The Costs of Entering through Tying: Experimental Evidence*

Hyunjin Kim† and Michael Luca‡

April 2018

Abstract

In recent years, many platforms have developed new products to complement their core business. For example, Google’s has expanded into Google Flights, Google Shopping, and Google Local in an effort to complement its search engine. One common approach used to gain traction in these adjacent markets has been to tie the complementary products to the primary good. In this paper, we empirically explore the strategic tradeoffs of pursuing a tying strategy to enter complementary markets, by examining Google’s decision to tie Google Local—its reviews product. We experimentally vary the content displayed on top of organic search results (called the “Onebox”), to show either exclusively Google Local reviews (the status quo), or reviews from multiple platforms that are determined to be the best-performing by Google’s own organic search algorithm. We find that tying increases search frictions, thereby reducing the likelihood of engagement with the OneBox. Overall, these results shed light on a strategic tradeoff in deciding whether to enter a complementary market using a tying strategy: while companies have successfully used this strategy, it can also come at a cost to consumers in the form of a lower-quality product.

* Byron Perpetua and Patrick Rooney provided excellent research assistance. Travis Brooks, Sebastien Couvidat, Dan Frank, and Will Seltzer provided valuable support and feedback on the design and implementation of the experiment. We thank Susan Athey, Benjamin Edelman, Joshua Gans, and Timothy Wu for constructive comments on this paper. This project received financial support from Yelp. Kim and Luca have consulting arrangements with Yelp, but their compensation and ability to publish are not tied to the results of this paper. All remaining errors are our own.

† Harvard Business School; hkim@hbs.edu

‡ Harvard Business School; mluca@hbs.edu
1. Introduction

In 2010, Google—a company known for its search engine—launched “Google Places,” an ecosystem for listing and reviewing local businesses. At the time, Google was the dominant general search engine, with over 70% market share in search, but a relative newcomer to the reviews market where more specialized platforms, like TripAdvisor and Yelp, had already amassed millions of reviews. Entry into this market was further complicated by the presence of strong network effects: the number and quality of reviews largely determine the value of a review platform to readers, while reviewers may simultaneously be more inclined to write reviews on platforms with established readers. Obstacles to entry were heightened by a commonly observed dynamic on user-generated-content websites that much of the content is provided by a small proportion of users.

Given these challenges, Google first pulled in content from multiple reviews websites to populate its review product, often not distinguishing where the reviews had initially been posted. However, after a Federal Trade Commission (FTC) case filed in January 2013 by other reviews websites claiming that Google was not providing adequate attribution or click-through-rates, Google agreed to stop using its competitors’ content without permission. To continue to gain traction in the reviews market, Google began pursuing a tying strategy in which its reviews product—now renamed “Google Local”—was tied to its search engine, thereby placing Google reviews above all organic search results in a separate box called the “OneBox” that excluded all competitor content. This tying strategy was also used by Google when launching Google Shopping, Google News, Google Images, and Google Flights, and has been commonly pursued by other platform businesses—Microsoft entered into browsers by pre-installing Internet Explorer in Windows, and Facebook entered into messaging through Facebook Messenger.

Tying can be a successful approach to entering an adjacent market (see Carlton and Waldman (2002) and Carlton, Gans, and Waldman (2010) for a review). Yet, tying can also impose costs on customers. Prior work has focused on mechanisms of pricing to assess tying strategies, examining

---

2 While Google had initially allowed users to write a review in 2007, they made significant investments in local reviews as a standalone product in 2010, with an overhaul of Google Places (information on local businesses), launch of Google Hotpot (local search results for “places” on Google powered by reviews and ratings). See https://techcrunch.com/2011/07/21/google-places-stops-stealing-reviews/ and https://techcrunch.com/2010/04/20/google-places/
3 https://techcrunch.com/2011/06/01/google-places-borrowing-yelp-iphone-app/
whether tying may lead to price rises that may result in lower consumer welfare (Whinston (1990), Nalebuff (2004), Carlton, Gans, and Waldman (2010)). However, another dimension of consumer cost—other than the price consumers pay—is the quality of the product they can obtain. In fact, in the context of Google Local reviews, where reviews platforms and search engines do not charge consumers, quality degradation is the main dimension through which costs to the consumer might manifest.

As described by the growing literature on two-sided markets, Google acts as an intermediary to bring together buyers and sellers on two sides of the market (Rochet and Tirole (2003)). On one side are consumers who have preferences for goods and services, and on the other side are web merchants who are seeking consumers. Google provides utility by matching both sides of the market. In particular, Google’s utility to consumers comes through reduced search frictions they experience from being matched to their preferred goods or services. If Google were to provide low quality search results, it would reduce the value of search for both sides of the market—for merchants who lose consumers they may have sold to, and for consumers who do not find the goods and services they value most highly. This cost to consumers could manifest in various ways. Some consumers may eventually find their desired target, but suffer higher search costs in the process. Others may simply fail to find what they are looking for within their time constraint, leading to unfulfilled transactions. Others still may end up buying goods or services that are a worse match.

The extent to which consumers are left with a better or worse product when a firm ties a complementary product to its primary product is a priori ambiguous. On the one hand, there are clear instances where tying is not associated with any costs to consumers in terms of quality. For example, when consumers search for the current time or a solution to an arithmetic problem, Google’s tying of their clock or calculator, by showing the exact time or arithmetic solution above search results, likely benefits consumers. The search has one correct answer, meaning that consumer preferences do not play a role, and displaying that answer at the top of results can minimize search frictions.

On the other hand, consider a search for a local product or service, such as a hotel or a dentist. Such searches based on local intent comprise the largest category of search, representing roughly one third\(^5\) of total desktop search volume and over one-half of mobile search volume.\(^6\) In these searches,


there is no single objective answer, and an intermediary’s matching of consumers to their preferred goods and services plays an important role. As Eric Schmidt recently described the OneBox, it may be the case that by displaying Google-only content, “Your search just gets better and better.”

In principle, the Google Local product could be universally better than existing alternatives due to superior engineering solutions, which might then provide more useful content to users than alternative providers. In this case, users would experience lower search frictions in finding the goods and services they prefer. Alternatively, if Google Local does not provide superior content, users would incur higher search costs in sifting through what they are shown in order to find goods and services they prefer, and a worse match on average.

One might expect that market pressure would lead Google to exclude review competitors only if this leaves users better off. For example, if a worse product led users to switch to Bing for general search, then Google may have the incentive to compete purely on quality. However, there is increasing recognition that the Internet does not eliminate all search frictions. For example, users appear to display widespread inertia in product choice, with few users switching between browsers or other pre-installed programs. Furthermore, Ellison and Ellison (2009) demonstrate that firms can engage in deliberate obfuscation to make it more difficult for consumers to search and acquire information, in order to maintain market power. In a similar vein, a platform may have the incentive to increase search frictions for their own benefit. In spite of potential user costs, if these costs do not lead to sufficiently lower demand, Google could profit from this decision. Beyond the revenue opportunity from keeping users on its pages (resulting from advertising revenues, which generally increase with time spent on the platform), Google might benefit from debilitating its rivals. As Google’s content is given a favorable position over organic search results, competitors would lose their ability to gain consumers through organic search, leading to not only reduced profits but also deterioration of their products as they lose additional content that they could generate.

In this paper, we empirically investigate Google’s strategic decision to tie its local reviews product to the search engine as it entered the local reviews market. In this setting, Google has (at least) three easily implementable options in placing a OneBox on top of search results. For example, if a user searches for hotels, Google can: (1) show the OneBox with Google-only reviews, excluding any competitor content; (2) show the OneBox with the best reviews results determined by its own organic

---

8 https://www.theguardian.com/technology/2013/dec/01/default-settings-change-phones-computers
search algorithm by ingesting content from services like TripAdvisor and Yelp without licensing, a practice referred to as “scraping,” or; (3) show the OneBox with the best reviews results by seeking licensing permission from other services to use their content. It is worth noting that option (2), scraping content, was pursued by Google at the outset of Google Local but has now been prohibited by the FTC until the end of 2017, and that option (3), licensing, is that most pursued by search engines with lower market share, such as Bing and Yahoo, suggesting that there may be a relationship between platform dominance in search—the primary market—and the value of tying to the platform (see Figure 1 for a display of strategies used by main search engines in the US). However, if consumers prefer competitor review platforms but Google chooses to exclude competitor content through tying, consumers would be directed to lower quality reviews products and experience higher frictions in their search.

We leverage UsabilityHub and Mechanical Turk (MTurk) to explore. UsabilityHub is a user testing platform that provides companies with tools to test their website designs and receive feedback before launching it in the market. MTurk is an online labor market where companies can hire freelancers for short tasks that can be completed online, and is an increasingly common setting for experiments. We conducted an experiment on UsabilityHub and recruited 15,166 subjects from MTurk. Prompted with instructions that asked how they would search for a local business to visit if they were in a new city, participants were presented with a corresponding screenshot of a Google search result. All screenshots shared the same display structure of Google’s search results, where a OneBox showcasing three restaurants relevant to the search sat on top of Google’s organic search results. The experimental conditions only varied what information was shown inside this OneBox, and participants were randomly assigned to one of three experimental conditions. The “Google” condition represented Google’s tying strategy, where its reviews product was tied to its primary product – the search engine. It reflected Google’s current search results where the OneBox is populated exclusively by Google Local results. By the time of this experiment (2016), Google had begun to show information snippets for restaurants along with the review content in the OneBox. In order to make the experimental conditions comparable, any information snippets that Google displayed were stripped out in the Google Condition. The second “FOTU” condition represented a counterfactual strategy where the results shown in the OneBox were modified using a Chrome Browser Extension called “Focus on the User,” which pulled review content from multiple platforms

---

determined to be the highest ranking by Google’s organic search algorithm. The plugin detected the presence of a OneBox in the Google search results and conducted a separate search in the background to extract the most highly ranked local review results. The FOTU condition showed the same Google search results as the Google screen, only modifying the review content shown in the OneBox by displaying content from multiple platforms determined to be the highest ranking by Google’s organic search algorithm. Lastly, a third condition (“Google Snippet”) showed subjects a screen of the actual Google OneBox that did not strip out information snippets, as a check for the validity of the tasks as well as the measure of clicks to assess user preferences. After each participant read the instructions and viewed their assigned screenshot, we recorded where on the screenshot they clicked as our outcome measure.

Our findings show that Google’s strategy degrades search quality. We find that users are less likely to click on the OneBox in the Google condition where only Google Local reviews are shown, suggesting that they are directed to less useful links and incur higher search frictions to obtain the results that they want. We find a 2.2 percentage point difference in the likelihood of clicking in the OneBox (representing a 5% decrease when showing Google-only content), which is statistically significant at the 5% level.

Overall, these results show that in pursuing a tying strategy, Google faces a potential tradeoff. On the one hand, tying has led to a worse primary product at least in the short run, by providing less useful search results to users of its search engine. On the other hand, tying may lead to other long-term benefits, such as entry deterrence, customer acquisition, and economies of scale and scope. Thus, the extent to which tying is a good strategic decision depends on whether the short-term costs are countered by long-run benefits that can be gained during this time. While an assessment of the benefits of tying is outside the scope of this paper, Google Local reviews has grown significantly—at almost double the rate of its competitors—since its launch in 2010, amassing 3 million reviews within its first few months (which represented one-fifth of Yelp’s content at the time). It is difficult to imagine this level of growth in a world without tying.

---

10 This plug-in is open sourced and the code may be inspected https://github.com/FocusOnTheUserLocal/FocusOnTheUserLocal
11 The mechanism for populating the experimental box was the usage of a Wilson score interval on the set of results identified by Google’s own organic algorithm.
Our paper relates more broadly to the literature on incumbents and their incentives to invest in product quality. In some contexts, incumbents may have an incentive to increasingly invest in product quality due to their reputation as market leaders (Shapiro (1982), Rob and Fishman (2005)). In other contexts, firms may coast on their reputation over time, investing less in quality. Our results provide an empirical example in which the quality of a company’s core product is intentionally degraded in the short-run in the hopes of gaining an advantage in complementary markets.

The rest of our paper proceeds as follows: the next section describes the experimental setting and design; Section 3 presents the empirical results; Section 4 concludes.

2. Experimental design

Using UsabilityHub and Mechanical Turk, we ran an online experiment on a sample of over 15,000 users across 100 US cities from October 27 to November 3, 2016. We randomly assigned subjects to one of three experimental conditions. In the Google Snippet condition, subjects were shown Google’s actual search results screen, with Google reviews and information snippets highlighted in the Onebox above organic search results. In the Google Condition, subjects were shown Google’s search results screen where Google reviews were tied to the search engine (same as Google Snippet), with the information snippets removed. In the FOTU condition, subjects were shown Google’s search results screen modified by a plugin that populated the OneBox with review content from multiple platforms determined to be the best by Google’s organic algorithm. By comparing user behavior across the Google and FOTU conditions, our goal is to understand the extent to which user preferences are affected by Google’s favoring of its own content by tying its reviews product to its search engine.

The main approach we use to measuring preferences is user clicks, which asks users to click on the screen as if they were pursuing a real search. We use an additional approach to validate this click measure, which shows users the two sets of search results side by side and asks which they prefer (TBD – need to add).

2.1 Using online labor markets to test platform design changes

The optimal way to run our experiment would be by partnering with the platform we are interested in and running the experiment directly on their users. However, in many situations,
companies may not have the incentive to either run these experiments or reveal results even after experimentation. For example, Google may not want users to know whether they are providing a degraded product. One option that we propose in this paper is to leverage user-testing tools on platforms like UsabilityHub. In this section, we discuss the potential value and limitations of using such platforms for this purpose.

UsabilityHub is a user testing platform that allows companies to create web design tests to source feedback from users before launching. It offers five types of tests: (1) click tests that ask users where they would click on a website to measure how effective a webpage design is to accomplish a task; (2) preference tests that show users two designs side by side and ask them which webpage design they prefer; (3) recall tests that ask users what they recall after viewing a webpage design for 5 seconds; (4) survey tests that ask users survey questions about the webpage design; (5) navigation tests that investigate how users navigate a webpage design. These tests represent different forms of user testing, which has become a standard part of product development at many web companies that look to improve user experiences and increase conversion by identifying improvements before launching their website. As an example, another standard user testing website, usertesting.com, boasts over 34,000 companies as customers. UsabilityHub outsources these user testing tasks to designers involved on their website and also allows the option of recruiting subjects using Mechanical Turk (MTurk)—an online labor market where employers can hire independent contractors from around the world for short tasks that can be completed online. Prior studies have illustrated the many uses of MTurk, including psychology studies, surveys, mechanisms testing, and pilot design (Kuziemko, Norton, Saez, and Stantcheva (2015)).

There are many instances in which the implications of a platform design choice may be interesting and economically important, but the relevant company does not have an incentive to pursue experimentation or allow publication of the results. Tools like UsabilityHub and MTurk can provide one way to gain some insights in these instances without a company’s direct involvement. Since UsabilityHub offers standardized tests that are used by real companies to test their webpage design, experiments can be created on their platform that mirror these tests to explore features of interface design. Many tech companies, including Google, as well as companies ranging from Amazon and eBay to TripAdvisor and Trulia, have used UsabilityHub. Furthermore, these tools provide the

---

13 Usabilityhub.com
14 https://www.usertesting.com/who
15 https://usabilityhub.com/customers
ability to rapidly and affordably scale these types of experiments.

However, one concern that such tools introduce is that of external validity. Prior studies have found that results from experiments on online platforms like MTurk are qualitatively generalizable, while point estimates may not be, depending on the question explored (Horton, Rand, and Zeckhauser (2011), Pallais (2014), Pallais and Sands (2016)). Horton, Rand, and Zeckhauser (2011) replicated three classic lab experiments using MTurk, and concluded that subjects on MTurk behave similarly to subjects in physical laboratories. They tested the following: (1) the level of cooperation in a prisoner’s dilemma; (2) the response of individuals to priming by religious texts, measured by another prisoner’s dilemma experiment; (3) individual levels of risk-aversion depending on a loss versus gain frame. In each of these experiments, they confirmed the classic results that had been found in prior studies and found no significant qualitative differences between the lab and MTurk results. Finally, they also conducted a natural field experiment (in the sense of Harrison and List (2004)), by acting as an employer and manipulating the offered wage to hired workers. They found that as expected, workers’ labor supply curves sloped upward. These findings suggest that although point estimates are difficult to generalize between any two settings, online platforms like MTurk could provide valid setting to test for economic insights, which we leverage in our paper to test hypotheses about platform choices.

2.2 Recruitment of Experimental Samples

Jobs were posted on MTurk with the title “Quick user study,” accompanied by a description stating that the task would take 1 to 3 minutes and pay $0.10. This posting was comparable to other MTurk postings, where tasks take less than an hour and pay an average effective hourly wage of $4.80 (Kuziemko, Norton, Saez, and Stantcheva (2015)). Respondents were informed that they had 3 minutes to finish the task.

Job postings limited applicants to those who spoke English. We recruited a total of 15,166 users for the experimental sample. Subjects were randomly assigned to a city-treatment condition in a single step using UsabilityHub, which assigned each user to one of three conditions through an independent draw. We selected the hundred largest US cities by population to run our experiment.

2.3 Task and Treatments

All subjects were first presented with a description of the task: “We are conducting a study on
how people use Google. When you click the survey link, you'll be prompted with a scenario asking you to imagine you've just performed a specific kind of search on Google.” They were then directed to a link at UsabilityHub where they were given the following instructions: “You are planning a meal with friends at a pizza restaurant in [city name]. You decide to conduct research online. You type "pizza [city, state]" into Google and come upon the following page. What's your first click?” The city and state named depended on which city and state each subject had been randomly assigned to out of the full set of one hundred. After confirming reading the instructions, subjects were directed to a screen of a Google search result. We then recorded where on the search page they clicked. This task was designed to be similar in nature to user testing experiments that companies widely use to improve their web page design based on user behavior.

Each subject was randomly assigned to one of three conditions. We verified that subjects were balanced on demographic traits across the treatment conditions, as shown in Appendix Table 2. The first group was shown the “Google” screen, where search results reflect Google’s current OneBox exclusively populated by Google Local reviews. While Google generally adds information snippets for each restaurant displayed in the OneBox, we removed these in order to make the comparison between the control and treatment groups comparable. The second group was shown the “FOTU” screen, where search results were modified by the FOTU plugin to show review content from multiple platforms determined by Google’s organic algorithm to be the best sites for providing the information. For example, the query for “pizza albuquerque nm” shows 182 Google Local reviews in the Google condition, while the FOTU condition shows 569 reviews from Yelp and TripAdvisor (see Figure 2). Lastly, a third condition (“Google Snippet”) showed subjects a screen of Google’s current OneBox along with the information snippets to ensure the validity of the task as well as of clicks as a measure of preferences (see Figure 2). This condition also enabled us to have an approximate comparison point on the effect of removing the snippets on user behavior. Other than the review information presented in the OneBox, all three conditions provided identical screenshots.

Across the 100 cities, the number of reviews shown in the Google and the FOTU conditions differed greatly. The FOTU screenshots on average showed three times the total number of reviews than the Google screenshots, with a median result in the Google screenshot showing 96.75 reviews compared to a median result in the FOTU screenshot with 309 reviews (see Table 1). Relative to the number of reviews that would be shown, this difference implies that Google discarded about two-thirds of the reviews in the process of excluding competitors by not using its organic search algorithm.

We recorded where on the screen subjects clicked in order to search for a local pizza place to
visit. Clicks were coded into one of three categories: OneBox, Organic, and Other. OneBox represented clicks on the OneBox displayed at the top of the screen, while Organic counted clicks on any of the organic search results displayed below the OneBox, and Other captured all clicks elsewhere on the screen. Once subjects finished clicking on the screen, they were presented with an easy addition problem to ensure that they were actively completing the tasks.

3. Results

In our experiment, one city (Corpus Christi, TX) was dropped due to an error that led to an incorrect image being shown for one of the conditions. This resulted in a sample of 99 city experiments across 15,014 subjects. Pooling the entire set of city-level search experiments, we find that Google’s favoring of its own reviews reduces users’ probability of clicking in the OneBox by 5% (significant at 5% level).

Subjects took 15 to 20 seconds to complete the task. 42.3% of subjects in the Google condition clicked in the OneBox, compared to 44.5% of subjects in the FOTU condition (see Figure 3 and Table 2). This 2.2 percentage point difference represents a decrease of 5% in clicking in the OneBox from Google’s favoring of its own reviews (see Appendix Table 3 for regression results). This result is statistically significant at the 5% level. The higher number of clicks on the OneBox in the FOTU condition suggests that users are more likely to engage with and derive use from local search results in the OneBox (thereby reducing their search frictions) when results are drawn from Google’s merit-based organic algorithm (FOTU condition) rather than systematic favoring of Google’s own content.

Table 2 shows user behavior across all conditions. We find that the probability of clicking on the OneBox in the Google Snippet condition is in the ballpark of the other conditions, which helps validate our task. Furthermore, the Google Snippet condition provides an approximate comparison point. The difference between the Google Snippet condition and the FOTU condition is statistically insignificant. Given that the Google Snippet condition provides restaurant address, hours, and description in addition to the number of reviews, the results provide suggestive evidence that providing reviews from multiple platforms has approximately a comparable effect size as providing useful information about business traits.

Table 2 also shows that while most subjects clicked on either the OneBox or the Organic results (10 blue links), a small percentage (2.8 – 3.6%, statistically insignificant difference across conditions) clicked elsewhere on the page. This behavior may be interpreted in two ways. First, it may
be seen as a meaningful data point on the effect of different presentations of OneBox information, indicating when users are likely to not engage at all with the search results. Our main results above assume this view. However, this set of subjects who clicked elsewhere on the page may have also simply ignored the task in order to quickly complete the work and receive payment. In this case, the behavior provides little meaningful signal and may distort our results. As a check, we compute our main results leaving out this set of subjects and find that our results do not change meaningfully (see Appendix Table 1).

Taken together, these results suggest that in the short run, Google’s tying strategy that systematically favors its own content above content from other platforms comes at a cost to its users and thus degrades its primary search product. These results suggest that there may be thus a strategic tradeoff in pursuing a tying strategy between costs to the consumer and the long-run benefits that it may bring. These tradeoffs may depend on the metric that the platform wants to optimize: if Google had wanted to maximize consumer surplus, an optimal strategy may have shown Google content along with content from other platforms.

4. Discussion

Our results suggest that managers face a strategic trade-off in deciding whether to enter a complementary market by using a tying strategy. Tying a new complementary product can lead to a worse primary product, at least in the short run, as Google’s decision to favor its own content directs users to less preferred content and increases their search frictions.

Despite this cost, a number of reasons may make Google better off in the long run by pursuing this strategy. First, by favoring its own content, Google may effectively deter new entrants by pushing their links further down the search page and lowering the probability that consumers will discover them. The Onebox appears to capture a substantial proportion of user clicks, capturing at least 42% of clicks in any condition. In doing so, it significantly shrinks the market for all links displayed outside the Onebox, making it increasingly difficult for these links to attract consumers. Second, Google can further leverage its existing network effects by increasing readership, which provides an important mechanism for growth in this market. Lastly, Google may be able to encourage user substitution from organic search clicks to paid advertising clicks through universal search, as Edelman and Lai (2016) find. The combination of these benefits may allow Google to effectively enter and grow its reviews product more quickly while excluding competitors. Thus, the extent to which tying is an optimal
strategic decision depends on whether the short-term costs during this time can be countered by the long-run benefits that can be gained.

Broader trends in the review industry suggest that Google may have also received significant longer-term benefits from this strategy. At the time that Google entered the reviews market, its main competitors, TripAdvisor and Yelp, had already acquired 30 million and 15 million reviews, respectively. While Google has not officially released the current number of its total reviews, it appears that Google experienced rapid growth even in its first year of its entry into the reviews market. In the spring of 2011, Marissa Mayer (the head of Google’s local product) noted that they had already acquired 3 million total reviews since launching in 2010 – already one-fifth of Yelp’s content, which had a six-year head start. She also announced that Google Local was growing at more than 1 million reviews per month.16 In comparison, Yelp was growing at approximately 80,000 reviews per month.17 In spring 2015, an analyst report from Piper Jaffray examined a snapshot of 500 places in 20 major US cities on a monthly basis, and found that Google Local reviews were still growing more quickly at 6.0% per quarter, while Yelp was growing at 3.9% per quarter.18

This rate of growth in Google Local appears to be accompanied by a decrease in traffic to Yelp from Google over this time. Appendix Figure 1 shows Desktop traffic levels to Yelp from search engines, where the share of Desktop traffic from Google is decreasing at a faster rate compared to other search engines beginning in 2012. Over this time period, Google’s overall share of search does not appear to be decreasing (see Appendix Figure 2). Given that these plots cannot tease apart the effect of a number of other actions that Google took over this timeframe, such as buying Zagat and providing tangible incentives for Google users to write reviews, as well as any actions by competitors like Yelp and TripAdvisor, no firm conclusions can be drawn from these plots. However, these plots provide suggestive evidence that as Google entered the reviews market by tying its product to its search engine, it was able to leverage its network effects and reduce the growth of its incumbent competitors.

This also relates to an emerging discussion about dominant online platforms, such as Facebook, Google, and Amazon, which have been drawing increasing scrutiny from policymakers. This scrutiny has come as a part of mounting concerns about rising industry concentration in the

18 http://www.smarteranalyst.com/2015/06/04/yelp-inc-versus-google-inc-plus-review-growth-piper-jaffray/
United States – that large firms are accounting for higher shares of industry profits relative to their investment. On the one hand, higher profits and increased concentration may simply be a result of technological innovation and competition, where “superstar firms” with higher productivity capture a larger share of the market due to their ability to innovate and offer better products. On the other hand, concentration may also arise from (and further enable) anticompetitive forces where incumbent firms are increasingly able to prevent rivals from entering and growing, even when they can offer better quality products to consumers. While there has been increasing evidence that industry concentration is on the rise, there has been less insight on whether or to what extent consumers are being harmed by this consolidation (Autor et al. 2017, Rognlie 2015, Barkai 2016, de Loecker and Eeckhout 2017). Our findings suggest that dominant platforms may, at times, be degrading products for strategic purposes, such as excluding competitors in adjacent markets that they are looking to enter or grow in.

One further direction for future research may be to investigate the extent to which switching costs determine whether these consumer costs from tying accrue to the platform. In principle, faced with a lower quality product, consumers could simply leave Google and use a different search engine. However, in practice, search frictions appear to be artificially high on the internet, due to deliberate obfuscation by producers and cognitive costs that lead to a larger amount of consumer inertia in product choice than what switching costs would indicate (Ellison and Ellison (2009)). In the context that we study, Google is by far the dominant platform and is often seen as a monopolist due to its high market share (Google has retained over 70% of market share on desktop and a durable 98% market share on smartphone since 2008). In the search market, multi-homing is limited, with users performing over 90% of their search queries in a given month on a single search engine, and rarely updating their browsers or default search engines accompanying their browsers (European Commission (2010)). Furthermore, the user side of the complementary reviews market Google is entering has low switching costs and higher levels of multi-homing. Given these particularities, Google likely faces lower risks of losing users to competitors by incrementally degrading its product. However, this outcome may not apply to all firms that consider pursuing a tying strategy. Rather, the dynamics of tying and its implications for users, as well as the extent to which costs to the consumer translates into costs to the firm, depends on the market context. The idea of who the dominant platform is may vary depending on the particular industry vertical (e.g. even in search, Google may not be the dominant

---

19 http://gs.statcounter.com/press/yahoo-gains-further
platform in voice search), and the implications of tying may depend on the format of different ecosystems such as desktop or smartphone browsers or applications, which can impact the extent of consumer costs incurred as well as the degree to which small inconveniences can cause customers to stay or defect.

Furthermore, future research that ties together insights from legal analyses to understand the extent to which short-run costs from product quality degradation relates to antitrust concerns may be a fruitful avenue for future research. Google’s practice of showing its own content in priority over organic search results has attracted scrutiny from competition law enforcers for unfairly excluding competitors, which prior research on antitrust policy has analyzed (Edelman (2015), Wu et al (2015)). In its defense, Google has argued that their practices serve the interests of users, doing nothing more than creating a better product. While an antitrust analysis is beyond the scope of this paper, our findings on costs to the consumer may be related to the concept of consumer harm in antitrust research and policy, and examining their implications may provide important insights for regulating competition. Our findings show how Google’s actions might affect costs to the consumer in one category across U.S. cities, but present only one data point at a particular point in time. While questions of generalizability may be difficult to tackle outside the platform, Google could run multiple experiments across broad swaths of users, geographies, and domains, which would provide insight on whether, and to what extent, consumer costs might be incurred. We hope that our findings might inform future research to further explore these questions.
References


These screenshots show the results of searching “coffee near me” and “hotel near me” in Google, Yahoo, and Bing, the three leading search engines in the US (results from October 5, 2017). Google displays its own reviews results in a Onebox placed above organic search results. Yahoo displays a box similar to Google’s Onebox, which displays content licensed from and attributed to Yelp or TripAdvisor, depending on the search. The Bing shows a set of reviews results placed above organic search results, pulling in content that is licensed from and attributed to Yelp or TripAdvisor, depending on the search. The hotel results focus on the box used by all search engines.
The first screenshot shows an example of a screen shown to subjects in the Google condition, which represents Google's entry strategy using tying. It displays the current version of Google's OneBox with only Google Local reviews, removing the information snippets about the restaurant to make the two conditions comparable. The second screenshot shows an example of a screen shown to subjects in the FOTU condition. It displays the OneBox with reviews sourced from the top results in Google's organic algorithm using the FOTU plugin. The third screenshot shows an example of a screen shown to subjects in the Google Snippet condition. It displays Google search results exactly as Google currently displays them, along with information snippets for each OneBox entry. The only difference between this condition and the Google condition is the information snippets with restaurant description, address, and hours.
This figure shows our main result comparing Google and FOTU conditions, where showing reviews from multiple platforms in the FOTU condition increases users’ probability of clicking in the OneBox. The bars represent the percentage of users that clicked in the OneBox and the confidence bands show 95% confidence intervals around the mean.
<table>
<thead>
<tr>
<th></th>
<th>Google</th>
<th>FOTU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Reviews</td>
<td>335.85</td>
<td>1063.79</td>
</tr>
<tr>
<td></td>
<td>(31.79)</td>
<td>(102.07)</td>
</tr>
<tr>
<td>Min Reviews</td>
<td>58.28</td>
<td>160.77</td>
</tr>
<tr>
<td></td>
<td>(6.42)</td>
<td>(17.97)</td>
</tr>
<tr>
<td>Median Reviews</td>
<td>96.75</td>
<td>308.98</td>
</tr>
<tr>
<td></td>
<td>(8.75)</td>
<td>(31.06)</td>
</tr>
<tr>
<td>Max Reviews</td>
<td>180.83</td>
<td>594.05</td>
</tr>
<tr>
<td></td>
<td>(18.32)</td>
<td>(67.49)</td>
</tr>
<tr>
<td>Number of Yelp Results</td>
<td>0</td>
<td>1.85</td>
</tr>
<tr>
<td>(out of 3 results shown</td>
<td></td>
<td></td>
</tr>
<tr>
<td>in OneBox)</td>
<td></td>
<td>(0.10)</td>
</tr>
<tr>
<td>Number of TA Results</td>
<td>0</td>
<td>0.87</td>
</tr>
<tr>
<td>(out of 3 results shown</td>
<td></td>
<td></td>
</tr>
<tr>
<td>in OneBox)</td>
<td></td>
<td>(0.093)</td>
</tr>
<tr>
<td>Observations</td>
<td>99</td>
<td>99</td>
</tr>
</tbody>
</table>

This table compares the distribution of review content shown in screenshots of Google and FOTU conditions across the 100 cities in the experiment. Means and standard errors of the mean in parentheses are displayed for each variable.
<table>
<thead>
<tr>
<th></th>
<th>Google</th>
<th>FOTU</th>
<th>Google Snippet</th>
</tr>
</thead>
<tbody>
<tr>
<td>OneBox</td>
<td>0.42</td>
<td>0.45</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>Organic Search Results</td>
<td>0.55</td>
<td>0.52</td>
<td>0.52</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>Other</td>
<td>0.03</td>
<td>0.04</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>Observations</td>
<td>5011</td>
<td>5006</td>
<td>4997</td>
</tr>
</tbody>
</table>

This table shows the percentage of subjects that clicked within the OneBox, Organic Search Results, and elsewhere on the page (Other) across all three conditions, FOTU, Google, and Google Snippet. The standard error of the mean is presented in parentheses. The number of observations across the treatment and control conditions was determined by UsabilityHub, which randomized each individual into one of three conditions through an independent draw.
Appendix

Table 1 100 Cities Experiment: User Clicks across Experimental Conditions Excluding “Other”

<table>
<thead>
<tr>
<th></th>
<th>Google</th>
<th>FOTU</th>
<th>Google Snippet</th>
</tr>
</thead>
<tbody>
<tr>
<td>OneBox</td>
<td>0.44(0.01)</td>
<td>0.46(0.01)</td>
<td>0.47(0.01)</td>
</tr>
<tr>
<td>Organic</td>
<td>0.56(0.01)</td>
<td>0.54(0.01)</td>
<td>0.53(0.01)</td>
</tr>
<tr>
<td>Observations</td>
<td>4865</td>
<td>4823</td>
<td>4843</td>
</tr>
</tbody>
</table>

Means and standard errors of the mean (in parentheses) are displayed for each variable.

Table 2 100 Cities Experiment: Balance of Demographic Variables across Conditions

<table>
<thead>
<tr>
<th></th>
<th>Google</th>
<th>FOTU</th>
<th>Google Snippet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>33.03(0.14)</td>
<td>32.52**(0.14)</td>
<td>32.74(0.14)</td>
</tr>
<tr>
<td>Female indicator</td>
<td>0.42(0.01)</td>
<td>0.43(0.01)</td>
<td>0.42(0.01)</td>
</tr>
<tr>
<td>Education level</td>
<td>2.87(0.02)</td>
<td>2.87(0.02)</td>
<td>2.85(0.02)</td>
</tr>
<tr>
<td>College graduate indicator</td>
<td>0.64(0.01)</td>
<td>0.63(0.01)</td>
<td>0.62(0.01)</td>
</tr>
<tr>
<td>In US indicator</td>
<td>0.80(0.01)</td>
<td>0.80(0.01)</td>
<td>0.81(0.01)</td>
</tr>
<tr>
<td>Observations</td>
<td>5011</td>
<td>5006</td>
<td>4997</td>
</tr>
</tbody>
</table>

Means and standard errors of the mean in parentheses are displayed for each variable. Education level is coded from 0 to 5 based on whether the individual’s highest degree of schooling is some high school (0), a high school degree (1), some college (2), a bachelor’s degree (3), some postgrad (4), or a graduate degree (5). The number of observations across the treatment and control conditions was determined by UsabilityHub, which randomized each individual into one of three conditions through an independent draw. Stars indicate significant statistical difference from the Google condition.
Table 3  
100 Cities Experiment: The Effect of Showing Reviews from Multiple Platforms on Probability of Clicks on the OneBox

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LPM</td>
<td>LPM</td>
<td>Logistic</td>
<td>Logistic</td>
</tr>
<tr>
<td>FOTU</td>
<td>0.02**</td>
<td>0.02**</td>
<td>0.09**</td>
<td>0.09**</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.04)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.42***</td>
<td>0.40***</td>
<td>-0.31***</td>
<td>-0.40**</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.05)</td>
<td>(0.03)</td>
<td>(0.20)</td>
</tr>
<tr>
<td>City FE</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>N</td>
<td>10017</td>
<td>10017</td>
<td>10017</td>
<td>10017</td>
</tr>
</tbody>
</table>

The logistic regression results display the odds ratio. City FE indicates a full set of city dummies. Robust standard errors are in parentheses.

* p<0.10, ** p<0.05, *** p<0.01
These figures plot user sessions across 2010 to 2016, where a “session” is defined (by Google Analytics) as a collection of interactions or hits from a specific user during a defined period of time. A session is defined to end after 30 minutes of inactivity. This metric provides one of the cleanest and most-often used measure of user traffic.
This plot shows the market share of Desktop search queries conducted across the three largest search engines, Google, Bing, and Yahoo, as measured by ComScore.