} - s^C_{-o,t\geq 0} = \phi S(\Delta\zeta^*) = \phi s^A_{-o,t\geq 0}$. Then note that permanent inattention affects both expressions as follows:

$$\frac{\partial}{\partial \phi}(s^A_{-o,t\geq 0} - s^C_{-o,t\geq 0}) = s^A_{-o,t\geq 0} \qquad \text{and} \qquad \frac{\partial}{\partial \phi} s^D_{-o\infty} = s^D_{-o*} - s^*$$

Thus, it affects the gap between the A and C groups to the extent that a lot of people want to use $-o$. As we saw in Section 4, few people want to use Bing among Chrome users (less than 3%), so permanent inattention will have little effect on this difference. On the other hand, ϕ has a large impact on the long-run D share as long as (i) our treatment induces a large fraction of people to use $-o$ in return for a payment, and (ii) a large fraction would not want to use it were it not for the fact that they are being paid. Conditions (i) and (ii) are both true in the data, as we saw in Section 4: over 75% of users switch in response to our payment, while the fraction who actually want to use is around 20%.Based on these numbers, we should expect the effect of permanent inattention on $s^D_{-o\infty}$ to be on the order of ten times larger than the effect on $s^A_{-o,t\geq 0} - s^C_{-o,t\geq 0}$.

Now suppose that there is no permanent inattention. Then $s^A_{-o,t\geq 0} - s^C_{-o,t\geq 0} = S(\Delta\zeta^*) - S(\Delta\zeta^* - \sigma(1-\delta))$ and $s^D_{-o\infty} = s^*$. Switching costs affect both expressions as

$$\frac{\partial}{\partial \sigma(1-\delta)}(s^A_{-o,t\geq 0} - s^C_{-o,t\geq 0}) = S'(\Delta\zeta^* - \sigma(1-\delta)) \qquad \text{and} \qquad \frac{\partial}{\partial \sigma(1-\delta)} s^D_{-o\infty} = S'(\Delta\zeta^* + \sigma(1-\delta)).$$

Therefore, the effect on both expressions should be roughly similar as long as the density of marginal users does not change too much.

The main takeaway from this analysis is that switching costs roughly have the same impact on $s^A_{-o,t\geq 0} - s^C_{-o,t\geq 0}$ and $s^D_{-o\infty}$, whereas permanent inattention has a much larger impact on $s^D_{-o\infty}$. This enables us to identify both parameters.

[19] We do not observe choices at $t = \infty$ as our sample ends after eight weeks, so our estimation (Section 5.2) uses an adjusted expression. However, given our estimates from Section 5.3, the probability of paying attention after eight weeks is on the order of one thousandth, so the difference between these two expressions is negligible.

Quality preferences α and ρ. Comparing the \$10 Switch Bonus at time $t = 2$ across the Ranking Degradation and Ad Blocking conditions identifies preferences for the components of quality. Similar to equation 21, the market share for a Switch Bonus user in the Search Extension Intervention $I \in \{RC, CA, RA\}$ is given by

$$s^{S10I}_{-o,t\geq 2} = S\left(\Delta\zeta^I\right). \tag{29}$$

where $\Delta\zeta^I$ is the quality implied by intervention I. Note that if $I = CC$ – that is, if the user was assigned to the control group for both Ranking Degradation and Ad Blocking –, then $\Delta\zeta^I = \Delta\zeta^*$.

Recall that search engine quality is given by $\zeta_j = \alpha a_j + \rho r_j + \xi_j$. The effect of Ranking Degradation on quality difference is $\zeta^{RC}_{-o} - \zeta^{CC}_{-o} = \zeta^{RA}_{-o} - \zeta^{CA}_{-o} = \rho\left(r^{RC}_{-o} - r^{CC}_{-o}\right)$. Thus, comparing Ranking Degradation relative to its control on $t \geq 2$ market shares (equations 21 and 29) and rearranging gives

$$\rho = \frac{S^{-1}\left(s^{S10RC}_{-o,t\geq 2}\right) - S^{-1}\left(s^{S10CC}_{-o,t\geq 2}\right)}{r^{RC}_{-o} - r^{CC}_{-o}}. \tag{30}$$

A similar expression can be obtained by comparing $s^{S10RA}_{-o,t\geq 2}$ and $s^{S10CA}_{-o,t\geq 2}$. Let r_j be defined in units of click through rates. Then, the right-hand side of equation (30) is observed in the data: it is the ratio of two treatment effects. Analogous equations also hold for the Ad Blocking condition, where we define a_j such that observed Bing ad load is $a_j = 1$.

5.2 Estimation

We estimate our model parameters by the generalized method of moments (GMM). The parameters that we estimate are the perceived quality difference $\Delta\tilde{\zeta}$, learning $\zeta^*_o - \tilde{\zeta}_o$, the price response η, the amortized switching cost $\sigma(1-\delta)$, the fraction of inattentive consumers ϕ, and the per-period probability of paying attention among attentive consumers π.

To estimate the error distribution, we assume it follows a lognormal distribution that we normalize so that $S(-1) = 0$ and $S(0) = 1/2$.[20] Since $S(0) = 1/2$, the interpretation of preference shifters $\Delta\tilde{\zeta}$ and $\Delta\zeta^*$ is that they represent differences from the level of quality that would result in market shares of one half. After these two normalizations, the distribution has one remaining parameter $a \in (0, \infty)$ that captures the skewness of the error distribution. As a approaches zero, the error distribution converges to a normal (i.e., symmetric) distribution. As a approaches infinity, the distribution becomes skewed with a long tail of customers that have a strong preference in favor of Google. We estimate the quality difference $\Delta\tilde{\zeta}$ and learning $\zeta^*_o - \tilde{\zeta}_o$ separately for Chrome and Edge users. We set the rest of the parameters to be the same for all users.

The moments we use are based on the market shares that we observe in the data, which we present in Table 7. The first set of moments are simply market shares: the baseline market share s_{-o0}, the Active Choice market share $s^A_{-o,t\geq 1}$, the market shares for the Switch Bonus group during the incentivized period at

[20] The PDF of the normalized error distribution is $\exp\left(-\log(x+1)^2/(2a^2)\right) / \left(\sqrt{2\pi}a(1+x)\right)$.

different prices $s^{S1}_{-o,t=1}$, $s^{S10CC}_{-o,t=1}$, and $s^{S25}_{-o,t=1}$, and the post-survey 2 market shares of the Switch Bonus group under different interventions $s^{S10CC}_{-o,t\geq 2}$, $s^{S10RC}_{-o,t\geq 2}$, $s^{S10CA}_{-o,t\geq 2}$, and $s^{S10RA}_{-o,t\geq 2}$. Following the arguments in Section 5.1, these shares identify the error distribution, perceived differences in quality, learning, the price response, and quality preferences.

To write out these nine moment conditions, we use m to index the moments that we target. For example, m can represent baseline choices for Chrome users, S10CC choices during the incentivized period for Edge users, or Active Choice choices at time $t \geq 2$ for Chrome users. We denote by $s_m(\theta)$ the market share predicted by our model for moment m when the model parameters are θ. We also define y_{mi} to be subject i's choice corresponding to moment m. Our first nine moment conditions take the form

$$g_{mi}(\theta) = y_{mi} - s_m(\theta), \qquad \mathbb{E}[g_{mi}(\theta^*)] = 0,$$

where θ^* is the vector of true parameters.

To identify the attention probability π, we exploit the market shares of the Default Change group right after Survey 1, after one week, and after two weeks. Rather than using these four market shares directly, we exploit our expression for the identification of π (equation 24). The moment condition that we use is

$$g_{mi}(\theta) = (1-\pi)^{1/2}(y_{m,i,\text{week}=1} - y_{m,i,*}) - (y_{m,i,\text{week}=2} - y_{m,i,\text{week}=1}), \qquad \mathbb{E}[g_{mi}(\theta^*)] = 0$$

where $y_{m,i,*}$ denotes D group choices right after survey 1, and $y_{m,i,\text{week}=2}$ and $y_{m,i,\text{week}=2}$ denote D group choices at the end of weeks 1 and 2.

To identify switching costs and inattention, we need moments for $s^A_{-o,t\geq 0} - s^C_{-o,t\geq 0}$ and $s^D_{-o\infty}$. We already included moments corresponding to s_{-o0} (which is the same as $s^C_{-o,t\geq 0}$) and $s^A_{-o,t\geq 1}$, so we must include an additional moment for $s^D_{-o\infty}$. We use an empirical analogue of equation (28),

$$g_{mi}(\theta) = y_{m,i,\infty} - \phi y_{m,i,*} - (1-\phi)\left[\tilde{s}^*(\theta) + (1-\pi)^2(s^*(\theta) - \tilde{s}^*(\theta))\right], \qquad \mathbb{E}[g_{mi}(\theta^*)] = 0$$

where $y_{m,i,*}$ denotes D group choices right after survey 1, and $y_{m,i,\infty}$ denotes D group choices after a long period has occured. In practice, we do not observe choices more than two monts after the experiment starts, so our actual estimation uses an adjusted version of this moment that uses D group choices at the end of our experiment.[21] However, given our estimates from Section 5.3, the probability of not having payed attention after eight weeks is on the order of one thousandth, so the difference between these two moments are negligible.

[21] We now derive the expression for $s^D_{-o,\text{week}=8}$ that we use for estimation. After two weeks (that is, after learning) the fraction of people that would like to switch if attentive is no longer $s^D_{-o*} - \tilde{s}^*$ but $s^D_{-o*} - s^*$. The geometric decay process thus resembles equation 23, but it goes from s^D_{-o*} to s^* (and not from s^D_{-o*} to $\tilde{s}^*$). After accounting for the fraction $\left[1-(1-\tilde{\pi})^2\right](s^* - \tilde{s}^*)$ of users that switched back too early, we obtain the following expression for the market shares after week 2:

$$s^D_{-o,\text{week}=w>2} = \phi s^D_{-o*} + (1-\phi)\left[s^* + \left(1-(1-\tilde{\pi})^2\right)(s^* - \tilde{s}^*) + (1-\tilde{\pi})^w\left(s^D_{-o*} - s^*\right)\right].$$

There are two important issues we must deal with before computing these moment conditions and the GMM objective function. First, given the nature of our experiment, we don't observe all moments for every participant. For a participant that was randomized into S10CC, for instance, we observe the moments corresponding to S10CC choices at times $t = 1$ and $t \geq 2$, but we don't observe any of the Default Change or Active Choice choices. Second, we used different randomization probabilities for original Google and Bing users, so unconditional means would overweight Bing users in most of our treatments. Appendix B explains how we address these two issues by using different weights for different participants.

Our main demand estimates are given by

$$\hat{\theta} = \text{argmin}_{\theta} G(\theta)' \Omega G(\theta)$$

where θ is the full vector of the parameters that we estimate, $G_m(\theta)$ is the average of $g_{im}(\theta)$ across users, and Ω is a weighting matrix. We use two-step GMM, where we first set $\Omega = Cov(G(\theta))^{-1}$ for some arbitrary θ, and then, in our second step, we set Ω to be the optimal weighting matrix given the initial set of estimates.

5.3 Demand Parameter Estimates

Table 8 presents the demand parameter estimates. We find that the parameter a for the error distribution is quite high—that is, the error distribution is very skewed. This means that there is a large density of users that are close to indifferent between both search engines as well as a fat tail of users with strong preferences in favor of Google. On the other hand, there are essentially no users with strong preferences in favor of Bing. We estimate such a skewed distribution because a small price increase from \$0 to \$1 in the S group had a large impact on market shares (from 2.58 percent to 33.7 percent) whereas a large price increase from \$10 to \$25 only increased market shares from 67.7 percent to 75 percent.

We now discuss two important implications of a skewed error distribution. First, a given change in price or quality has a large impact on market shares when the marginal users are those that are close to indifferent between both search engines (when the market share of Google is high), but it only has a modest impact when the marginal users belong to the fat tail of users with strong preferences in favor of Google (when the market share of Google is high). This feature drives many of our findings below. Second, the combination of a fat tail and permanent inattention implies that there are large consumer surplus losses due to agents with strong preferences that nevertheless use an undesired search engine because they are inattentive. For that reason, when computing consumer surplus, we censor the error distribution at \$25 to avoid such large changes in surplus.

We now discuss parameters for Chrome users as well as the rest of the parameters for all users—which are mainly identified from data on Chrome users because they account for most participants in our sample. We use our estimate of the price response η to interpret all other parameters in units of dollars per two-week period. For Chrome users who have not used Bing, we observe a negative Bing preference shifter $\Delta\tilde{\zeta}$ (equivalent to a payment of \$3.11 per two-week period), consistent with the low market share of Bing in the

data. After learning Bing's quality, its perceived utility increases by \$0.27 per two-week period. Learning about Bing affects those users who are close to indifferent between both search engines, where there is a large density of users. Thus, as we note above, a small change in the perception of quality is enough to generate a large change in market shares.

The estimated switching costs $\sigma(1-\delta)$ are small (0.4 cents), given the small effect on market shares of the Active Choice intervention—which, as we explain in Section 5.1, are mainly driven by switching costs rather than inattention. One the other hand, we find that inattention plays an important role. 33.4 percent of users are permanently inattentive. The per-period attention probability π =0.824 means that during every two-week period most users that are not permanently inattentive make an attentive choice, whereas less than one fifth stay with their previous search engine. Our estimates for the preferences for quality indicate that our ad block intervention decreased utility as much as a price decrease of \$0.131, and the ranking degradation resulted in a decrease in utility of \$0.109. Neither of these two effects is significant.

For Edge users, we find a negative Google preference shifter $\Delta\tilde{\zeta}$ that is equivalent to a payment of \$7.61 per two-week period, consistent with the high share of Bing among Edge users. Based on our estimate of $\zeta^* - \tilde{\zeta}$, we also find that there is learning about Google, which is equivalent to a payment of \$1.15 per two-week period. This is larger than the dollar value of learning we find among Chrome users. To understand why, recall that the marginal consumer among Edge users belongs to the fat tail of users that have strong preferences for Google; hence, a large price difference is necessary to justify the change in shares from 36.8 percent to 38.9 percent due to learning.

6 Economies of Scale in Data

We suspect that there are economies of scale in providing users with well-ranked search results. In particular, as a search engine responds to more searches for a search term it learns which results users prefer and can subsequently rank these results higher. For the analysis in this section we use data on search terms and clicks from Bing.

We are interested in estimating the causal effect of the number of past searches of a given search term on the current relevance of results for that term. As is standard industry practice, we will use click-related outcomes as our measure of the relevance of the links presented in a search results page. In particular, our measure of relevance is the probability that a user clicks on the top-ranked result. Henceforth, we will refer to search terms as queries and instances where a user enters a query as searches.

Google argues that the effect of additional search data on improving result rankings diminishes quickly, implying that increased scale has little effect on the relevance of results (e.g. Varian 2015). This argument is less convincing if there are many *tail* queries (i.e., queries with few searches) for which additional click and query data may still be valuable. Indeed, examining all searches made on Bing in 2021, we find that the distribution of queries exhibits a long tail: Figure 8 shows what fraction of searches accrues to queries with different occurrence rate. More than 38.7 percent of searches are for rare queries that are searched less

than 100 times. For this reason and for reasons of identification that we explain below, we will focus on new queries.

For our analysis, we randomly sample 43,991 new search terms, i.e., search terms for which Bing has no search record between January 2021 and January 2022. For these search terms, we record each search in the subsequent year (i.e., between January 2022 and January 2023).[22]

6.1 Empirical Strategy

We are interested in recovering the relationship between result relevance r_{qt} for query q at time t and the number of searches n_{qt}:

$$r_{qt} = f(n_{qt}; \theta) + \varepsilon_{qt}$$

In estimating this relationship we need to address two potential confounders. First, there could be persistent differences in the difficulty of serving results across queries. For instance, navigational queries ("facebook.com") are both very common and easy to answer *at any scale*. Thus, a simple cross-sectional comparison may bias results in favor of finding returns to scale. To control for such persistent differences in the difficulty of serving results across queries, we use previously unseen queries over time as they gather their first searches and control for query fixed effects (as in He et al. (2017)). However, this strategy does not address time-varying confounders.

To deal with such time-varying confounders we rely on a key identifying assumption, which seems plausible in this context: any *causal* effect of more search data on clicks has to be due to the results that the search engine serves.[23] Before we describe how we formalize this identification strategy, we conduct an event study to motivate our approach.

Figure 9 shows click-through rates against the number of searches a given query has had since its first or second ranking change. Updates to the ranking lead to an increase in click-through-rates, which indicates that over time the search engine learns to serve more useful results. We also observe a general downward trend in click-through rates that is unrelated to the results the search engine serves. This may be due to compositional changes in the types of users over the lifetime of a query and could confound the returns to scale estimates. For instance, suppose a query is concerned with a news event. For users that arrive early in the query's lifetime the event is, well, news – and hence they are likely to click on results. By contrast, users that arrive later may already know of the event, and just want to verify the date on which it occurred – something that can be accomplished without the need to click. If we did not account for time-varying

[22] The resulting panel dataset at the (query, search)-level records (i) the date of each search, (ii) various click-related outcome measures (was there any click, was the top result clicked, was there any click from which the user did not immediately return) for each search, and (iii) an identifier for the URL that was top-ranked for these searches. We report summary statistics for this dataset in Table A4 in Appendix D.1. Optimally, we would like to subset to queries which Bing has never seen, but we are limited by Bing's retention period of 24 months.

[23] This assumption would be violated, for example, if the search engine also uses data to improve the interface of the search results. However, to cause concern such interface adjustments would have to be query specific, which seems less plausible.

confounders, this difference in user types over a query's lifetime could bias our estimation.

Our identifying assumption — that the causal effect of more data can only operate through the search ranking — allows us to apply the front door criterion (Pearl and Mackenzie, 2018; Imbens, 2020; Bellemare et al., 2024) to purge confounding variation in click-through rates, such as the observed secular trends. The directed acyclical graph (DAG) in Figure 10 illustrates our strategy. In this graph, n_t refers to the number of prior searches for a given query, ℓ_t to the links that the search engine serves, r_t to the click-through rate (i.e., our measure of result relevance), and u_t is a confounder that affects both the number of prior searches and the click-through rate, such as a changing composition of users over the lifetime of a query. The causal effect of additional prior searches (i.e., additional data) on the links served is given by γ and the causal effect of the links served on click-through rate is δ.

Regressing r_t directly on n_t would be biased by the confounding variation in u_t. The key insight from this graph is that by regressing r_t on ℓ_t after conditioning on n_t, one can isolate the causal variation in data that leads to changes in click-through rates. The intuition is that we stack many event studies like Figure 9: like in the event-study, we isolate the effect on CTR that comes with a change in the ranking. We then regress the predicted value $\hat{r}_t = \hat{\delta} \cdot \ell_t$ on n_t and obtain the causal effect as the product $\gamma \cdot \delta$.[24] In Appendix D.3, we explain how we combine this identification strategy with controls for heterogeneity across queries and how we address the fact that ℓ_t is not a scalar in our setting.

6.2 Estimation

We now move on to estimating the relationship between search result relevance (as proxied by the click-through rate) and the number of prior searches for a given query q on date t. While we investigate the relationship non-parametrically in Appendix D.2, for simplicity's sake we specify a functional form for the main analysis. We choose a simple functional form that allows for different relationships between data and CTR, including linear ($\theta = 0$), logarithmic ($\theta = 1$), and more concave than logarithmic ($\theta > 1$):

$$\hat{r}_{qt} = \alpha_q + \beta \frac{(n_{qt})^{1-\theta} - 1}{1-\theta} + \varepsilon_{qt} \tag{31}$$

This specification requires estimation of a non-linear model with high-dimensional fixed effects for which we develop a simple procedure in Appendix D.4. Since our equilibrium model does not distinguish between different queries, we choose an overall intercept α by matching the average click-through rate predicted by our model to that in our experimental data.[25]

We exhibit the resulting estimates in Table 9. We can strongly reject the null that data does not matter

[24] Unlike in an instrumental variables setting, this product of two reduced-form causal effects is the object of interest. In the case of instrumental variables, one is interested in estimating the causal effect of a mediator X on Y and exploits the exclusion restriction that the effect of the instrument can only operate through the mediator. By contrast, we are interested in the total effect of n_t on c_t, and we exploit the exclusion restrictions that the relationships between n_t and r_t and r_t and c_t are each unconfounded.

[25] This average is computed by integrating over the complete query frequency distribution and weighting each query by its number of searches. While we would optimally use lifetime query frequencies (i.e., query q has seen n views since inception), we actually observe query frequencies over a one year period.

($\beta = 0$). As to the shape of the returns from data, the estimates strongly point towards a log-linear relationship between CTR and number of searches. Thus, doubling the amount of searches will lead to a fixed increase in CTR, no matter the starting point. Our estimates imply that, if Bing had access to Google's data, CTR would increase from 23.50% to 24.79%.

We exhibit the fit of the model in Figure 11, where we compare the model predictions to a binscatter plot of the data. The model fits well across many orders of magnitudes. Furthermore, our estimates are not too different from those in prior studies. He et al. (2017, Fig. 4 and 5) find that going from 300 to 1,900 searches yields a CTR improvement of about 2pp for both (anonymous) search providers they have data for.[26] Our estimates imply an improvement of 1.49pp, coming in slightly below their estimates. However, Schaefer and Sapi (2023, Fig 5(c)) find that going from 1 to 5,000 searches yields a CTR improvement of 3-5pp. Our estimates imply a slightly larger improvement of 6.08pp.

We can use our fitted model to calculate counterfactual average click-through rates on Bing if its market share were to be multiplied by λ: the new click-through rate r' after such an increase in market share is given as a function of the old click-through rate r by

$$r' = \alpha - \frac{\beta}{1-\theta} + \lambda^{1-\theta} \cdot \left(r - \alpha + \frac{\beta}{1-\theta} \right).$$

This expression will be used in the computation of long-run counterfactuals below.

We close this section by noting several limitations in our approach. First, our model is estimated exclusively off new queries. This could be problematic if the effect of data differs between new queries and other types of queries. Second, extending our model to queries with many views requires extrapolation as new queries typically do not reach such high levels of popularity. Thus, if our functional form is not flexible enough, any specification error could be amplified by extrapolation outside of the support of our data. Still, the fit of our model ameliorates this concern. Furthermore, a priori, it is most plausible that the returns from additional data are strongest for queries that are not frequently searched.

7 Counterfactuals

Using our model estimates, we now evaluate the effect of proposed anti-trust policy interventions. With this goal, we consider several counterfactuals. As a benchmark, we compare them to a *Status Quo* scenario, in which demand behaves just as in the control group, and market shares are given by $s_{-o} = (1 - \phi) S\left(\Delta\tilde{\zeta} - \sigma(1-\delta)\right)$.

In our first counterfactual, *No Frictions*, all demand frictions are eliminated: users are perfectly attentive, they are informed about the true quality of search engines, and there is no switching cost. Therefore, market shares are given by $s_{-o} = S(\Delta\zeta^*)$: the only determinant of demand is the difference in the true quality of

[26] We convert their results on the effect of the number of clicks to estimates of the effect of the number of searches by dividing by their average CTR of 0.5.

search engines. Although this counterfactual is not likely to occur in practice, it is useful as a benchmark in which, when not accounting for economies of scale in data, the market achieves the social optimum.

The second counterfactual represents the introduction of a *Choice Screen* like the one introduced by Android in Europe after the European Commission's 2018 decision. At the time of installing their browser, users are presented with a screen that allows them to choose their default search engine. This eliminates frictions due to inattention and switching costs, but users may still be imperfectly informed about search engine quality. Therefore, demand behaves just as in the Active Choice group, with $s_{-o} = S(\Delta\tilde{\zeta})$.

In the third counterfactual, Correct Perceptions, all users are perfectly informed about the true quality of both search engines. However, unlike the *No Frictions* counterfactual, users are still subject to switching costs and there is inattention. The market share is thus given by $s_{-o} = (1-\phi)S(\Delta\zeta^* - \sigma(1-\delta))$.

We then consider two counterfactuals that consider the effects of changing default search engines. In *Default Change,* the original default search engine on Chrome becomes Bing and the original search engine on Edge becomes Google. Thus, switching costs operate in the opposite direction as in the status quo, and the market share is given by $s_{-o} = \phi + (1-\phi)S(\Delta\zeta^* + \sigma(1-\delta))$.In *Bing Default*, the default search engine upon installing a new browser is Bing on all browsers. This approximates proposed remedies that would not allow Google to bid for default positions. Presumably, Bing would be able to outbid its competitors, so it would become the default on all browsers. For Edge users, the market will operate as in the status quo, whereas it would operate as in Default Change for Chrome users.

We also consider a *Bing Default + Delayed Choice Screen* counterfactual that aims to achieve the benefits of changing defaults while mitigating consumer surplus losses. Bing becomes the default search engine on all platforms, as in the Bing Default counterfactual. Additionally, a choice screen is presented to consumers to avoid welfare losses due to inattention. However, showing a choice screen upon installation would render the default change ineffective, as users would choose their preferred search engine before gaining any experience with the default. To address this, we consider a choice screen that is shown two weeks after installation, giving users time to learn the true quality of Bing. During the first two weeks, market shares are the same as in the Bing Default counterfactual, but afterward, they shift to those in the No Frictions counterfactual for Chrome users and to those in the Choice Screen counterfactual for Edge users. To compute consumer surplus, we give weight $2/52$ to the consumer surplus before the choice screen and weight $50/52$ to the consumer surplus after the choice screen. This approximates what would happen if consumers reinstalled their browser every year, with the default resetting to Bing.

We then consider a *Bing Payments* counterfactual in which Bing pays users \$10 (on top of their existing Bing rewards program) to induce them to use Bing more. This allows us to determine the extent to which payments can level the playing field in favor of Bing. Shares are therefore given by $s_{-o} = (1-\phi)S\left(\Delta\tilde{\zeta} - \sigma(1-\delta) + \eta p_B\right)$, where the payment p_B is positive for Chrome users but negative for Edge users.

In our final two counterfactual, regulators mandate data sharing between search engines, as proposed by certain antitrust authorities. This allows both search engines to exploit the data from all users, so the

relevance of the results presented by both search engines is what it would be if each search engine was a monopolist and was able to observe data from all users in the market. As data sharing can only have an impact on market shares if consumers are aware of the quality of the alternative search engine, we consider two scenarios: *Correct Perceptions + Data Sharing* and *No Frictions + Data Sharing.*

Direct effects and equilibrium effects We present two results for each counterfactual. First, we analyze the direct impact of these policies on demand, without accounting for returns to scale from data. These effects should be thought of as the short-run effect of these policies, before the relevance of the results presented by both search engines changes in response to changes in market shares. To obtain these results, we compute market shares and consumer surplus based on the following demand function (at prices equal to zero for all counterfactuals except for *Bing Payments*):

$$s = D(p, \zeta^*, \tilde{\zeta}; \mathscr{C}). \tag{32}$$

In this expression, s is the vector of market shares, p is the vector of prices, ζ^* is the vector of true qualities, and $\tilde{\zeta}$ is the vector of the qualities perceived by agents who have not experienced search engines. This demand function differs across counterfactuals $\mathscr{C}$ as described above. For the *Status Quo,* for instance, market shares are given by $s_{-o} = (1-\phi)S(\Delta\tilde{\zeta} - \sigma(1-\delta))$, and they are $s_{-o} = S(\Delta\tilde{\zeta})$ for the *Choice Screen* counterfactual. Appendix C gives full expressions for all counterfactuals.

In our second set of results, we analyze the effect of these counterfactuals in equilibrium, accounting for economies of scale in data. We use our model of economies to scale from Section 6 to endogenize search result relevance, capturing the fact that search engines with a larger market share will be able to use data to improve the relevance of their results. The true quality is thus described by the following relationship:

$$\zeta^* = Z(s; \mathscr{C}). \tag{33}$$

In counterfactuals in which there is no data sharing, the function $Z(s;\mathscr{C})$ is given by the expression $\zeta_j^* = \alpha a_j + \rho \times C\tilde{T}R(s_j/\hat{s}_j) + \xi_j$. This follows from substituting our results about click-through rates from Section 6 into our definition of quality $\zeta_j^* := \alpha a_j + \rho r_j + \xi_j$. In counterfactuals in which there is data sharing, each search engine's quality is as if they had a market share of one, so $Z(s;\mathscr{C})$ is as derived from the expression $\zeta_j^* = \alpha a_j + \rho \times C\tilde{T}R(1/\hat{s}_j) + \xi_j$.

An equilibrium consists of a joint solution of equations 32 and 33 in s and ζ^*. Let $s^{\text{eq}}(\mathscr{C})$ and $\zeta^{\text{eq}}(\mathscr{C})$ denote the equilibrium shares and qualities given counterfactual $\mathscr{C}$. From these equilibrium quantities, we can compute consumer surplus in equilibrium from the the demand function $D(p, \zeta^{\text{eq}}(\mathscr{C}), \tilde{\zeta}; \mathscr{C})$. Appendix C presents the exact expressions we use to compute counterfactuals.

Since our estimation of returns to scale in data relies on strong assumptions, this second set of counterfactuals should be interpreted as a speculative exercise that attempts to get a sense of the magnitude of the effects that may arise in equilibrium.

Results Table 10 presents our counterfactual results. Panel A presents counterfactuals that do not account for equilibrium effects. The *No Frictions* counterfactual leads to an important increase in the market shares of alternative search engines. The market-wide effect is an increase in Bing's market share from 4.73 percent to 22.2 percent, which increases consumer surplus by $2.83 per consumer per year. The *Choice Screen* and *Correct Perceptions* counterfactuals decompose these effects into information frictions due to misperceptions of quality and due to permanent inattention and switching costs (although switching costs negligible, so they have almost zero impact). We find that the *Choice Screen* only decreases Google's overall market share by 0.87 percentage points: while it causes some Chrome users to switch to Bing, this is offset by some Edge users switching to Google. Despite the small effect on shares, it does have a positive impact on consumer surplus (73.8 percent of the gains from No Frictions), since more people end up using their preferred search engine. These numbers suggest that regulators' focus on choice screens has been misguided, since they can only achieve minor effects on market shares. These results are consistent with Decarolis et al. (2023), who find that introducing choice screens in the EU had a limited impact on market shares.

In the Correct Perceptions counterfactual, we can see that Google's overall market share decreases by 11 percentage points. This highlights that a key element that is necessary to achieve a significant reduction in Google's market share is to correct consumers' misperceptions. We now look at policies that expose a larger fraction of the population to the alternative search engine, thus reducing those misperceptions.

In the *Default Change* counterfactual, we find that Bing's overall market share goes up by 42.1 percentage points. However, this intervention reduces consumer surplus by a significant $77.1 per consumer per year because an important fraction of users are now defaulted into an undesirable search engine. The effect is so big because there are many users with strong preferences for Google who nevertheless use Bing because they are permanently inattentive.[27] The Bing Default counterfactual achieves very similar aggregate results.

Our counterfactuals suggest that two of the most commonly proposed policies to curb Google's dominance have important drawbacks. Choice screens increase consumer surplus, but they barely move the needle in terms of market shares. Default changes have a big impact on market shares, but they come at the cost of a large reduction in consumer surplus. However, the *Bing Default + Delayed Choice Screen* intervention might achieve the best of both approaches. By setting Bing as the default for two weeks, it corrects people's perceptions. After two weeks, it allows them to make an active choice, thus avoiding the large consumer surplus losses. Ultimately, this policy reduces Google's market share by 18.6 percentage points, while limiting consumer surplus losses to less than $1 per customer per year.

Finally, the *Bing Payments* counterfactual achieves a large increase in the market share of Bing. While it also achieves a large increase in consumer surplus, this is mainly because of the payment consumers receive.

Panel B presents results from our equilibrium counterfactuals. In the first eight rows, we find that

[27] As we explain in Section 5.3, we censor the error at $25, so this value is not driven by the fat tail of the error distribution that we estimate. These numbers should thus be thought of as a conservative lower bound for what the actual welfare effects could be.

economies of scale from data reinforce the direct effects of interventions on market shares by around one tenth. Additionally, we observe that economies of scale only have a limited effect on consumer surplus: there are winners (Bing users) and losers (Google users), so the net effects are small. In the last two rows, we see that *Data Sharing* affects market shares by only a few percentage points, and it only affects consumer surplus by a few cents per consumer per year. These findings suggest that economies of scale play a role, but it is secondary in comparison to the direct effects from demand frictions. This is the case, first, because we measure modest economies of scale, and second, because we find a small demand response to the relevance of the search results presented by search engines.

8 Conclusion

Google's large market share in web search has caused concern among regulators in the US and the EU. This paper sheds light on this debate, focusing particularly on the role of browser defaults and economies of scale. Our results highlight that browser defaults are partially responsible for Google's large market share in web search. However, this does not arise only because of switching costs and users' inattention. Our findings show that consumer's lack of exposure to Bing—which is, in part, driven by browser defaults—is a key channel through which Google maintains a higher share than what it would have absent any frictions. We also find that sharing Google's click-and-query data with Microsoft may only have a minor effect on market shares. Our findings suggest that the best way for antitrust authorities to ensure competition in web search may be to implement policies that ensure that a large number of consumers are exposed to more than one search engine.

Table 1: **Sample Sizes**

Description	Sample size
Saw study advertisement	5,280
Passed screening questions	2,737
Consented	2,736
Finished Survey 1 & Willing to commit	2,518
Not rejected for multiple responses	2,390
Consistent baseline search engine	1,726
Finished Survey 2	1,660
Balanced panel	1,461

Notes: This table presents sample sizes at each stage of the experiment. "Not rejected for multiple responses" drops users that attempted to complete the survey multiple times to earn additional compensation. "Consistent baseline search engine" requires users to have at least 10 searches were observed during the 20 days preceding the experiment. Furthermore, the same search engine (Bing or Google) needs to be (i) reported on Survey 1 as the search engine they "usually use," (ii) reported on Survey 1 as the search engine used for address bar search, (iii) the search engine for more than half of searches recorded by Search Extension before installation, and (iv) the address bar default recorded by Search Extension before installation. The "Balanced panel" refers to the users whose Search Extension was active 8 weeks after the beginning of the experiment. The sample in each row is a strict subset of the row above.

Table 2: **Sample Demographics**

	(1) Analysis sample	(2) U.S. adults
Income ($000s)	56.45	40.86
College	0.58	0.33
Male	0.45	0.49
Age	36.39	48.16
White	0.60	0.32
Baseline Google user	0.96	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 (2024a) 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: **Summary Statistics**

Description	Sample	Mean	SE
Fraction of days with at least one search	All users before experiment	0.620	0.035
Fraction of days with at least one search	All users during experiment	0.652	0.033
Average number of daily searches	All users before experiment	11.105	1.779
Average number of daily searches	All users during experiment	10.911	1.159
Average daily share of Google searches	Google users before experiment	0.978	0.001
Average daily share of Google searches	Bing users before experiment	0.035	0.008

Notes: This table provides summary statistics of participants' search behavior both before (pre Survey 1) and during the experiment (between Survey 1 and eight weeks after Survey 1). "All users" refers to participants across all treatment groups. The statistics show the mean and standard error of individual-level averages.

Table 4: **Browser and Search Engine Statistics**

Operating system	Browser	Self-reported default search engine	Share (Sample)
Mac	Chrome	Google	0.227
		Bing	0.001
	Edge	Google	0.001
		Bing	0.001
Windows	Chrome	Google	0.663
		Bing	0.008
	Edge	Google	0.013
		Bing	0.033

Notes: This table provides the unconditional share of (Operating System, Browser, Self-Reported Default Search Engine)-triplets in our sample; the shares add up to 1.

Table 5: Average Change in Ratings in Switch Group

	Change in the relative preference for Bing							
	(1) Overall quality	(2) Relevant links	(3) Result page features	(4) Relevant ads	(5) AI chat	(6) Privacy	(7) Rewards	(8) Ads volume
Panel A: Baseline Google users								
Δ Rating	0.403***	0.231***	0.511***	0.452***	0.489***	0.425***	0.353***	-0.077
	(0.074)	(0.078)	(0.090)	(0.081)	(0.077)	(0.063)	(0.066)	(0.058)
Rating at baseline: Mean	-1.520	-1.337	-1.143	-0.835	-0.443	-0.648	0.183	0.729
Rating at baseline: Sd	0.748	0.829	0.886	0.835	1.060	0.936	1.086	0.717
N	221	221	221	221	221	221	221	221
Panel B: Baseline Bing users								
Δ Rating	-0.394*	-0.121	-0.303	-0.121	-0.303*	-0.121	-0.152	-0.091
	(0.213)	(0.212)	(0.211)	(0.161)	(0.171)	(0.155)	(0.152)	(0.118)
Rating at baseline: Mean	0.364	0.121	0.333	0.091	1.030	0.242	1.576	0.848
Rating at baseline: Sd	1.295	1.244	1.021	0.843	0.883	0.663	1.001	0.712
N	33	33	33	33	33	33	33	33

Notes: This table presents the average change in the relative preference for Bing (ΔRating) after fourteen days of exposure to Bing in the Switch Group. When ΔRating is positive, participants rated Bing more favorably after exposure to it. The table also displays the mean and standard deviations of the baseline relative rating of search engines where the baseline ratings were measured during Survey 1. The sample includes all baseline Google and Bing users in the $10 Switch Bonus Control (S10CC) group who passed the attention checks, including those who did not accept the offer to switch during Survey 1. The $10 Switch Bonus Control group participants were offered $10 to switch to Google for the 14 days between Survey 1 and Survey 2. Define Y_{i1} and Y_{i2} as responses on Surveys 1 and 2. The regression used to compute the average change is $(Y_{i2} - Y_{i1}) = \tau^C + \varepsilon_i$. 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. Standard errors clustered at the participant level.

Table 6: **Effects of Ranking Degradation and Ad Blocking on Bing Clicks and Quality Ratings**

	Clicks on Bing					**Change in the relative preference for Bing**						
Dep. var:	(1) Prior search rank	(2) Ad clicks	(3) CTR all	(4) CTR organic	(5) CTR organic (top)	(6) Overall quality	(7) Relevant links	(8) Result page features	(9) Relevant ads	(10) AI chat	(11) Privacy	(12) Rewards
Ad Block	0.033	-2.996***	-0.013	0.012	-0.017*	-0.038	-0.050	-0.152*	-0.169**	-0.087	-0.071	-0.118*
	(0.096)	(0.367)	(0.027)	(0.026)	(0.010)	(0.070)	(0.072)	(0.086)	(0.072)	(0.076)	(0.061)	(0.065)
Rank Degrade	3.422***	1.867***	-0.057**	-0.075***	-0.153***	-0.287***	-0.311***	-0.328***	-0.165**	-0.037	-0.099	0.065
	(0.090)	(0.354)	(0.027)	(0.027)	(0.010)	(0.070)	(0.072)	(0.086)	(0.072)	(0.076)	(0.061)	(0.065)
Constant	1.640***	2.082***	0.361***	0.343***	0.237***	0.425***	0.241***	0.583***	0.432***	0.479***	0.425***	0.365***
	(0.054)	(0.229)	(0.017)	(0.017)	(0.008)	(0.061)	(0.063)	(0.075)	(0.063)	(0.066)	(0.053)	(0.057)
R^2	0.666	0.123	0.008	0.014	0.294	0.019	0.021	0.020	0.012	0.002	0.005	0.005
N	646	646	636	636	636	895	895	895	895	895	895	895

Notes: This table presents the effects of Ranking Degradation and Ad Blocking on (i) Bing click outcomes and (ii) the relative preference for Bing as measured by survey ratings. In the left panel, we limit the sample to users who accepted the switch offer at the time of Survey 1 and measure the effect of Ranking Degradation and Ad Blocking on the average initial rank number of the clicked Bing search links before ranking degradation (1), the count of ad clicks on Bing (2), Bing click-through rate (CTR) including both ad clicks and search link clicks (3), Bing CTR including only search link clicks (4), and Bing CTR including only the clicks associated with the "first-ranked" Bing search result (5). Define Y_i as the outcome variable of interest observed during the first two weeks since Survey 1, and define w_i^R and w_i^A as indicators for the Ranking Degradation and Ad Blocking groups. The regression model underlying columns (1)-(5) is $Y_i = \tau^A w_i^A + \tau^R w_i^R + \mu + \varepsilon_i$. In the right panel, we measure the impact of Ranking Degradation and Ad Blocking on quality ratings. The sample associated with the right panel includes all baseline Google users in the $10 Switch Bonus group who passed the attention checks including those who did not accept the offer to switch at the time of Survey 1. Define Y_{i1} and Y_{i2} as responses on Surveys 1 and 2. The regression underlying the left panel is $(Y_{i2} - Y_{i1}) = \tau^A w_i^A + \tau^R w_i^R + \mu + \varepsilon_i$. The survey questions in column 6 and columns 7–12, 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. *, **, ***: statistically significant with 90, 95, and 99 percent confidence, respectively. Standard errors clustered at the participant level.

Table 7: **Empirical Moments for Demand Estimation**

Description	(1) Formula	(2) Estimate	(3) SE
Baseline Chrome users			
Baseline and Control, $t \geq 0$	$\hat{s}^{C}_{n,t\geq 0}$	0.012	0.003
Active Choice group Bing share, $t \geq 1$	$\hat{s}^{A}_{n,t\geq 1}$	0.026	0.010
Default Change group Bing share, $t = *$	$\hat{s}^{D}_{n,t=*}$	0.783	0.036
Default Change group Bing share, $t = 1$	$\hat{s}^{D}_{n,t=1}$	0.510	0.043
Default Change group Bing share, $t = 2$	$\hat{s}^{D}_{n,t=2}$	0.436	0.043
\$1 Switch Bonus group Bing share, $t = 1$	$\hat{s}^{S1}_{n,t=1}$	0.337	0.056
\$25 Switch Bonus group Bing share, $t = 1$	$\hat{s}^{S25}_{n,t=1}$	0.750	0.046
\$10 Switch Bonus (CC) group Bing share, $t = 1$	$\hat{s}^{S10}_{n,t=1}$	0.677	0.032
\$10 Switch Bonus (CC) group Bing share, $t \geq 2$	$\hat{s}^{S10}_{n,t\geq 2}$	0.206	0.027
\$10 Switch Bonus (CR) group Bing share, $t \geq 2$	$s^{R}_{n,t\geq 2}$	0.158	0.023
\$10 Switch Bonus (BC) group Bing share, $t \geq 2$	$s^{B}_{n,t\geq 2}$	0.146	0.024
Baseline Edge users			
Baseline Google share, $t = 0$	$\hat{s}_{n,t=0}$	0.273	0.000
Active Choice group Google share, $t \geq 1$	$\hat{s}^{A}_{n,t\geq 1}$	0.368	0.052
\$10 Switch Bonus (CC) group Google Share, $t = 1$	$\hat{s}^{S10}_{n,t=1}$	0.709	0.073
\$10 Switch Bonus (CC) group Google Share, $t \geq 2$	$\hat{s}^{S10}_{n,t\geq 2}$	0.389	0.054

Notes: This table presents the empirical moments used for the demand estimation procedure described in Section 5. Standard errors clustered at the participant level.

Table 8: **Demand Parameter Estimates**

Description	(1) Formula	(2) Estimate	(3) SE
All users			
Distribution shape	a	2.877	0.529
Price response	η	0.321	0.082
Permanent inattention	ϕ	0.334	0.059
Attention probability	π	0.824	0.151
Amortized switching cost	$\sigma(1-\delta)$	\$0.004	0.007
Baseline Chrome users			
Bing preference shifter	$\Delta\tilde{\xi}$	-\$3.106	0.014
Learning	$\zeta^*_{-o} - \tilde{\xi}_{-o}$	\$0.270	0.139
Ad load response	$\alpha(a^{CA}_{-o} - a^{CC}_{-o})$	-\$0.131	0.099
Quality response	$\rho(r^{RC}_{-o} - r^{CC}_{-o})$	-\$0.109	0.097
Baseline Edge users			
Google preference shifter	$\Delta\tilde{\xi}$	-\$7.612	2.488
Learning	$\zeta^*_{-o} - \tilde{\zeta_{-o}}$	\$1.149	6.566

Notes: This table presents the parameter estimates from the demand estimation procedure described in Section 5. Standard errors clustered at the participant level.

Table 9: Economy of Scale Estimates

Description	Parameter	Estimate	SE
Click-through rate at inception	α	0.1811	-
Value of additional data	β	0.0056	(0.0007)
Shape of returns from data	θ	0.9458	(0.0318)

Notes: This table presents our parameter estimates for the relationship between the amount of data and the relevance of search results, as measured by the click-through-rate. See 31 for the functional form specification.

Table 10: **Counterfactual Simulations**

Panel A: Direct Effects

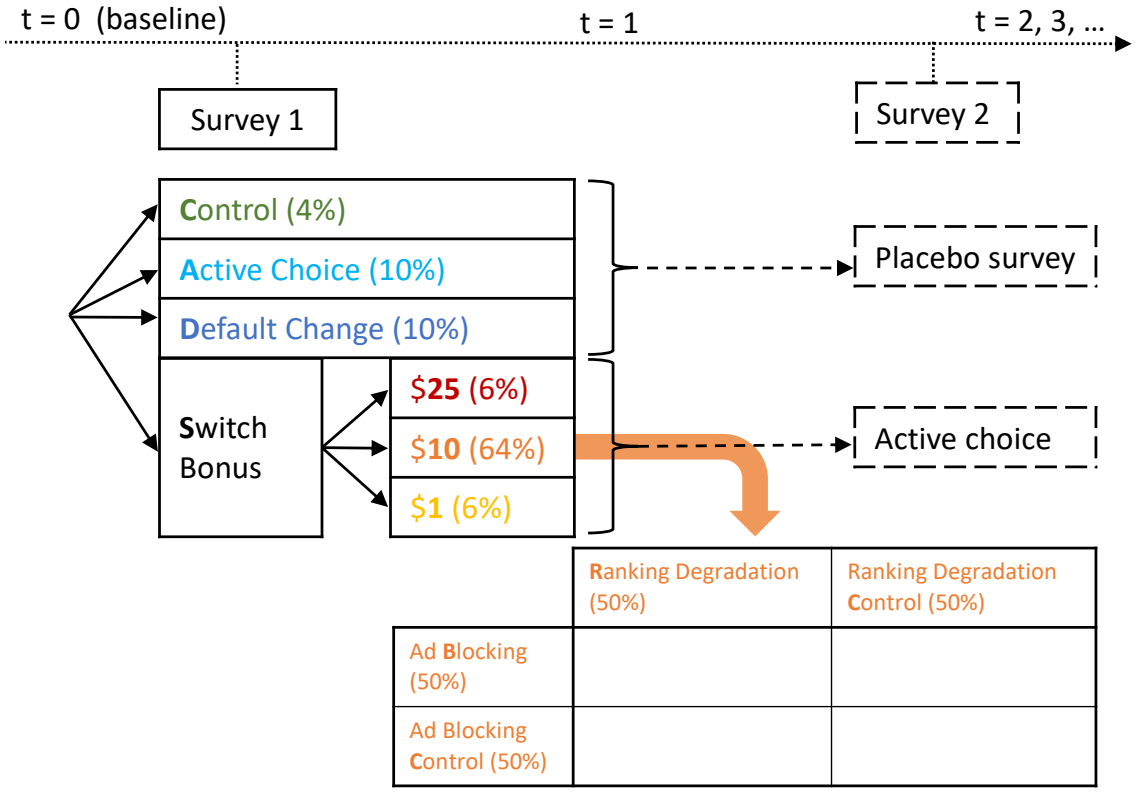
	Chrome		Edge		Combined	
	(1)	(2)	(3)	(4)	(5)	(6)
Description	Google share	CS gain	Google share	CS gain	Google share	CS gain
Status Quo	0.99	0.00	0.22	0.00	0.95	0.00
No Frictions	0.80	0.87	0.35	43.95	0.78	2.83
Choice Screen	0.97	0.10	0.33	43.74	0.94	2.09
Correct Perception	0.87	0.54	0.23	0.14	0.84	0.52
Default Change	0.53	-80.76	0.57	0.60	0.53	-77.06
Bing Default	0.53	-80.76	0.22	0.00	0.52	-77.08
Bing Default + Delayed Choice Screen	0.79	-2.27	0.33	42.06	0.77	-0.25
Bing Payments ($10)	0.56	94.13	0.17	208.34	0.54	99.34

Panel B: Equilibrium Effects

	Chrome		Edge		Combined	
	(1)	(2)	(3)	(4)	(5)	(6)
Description	Google share	CS gain	Google share	CS gain	Google share	CS gain
Status Quo	0.99	0.00	0.22	0.00	0.95	0.00
No Frictions	0.77	0.97	0.35	45.03	0.75	2.98
Choice Screen	0.97	0.10	0.33	43.91	0.94	2.09
Correct Perception	0.85	0.54	0.23	1.22	0.82	0.57
Default Change	0.50	-80.00	0.56	1.20	0.51	-76.30
Bing Default	0.50	-80.01	0.22	1.83	0.49	-76.28
Bing Default + Delayed Choice Screen	0.77	-2.15	0.33	43.18	0.75	-0.09
Bing Payments ($10)	0.56	94.75	0.17	210.30	0.54	100.02
Data Sharing	0.84	1.10	0.23	2.83	0.81	1.18
Choice Screen + Data Sharing	0.76	1.69	0.35	46.25	0.74	3.72

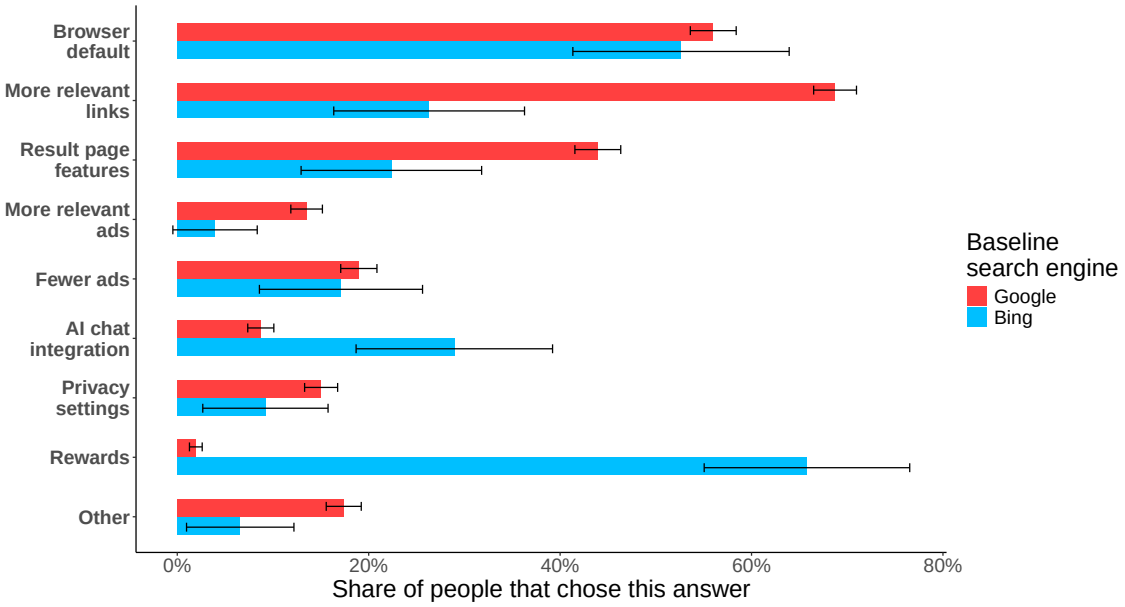
Notes: This table presents the counterfactual simulation results from the procedure described in Section 7. CS gain means consumer surplus gain, in $/consumer-year. Panel A computes the direct effects, i.e., the effects before taking into account any potential change in result relevance implied by our economy of scale analysis. Panel B additionally takes these changes into account.

Figure 1: **Experimental Design**



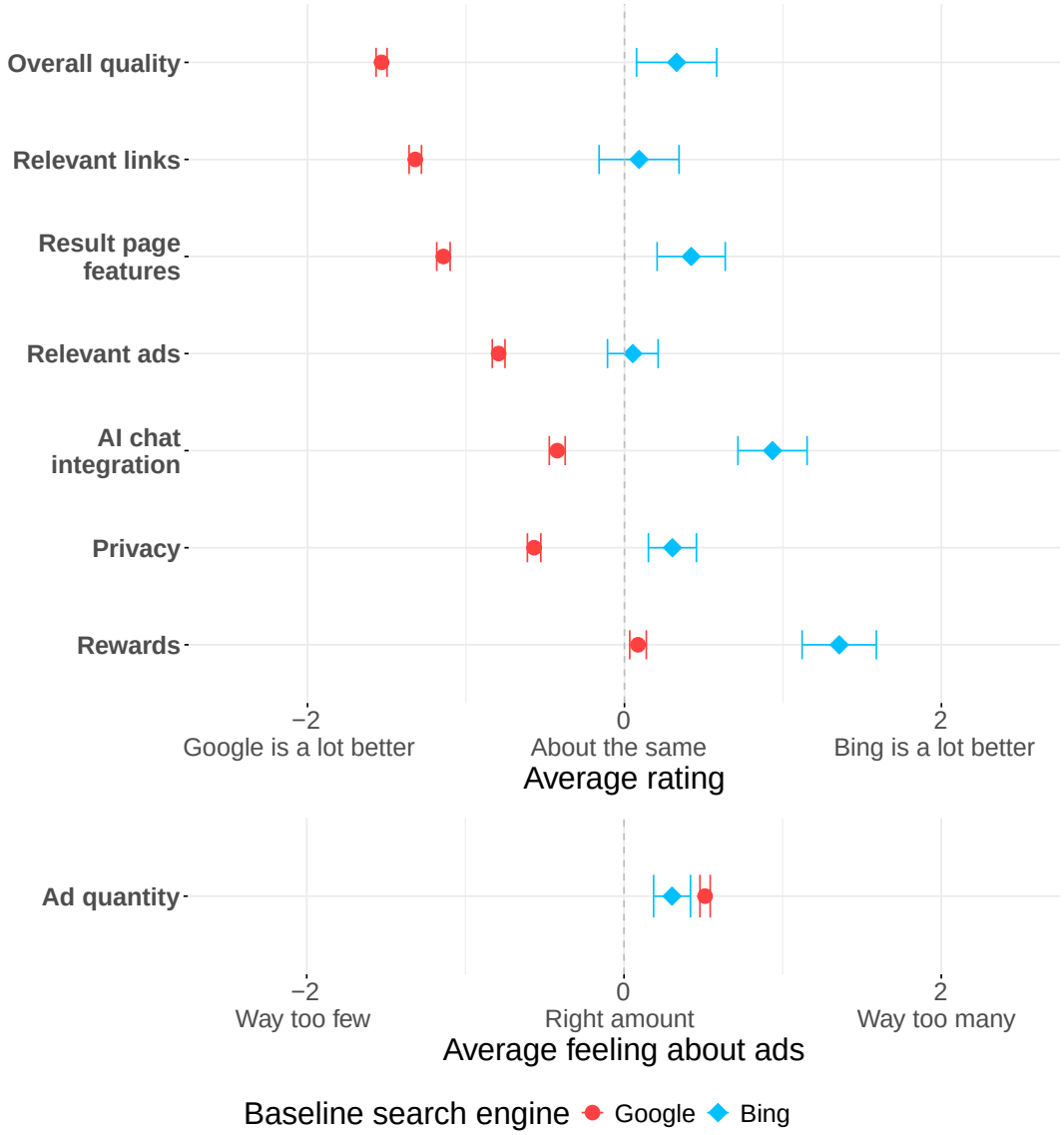
Notes: This figure illustrates our design. There are four treatment groups. The control group was guided through how to change their bookmarks. The default group was paid $10 to switch their default from Google to Bing for two days. The active choice group was asked about their preferred default and then guided to implement this choice. The switch-bonus treatment group was asked to switch search engines in return for a bonus payment of either $1, $10, or $25. The $10-group was further randomized in a two-by-two factorial design, which varies whether ads were blocked and whether search results are degraded. The switch bonus group made an active choice after the fourteen days incentive period.

Figure 2: **Why People Use Google or Bing**



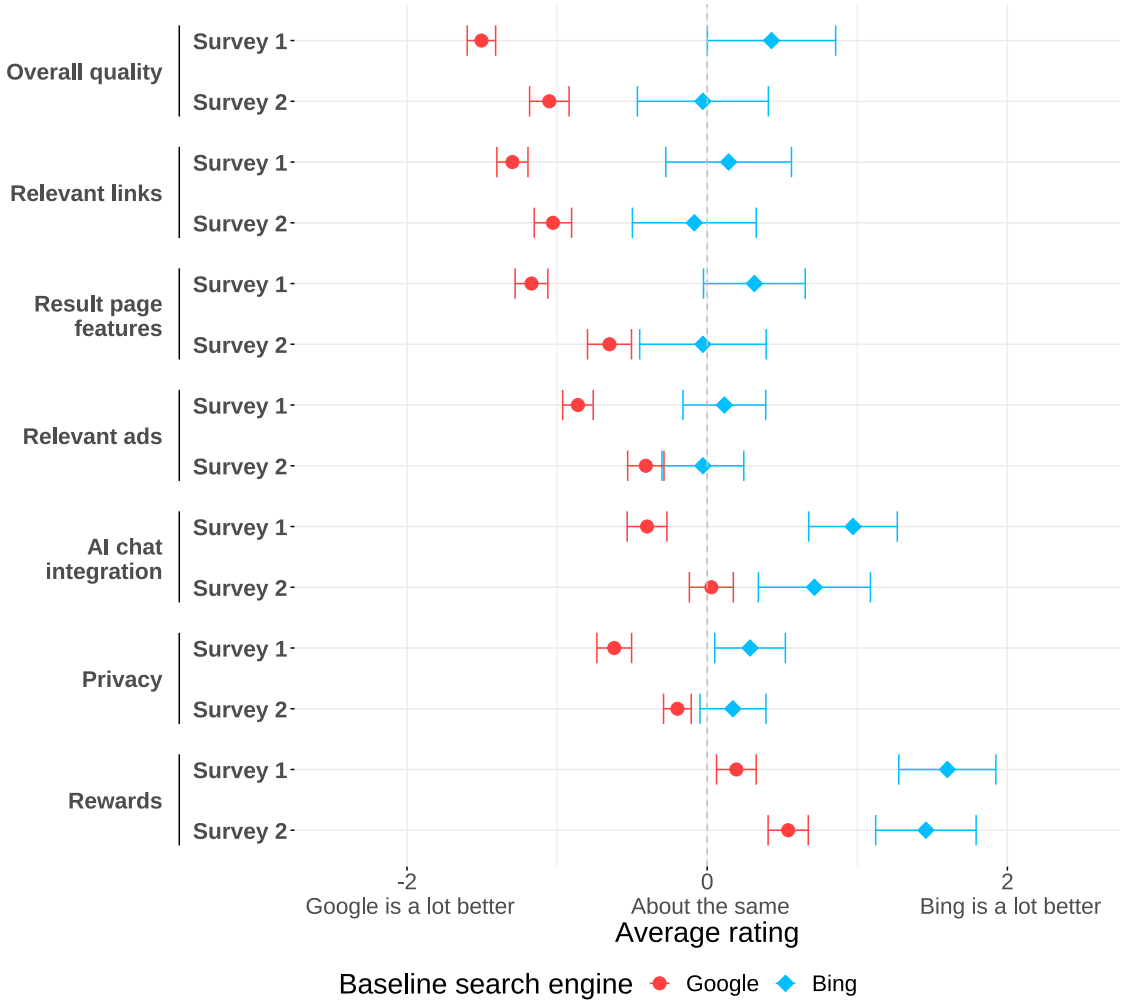
Notes: This figure presents the share of baseline Google and Bing users that chose each answer to the following question: "Why do you use [baseline search engine used] instead of [other search engine] for your searches on this web browser? Choose all that apply." Whiskers indicate 95 percent confidence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

Figure 3: **Initial Ratings of Google and Bing**



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.

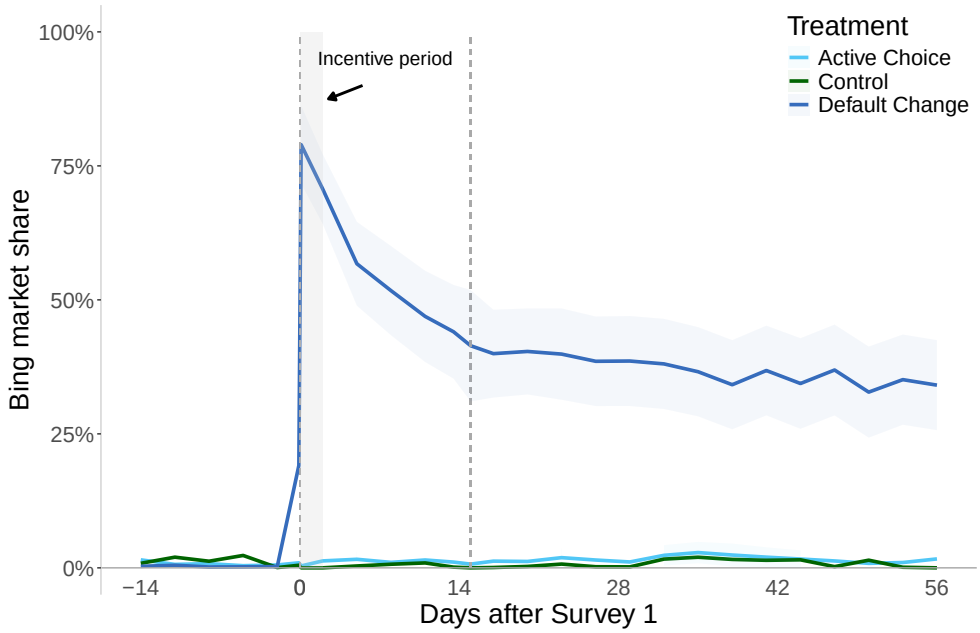
Figure 4: **Switch Treatment ($10) Search Engine Rating Change**



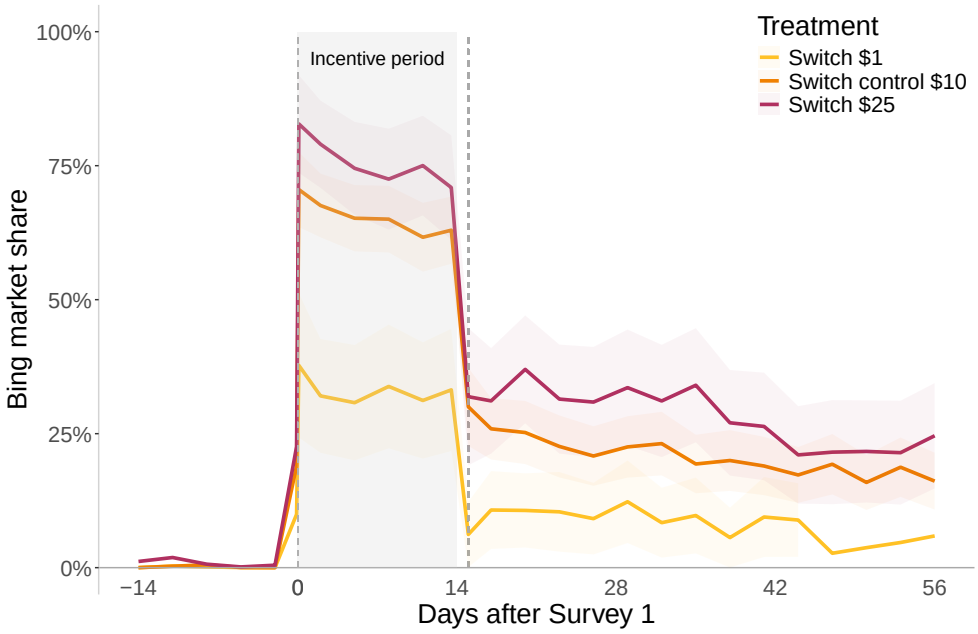
Notes: This figure presents average responses to the search engine rating questions for baseline Google and Bing users who were assigned to the $10 Switch Bonus Control group (S10CC). The responses to the rating questions were collected from both Survey 1 and Survey 2. The figure presents the average ratings 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. Whiskers indicate 95 percent confidence intervals.

Figure 5: **Search Market Shares:**

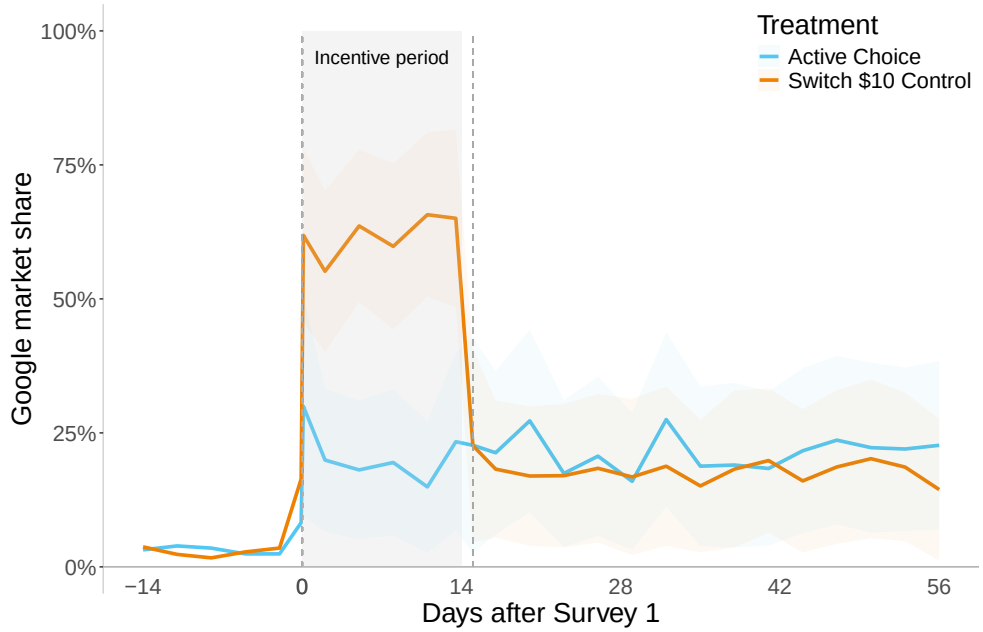
(a) **Baseline Google Users: Active Choice, Control and Default Change**



(b) **Baseline Google Users: Switch Bonus**



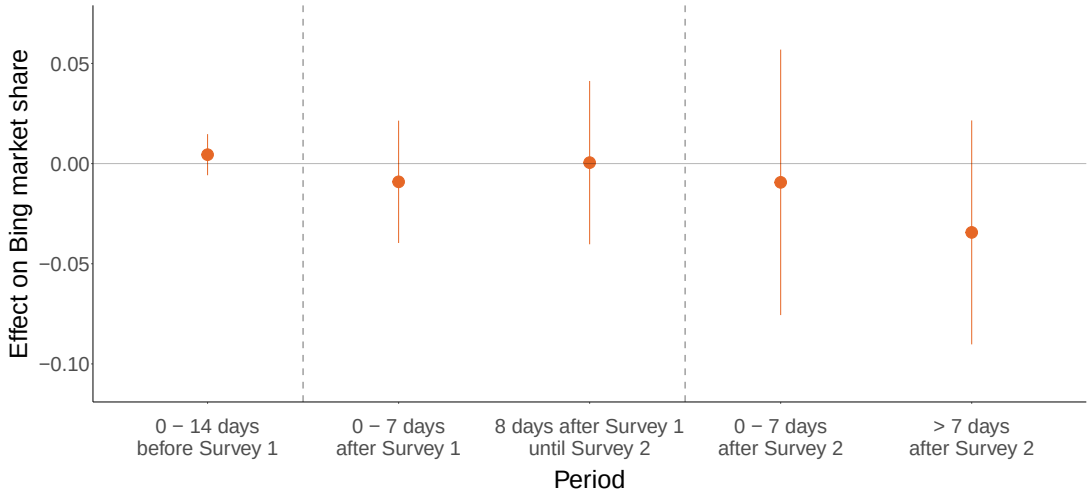
(c) **Baseline Bing Users: Active Choice and $10 Switch Bonus Control**



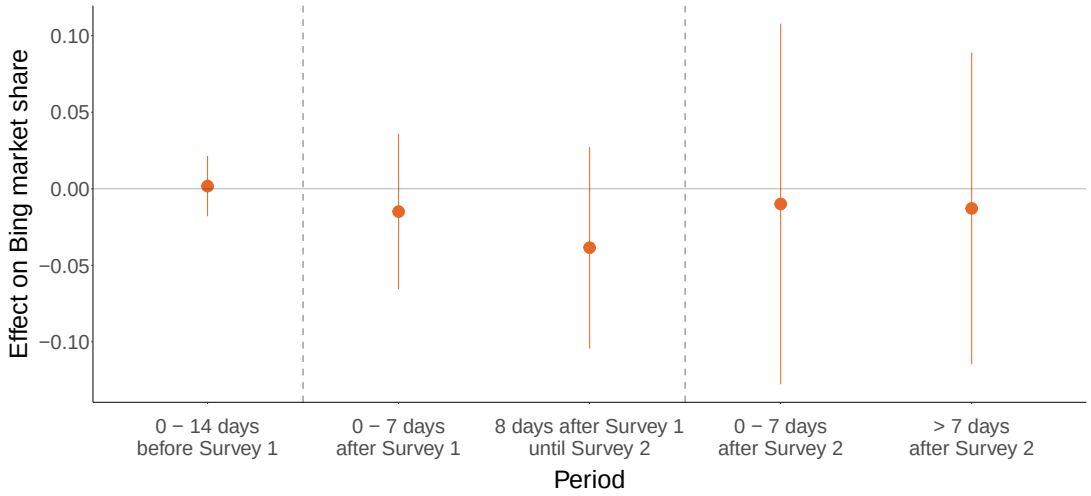
Notes: This figure presents the alternative search engine market share by treatment group for each day of the experiment; Days $t \geq 15$ refer to Day $t - 15$ after Survey 2 (for instance, Day 16 refers to the first day after the second survey). As participants can take Survey 2 after the day on which they are invited, the plot skips days for some participants. Panels (a) and (b) present results for baseline Google users, while Panel (c) presents results for baseline Bing users. The dashed vertical lines mark the dates of the two surveys, and the shaded area indicates the incentive period (which varies across treatments). To arrive at daily market shares, we first compute the daily market share for each subject separately and then average shares across subjects. To smooth the plot, we do not directly plot these estimates, but take averages within bins (0-, 0+, [1-3], [4-6], [7-9], [10-12], [13-14], [15-15], [16-18], [18-20], [20-22], etc) and plot the line connecting the dots given by within-bin averages of the search engine market share (on the vertical axis) and Days after Survey 1 (on the horizontal axis.) 0- (respectively 0+) refers to searches on the day of the first survey but before (respectively after) this survey.

Figure 6: **Effects of Ranking Degradation and Ad Blocking on Market Share**

(a) **Effects of Ranking Degradation**

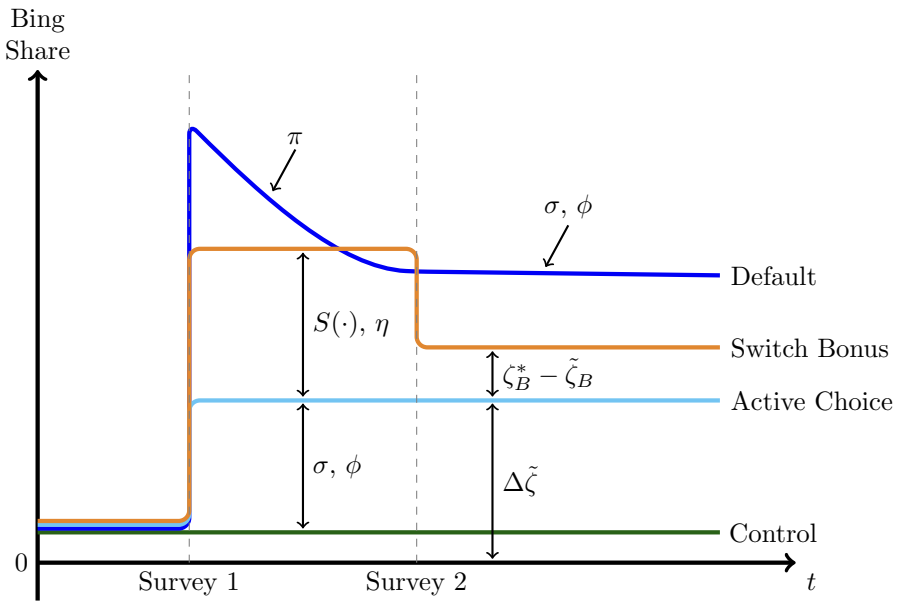


(b) **Effects of Ad Blocking**



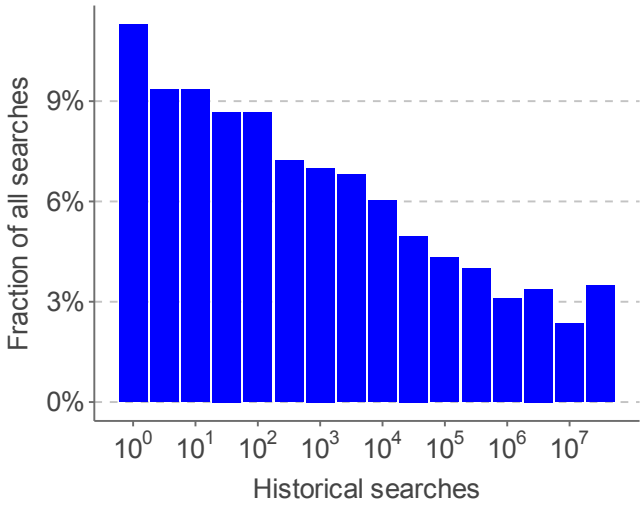
Notes: Panels (a) and (b) present the effects of the Ranking Degradation and Ad Blocking treatments on Bing market share by period of the experiment, estimated using equation (12). The sample includes only baseline Google users who were offered $10 to switch to Bing for the 14 days between Survey 1 and Survey 2. Some of the participants in the sample received ranking degradation or/and ad blocking treatment. Whiskers indicate 95 percent confidence intervals. The dashed vertical lines mark the dates of the two surveys.

Figure 7: Connection between experimental market shares and parameters.



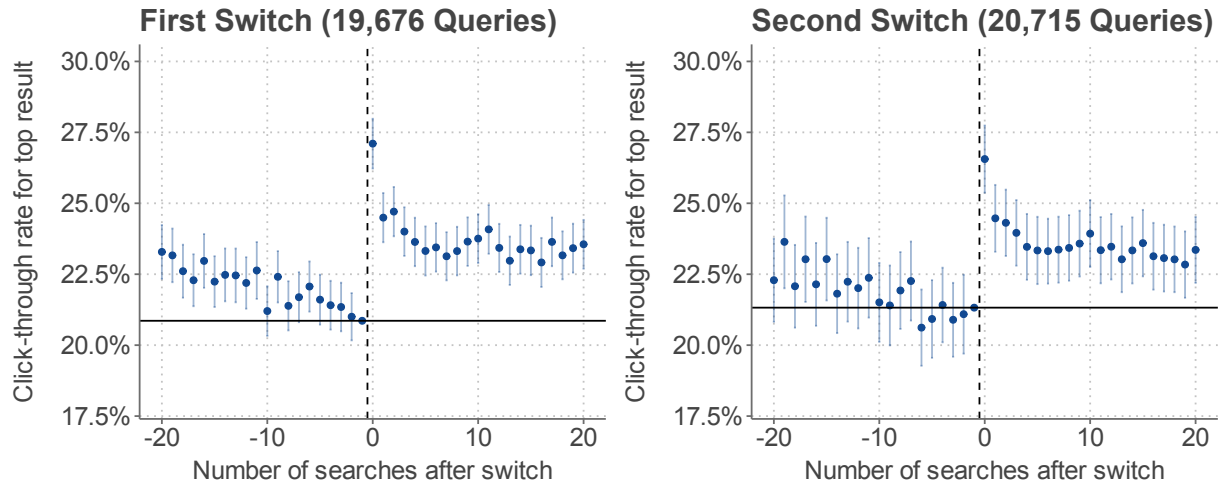
Notes: This figure illustrates, for Chrome users, how the market shares of Bing behave over time in each treatment condition. It also shows which elements identify each of the model parameters.

Figure 8: Distribution of Views Across Queries



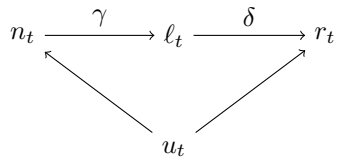
Notes: This graph provides the fraction of all searches for queries of a certain popularity. We count all searches made on Bing in the period between 10/1/21 and 10/1/22, and group queries by how often they were searched for. The shape of the distribution supports the commonly-held assumption that there are a large number of tail queries each associated with a lower number of searches but jointly responsible for a considerable share of searches on Bing.

Figure 9: Event Study Illustrating Effect of Ranking Change on CTR



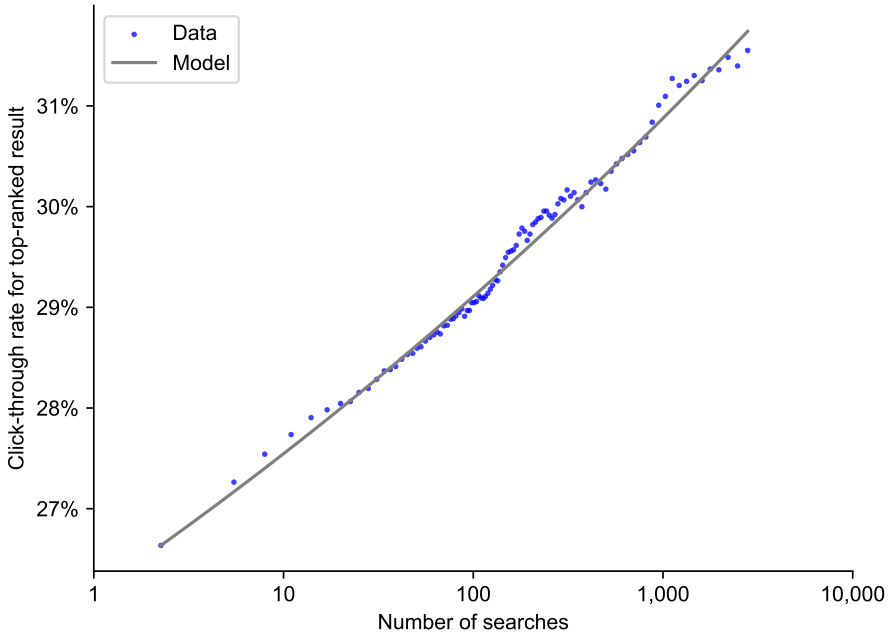
Notes: We regress the CTR on query fixed effects and a set of dummies measuring how many searches happened since the k-th switch in the top-ranked result. The figures report the these dummies, providing evidence that CTR increases at the time of a ranking change but otherwise follows a secular decline. The secular decline motivates our identification strategy (as it suggests the presence of a potential confounder), and the positive effect of ranking changes suggests that rankings do affect CTR.

Figure 10: Directed Acyclic Graph Underlying Estimation Strategy



Notes: This DAG represents how we identify the causal effect of increasing the number of searches (n_t) on the result relevance as measured by CTR (r_t) while accounting for an unobserved confounder (u_t) that could (e.g.) represent how different types of users arrive over the lifecycle of a query. The key to our identification strategy is the mediator ℓ_t, which represents the search engine's ranking of links. In particular, the search engine only learns from user data in the form of additional searches, and hence the relationship between n_t and ℓ_t is not confounded. Similarly, the relationship between ℓ_t and r_t is not confounded conditional on n_t.

Figure 11: Model Fit for Returns to Scale



Notes: We exhibit the model fit for (31) by exhibiting both the model predicted CTR as a function of the number of prior searches for a query as well as a binscatter of the actually observed CTR as a function of the number of prior searches. The model estimates a linear relationship between CTR and the logarithm of the number of prior searches, and this relationship fits the observed data well.

References

Allcott, Hunt, Luca Braghieri, Sarah Eichmeyer, and Matthew Gentzkow, "The welfare effects of social media," *American economic review*, 2020, *110* (3), 629–676.

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

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

Athey, Susan and Glenn Ellison, "Position auctions with consumer search," *The Quarterly Journal of Economics*, 2011, *126* (3), 1213–1270.

Azevedo, Eduardo M, Alex Deng, José Luis Montiel Olea, Justin Rao, and E Glen Weyl, "A/B Testing with Fat Tails," *Journal of Political Economy*, 2020, *128* (12), 4614–000.

Bajari, Patrick, Victor Chernozhukov, Ali Hortaçsu, and Junichi Suzuki, "The Impact of Big Data on Firm Performance: An Empirical Investigation," in "AEA Papers and Proceedings," Vol. 109 American Economic Association 2019, pp. 33–37.

Bellemare, Marc F., Jeffrey R. Bloem, and Noah Wexler, "The Paper of How: Estimating Treatment Effects Using the Front-Door Criterion*," *Oxford Bulletin of Economics and Statistics*, 2024.

Blake, Thomas, Chris Nosko, and Steven Tadelis, "Consumer heterogeneity and paid search effectiveness: A large-scale field experiment," *Econometrica*, 2015, *83* (1), 155–174.

Blake, Tom, Sarah Moshary, Kane Sweeney, and Steve Tadelis, "Price salience and product choice," *Marketing Science*, 2021, *40* (4), 619–636.

Brynjolfsson, Erik, Avinash Collis, and Felix Eggers, "Using massive online choice experiments to measure changes in well-being," *Proceedings of the National Academy of Sciences*, 2019, *116* (15), 7250–7255.

Bursztyn, Leonardo, Benjamin R Handel, Rafael Jimenez, and Christopher Roth, "When product markets become collective traps: The case of social media," Technical Report, National Bureau of Economic Research 2023.

Carroll, Gabriel D, James J Choi, David Laibson, Brigitte C Madrian, and Andrew Metrick, "Optimal Defaults and Active Decisions," *Quarterly Journal of Economics*, 2009, *124* (4), 1639–1674.

Chiou, Lesley and Catherine Tucker, "Search Engines and Data Retention: Implications for Privacy and Antitrust," *Available at National Bureau of Economic Research 23815*, 2017.

Decarolis, Francesco, Muxin Li, and Filippo Paternollo, "Competition and Defaults in Online Search," *Available at SSRN 4660406*, 2023.

DellaVigna, Stefano and Ulrike Malmendier, "Paying Not to Go to the Gym," *American Economic Review*, 2006, *96* (3), 694–719.

Department of Justice, "United States vs Google LLC," 2020.

Dinielli, David, Fiona M Scott Morton, Katja Seim, Michael Sinkinson, Amelia Fletcher, Gregory S Crawford, Jacques Crémer, Paul Heidhues, Michael Luca, Tobias Salz et al., "Consumer Protection for Online Markets and Large Digital Platforms," 2023.

Duque, Vásquez Omar, "The Potential Anticompetitive Stickiness of Default Applications: Addressing Consumer Inertia with Randomization," *Available at SSRN 4077132*, 2022.

Ecommercedb, "Customer Journey in Online Shopping: Most Start on Search Engines," 2024.

Edelman, Benjamin, Michael Ostrovsky, and Michael Schwarz, "Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords," *American economic review*, 2007, *97* (1), 242–259.

Einav, Liran, Benjamin Klopack, and Neale Mahoney, "Selling Subscriptions," *Available at National Bureau of Economic Research 31547*, 2023.

Ericson, Keith M Marzilli, "Consumer inertia and firm pricing in the Medicare Part D prescription drug insurance exchange," *American Economic Journal: Economic Policy*, 2014, *6* (1), 38–64.

European Commission, "CASE AT.40099, Google Android," 2018.

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

Fowlie, Meredith, Catherine Wolfram, Patrick Baylis, C Anna Spurlock, Annika Todd-Blick, and Peter Cappers, "Default Effects and Follow-on Behaviour: Evidence from an Electricity Pricing Program," *The Review of Economic Studies*, 2021, *88* (6), 2886–2934.

Handel, Benjamin R, "Adverse selection and inertia in health insurance markets: When nudging hurts," *American Economic Review*, 2013, *103* (7), 2643–2682.

He, Di, Aadharsh Kannan, Tie-Yan Liu, R Preston McAfee, Tao Qin, and Justin M Rao, "Scale Effects in Web Search," in "Web and Internet Economics: 13th International Conference, WINE 2017, Bangalore, India, December 17–20, 2017, Proceedings 13" Springer 2017, pp. 294–310.

Heidhues, Paul, Alessandro Bonatti, L Elisa Celis, Gregory S Crawford, David Dinielli, Michael Luca, Tobias Salz, Monika Schnitzer, Fiona M Scott Morton, Michael Sinkinson et al., "More Competitive Search Through Regulation," *Yale J. on Reg.*, 2023, *40*, 915.

Ho, Kate, Joseph Hogan, and Fiona Scott Morton, "The impact of consumer inattention on insurer pricing in the Medicare Part D program," *The RAND Journal of Economics*, 2017, *48* (4), 877–905.

Imbens, Guido W, "Potential outcome and directed acyclic graph approaches to causality: Relevance for empirical practice in economics," *Journal of Economic Literature*, 2020, *58* (4), 1129–1179.

Ostrovsky, Michael, "Choice screen auctions," in "Proceedings of the 22nd ACM Conference on Economics and Computation" 2021, pp. 741–742.

Patterson, Mark R, "Google and search engine market power," *Fordham Law Legal Studies Research Paper*, 2012, (2047047).

Pearl, Judea and Dana Mackenzie, *The book of why: the new science of cause and effect*, Basic books, 2018.

Schaefer, Maximilian and Geza Sapi, "Complementarities in Learning from Data: Insights from General Search," *Information Economics and Policy*, 2023, *65*, 101063.

SEO Statistics, "124 SEO Statistics for 2024," 2024.

StatCounter, "Desktop Search Engine Market Share United States Of America," 2024.

_ , "Desktop vs. Mobile vs. Tablet Market Share: United States of America," 2024.

Stigler Committee on Digital Platforms, "Stigler Committee on Digital Platforms: Final Report," *Stigler Center News*, 2019.

The U.S. House Judiciary Subcommittee on Antitrust, "Investigation of Competition in Digital Markets," 2020.

Tucker, Catherine, "Digital Data, Platforms and the Usual [Antitrust] Suspects: Network Effects, Switching Costs, Essential Facility," *Review of Industrial Organization*, 2019, *54* (4), 683–694.

UK Competition and Markets Authority, "Online Platforms and Digital Advertising Market Study," 2020.

Varian, Hal, "Is there a Data Barrier to Entry?," *Google*, 2015.

Varian, Hal R, "Position auctions," *international Journal of industrial Organization*, 2007, *25* (6), 1163–1178.

Online Appendix

Default Effects and Economies of Scale in Web Search

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

Table of Contents

A Data Appendix **2**

B Details about Estimation **7**

C Computation of Counterfactuals and Consumer Surplus **7**

C.1 Direct effects 8

C.2 Equilibrium effects 9

D Economy of Scale **12**

D.1 Summary Statistics 12

D.2 Descriptives 12

D.3 Implementation Details for Identification Argument 13

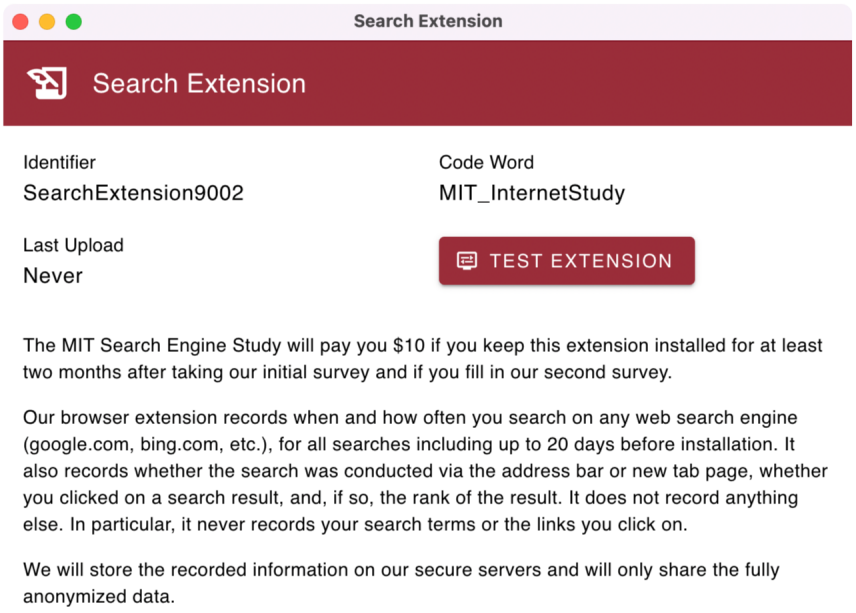
D.4 Taylor Expansion to Address HDFE in NLLS 15

D.5 Effect of Data on Result Relevance 15

D.6 Cross-Query Learning 17

A Data Appendix

Figure A1: **Search Extension Window**



Notes: This figure shows the image of Search Extension. The participants click "Test Extension" button to ensure that the extension sends the anonymized data to our data repository.

Table A1: **Covariate Balance**

(a) **Baseline Google Users**

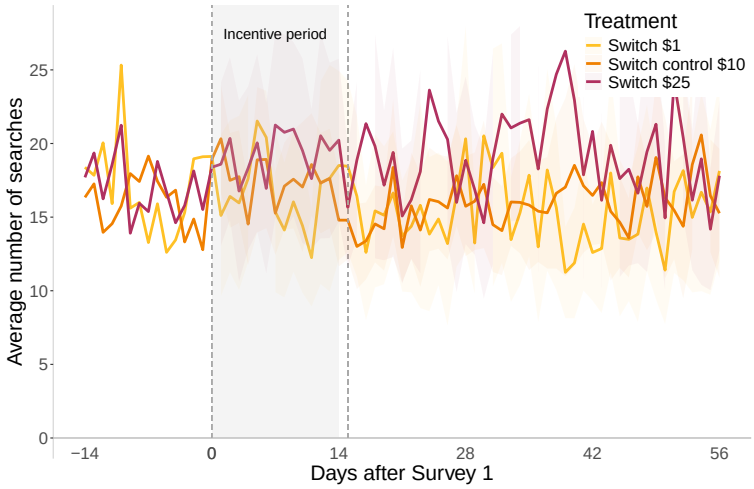
	(1) Control	(2) Active Choice	(3) Default Change	(4) Switch ($1)	(5) Switch ($25)	(6) Switch ($10 & CC)	(7) Switch ($10 & BC)	(8) Switch ($10 & CR)	(9) Switch ($10 & BR)	F-test p-value
Income ($000s)	52.30	62.48	59.64	59.07	61.94	53.80	52.18	55.96	57.59	0.40
College Degree	0.46	0.57	0.63	0.57	0.57	0.57	0.59	0.58	0.61	0.58
Male	0.33	0.44	0.52	0.47	0.46	0.42	0.39	0.47	0.48	0.13
Age	35.89	35.64	37.22	34.52	35.89	35.48	36.63	36.80	37.49	0.43
White	0.68	0.67	0.73	0.62	0.69	0.62	0.65	0.68	0.66	0.52

(b) **Baseline Bing Users**

	(1) Active Choice	(2) Switch ($10 & CC)	(3) F-test p-value
Income ($000s)	45.00	54.21	0.43
College Degree	0.47	0.45	0.82
Male	0.53	0.66	0.25
Age	35.13	37.61	0.37
White	0.66	0.63	0.81

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. The sample underlying this table includes all participants (inlcuding participants who did not stay with us until endline.)

Figure A2: **Search volume in Switch Bonus groups**



Notes: This figure presents the average number of searches on any search engine by participants of the Switch Bonus groups for each day of the experiment. The dashed vertical lines mark the dates of the two surveys. To arrive at daily market shares, we first compute the daily market share for each subject separately and then average shares across subjects.

Table A2: **Completion Rates**

(a) **Baseline Google Users**

	(1) Control	(2) Active Choice	(3) Default Change	(4) Switch ($1)	(5) Switch ($25)	(6) Switch ($10 & CC)	(7) Switch ($10 & BC)	(8) Switch ($10 & CR)	(9) Switch ($10 & BR)	F-test p-value
Survey 2	0.937	0.994	0.968	0.977	0.92	0.962	0.943	0.935	0.979	0.012
2 days after	0.966	0.97	0.98	0.988	0.981	0.968	0.961	0.981	0.967	0.886
2 weeks after	0.932	0.97	0.974	0.952	0.932	0.956	0.927	0.969	0.956	0.302
After Day 56	0.825	0.898	0.861	0.86	0.804	0.847	0.829	0.834	0.851	0.6

(b) **Baseline Bing Users**

	(1) Active Choice	(2) Switch ($10 & CC)	F-test p-value
Survey 2	0.921	0.921	1
2 days after	1	0.971	0.321
2 weeks after	1	0.943	0.156
After Day 56	0.842	0.868	0.748

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 sample in each row is a strict subset of the row above. The first row displays the share of the users who completed Survey 2. The second row displays the share of the users whose extension record is observed 2 days after Survey 2 completion time. The third row displaysthe share of the users whose extension record is observed 14 days after Survey 2 completion time. The last row shows the share of the users who was observed 8 weeks after the users completed the initial survey. The rightmost column presents the p-value of an F-test of a participant-level regression of completion indicators on the treatment group indicators.

Table A3: **Search Volume**: Searches Per Day

(a) **Baseline Google Users**

	(1) Control	(2) Active Choice	(3) Default Change	(4) Switch ($1)	(5) Switch ($25)	(6) Switch ($10 & CC)	(7) Switch ($10 & BC)	(8) Switch ($10 & CR)	(9) Switch ($10 & BR)	F-test p-value
t = 0	11.041	14.453	13.002	11.165	12.791	12.707	12.874	13.047	13.336	0.853
t = 1	10.036	14.220	12.564	12.995	15.157	16.194	14.463	14.177	14.905	0.141
t = 2	10.645	13.472	13.685	12.444	13.102	11.959	12.615	11.755	12.625	0.791
p-value (t = 0, t = 1)	0.402	0.665	0.708	0.308	0.089	0.000	0.037	0.085	0.018	-
p-value (t = 0, t = 2)	0.724	0.130	0.223	0.511	0.998	0.104	0.506	0.066	0.213	-

(b) **Baseline Bing Users**

	(1) Active Choice	(2) Switch ($10 & CC)	F-test p-value
t = 0	15.053	16.855	0.678
t = 1	14.278	14.872	0.872
t = 2	13.689	14.337	0.863
p-value (t = 0 vs t = 1)	0.576	0.258	-
p-value (t = 0 vs t = 2)	0.251	0.192	-

Notes: Panels (a) and (b) present the average number of searches per day on any search engine on each phase of the experiment for each of the treatment groups for baseline Google user and baseline Bing user samples, respectively. The phases are defined as follows: $t = 0$, $t = 2$, and $t = 3$ refer to the the days before Survey 1, the days between Survey 1 and Survey 2, and the days after Survey 2, respectively. The two bottom rows present the p-value of paired t-tests between the pre-Survey 1 average per day and the two other periods. The rightmost column presents the p-value of an F-test of a participant-level regression of the average search volume per day on the treatment group indicators.

B Details about Estimation

To translate the data from our experiment into moment conditions that fit a standard GMM framework, we think of our experiment using a potential outcomes setup. Hypothetically, there is a realized value y_{mi} of participant i's choice for every condition m. However, because of randomization, we do not observe many of these choices and the corresponding moments. To address this issue, we rewrite our moment conditions in the form

$$g_{mi}(\theta) = w_{mi} \cdot (y_{mi} - s_m(\theta)), \qquad \mathbb{E}[g_{mi}(\theta^*)] = 0$$

and

$$g_{mi}(\theta) = w_{mi} \cdot \left[(1-\pi)(y_{mi} - y_{mi}^-) - (y_{mi}^+ - y_{mi})\right], \qquad \mathbb{E}[g_{mi}(\theta^*)] = 0$$

where w_{mi} are weights that allow us to account for the fact that some of the choices y_{mi} are unobserved. Whenever y_{mi} is not observed, we simply set $w_{mi} = 0$. When y_{mi} is observed, we set w_{mi} to be the inverse of the (empirical) probability that we observe y_{mi} conditional on i's baseline search engine. Under these weights, it is still the case that $\mathbb{E}[g_{mi}(\theta^*)] = 0$ despite the fact that some of these choices are unobserved and that this occurs with different probabilities for baseline Bing and Google participants.

C Computation of Counterfactuals and Consumer Surplus

Consider the utility of agent i in some counterfactual $\mathscr{C}$.The difference in the user's perceived utilities—the utility that drives choices—can be written as

$$\Delta u_{i,\mathscr{C}} = \Delta v_{\mathscr{C}} + \Delta b_{\mathscr{C}} + \Delta \varepsilon_i,$$

where $\Delta v_{\mathscr{C}}$ denotes differences in true mean utilities, $\Delta b_{\mathscr{C}}$ denotes additional differences due to biases like misperceptions or non-welfare relevant switching costs, and $\Delta \varepsilon_i$ denotes differences in idiosyncratic preferences. To illustrate these terms, we now consider what they look like in the status quo. If we assume that switching costs are welfare relevant, the term for true utilities is $\Delta v_{\mathscr{C}} = \Delta\zeta - \sigma(1-\delta)$ to account for the difference in the quality of the search engine and for the switching cost. The bias term is $\Delta b_{\mathscr{C}} = \zeta_{-o}^* - \zeta_{-o}$ since the user is not aware of the true quality of the alternative search engine. We also consider the more general case in which a fraction $\omega \in [0,1]$ is welfare relevant. In that case, the remaining fraction of the switching cost $(1-\omega)\sigma(1-\delta)$ is part of $\Delta b_{\mathscr{C}}$.

The market share of search engine n is given by

$$s_{-o,\mathscr{C}} = S(\Delta v_{\mathscr{C}} + \Delta b_{\mathscr{C}}),$$

where $S(\cdot)$ is the CDF of the difference in idiosyncratic preferences. If that difference is uniformly distributed, $S(x) = \begin{cases} 0 & x \leq \frac{1}{2} \\ x + \frac{1}{2} & -\frac{1}{2} < x \leq \frac{1}{2} \\ 1 & x > \frac{1}{2} \end{cases}$, and if it follows a logistic distribution, $S(x) = \frac{1}{1+e^{-x}}$.

We now derive an expression for consumer surplus. In the absence of any misperceptions, the consumer surplus relative to the utility of the original search engine is given by $\frac{1}{\eta} V(\Delta v_{\mathscr{C}})$, where $V(x) = \int_{-\infty}^{x} S(x')dx'$. If the difference between idiosyncratic preferences is uniformly distributed, $V(x) = \begin{cases} 0 & x \leq \frac{1}{2} \\ \frac{(x+\frac{1}{2})^2}{2} & -\frac{1}{2} < x \leq \frac{1}{2} \\ x & x > \frac{1}{2} \end{cases}$, and if it follows a logistic distribution, $V(x) = \log(1 + \exp(x))$. The consumer surplus relative to the utility of the original search engine *in the status quo* is given by

$$CS_{\mathscr{C}} = \frac{1}{\eta}\left[V(\Delta v_{\mathscr{C}} + \Delta b_{\mathscr{C}}) - \Delta m_{\mathscr{C}} \cdot S(\Delta v_{\mathscr{C}} + \Delta b_{\mathscr{C}}) + (v_{o,\mathscr{C}} - v_{o,SQ})\right]$$

The first term in this expression gives the consumer surplus, relative to the utility of the original search engine in $\mathscr{C}$, if users's true utility is indeed described by their perceived utility. The second term is an adjustment due to biases that lead to suboptimal choices. For the share $S(\Delta v_{\mathscr{C}} + \Delta b_{\mathscr{C}})$ of users that choose the alternative search engine, their true utility is lower (or higher) to the extent that they have biases, $-\Delta b_{\mathscr{C}}$. The final term accounts for the fact that we want to measure utility relative to the status quo to be able to compare consumer surplus across counterfactuals. We therefore must adjust consumer surplus by the degree to which utility at the anchor point–the original search engine–changed relative to the status quo, $v_{o,\mathscr{C}} - v_{o,SQ}$.

C.1 Direct effects

We now give expressions for $\Delta v_{\mathscr{C}}$, $\Delta b_{\mathscr{C}}$, and $(v_{o,\mathscr{C}} - v_{o,SQ})$ in each of our counterfactuals. Note that the Data Sharing and Data Sharing + Choice Screen counterfactuals are only relevant in equilibrium, since they involve a change in the quality of search engines that arises from the use of data.

Status Quo For the status quo, $\Delta v_{SQ} = \Delta\zeta - \omega\sigma(1-\delta)$ to account for differences in utilities and for welfare-relevant switching costs. The bias term $\Delta b_{SQ} = \tilde{\zeta}_{-o} - \zeta^*_{-o} - (1-\omega)\sigma(1-\delta)$ accounts for misperceptions and for the fraction of switching costs that is not welfare-relevant. Trivially, $(v_{o,SQ} - v_{o,SQ}) = 0$.

No Frictions Since there are no switching costs, $\Delta v_{NF} = \Delta\zeta$. Since there are no biases, $\Delta b_{NF} = 0$. And the true quality of the original search engine is unchanged, so $(v_{o,NF} - v_{o,SQ}) = 0$.

Choice Screen Since there are no switching costs, $\Delta v_{CS} = \Delta\zeta$. Customers still have misperceptions about search engines, so $\Delta b_{NF} = \tilde{\zeta}_{-o} - \zeta^*_{-o}$. And the true quality of the original search engine is unchanged, so $(v_{o,NF} - v_{o,SQ}) = 0$.

Correct Perceptions True utilities do not change, so $\Delta v_{CP} = \Delta\zeta - \omega\sigma(1-\delta)$. There are no misperceptions, so $\Delta b_{CP} = -(1-\omega)\sigma(1-\delta)$. And the true quality of the original search engine is unchanged, so $(v_{o,CP} - v_{o,SQ}) = 0$.

Bing Payments Utilities change due to payments, so $\Delta v_{BP} = \eta\Delta p + \Delta\zeta - \omega\sigma(1-\delta)$. Customers still have the same biases as in the status quo, so $\Delta b_{CP} = \tilde{\zeta}_{-o} - \zeta^*_{-o} - (1-\omega)\sigma(1-\delta)$. The true quality of the original search engine is unchanged, so $v_{o,NF} - v_{o,SQ} = 0$ for Chrome users. For Edge users, we must account for the fact that the utility of using Bing increased by \$10, so $v_{o,BP} - v_{o,SQ} = \eta\Delta p$.

Default Switch Switching costs are still present but they go the other way around, so $\Delta v_{DS} = \Delta\zeta + \omega\sigma(1-\delta)$. We assume there are no misperceptions about search engines, so $\Delta b_{CP} = (1-\omega)\sigma(1-\delta)$. Finally, the true quality of the original search engine changes because it is now subject to switching costs so $(v_{o,NF} - v_{o,SQ}) = -\omega\sigma(1-\delta)$.

Bing Default Edge users behave just as in the status quo: $\Delta v_{BD,E} = \Delta\zeta - \omega\sigma(1-\delta)$, $\Delta b_{BD,E} = \tilde{\zeta}_{-o} - \zeta^*_{-o} - (1-\omega)\sigma(1-\delta)$, and $(v_{o,SQ} - v_{o,SQ}) = 0$.Chrome users behave just as in the Default Switch counterfactual:$\Delta v_{BD,C} = \Delta\zeta + \omega\sigma(1-\delta)$, $\Delta b_{BD,C} = (1-\omega)\sigma(1-\delta)$, and $(v_{o,BD,C} - v_{o,BD,C}) = -\omega\sigma(1-\delta)$.

Bing Default + Delayed Choice Screen During the first two weeks, the market behaves just as in the Bing Default counterfactual. Starting on week 3, Edge users behave as in the Status Quo. Chrome users behave as in the Perfect Information counterfactual: $\Delta v_{DCS,C} = \Delta\zeta - \omega\sigma(1-\delta)$, $\Delta b_{DCS,C} = -(1-\omega)\sigma(1-\delta)$, and $(v_{o,DCS,C} - v_{o,SQ}) = 0$. We compute market shares and consumer surplus over a span of one year: we give weight 1/26 to the first two weeks, and we give weight 25/26 to the market after week 3.

C.2 Equilibrium effects

To account for equilibrium effects, we must account for the fact that true qualities ζ_j are now a function of the share of people using search engines. Suppose that a share $\bar{s}_j$ of people use search engine j *across all browsers*. Following our economies of scale model, they are given by

$$\zeta_j(\bar{s}_j) = \rho\left[\alpha - \frac{\beta}{1-\theta} + \left(\frac{\bar{s}_j}{\bar{s}_{j,SQ}}\right)^{1-\theta}(C\hat{T}R_j - \alpha + \frac{\beta}{1-\theta})\right] + \xi_j$$

We can thus derive the following expressions for $\Delta v_{\mathscr{C}}$, $\Delta b_{\mathscr{C}}$, and $(v_{o,\mathscr{C}} - v_{o,SQ})$ in equilibrium.

No Frictions Since there are no switching costs, $\Delta v_{NF} = \Delta\zeta = \zeta^*_{-o}(\bar{s}_{-o,NF}) - \zeta^*_o(1-\bar{s}_{-o,NF})$. Since there are no biases, $\Delta b_{NF} = 0$. The true quality of the original search engine changes with the new market shares, so $v_{o,NF} - v_{o,SQ} = \zeta^*_o(1-\bar{s}_{-o,NF}) - \zeta^*_o(1-\bar{s}_{-o,SQ})$.

Choice Screen Since there are no switching costs, $\Delta v_{CS} = \Delta\zeta = \zeta^*_{-o}(\bar{s}_{-o,CS}) - \zeta^*_o(1-\bar{s}_{-o,CS})$. Customers still have misperceptions about search engines, so $\Delta b_{NF} = \tilde{\zeta}_{-o} - \zeta_{-o}(\bar{s}_{-o,CS}) - (1-\omega)\sigma(1-\delta)$. The true quality of the original search engine changes with the new market shares, so $v_{o,NF} - v_{o,SQ} = \zeta^*_o(1-\bar{s}_{-o,NF}) - \zeta^*_o(1-\bar{s}_{-o,SQ})$.

Correct Perceptions True utilities change due to new market shares, so $\Delta v_{CP} = \zeta^*_{-o}(\bar{s}_{-o,CP}) - \zeta^*_o(1-\bar{s}_{-o,CP}) - \omega\sigma(1-\delta)$. There are no misperceptions, so $\Delta b_{CP} = -(1-\omega)\sigma(1-\delta)$. The true quality of the original search engine changes with the new market shares, so $v_{o,CP} - v_{o,SQ} = \zeta^*_o(1-\bar{s}_{-o,CP}) - \zeta^*_o(1-\bar{s}_{-o,SQ})$.

Data Sharing True utilities change due to full data sharing, so $\Delta v_{DS} = \zeta^*_{-o}(1) - \zeta^*_o(1) - \omega\sigma(1-\delta)$. There are no misperceptions, so $\Delta b_{DS} = -(1-\omega)\sigma(1-\delta)$. The true quality of the original search engine changes because of data sharing, so $v_{o,DS} - v_{o,SQ} = \zeta_o(1) - \zeta_o(1-\bar{s}_{-o,SQ})$.

Data Sharing + Choice Screen True utilities change due to full data sharing, and there are no switching costs, so $\Delta v_{DS+CS} = \zeta_{-o}(1) - \zeta_o(1)$. There are no misperceptions, so $\Delta b_{DS+CS} = -(1-\omega)\sigma(1-\delta)$. The true quality of the original search engine changes because of data sharing, so $v_{o,DS+CS} - v_{o,SQ} = \zeta_o(1) - \zeta_o(1-\bar{s}_{-o,SQ})$.

Bing Payments Utilities change due to payments. Thus, $\Delta v_{BP} = \eta\Delta p + \zeta^*_{-o}(\bar{s}_{-o,BP}) - \zeta^*_o(1-\bar{s}_{-o,BP}) - \omega\sigma(1-\delta)$ for Chrome users and $\Delta v_{BP} = -\eta\Delta p + \zeta^*_{-o}(\bar{s}_{-o,BP}) - \zeta^*_o(1-\bar{s}_{-o,BP}) - \omega\sigma(1-\delta)$ for Edge users. Customers still have the same biases as in the status quo, so $\Delta b_{BP} = \tilde{\zeta}_{-o} - \zeta^*_{-o}(\bar{s}_{-o,BP}) - (1-\omega)\sigma(1-\delta)$. The true quality of the original search engine changes with the new market shares, so $v_{o,BP} - v_{o,SQ} = \zeta^*_o(1-\bar{s}_{-o,BP}) - \zeta^*_o(1-\bar{s}_{-o,SQ})$ for Chrome users. For Edge users, we also need to account for the fact that the utility of using Bing changed by \$10, so $v_{o,BP} - v_{o,SQ} = \zeta^*_o(1-\bar{s}_{-o,BP}) - \zeta^*_o(1-\bar{s}_{-o,SQ}) + \eta\Delta p$.

Default Switch Switching costs are still present but they go the other way around, so $\Delta v_{DS} = \zeta^*_{-o}(\bar{s}_{-o,DS}) - \zeta^*_o(1-\bar{s}_{-o,DS}) + \omega\sigma(1-\delta)$. We assume there are no misperceptions about search engines, so $\Delta b_{CP} = (1-\omega)\sigma(1-\delta)$. Finally, the true quality of the original search engine changes due to quality and because it is now subject to switching costs so $(v_{o,DS} - v_{o,SQ}) = \zeta^*_o(1-\bar{s}_{-o,DS}) - \zeta^*_o(1-\bar{s}_{-o,SQ}) - \omega\sigma(1-\delta)$.

Bing Default Edge users behave just as in the status quo, but qualities do change because of economies of scale: $\Delta v_{BD,E} = \Delta\zeta - \omega\sigma(1-\delta)$, $\Delta b_{BD,E} = \tilde{\zeta}_{-o} - \zeta^*_{-o}(\bar{s}_{-o,BD}) - (1-\omega)\sigma(1-\delta)$, and $(v_{o,SQ} - v_{o,SQ}) = 0$.

The true quality of Bing changes because it is now subject to switching costs, so $(v_{o,BD,E} - v_{o,BD,E}) = \zeta_o^*(1-\bar{s}_{-o,BD}) - \zeta_o^*(1-\bar{s}_{-o,SQ})$. Chrome users behave just as in the Default Switch counterfactual:$\Delta v_{BD,C} = \zeta_{-o}^*(\bar{s}_{-o,BD}) - \zeta_o^*(1-\bar{s}_{-o,BD}) + \omega\sigma(1-\delta)$, $\Delta b_{BD,C} = (1-\omega)\sigma(1-\delta)$, and $(v_{o,BD,C} - v_{o,BD,C}) = \zeta_o^*(1-\bar{s}_{-o,BD}) - \zeta_o^*(1-\bar{s}_{-o,SQ}) - \omega\sigma(1-\delta)$.

Bing Default + Delayed Choice Screen During the first two weeks, the market behaves just as in the Bing Default counterfactual. Starting on week 3, Edge users behave as in the Status Quo. Chrome users behave as in the Perfect Information counterfactual: $\Delta v_{DCS,C} = \zeta_{-o}^*(\bar{s}_{-o,DSC}) - \zeta_o^*(1-\bar{s}_{-o,DSC}) - \omega\sigma(1-\delta)$, $\Delta b_{DSC,C} = -(1-\omega)\sigma(1-\delta)$, and $(v_{o,DSC,C} - v_{o,SQ}) = \zeta_o^*(1-\bar{s}_{-o,DSC}) - \zeta_o^*(1-\bar{s}_{-o,SQ})$. We compute market shares and consumer surplus over a span of one year: we give weight 1/26 to the first two weeks, and we give weight 25/26 to the market after week 3.

C.2.1 Computing equilibria

In each of the above counterfactuals, we can plug in the above expressions into our expression for market shares to obtain the following expression for market shares among users of browser b:

$$s_{b,-o} = S(\Delta v_{b,\mathscr{C}}(\bar{s}_{-o}) + \Delta b_{b,\mathscr{C}}(\bar{s}_{-o})).$$

We can aggregate those market shares to obtain Google's total market share

$$\bar{s}_G = \frac{n_{CH} S(\Delta v_{CH,\mathscr{C}}(1-\bar{s}_G) + \Delta b_{CH,\mathscr{C}}(1-\bar{s}_G)) + n_{ED} S(\Delta v_{ED,\mathscr{C}}(\bar{s}_G) + \Delta b_{ED,\mathscr{C}}(\bar{s}_G))}{n_{CH} + n_{ED}}.$$

Finding an equilibrium consists of computing a solution $\bar{s}_{G,\mathscr{C}}$ to the above equation. Once we obtain such solution, it is straightforward to compute qualities, which we can use to compute $\Delta v_{\mathscr{C}}$, $\Delta b_{\mathscr{C}}$, and $(v_{o,\mathscr{C}} - v_{o,SQ})$. This allows us to compute equilibrium market shares and consumer surplus.

D Economy of Scale

D.1 Summary Statistics

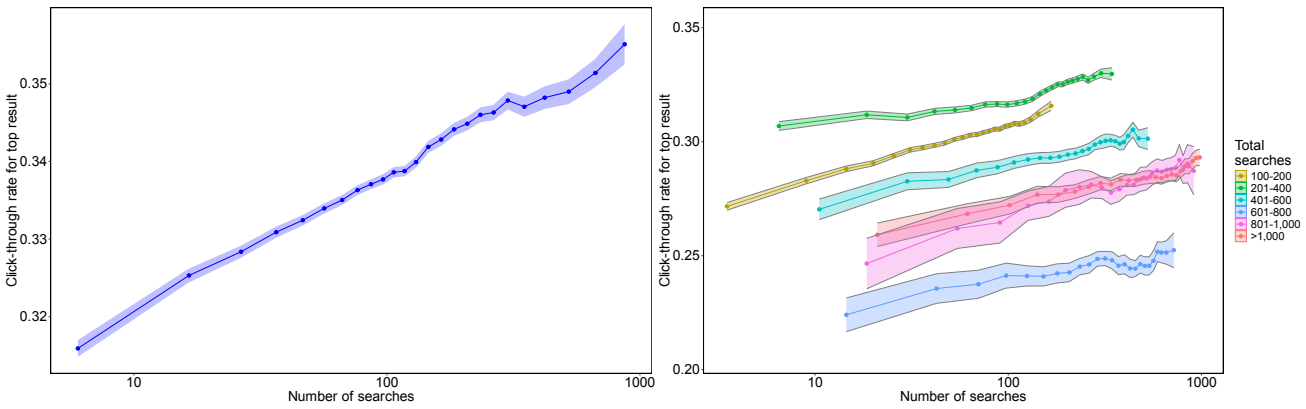
Table A4: Summary Statistics for Economy of Scale

Description	Variable	Mean	Min	Median	Max
Number of Searches	n_{qt}	374	1	155	2,980
Dummy: First Result Clicked?	r_{qt}	0.23	0	0	1
Predicted Dummy: First Result Clicked?	$\hat{r}_{qt}$	0.31	0.02	0.18	1.14
Number of Observations		12,194,034			
Number of Queries		43,991			
Number of Top-Ranked URLs		244,136			

D.2 Descriptives

Before moving on to imposing a functional form, we use binscatters to present nonparametric plots of the relationship between the predicted click-through rate and the number of searches, controlling for query fixed effects. The left plot of Figure A3 exhibits the overall relationship, which seems to be roughly log-linear: each additional doubling in the number of searches leads to an about equal increase in click-through rate. The right plot separately analyzes this relationship for queries of differing popularity: while we find that the average level of click-through rate varies by query popularity, the relationship seems to be robustly well-described as linear in the log of the number of views.

Figure A3: Non-Parametric Estimates of Returns to Scale



Notes: These graphs show the over-time relationship between the number of searches for a particular query and the rate of clicks on the top result (as disciplined by changes in the result shown first.) Formally speaking, we follow Section 6.1 and first regress the click-through rate on fixed effects for the top-ranked URL. We then use binsreg to analyze the relationship between the predicted click-through rate and the number of searches, controlling for query FE. The left plot exhibits the overall relationship, which seems to be roughly log-linear: each additional doubling in the number of views leads to an about equal increase in clickthrough-rate. The right plot separately analyzes this relationship for queries of differing popularity: while we find that the average level of clickthrough-rate varies by query popularity, the relationship seems to be robustly well-described as linear in the log of the number of views.

D.3 Implementation Details for Identification Argument

In this appendix we discuss implementation details and explain how we combine this identification strategy with controls for heterogeneity across queries and how we address the fact that ℓ_t is not a scalar in our setting.

Indexing queries by q and time by t, we first project CTR on a fixed effect for the top-ranked URL while flexibly controlling for the number of searches a query has received so far, i.e.,

$$r_{qt} = \delta_{q,u(q,t)} + \eta_{n(q,t)} + \varepsilon_{qt}, \tag{34}$$

where $u(q,t)$ gives the index of the top-ranked result served on the search result page and $n(q,t)$ gives the number of searches that query q has seen by time t. As the regression includes a fixed effect $\eta_{n(q,t)}$, we are flexibly controlling for the number of searches a query has received so far. We then use the fitted estimates $\widehat{r}_{qt} = \delta_{q,u(q,t)}$ from just the query-by-URL fixed effect as our dependent variable in estimating the relationship between searches and click-through rate. Intuitively, these fitted estimates will capture systematic improvements in CTR that are driven by Bing changing the order in which it serves search results; by contrast, they will ignore changes due to pure temporal patterns (such as a secular trend in CTR.)

We now describe an identification strategy that follows the DAG in Figure 10. We focus on a single

query and suppress the query subscript q so that observations are indexed by time t. In the DAG, n_t represents the number of searches for a given search term by time t, ℓ_t the rankings of links on the results page at t, r_t the resulting CTR, and u_t the unobserved confounders. For now, we assume these are all scalars (though they will not be in our eventual implementation.) The identification challenge is that u_t may introduce a correlation between r_t and n_t, since $u_t \to n_t$ and $u_t \to r_t$. The regression

$$r_t = \alpha + \beta n_t + \varepsilon_t$$

would therefore lead to biased estimates of β, i.e. $\lim_{n\to\infty} \hat{\beta} \neq \gamma \times \delta$.

However, the confounder does not affect the ranking quality directly: there is no arrow from u_t to ℓ_t. Hence, the following regression (Bellemare et al., 2024, eq. 7)

$$\ell_t = \kappa + \gamma n_t + \omega_t$$

will yield an unbiased estimator $\hat{\gamma}$ of the effect of additional data (n_t) on ranking of links (ℓ_t). Similarly, we can run (Bellemare et al., 2024, eq. 8)

$$r_t = \lambda + \delta \ell_t + \phi n_t + \nu_t$$

to get an unbiased estimator $\hat{\delta}$ of the effect of the ranking of links (ℓ_t) on CTR (r_t). Multiplying together these two numbers we get an unbiased estimator $\hat{\gamma} \times \hat{\delta}$ of the effect of additional data on CTR.

In practice, we do not literally follow this recipe because of the complication introduced by the fact that the ranking of links ℓ_t is not a scalar. Keeping with it being a scalar, we now explain our alternative recipe. We predict CTR from ranking quality, i.e., we run

$$r_t = \lambda + \delta \ell_t + \phi n_t + \nu_t$$

and form a prediction $\hat{r}_t = \hat{\delta} \ell_t$ of CTR based just on the current ranking quality. Then we regress this prediction on the number of prior searches n_t, i.e. we effectively run

$$\hat{r}_t = \psi + \eta n_t + \varepsilon_t.$$

As $\hat{r}_t$ here is just ℓ_t multiplied by $\hat{\delta}$, this regression must yield $\hat{\eta} = \hat{\delta} \times \hat{\gamma}$, i.e., our estimator is numerically equivalent to what Bellemare and Bloem's recipe would find.

Now, we introduce our complication: ranking quality is measured by the identity of the top-ranked URL. Let $u(t)$ index the URL top-ranked as of search t (recall we still assume there is just one search term, so we do not need indices for search terms.) Then effectively ℓ_t is a vector of dummies: assuming there are

U possible URLs that could be ranked first for this query,

$$\ell_t = (1(u(t) = 1), \dots 1(u(t) = U))'.$$

This multidimensionality of ℓ_t makes the regression that Bellemare et al. (2024) propose hard to interpret and implement. Still, our alternative way of first forming predictions of CTR works, and as argued above, in the scalar setting it would be exactly equivalent to employing the front-door criterion.

Thus, reintroducing the fact that there are multiple queries q to keep track of, in practice we follow (34). That is, we first project CTR on a fixed effect for the top-ranked URL while flexibly controlling for the number of searches a query has received so far. Subsequently, we use the fitted estimates from just the query-by-URL fixed effect in this regression as our dependent variable in estimating the relationship between searches and click-through rate.

D.4 Taylor Expansion to Address HDFE in NLLS

Our main estimating equation (35) describes a non-linear relationship between the number of previous searches for a query and its click-through rate. However, the estimating equation also features a high-dimension fixed-effect, which results in an incidental parameter problem. To address this concern, this appendix develops a methodology that utilizes repeated Taylor expansions of an estimating equation to derive exact estimates of non-linear parameters in the presence of fixed-effects.

To begin with, we find initial estimates $(\hat{\beta}^0, \hat{\theta}^0)$ by regressing demeaned c_{qt} on demeaned $\frac{\beta}{1-\theta} n_{qt}^{1-\theta}$ (computing this term and then demeaning for any trial value of the parameters.) However, as the underlying regression is not linear, the resulting estimates from this exercise are only possibly poor approximations to the true parameter values. To make progress, we turn the regression into a linear problem by utilizing a Taylor series expansion of 31 around initial estimates $(\hat{\beta}^0, \hat{\theta}^0)$. In particular,letting $\hat{\gamma}^0 = \frac{\hat{\beta}}{1-\hat{\theta}^0}$, we have

$$\hat{r}_{qt} = \alpha_q + \gamma n_{qt}^{1-\hat{\theta}^0} + \gamma(\hat{\theta}^0 - \theta)\log(n_{qt}) n_{qt}^{1-\hat{\theta}^0} + O(\theta - \hat{\theta}^0)^2 + O(\gamma - \hat{\gamma}^0)^2 + \varepsilon_{qt}.$$

As this equation is linear in easily constructed regressors $n_{qt}^{1-\hat{\theta}^0}$ and $\log(n_{qt}) n_{qt}^{1-\hat{\theta}^0}$, it can be estimated while correctly accounting for the FE α_q, thus yielding new estimates $(\hat{\beta}^1, \hat{\theta}^1)$. We can then form a new Taylor expansion around those estimates, yielding$(\hat{\beta}^2, \hat{\theta}^2)$ and so on. We iterate until convergence, and obtain standard errors via block-boostrapping (resampling at the query-level.)

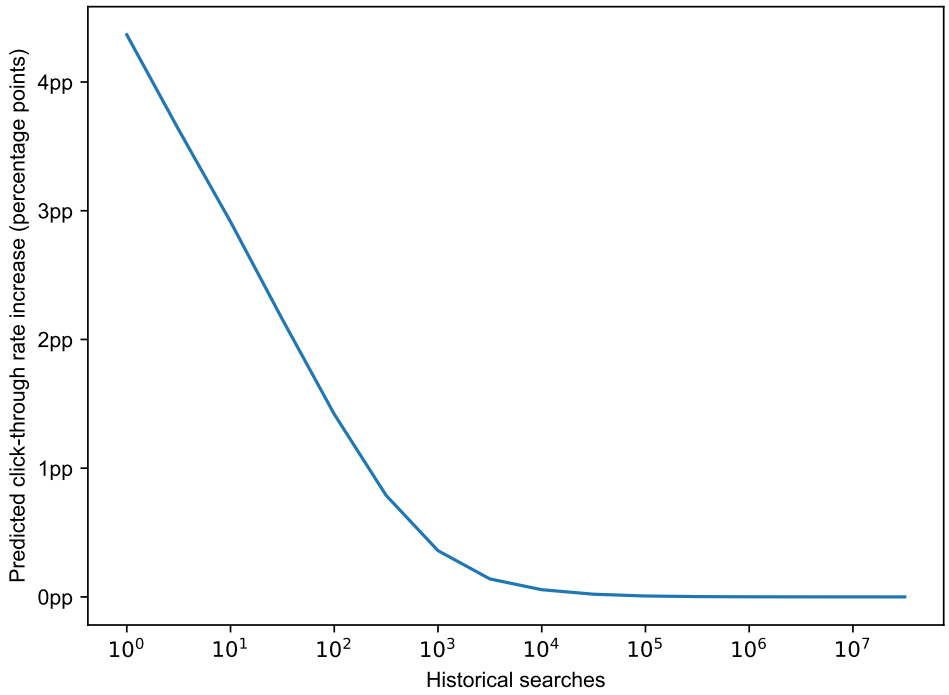
D.5 Effect of Data on Result Relevance

We now use our estimates in Table 9 to anticipate by how much Bing's CTR would increase if it were to obtain additional data, which could come either from an increase in its market share or from regulatory provisions that require the sharing of click and query data. Note that all estimates in this subsection take a partial equilibrium approach, i.e., they do not consider the effect that an improve in Bing's CTR may have

on its market share and the feedback loop that could potentially result from this effect. When moving to the full model below, we will take into account this flywheel.

Suppose Bing was to obtain an additional 1000 searches on each query. This would result in an increase of CTR from 23.50% to 25.05%, an increase of 1.55pp. We can see in Figure A4 that this increase mostly comes from an improvement in serving results on tail queries. Similarly, what if Bing multiplied it's market share by 7, making it roughly equal to Google's market share? In this case, our estimates imply that Bing's CTR would increase from 23.50% to 25.25%, an increase of 1.75pp.

Figure A4: Marginal Effect of Additional Impressions on Bing CTR by Query Popularity



Notes: This figure shows the effect (in percentage points) on CTR of increasing the number of searches that Bing observes for each query by 1000. The resulting improvement in CTR is concentrated on tail queries, which benefit the most from additional data; queries that already had large quantities of data improve less.

More generally, we can use our parameter estimates to calculate counterfactual average click-through rates on Bing if its market share were to be multiplied by λ: the new click-through rate c' after such an increase in market share is given as a function of the old click-through rate by

$$r' = \alpha - \frac{\beta}{1-\theta} + \lambda^{1-\theta} \cdot \left(r - \alpha + \frac{\beta}{1-\theta} \right)$$

D.6 Cross-Query Learning

A complication that we do not address in the main text is that learning how to rank results on query q may not be limited to using data from query q – customer behavior on impressions on related queries q' are also helpful. Denote the set of all queries (including the focal query q) as Q, and define a distance metric $d(q,q')$ that measures the distance between any focal query q and potentially related query q'. The distance between any query and itself is zero, i.e., $d(q,q) = 0$ for all q. Similar to in the main text, we will assume that the CTR on query q is given

$$\hat{r}_{qt} = \alpha_q + \beta \frac{1}{1-\theta} \left(\sum_{q' \in Q} f(d(q,q');\gamma) \times n_{q't} \right)^{1-\theta} + \varepsilon_{qt} \tag{35}$$

As before, α_q is a fixed effect that captures the fact that different queries may have different baseline CTR. More importantly, β is our measure of the value of gathering additional data, and θ measures economies of scale, i.e., the speed at which this value declines with the amount of data already gathered. In particular, $\theta \approx 0$ implies linear returns to additional data, $\theta \approx 1$ implies logarithmic returns, and $\theta > 1$ implies worse-than-logarithmic returns.

Finally, $f(\cdot)$ is a function (parameterized by γ) that maps the distance $d(q,q')$ between a focal query q and a related query q' into a monotonically declining weight. The speed at which these weights decay as we consider more and more distant queries play an important role in the economies of scale: if the weight decays only slowly, it may not matter if a search engine has never seen a query before as it can apply it's learnings from other, related queries. If the weight decays quickly, on the other hand, then not having seen a particular query before would be a strong disadvantage in trying to serve its results. Given our limited data, we will parameterize

$$f(d) = \exp(-exp(\gamma)d), \tag{36}$$

so that $\gamma = -\infty$ corresponds to no decay with distance (i.e., all views on all queries matter to CTR on any focal query) and $\gamma \to \infty$ corresponds to the case of no cross-query learning (but still allows views on the focal query to matter as $\lim_{\gamma \to \infty} \exp(-\exp(\gamma) \times 0) = 1$.)

To estimate (35), we supplement the data on new queries by obtaining, for each query in the original dataset, similar data on the 50 other queries most related to the original query, as reported by Bing's internal metrics. We emphasize that this means we have data only on the most related queries in Q; to the extent that there is little cross-query learning, we would expect this to not bias our results as searches for less related queries would not yield additional learning on Bing's side. The supplemental dataset contains, aggregated to the daily level, the total number of searches for and clicks on results pages of each of these related queries, again between 2022-01-24 and 2023-01-23. We note that, by construction, these related queries are not necessarily new; hence, we also obtain the total number of impressions between 2021-01-24 and 2022-01-23 for the related queries (this number is by definition zero for the focal queries.) Whenever we consider a

running tally of searches in our estimation below, we consider the period between 2022-01-24 and 2023-01-23 (when we see all searches) and add to the searches that have ocurred by any given date during this period the searches that occurred in the prior year, i.e., from 2021-01-24 and 2022-01-23. However, we cannot account for searches even further in the past due to Bing's retention policy for query data. Finally, we have access to Bing's internal distance measure between the related queries and the original focal query.

As before, we use the fitted estimates $\widehat{r}_{qt}$ using just the $\delta_{q,u(q,t)}$ fixed effect from regression (8) as our dependent variable in the estimation of (35). Computationally speaking, we obtain our estimates of (35) via a non-linear least squares procedure and standard errors from a block bootstrap (where a block is a focal query.) As our non-linearity correction from Section 6.2 did not yield substantively different estimates there, we avoid implementing a more complicated version of this procedure in this robustness check and simply report the parameters estimated via our demeaned non-linear least squares procedure, noting that these estimates should be interpreted with caution.

As before, we calibrate the intercept α such that the our average predicted CTR matches that from our experiment. However, a complication emerges: to take this average, we need to know for each query in the query frequency distribution how many views there are on related queries. However, we also require information on the number of views on related queries. While we have this information for the new queries (used in estimation above), we do not have this information for all queries that Bing sees. Hence, we need to predict the value of the term in parentheses, i.e., $\sum_{q' \in Q} f(d(q,q');\gamma) \times searches_{q't}$ from just the number of views on the focal query. We use the model

$$log(\sum_{q' \in Q} f(d(q,q');\gamma) \times searches_{q't}) = \beta_0 + \beta_1 log(searches_q) + u_q \tag{37}$$

We fit this equation on our sample of new queries (for which we observe views on related queries), and find $\hat{\beta}_0$ =0.0509(0.0058) and $\hat{\beta}_1$ =0.9951(0.0010) with R^2 =0.94, suggesting that we can predict this quantity very well. We can thus use

$$r_{qt} = \alpha_q + \beta \frac{1}{1-\theta} (\exp(\beta_0 + \beta_1 log(n_q)))^{1-\theta} + \varepsilon_{qt}$$

to predict CTR from just an observation of the number of views on a particular (focal) query. This allows us to calibrate α.

We exhibit our results in Table A5. Most importantly, we still find returns that are essentially logarithmic, though once taking into account spillovers, the returns are slightly more convex than logarithmic (i.e., $\hat{\theta} < 1$). Furthermore, we can strongly reject the null that additional searches have no impact on performance (i.e., β is significantly different from zero.) Finally, our estimate of γ suggests a limited role of cross-query learning. This is illustrated by Figure A5, which plots the implied weight of searches on related queries against their distance from the focal query. In particular, the horizontal axis measures the distance to the focal query in units of Bing's internal distance metric; these units are restricted to lie between zero (only assigned for identical queries) and two (practically never assigned.) The solid black line indicates the

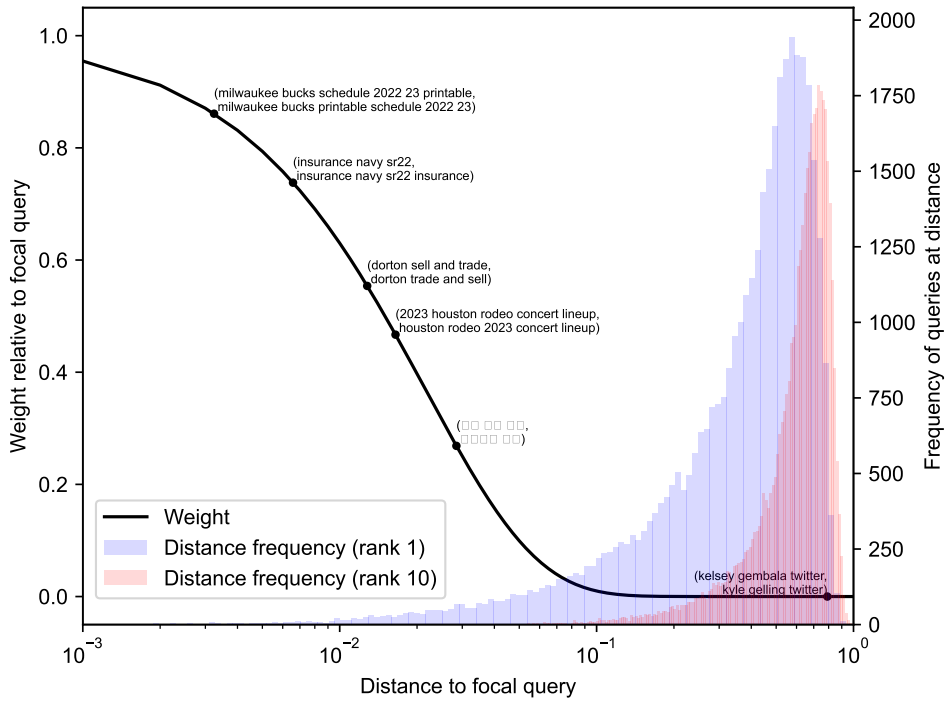
weight that our estimates imply for views on a query at a certain distance from a focal query: for instance, at distances of 0.01 our estimates imply a weight of about 0.4, suggesting that each search for a related query at this distance is worth about 40 percent of a search for the original query when it comes to learning how to rank search results. As query distances are hard to interpret, we exhibit the distribution of distances between focal queries and their top-related query (in blue) or their tenth-most related query (in red). We can see that it is rare for queries to have a related query at distance low enough to be assigned a significant weight. Finally, to aid the reader's interpretation of the query distance metric, we exhibit some randomly selected examples of pairs of queries, plotted at their distance and distance-implied weight.

Table A5: Economy of Scale Estimates

Description	Parameter	Estimate	SE
Click-through rate at inception	α	0.1744	-
Value of additional data	β	0.0056	(0.0007)
Shape of returns from data	θ	0.9272	(0.0292)
Relative weight on related queries	γ	3.8327	(0.6659)

Notes: This table provides the estimates of the parameters in (35), obtained via non-linear least squares. Standard errors are from a block-bootstrap clustered at the focal query level.

Figure A5: Cross-Query Learning



Notes: This graph illustrates the (limited) extent of cross-query learning implied by our parameter estimates. The horizontal axis measures the distance to the focal query in units of Bing's internal distance metric; these units are restricted to lie between zero (only assigned for identical queries) and two (practically never assigned.) The solid black line indicates the weight that our estimates imply for views on a query at a certain distance from a focal query: for instance, at distances of 0.01 our estimates imply a weight of about 0.4, suggesting that each view on a related query at this distance is worth about 40 percent of a view on the original query when it comes to learning how to rank search results. As query distances are hard to interpret, we exhibit the distribution of distances between focal queries and their top-related query (in blue) or their tenth-most related query (in red). We can see that it is rare for queries to have a related query at distance low enough to be assigned a significant weight. Finally, to aid the reader's interpretation of the query distance metric, we exhibit some randomly selected examples of pairs of queries, plotted at their distance and distance-implied weight.

According to the estimates in Table A5, if Bing were to increase its market share by 7x, its CTR would improve from 23.50% to 25.54%, an improvement of 2.04pp. As expected, this increase is slightly larger than the 1.75pp that we found for an increase in market share by 7 in the main text. In other words, accounting for cross-query spillovers slightly raises our estimates of the importance of economies of scale, but does not lead to any changes in qualitative conclusions.