

Smaller Slices of a Growing Pie: The Effects of Entry in Platform Markets

Oren Reshef *

University of California, Berkeley

November 2019

Job Market Paper

Abstract

The entry of firms into a platform has an ambiguous effect on the profitability of incumbent firms' operating on the platform: While entry increases competitive pressure on incumbents, supply-side expansion may attract new consumers—effectively increasing total platform size and presumably benefiting all firms. Guided by a simple model, this paper explores how firm entry affects incumbents' outcomes in a two-sided market. Specifically, I focus on Yelp Transactions Platform, an online platform that connects consumers with local services. I study a quasi-exogenous increase in firms on the platform and exploit geographic variation to employ a difference-in-differences research design. I find that, on average, market expansion favors incumbents, though the average effect masks substantial heterogeneities: High-quality incumbent firms experience a positive effect, whereas low-quality firms perform unambiguously worse. Using a structural model, my analysis finds a non-monotonic relationship between market expansion and firm performance. Lastly, I use YTP's granular data on consumer and incumbent behavior to explore other market outcomes, main mechanisms, and firms' strategic responses.

*Haas School of Business, UC Berkeley (orenre@berkeley.edu). I am thankful to Steve Tadelis and John Morgan for their endless support and guidance. I am additionally grateful for the comments and suggestions of Ben Handel, Ganesh Iyer, Abhishek Nagaraj, Reed Walker, and especially Mike Luca, whose support made this project possible. I have also immensely benefited from discussions with Abhay Aneja, Itai Ater, Ned Augenblick, Giovanni Compiani, Charu Gupta, Hyo Kang, Petr Martynov, Carl Shapiro, Avner Shlain, Gauri Subramani, and Miguel Villas-Boas. I thank Yelp for providing data for this analysis. The ability to publish is not tied to the results of this paper.

1 Introduction

Platform markets have now spread over diverse sectors of the economy, including retail (eBay, Amazon, Taobao), travel (Airbnb, Homeaway), services (TaskRabbit, Upwork, Uber), and finance (Kickstarter, Ant Financial).¹ As platforms spread to diverse areas of the economy, more firms and sellers begin to operate within a platform settings. It is thus important to understand the impact of platform growth and firm entry on the incumbent firms operating on the platform and platform dynamics. In traditional markets, entry of new firms exacerbates competitive pressure on incumbents and erodes their excess profits (Porter, 1989, 1997, or textbook analysis in Samuelson, 1951). However, in two-sided markets, entrants may have an indirect positive effect on incumbent firms because entrants can increase the appeal of the platform to potential consumers. If this sufficiently increases total demand, then market expansion can more than compensate for competitive pressure.

In this paper, I consider the net effect of these competing forces within a major platform market. Specifically, I address three related questions: First, how does new-firm entry affect the performance of incumbent firms in platform markets? Second, which types of incumbent firms benefit or lose the most from entry? And third, do incumbents readjust their strategies to respond to entry, and if so, how? Guided by a stylized model, I answer these questions using proprietary data from the Yelp Transactions Platform (YTP), an online platform for takeout and delivery from local restaurants. I study YTP’s partnership with Grubhub delivery service, which substantially and suddenly increased the number of sellers on the platform without directly affecting the number of consumers.

I use the sharp change and geographic variation in the effect of the partnership to employ a difference-in-differences research design and find that entry benefits high-quality incumbents and raises their revenue by up to 15.8%. In contrast, entry hurts low-quality incumbents and reduces their revenue by as much as 9.2%. Hence, entry in two-sided markets changes the nature of competition by increasing the returns to higher quality and the average quality purchased. These results have important implications for firms’ platform incentives and strategies, for platform designers, and for competition policy. I complement the reduced-form results with a structural model of consumer participation and product selection, which allows me to extrapolate the results to a broader set of market conditions.

The paper fills a substantial gap in knowledge because empirically studying entry, especially in two-sided markets, is challenging given that entry and competition are often hard

¹ In fact, out of the ten firms with highest market capitalization, seven are platforms: Apple, Amazon, Alphabet, Microsoft, Facebook, Alibaba, and Tencent. Moreover, of the 2018 most promising ‘unicorns’—start-ups with valuation of over \$1 billion—about 60 to 70 percent were platform (Cusumano et al., 2019).

to isolate and measure. One characteristic muddling empirical analysis is the fact that the relevant set of competitors depends heavily on the market’s definition, which may vary over time. Additionally, entrants make strategic decisions about where and when to enter, and each decision will correspond to the dynamics within and outside the immediate market. As such, entry into markets is endogenous and likely correlated with unobserved market characteristics or shocks that confound estimating the impact of entry. Consequently, while there is a huge theoretical literature on indirect network externalities and two-sided markets (Katz and Shapiro, 1985, Rochet and Tirole, 2006), the empirical literature is sparse.

I overcome these challenges using proprietary data from Yelp—the commonly used consumer-review website for local businesses—as well as YTP purchases to study all food ordering- and delivery-transactions finalized on YTP over a period of almost two years. Thanks to the type of services provided and the granular data that Yelp collects, I am able to clearly define the relevant market, market participants, and outcomes of interest. The core sample includes almost 50,000 incumbent establishments in almost 4,000 municipalities, and the main outcomes of interest are establishment-level weekly ordering quantities and weekly revenue from the platform. I also take advantage of Yelp’s core data, which includes business attributes and information on how consumers perceive establishments. Subsequent analysis uses granular data on item-level pricing, firms’ advertising expenditures on Yelp, and consumers’ search and review behavior.

Using these data, my empirical strategy exploits the event of YTP’s partnership with Grubhub, the largest food delivery service in the United States. Notably, YTP only *connects* users with delivery services (partners); these partners establish agreements with individual restaurants and handle the deliveries themselves. Thus, the number of restaurants on the platform depends on YTP’s delivery partners. Following the Grubhub partnership that commenced in February 2018, YTP users instantly gained access to Grubhub’s extensive network of affiliated restaurants, sharply increasing the number of businesses available on YTP. While the overall change in the number of businesses was substantial, there was significant variation in the impact of the partnership across geographical regions (For example, the number of restaurants available on the platform significantly increased in Berkeley, California, but remained largely unchanged in the neighboring city of Richmond). I use quasi-experimental variation in the intensity across geographic regions to employ a difference-in-differences research design in order to study the causal effect of market expansion through firm entry on the performance of incumbent businesses. Since Grubhub’s network of affiliated restaurant is not randomly assigned, I perform several tests to verify the validity of the research design, as described in Section 4.

To guide my empirical analysis, I develop a stylized model to demonstrate the counter-

vailing effects of entry on incumbents' performance. Initially, holding fixed the number of users on the platform, entrants compete with incumbents to capture participating consumers, and reduce the incumbents' market share. On the other hand, consumers choose whether they wish to join the platform based on their beliefs about how attractive the platform is. When more firms are present in the market, consumers' expected benefit from joining the platform increases, leading to higher participation, which increases the total market size. Increased overall demand may in turn benefit all firms in the market, offsetting the negative effects of the reduced market share on incumbents' profits. Simply put, when more firms enter, incumbents receive a smaller slice of a larger pie. Thus, whether incumbents benefit on net is an empirical question. The model in turn produces several testable predictions, most notably that higher quality firms will benefit more (or suffer less) from entry.

I find that the entry of new firms *increased* incumbent firms' weekly revenue on average by 4.5% in treated markets as compared to untreated markets. This result is best explained by increases in total platform demand: markets that experienced an increase in the number of restaurants were able to attract 36% more consumers to the platform and increase platform-level revenue by almost 60%, as compared to untreated markets. Notably, while the positive average effect is consistent with the hypothesis of countervailing forces affecting firms, the effect masks considerable heterogeneity in firms' outcomes. In particular, guided by predictions from the model, my analysis tests for heterogeneous effects across firm quality.² I find that entry *increased* weekly revenue of high-quality incumbent firms by 9.8%–15.8%. In contrast, entry *decreased* weekly revenue of low-quality firms by up to 9.2%. These results suggest that the total effect depends on firms' characteristics: Entry changes the nature of competition on the platform, increases returns to high ratings, and improves the average quality purchased by consumers. Additionally, I find that differentiation helps mitigate the negative impact of entry: Incumbents that were similar to entrants suffered a 3%–6% larger revenue loss.

In the second part of the analysis, I examine firms' responses to changes in their competitive environment and the forces driving the main result. Platform expansion intensified the positive relationship between firm quality and performance. Accordingly, I find suggestive evidence that firms responded to this change in incentives and increased their subsequent investments in quality: The average quality of firms in treated cities rose by about 1% more than control cities, an increase associated with a 3% growth in revenue. I find that entry also affected firms' advertising behaviors on Yelp. In particular, average advertising *decreased* in treated markets, and this effect was driven solely by highly-rated establishments. This

² Quality is defined by the relative Yelp star rating in the market on the eve of integration (Section 4.1).

finding suggests that market size and advertising operate as substitutes, because a firm’s improved performance reduces the benefits from advertising. Finally, while the partnership positively affected revenue on YTP, the effect on total firm revenue is unclear. One possibility is that new consumers may be substituting away from other platforms, leaving total revenue unchanged; however, using Yelp’s search data, I provide suggestive evidence that appears inconsistent with strong substitution patterns across platforms. I am also able to rule out that the main effects are driven by the ordering of search results.

I then turn to study the impacts of initial platform size. In the case analyzed here, YTP had (on average) only 4.2% of a city’s total number of restaurants even after Yelp and GrubHub partnered. It thus remains unclear how the main results extend to settings where total market size is larger—for instance, will the entry of new firms continue to increase high-quality incumbents’ revenue when 80% of a city’s restaurants are already present on the platform? To tackle this question, I impose additional functional-form assumptions on the theoretical model. I then estimate a structural model of discrete choice and consumer entry to characterize consumers’ utility function and entry-cost distributions. Using these structural estimates, I simulate firm performance for the full range of market saturation, up to the point where all firms participate on the platform.

The simulations yield several results. First, entry improves consumer welfare by 32% for the median affected market; welfare continues to increase as the fraction of establishments on the platforms grows, but at a decreasing rate. Second, high-rated businesses perform better than low-rated businesses at all levels of market entry. Finally, for almost all firms (i.e., excluding those with extremely low ratings), there is a non-monotonic, inverted U-shape relationship between the percentage of businesses on the platform and sales: When the fraction of firms on the platform is relatively small, the *market size* effect dominates and sales increase as more competitors enter the market. In contrast, when a sufficient number of businesses enter the platform, the *market share* effect dominates, and sales drop with each additional entry. Moreover, there is a monotonic positive relationship between firm rating and the number of firms on the platform needed to maximize sales (i.e., higher-rated firms prefer higher levels of firm participation on the platform). For instance, the main estimates suggest that the bliss point, the point at which profits are maximized, is about 46% participation for top-rated firms and only 11% for the median-rated firm.

This paper contributes to several strands of the literature. First, this work builds on the theory of network externalities and two-sided markets, as introduced by [Katz and Shapiro \(1985\)](#) and [Farrell and Saloner \(1985\)](#). Identifying indirect network externalities is extremely difficult and is usually estimated from just one market, by comparing two competing technologies, or by imposing structural assumptions on network development. (See, e.g., [Rysman](#),

2004, Springel, 2018, Ohashi, 2003, Corts and Lederman, 2009, Stremersch et al., 2007, Lee, 2013 and Nair et al., 2004). I add to this literature by providing clear evidence identifying the importance and magnitude of indirect network externalities across many markets.

Second, this paper focuses on the performance of the firms operating on the platform, and offers new insights on the forces effecting platform participants. Previously, the two-sided markets and platform-strategy literatures predominantly emphasized platform-level outcomes and strategy, such as platform pricing, platform compatibility, competition between platforms, and efficient platform size.³ More recently, a small literature has begun to explore the positive spillovers generated by entry in two-sided markets (see, e.g., Li and Agarwal, 2016, Cennamo et al., 2016, Mahajan et al., 1993, Shen and Xiao, 2014, Cao et al., 2018), with most work focusing on the spillovers generated across platforms. I extend this literature by studying the expansion of sellers *within* a platform; I offer a new mechanism to generate positive spillovers across competitors and examine heterogeneous effects by firm quality and platform maturity, which remains mostly unexplored in the previous literature.

Third, and more generally, this work contributes to the industrial organization literature on entry and its effect on incumbent firms. Starting with the seminal paper of Bresnahan and Reiss (1991), the empirical literature has mostly focused on the competitive effect of entry and the downward pressure on prices (see, e.g., Berry, 1992, Nickell, 1996, Syverson, 2004, Hortaçsu and Syverson, 2007, Jia, 2008) and product variety (Illanes and Moshary, 2018). Additionally, this work has normative implications to the analysis of barriers to entry (see, e.g., Goolsbee and Syverson, 2008, Hauser and Shugan, 2008, Seamans, 2013, Kadiyali, 1996, Fudenberg and Tirole, 1984, Ellison and Ellison, 2011). In particular, I argue that the profitability of barriers depends on the characteristics of firms and that, in certain cases—such as two-sided markets—firms can perform better by welcoming entry rather than deterring it.

Finally, following the seminal work of Melitz (2003), the trade literature studied the effects of entry into new foreign markets and, similar to this paper, finds that entry differentially impacts incumbents as a function of their quality (e.g., Aghion et al., 2009, Pavcnik, 2002). Though my work focuses on consumers' information frictions as the source of positive externalities, the main results apply to alternative mechanisms posited in prior literatures, which suggests that the forces described in this paper may be extended to other settings. Some prominent examples include (1) agglomeration effects (e.g., Marshall, 1890, Murphy

³ Some prominent theoretical and empirical analyses at the platform level include: Jin and Rysman (2015), Seamans and Zhu (2013), Kaiser and Wright (2006), Cullen and Farronato (2019), Zhu and Iansiti (2012), Farronato and Fradkin (2018), Gawer and Henderson (2007), Zhu and Liu (2018), Dubé et al. (2010), Caillaud and Jullien (2003), Rochet and Tirole (2006), Armstrong (2006), and Hagiu and Wright (2015).

et al., 1989); (2) new product discovery (e.g., Bass, 1969, Fosfuri and Giarratana, 2009); and (3) taste for variety (Dixit and Stiglitz, 1977). This paper also contributes to our understanding of online-rating mechanisms and how they develop as platforms grow (Luca, 2016, Chevalier and Mayzlin, 2006, and the review in Tadelis, 2016).

The remainder of the paper is organized as follows. Section 2 presents the conceptual framework that motivates the empirical analysis. Section 3 gives an overview of YTP and its partnership with Grubhub. Section 4 lays out the data and empirical strategy. Section 5 presents the results. Section 6 describes the structural model and simulation results. Section 7 provides conclusions and outlines new opportunities for future research.

2 Conceptual Framework

The conceptual framework develops a model of consumer choice and its effect on firms' outcomes. This theoretical analysis guides the empirical analysis and is the basis for the structural estimation presented in Section 6.

2.1 Setup

The market consists of a unit continuum of consumers. Each consumer is interested in purchasing exactly one unit of the product or service and may join the platform to shop for a product. Consumers who do not join the platform receive an outside utility ω , which is the same for all consumers.⁴

The platform has a menu of N firms from which consumers may choose. Firms are characterized by quality, which can take two values, $q_j \in \{0, q\}$ with equal probability, where $q > 0$. Purchasing a product generates a mean utility equal to the seller's quality, q_j . Thus, q captures vertical differentiation between sellers. In addition, firms are horizontally differentiated in the sense that consumer i buying from seller j also receives an idiosyncratic utility shock, $\epsilon_{ij} \sim G(\epsilon)$. The utility consumer i gains from buying from seller j is given by $U_i = q_j - p_j + \epsilon_{ij}$. I denote the distribution of the difference between two random variables distributed $G(\cdot)$ as $\tilde{G}(\cdot)$

In order to use the platform, consumers must pay a one-time user-specific fixed cost, $c_i \sim H(c)$. This cost may be either monetary or the hassle cost associated with setting up an account and learning how to use the platform efficiently. Users are fully rational and know

⁴ All of the results hold with heterogeneous ω_i . I impose this restriction for the structural analysis primarily because it is difficult to separately identify heterogeneous outside options, ω_i from the heterogeneous entry costs, c_i (presented below). The structural analysis, however, allows ω to differ across *markets*.

the fundamentals of the model as well as their individual entry cost, c_i . They form (correct) beliefs on the number of firms available on the platform, their prices, the number of high and low quality firms, and the distribution of entry costs and idiosyncratic shocks. They do not, however, know the realization of the idiosyncratic utility shock, ϵ_{ij} or the prices on the platform.⁵ The search engine of the platform is extremely efficient and, conditional on using the platform, consumers immediately and costlessly observe all the sellers available on platform as well as the idiosyncratic utility shock. Consumers always choose the seller (or the outside option) that maximizes their utility.

Firms know their quality, q_j , and the quality of all other firms in the market, q_{-j} , but do not know the idiosyncratic shock, ϵ_{ij} . Firms face a constant marginal cost of r to provide one unit of the product. They compete in Bertrand-Nash competition and set prices to maximize profits.

The timing of the model is as follows: First, firms form beliefs about other firms' strategies and consumers' entry decisions, and consumers form beliefs about firms' prices and the expected value from joining the platform. Second, firms set prices. Third, entry costs are realized and consumers choose whether to join the platform. Finally, random utility shocks are realized and consumers choose the best option available to them.

2.2 Analysis

Given that all firms of the same quality level a priori face the same demand and marginal costs, in a symmetric equilibrium, $p_j = p_h \quad \forall j \in N_h$ and $p_k = p_l \quad \forall k \in N_l$, where N_h and N_l are the sets of high- and low-quality firms, respectively.

Lemma 1 *Under some restrictions on $G(\cdot)$, high-quality firms charge higher prices but provide greater mean utility to consumers than low-quality firms, $p_h > p_l > p_h - q$.*

Proof All proofs are in Appendix B.

After joining the platform, consumers enjoy zero search costs and full information. Thus, they always choose the firm j that maximizes their utility, $j = \arg \max u_{ij}$. Consumers, however, will join the platform only if their (expected) gain from joining outweighs their outside option plus the realized entry cost: $E_\epsilon[\max(u_{ij})] > \omega + c_i$.

Given consumers' behavior, firms' profit functions for a high- and low-quality firm j can

⁵ Unobserved prices are consistent with the classical search literature (e.g., [Stigler, 1961](#), [Burdett and Judd, 1983](#)).

be written, respectively, as:

$$\begin{aligned}\Pi_j^h &= H(E_\epsilon[\max(u_{ij})] - \omega) * (p_j - r) * \\ &P(\{\epsilon_{ij} > \epsilon_{ik} - p_h + p_j, \forall k \in N_h\} \cap \{\epsilon_{ij} > \epsilon_{im} - p_l + p_j - q, \forall m \in N_l\} \\ &\cap \{\epsilon_{ij} > \omega + p_j - q\})\end{aligned}\tag{1}$$

$$\begin{aligned}\Pi_j^l &= H(E_\epsilon[\max(u_{ij})] - \omega) * (p_j - r) * \\ &P(\{\epsilon_{ij} > \epsilon_{ik} - p_l + p_j, \forall k \in N_l\} \cap \{\epsilon_{ij} > \epsilon_{im} - p_h + p_j + q, \forall m \in N_h\} \\ &\cap \{\epsilon_{ij} > \omega + p_j\})\end{aligned}\tag{2}$$

The profit functions have an intuitive interpretation: The first element captures the share of users actively searching on the platform and corresponds to total *market size*. This element has the same magnitude for high- and low-quality firms. The second element is the per-unit markup. By Lemma 1, this second element is strictly larger for high-quality firms. The third element represents the probability that users choose to purchase from firm j . In particular, firm j is chosen by consumer i if the firm generates higher utility than all high-quality firms, low-quality firms, and the outside option (the entry cost is sunk). This term corresponds to the firm's *market share* of the users actively searching on the platform. This term is strictly larger for high-quality firms, since they face less competition from both high- and low-quality firms. These intuitions are formalized in the following lemma:

Lemma 2 *In equilibrium, high-quality firms sell more and generate higher revenues and profits than low-quality firms, $\Pi_j^h > \Pi_j^l$.*

The main goal of the analysis is to study the impact that entry of new sellers has on incumbent firms' profits. The following proposition motivates the two main specifications of the empirical analysis:

Proposition 1 *The effect of entry on firms' profits and revenue is:*

- (1) *Ambiguous, $\frac{\partial \Pi^h}{\partial N}$ and $\frac{\partial \Pi^l}{\partial N}$ can be positive, negative, or zero.*
- (2) *More positive (or less negative) for high-quality firms compared to low-quality firms, $\frac{\partial \Pi^h}{\partial N} > \frac{\partial \Pi^l}{\partial N}$.*

The intuition behind the first part of Proposition 1 is straightforward: On the one hand, increasing the number of firms in the market will positively affect market size; as the number of firms increases, the expected value of the *best* product increases as well, drawing more consumers into the market. On the other hand, as the number of firms increases, the probability of any specific firm to provide the highest value decreases, eroding its market

share. The relative magnitudes of these two forces determine the total impact of entry on incumbent firms.

The second part of Proposition 1 stems from the fact that both high- and low-quality firms benefit from the increase in total market size, but there is an inherent asymmetry in the effect of entry on market share. To see this, consider the case where q is extremely large. In that case, adding many low-quality firms will hardly affect the market share of a high-quality firm. Inversely, adding just one high-quality firm is likely to substantially reduce the market share of the low-quality firms. The same intuition holds in general: low-quality firms hurt more from new competitors joining the market compared to high-quality firms.

3 Setting

I apply this framework to a portion of the food delivery–service industry in the United States, a \$35 billion dollar industry that is expected to grow at an average annual rate of more than 20% in the next 10 years.⁶ The main focus is the Yelp Transactions Platform, an online platform launched in 2013 by the consumers’ review website, Yelp. YTP enables users to order food delivery and pickup from local restaurants through several food-delivery services.⁷ The empirical analysis focuses on YTP as the relevant market in which restaurants compete for users interested in food-delivery services.

YTP operates as a part of the standard Yelp website and features a subset of restaurants available on Yelp. Shoppers are automatically directed to the platform by applying the “Delivery” or “Takeout” filters, or by using similar words in a search query.⁸ Figure 1a depicts the results of a search query on YTP: the shopper views a list of restaurants relevant to the query and user’s location as well as a map of the establishments. Restaurants’ data are pulled from the standard Yelp website and include the star rating, number of reviews, food category, and dollar rating. Shoppers can then go to the business page to learn more about the restaurant or initiate an order. Initiating an order redirects users to the restaurant menu page, presented in Figure 1b. Consumers can then choose the specific menu items they are interested in and finalize the transaction. Notably, the entire order process is native to YTP and is finalized without users being redirected to external websites.

⁶ According to a UBS Investment Bank [report](#).

⁷ YTP allows users to transact with a myriad of other local businesses, including hotels, home services, and local services (such as doctors and mechanics). Due to the nature of the institutional shock of interest, this paper restricts attention to food ordering.

⁸ There are alternative ways to access YTP: First, shoppers can start an order directly from the business page. Second, when a user performs a search on the standard Yelp website, businesses that are YTP affiliates will have an “Order Now” button next to their name on the search results.

Importantly, YTP is not a delivery service but a marketplace where consumers can find food delivery and specific restaurants; delivery and takeout transactions must be implemented and handled by third-party delivery services, referred to as “partners.” For example, as seen in the bottom-right corner of Figure 1b—the order is carried out by Grubhub. Accordingly, Grubhub processes the order and sends the details to the restaurant; in the case of delivery, a Grubhub employee will also pick up the prepared food and deliver the meal to the customer’s address. Other delivery services partnering with YTP include Delivery.com, ChowNow, Eatstreet and more. The identity of the delivery service is determined automatically by the YTP algorithm and cannot be changed by the user. YTP partners with delivery services and *not* with specific restaurants. Thus, in order for a restaurant to appear on YTP, the restaurant must first contract with a delivery service, which can, in turn, sign a partnership agreement with YTP. Thus, the supply of restaurants on YTP is determined by its partners and their network of affiliated restaurants.

In August 2017, Yelp entered into a new partnership with Grubhub (the largest food delivery service in the United States to date) that would allow users to begin ordering from Grubhub through YTP beginning in February 2018.⁹ In return, Grubhub agreed to pay Yelp a fixed fee for every order sent through YTP. Thanks to this partnership, YTP users were able to access Grubhub’s extensive network of restaurants, nearly doubling the number of restaurants on YTP. Grubhub, on the other hand, gained access to Yelp’s enormous user base.^{10 11}

The partnership between YTP and Grubhub launched mid-February 2018, with the systems gradually integrating across platforms (iOS, Android, www) and geographical areas thereafter. Integration was (formally) finalized on March 19, 2018. The research design exploits the sharp and dramatic increase in the number of restaurants available on YTP in order to analyze the impact on existing restaurants.¹²

⁹ As part of the agreement, Yelp sold its subsidiary delivery service, Eat24, to Grubhub. Yelp acquired the service in 2015 for almost half the sale price. While the service had substantial presence in several areas, such as San Francisco and Miami, its market share and network of affiliated restaurants were both relatively small compared to Grubhub’s network. This agreement essentially meant that Yelp was exiting the operational side of delivery and focusing on the online interface with consumers.

¹⁰ The strategic alliance was based on the notion that Grubhub would supply the restaurants and Yelp would supply the buyers, as is evident in the joint [press release](#) following implementation: “[...] combination of Grubhub’s unmatched restaurant network and efficient delivery infrastructure with Yelp’s large purchase-oriented audience. [...] Yelp users will be able to order from far more local restaurants [...] from Grubhub’s huge network of local favorites.”

¹¹ Note that the partnership did not imply exclusivity: Grubhub restaurants remained active on its own website, and YTP still featured businesses affiliated with other delivery services.

¹² To be conservative, I code February 19, a month before completing the integration, as the first week of treatment. In Section D.2 I discuss alternative definitions of the integration dates: First, defining the integration date as March the 19th, and second, excluding the period between February to March 19th. In

4 Methodology and Research Design

4.1 Data

This section provides an overview of the data used in the paper. I focus on the core pieces of data required for my findings here and relegate a more detailed account to Appendix C. Table 1 presents the descriptive statistics.

I use proprietary Yelp and YTP data covering a period of almost two years, from the beginning of 2017 to the end of 2018, with the platform integration occurring approximately at the middle of the period—beginning in February 2018. The data include all food orders completed on YTP during that period.¹³ For each transaction, I observe item-level description and price, the date on which the order was made, the identity of the user and business, and the delivery partner. I aggregate transactions to the business-week level; the main outcomes of interest are the number of weekly orders and total weekly revenue, excluding tips, taxes, and delivery fees. The final sample includes 56,493 establishments and over four million business-week observations.

The data represent all restaurants in cities where YTP is available. My business data include the Yelp business ratings, the dates on which the business joined and exited YTP, the type of food sold, the business address, and the platform’s Dollar Ratings.¹⁴ YTP ratings are based on Yelp’s Star Rating system, a user-generated rating on a one- to five-star scale. While the Yelp ratings presented to users are rounded to the nearest half star, in my analysis I use the underlying, continuous, rating. High- and low-quality businesses are defined based on their Yelp rating on the eve of integration, as opposed to the rating on a specific week, since subsequent ratings might react to treatment assignment. I define high- and low-ratings in two alternative ways: (1) *binary definition*: indicator for above- or below-median rating in the relevant geographical area; and (2) *sharp binary definition*: indicator for above the 75th percentile or below the 25th percentile rating in the relevant geographical area, i.e., above or below the median rating, excluding the interquartile range. It is valuable to note that both of these definitions also include new entrants when calculating the rating percentile, which

general, the main estimates remain statistically significant and are larger in magnitude when using either of these alternative definitions.

¹³ Unfortunately, as part of my agreement with Yelp, I am unable to disclose sensitive business information regarding the *levels* of platform or business performance. I cannot disclose, for instance, the total number of orders or users on the platform, the number of orders per business, revenue, or the precise number of users. Accordingly, all of the main results will be presented as a percentage change rather than absolute values.

¹⁴ Dollar Ratings are meant to approximate the overall cost per dinner, and are assigned by users and aggregated by Yelp. Dollar ratings on Yelp take on four discrete values: \$ = under \$10, \$\$=11-30, \$\$\$=31-60, and \$\$\$\$= over \$61.

I have chosen to do to control for the differences in quality of entering firms across markets. I also collect similar data on restaurants that do not participate in YTP; I use these data to conduct several placebo tests. Descriptive statistics on firms’ ratings, dollar ratings, and tenure on YTP are presented in Panel A for Table 1.

The data include documentation of consumers’ search processes. In particular, I use data on total usage of YTP and the number of users interested in delivery on the platform (see Section 6 for details). I also take advantage of the characteristics of search sessions that ended in a food order, including the number of searches, the number of business views, the session duration, and the search queries. Finally, I collect data on businesses’ daily advertising expenditures on Yelp.

For the structural estimation in Section 6, I use several external sources to collect city demographics: County-level data on total population, gender, age, and income come from the American Community Survey 5-Year Data for 2017. I use the IRS Individual Income Tax ZIP Code Data (2016) to provide zip code-level data on annual gross income. These zip code-level data are joined with business location to approximate the type of neighborhood the business occupies (downtown, suburbs, etc.).

Market Definitions Since fresh food can only be delivered within a reasonable distance, in this paper’s context, markets are naturally defined by geographical areas. Accordingly, in the main analysis, I define a market by a city-state combination (3,965 markets). There are several reasons to choose this definition: First, data are most complete for restaurants’ and users’ city, as opposed to county or zip-code. Second, users search on the platform by city when submitting a query. Third, delivery areas are often bounded by arbitrary limits as opposed to real distance. Finally, though I do see some deliveries across city limits, the vast majority of deliveries takes place within a given city.¹⁵

While such market definitions are logical, a potentially appealing alternative is to use city-food category combinations as the relevant market, i.e., San Francisco pizzerias will be considered a different market than sushi restaurants in San Francisco. I discuss these alternative market definitions as part of the robustness checks in Section 5.3.

Treatment Intensity Treatment intensity is defined as the change in percentage of restaurants on YTP out of the total restaurants in the city following the Grubhub integration. I approximate the total number of restaurants in the city by using the total number of

¹⁵ Nevertheless, the main results are robust to alternative definitions of local markets, including county, three-digit zip code, and five-digit zip code.

restaurants featured on Yelp. Formally, treatment intensity is defined as:

$$TI = \frac{\# \text{ restaurants on YTP Post} - \# \text{ restaurants on YTP Pre}}{\text{Total numbers of restaurants on Yelp}} \quad (3)$$

To give a concrete example, if a city has a total of 50 restaurants, and 20 were listed on YTP before the partnership with Grubhub, when 10 new businesses were added following the partnership, the treatment intensity would be coded as 20% (10 over 50). Thus, treatment intensity captures the *change in percentage points* of restaurants on YTP. It is important to standardize the absolute number of restaurants added by potential market size since we would expect an addition of 100 new restaurants to have different implications in very large metropolitan areas compared to small towns.

Table 1 presents descriptive statistics on YTP participation and the average treatment intensity. Prior to the partnership with Grubhub, only a small percentage (3.2% on average) of restaurants in a city were available on YTP. While the increase in the share of businesses on the platform is substantial relative to the baseline, even after integration, on average, only 4%–5% of restaurants in the city are available on the platform.^{16 17}

4.2 Research Design

The partnership between YTP and Grubhub provides a platform-level institutional shock to the number of restaurants on the platform. Figure 2 presents the change in the number of businesses over time. As in all following figures, the first week of treatment, February 19, 2018 (8th week of 2018) is normalized to nearly zero. We can see a substantial and discontinuous increase in the number of restaurants on the platform in a short period of time; the number of restaurants rises by over 60% in just a few weeks.¹⁸

While the aggregated effect on the number of restaurants available on YTP is substantial, there is significant variation in the impact across cities. In fact, the median number of

¹⁶ Notably, the average treatment intensity at the business level is substantially larger than at the city level, 2% and 1.2%, respectively. The reason for this difference is that treated cities tend to be larger and have more businesses, thus pulling the mean upwards. I address this issue in Section 5.3, Robustness Checks.

¹⁷ An alternative definition of treatment intensity uses the *percentage change* in the share of restaurants on the YTP, e.g., in the example above, treatment intensity will be coded as 50% (10 over 20). The main issue with this definition is that it mechanically introduces very large intensities in small places. In an extreme example from the data, a city with only two businesses prior to integration receives a treatment intensity of 600%. Section 5.3 shows that the main results are robust to the alternative definition of treatment intensity.

¹⁸ I attribute all new businesses added in the first eight weeks following week zero to the Grubhub integration. Over 98% of these businesses have Grubhub as their first delivery partner. I do not include businesses added after week 8, since later additions might be the result of different trajectories of market development.

businesses added to a city following the Grubhub integration is zero. Figure 3 presents the distribution of treatment intensities. Figure 3a presents the distribution of the change in the percentage of restaurants available on YTP in a city. A little less than half of the cities were not affected by the partnership, and the vast majority of cities experienced an increase of less than 5% in the percentage of firms available on the platform. Figure 3b displays the distributions of the percent of restaurants on the platform out of the total number of restaurants in the city before and after the YTP-Grubhub partnership, conditional on a non-zero change.

I use the regional variation in treatment intensity to employ a difference-in-differences analysis, wherein I compare cities with little or no change in the share of restaurants available on YTP to cities with larger changes. The key identifying assumption is that the treated cities would have had similar trends to the control cities in the absence of the Grubhub integration. Though the parallel trend assumption is not directly testable, I offer several pieces of evidence to suggest that it holds in my setting. Appendix D.1 presents a detailed discussion of the robustness checks as well as the results for all of the tests I performed, as summarized in the following paragraphs.

First of all, variation in entry originates from a platform-level institutional change, which is unlikely to be correlated with unobserved city-level trends. Second, the two main outcomes of interest, weekly revenue and number of orders, for businesses in treated and control cities trend similarly in the period preceding the partnership between Yelp and Grubhub (see Section 5), which is consistent with the parallel trends assumption. Moreover, I conduct a placebo test, in which I counterfactually set the integration date to the middle of the pre-treatment period; I do not find any differences in trends in the pre-partnership period when looking at all business, nor do I find differences when restricting attention only to high- or low-rated firms separately, which are the three main specifications used in the empirical analysis (results are presented in Appendix Table A1).

Third, if treatment effects are driven by some other, unobserved, shifts in trends that are unrelated to the Grubhub partnership, then we can expect to find significant differences in other city-level outcomes not directly related to food ordering. I conduct several placebo tests to examine whether the partnership is correlated with outcomes of *non-YTP* businesses, such as the number of businesses on Yelp, the average rating, and the number of reviews per business. I find null effects on all of these dimensions (Appendix Table A2). Finally, in this setting, treatment intensity is determined by the presence of YTP and Grubhub in a given city and the overlap between their networks of restaurants in that city, which are randomly assigned. I find that treatment assignment is correlated with city characteristics: Treated cities are, on average, larger and have more restaurants and a higher share of restaurants

on YTP. To address potential concerns that initial differences are driving the main results, I test several alternative specifications: (1) Inverse probability weighting, which accounts for the different probabilities of selection into treatment based on observables. (2) A more demanding analysis in which I reassign treatment by propensity score bins. The main results of the analysis are robust to all of these specifications (Appendix Table A6).

4.3 Empirical Specification

The primary empirical specification takes the difference-in-differences functional form:

$$Y_{jt} = \beta Post_t * Treat_j + \gamma_j + \delta_{st} + \epsilon_{jt} \quad (4)$$

Where t is the index for the week, j denotes the unit of observation (establishments in the main specification, and city when analyzing aggregate effects), and s is an index for the state. Y denotes the outcome of interest: weekly revenue and the number of orders, in the main specification. $Post$ is a binary indicator variable for whether the partnership came into effect. $Treat$ captures the treatment intensity of unit j and takes three forms: (1) *Binary treatment*, an indicator that signifies whether the treatment intensity in the city is above-median intensity across all cities. (2) *Sharp binary treatment*, a binary indicator variable that takes the value 1 whenever treatment intensity in that city is above the 75th percentile and 0 whenever treatment intensity is below the 25th percentile; since the median treatment intensity is almost zero, this sharp binary treatment definition effectively compares cities with no change to cities with treatment intensities above the 75th percentile.¹⁹ (3) The continuous measure of treatment intensity, as defined in equation 3; this quantity represents the (continuous) change in the percentage of firms on YTP out of the total number of firms in the city. γ_j denotes unit-level (business or city) fixed effect, and δ_{st} denotes state-week fixed effects.²⁰ Standard errors on all regressions are clustered at the city level, which is the level at which the treatment is administered. The parameter of interest, β , captures the causal impact of increasing the share of businesses on the outcome variable.

Motivated by the model, subsequent analyses estimate heterogeneous treatment effects by firms' quality, as measured by Yelp business ratings, using the difference-in-difference-in-

¹⁹ The rationale behind the last definition is to omit cities that are only weakly treated, which enables this research design to compare cities that received no treatment with cities that experienced meaningful treatment intensity. I use this last treatment definition in most figures and tables.

²⁰ The regression equation does not include any of the business-level covariates described in Section 4.1, because, apart from Yelp business ratings, all covariates are constant over time and are absorbed by the unit fixed effect. Ratings are not included in the regression since these are potentially affected by treatment assignment (see Section 5.4), which may consequently bias the estimates.

differences framework given by:

$$\begin{aligned}
 Y_{jt} = & \beta_1 Post_t * Treat_j + \beta_2 Post_t * Treat_b * Low_j \\
 & + \beta_3 Post_t * Low_j + \gamma_j + \delta_{st} + \epsilon_{jt}
 \end{aligned}
 \tag{5}$$

Here—beyond the indices and variables defined in equation 4— Low_j serves as the binary variable defined in Section 4.1, which captures the relative ranking of businesses at the eve of integration. β_1 captures the impact of entry on outcome Y for highly-rated businesses. Similarly, $\beta_1 + \beta_2$ captures the causal impact of entry on outcome Y for low-rated businesses. Intuitively, the heterogeneous impacts are estimated from the differential changes between treated and control cities for the relevant subset of businesses, i.e., comparing high-quality firms in treated markets to high-quality firms in untreated markets, and similarly for low-quality firms.

5 Results

5.1 Aggregate Outcomes

I begin by assessing the impact of the supply-side expansion on total market size. Figure 4 depicts the percent differences in weekly aggregated market-level outcomes between markets that are above the 75th percentile and below the 25th percentile of treatment intensity (*sharp binary treatment*). Panels 4a and 4b plot the event-time coefficients from a version of equation 4, with the market-level number of unique users and total revenue as dependent variables, respectively. Following implementation, we can see a large and steady increase in the number of weekly users as well as in revenue for affected markets compared to unaffected markets.

These figures are also useful to examine trends in the development of market outcomes prior to integration. I find similar trends between eventual treated and control markets in the period prior to integration. These pre-trends serve as suggestive evidence that the key identifying assumption of the difference-in-difference estimator, parallel trends in the absence of treatment, holds in this settings.

Table 2 presents the formal estimation results of equation 4. Column (1) uses the binary treatment definition, above- and below-median treatment intensity. Column (2) uses the sharp binary treatment definition, as describes in Section 4.3; specifically, this result ignores markets that were only ‘weakly’ treated and considers only markets that were not affected by the integration as compared with markets where treatment intensity exceeded

the 75th percentile. Column (3) uses the continuous underlying treatment intensity. Across the board, there are statistically significant (at 1 percent level) positive effects of increasing entry on all three measures of demand and total market size. We can see in Column (1) that a positive increase in the supply-side of the market leads to an increase of 36.4% in the number of users on the platform. To understand the magnitude of this effect, note that the mean percentage change in the share of businesses on YTP in treated markets is about 85% of the initial share. This outcome implies a demand elasticity of approximately 0.42 with respect to supply. The effect on the weekly number of orders is similar in magnitude (36.7%) and the effect on total revenue is slightly higher, 58.7%. As expected, Column (2) estimates, obtained by considering markets that received sharper treatment, are larger for all three outcomes. The estimate in Column (3) captures the effect of changing the share of businesses on the platform from 0% to 100%. In contrast, though, the mean increase in the share of restaurants on the platform following the partnership is only 2% (Table 1). Thus, to get a better understanding of the magnitude of the estimates in Column (3), we need to multiply the two numbers, suggesting an increase of about 20% in the number of users on the platform.

Taken together, these results support the main prediction of two-sided markets literature (Katz and Shapiro, 1985, Rochet and Tirole, 2006): An increase in the supply side of the market draws more consumers into the market and leads to subsequent increases in demand, quantity purchased, and total revenue. This finding supports the main forces described in the model and motivates the subsequent analysis.

5.2 Effect on Incumbent Firms

Average Effects After establishing that the supply-side expansion indeed increased total market size, I now turn to explore the main research question: How did entry of new firms affect incumbent firms' performance? Graphical results appear in Figure A2. Table 3 displays the estimation results of equation 4. I restrict attention solely to *incumbent businesses*, i.e., businesses that operated on YTP prior to the Grubhub integration. Table 3's structure is similar to Table 2, with Columns (1) and (2) displaying the binary and sharp binary treatment definition and Column (3) presenting the continuous measure.

Panel A showcases the effect of entry on the weekly number of orders per business. I find a null average effect of entry on the number of orders per business. The effect in Column (3) is statistically significant but is economically small, a little less than 1% in the average

city.²¹ In contrast, in Panel B, I find significant *positive* effects of about 4.5% on incumbent firms’ revenue. The positive effects on weekly revenue are statistically significant and are robust to the specific definition of treatment.

Taken together, Table 3 finds weak but positive effects of entry on firm performance. This result suggests that the countervailing forces of market size and market share described in Proposition 1 offset each other, with the market-size effect slightly dominating. This finding would have been extremely difficult to explain without taking into account the expansion in market size.

Heterogeneous Effects While the average effects are consistent with the countervailing forces affecting firms, it is difficult to assess the influence for each force separately. To address this challenge, I test the second prediction of Proposition 1, wherein entry will differentially affect high- and low-quality firms. As discussed in Section 4.1, I use the relative rating within city on the eve of integration, including for the newly added firms, as a measure of restaurant quality.

Figure 5 depicts the differences in weekly performance between treated and control markets by restaurant quality. Panel 5a plots the event-time coefficients from a version of equation 4 using the weekly number of orders for *highly-rated firms*. Similarly, Panel 5b plots the event-time coefficients for *low-rated firms*. The figure paints a clear picture: The average effect masks considerable heterogeneities in the impact of entrance on incumbent firms. In particular, entry has opposite effects on high- and low-quality businesses. For high-quality firms, we see a clear upward trend following integration, leading to substantial increases in the number of orders per week and vice versa for low-quality firms. Figure 6 presents similar and sharper results for the effect of weekly revenue. Finally, it is worth noting that in all of the figures above, we see similar trends between the eventually treated and the control markets in the period prior to integration. This observation provides further reassurance that the parallel trends assumption, central to the difference-in-difference design, holds even when restricting attention to the subset of high- or low-rated firms.

Table 4 presents the estimation results of equation 5 on businesses’ number of weekly orders and revenue as outcomes. Columns (1) and (2) present the results when using the binary definition, columns (3) and (4) use the sharp binary treatment definition, and columns (5) and (6) use the continuous measure of treatment intensity. The odd columns—(1), (3), and (5)—use the above-/below-median definition of quality, and the even columns—(2), (4), and (6)—drop the interquartile range in terms of ratings. The *Treat*Post* variable captures the effect of entry on high-quality incumbents. As we can see on Panel A Column (1), entry

²¹ The estimated effect is 38.3% and the mean treatment intensity is 2%: $38.3\% * 2\% \approx 0.8\%$.

increases the number of weekly orders by 3.6%. The positive effect of entry on high-quality incumbents is consistent across specifications and ranges from 3.6% to 6%, all significant at the one-percent level. As expected, sharper treatment and quality definitions yield stronger results: The estimated coefficient is largest in Column (4) (sharp treatment, sharp rating) and smallest in Column (1) (weak treatment, weak rating).²² The variable $Treat*Post*Low$ captures the differential effect of entry on low-rated compared to highly-rated incumbents. To get the total effect of entry on low-rated firms, we should add the coefficients for high and low types. This result is presented in the third line, marked as $\beta_1 + \beta_2$ (referring to the notation in equation 5). Estimates for low-quality sellers are the reverse mirror image of high-quality businesses: I find statistically significant *negative* effects of entry on low-quality incumbents, with the decreased weekly number of orders ranging from 2.6% to 5.4%. Similar to the effect on high-quality firms, these results increase in absolute magnitude as the treatment definition becomes sharper.

Panel B presents similar effects of entry on weekly firm revenue. Again, I find that entry leads to increases in weekly revenue for high-quality firms and vice versa for low-quality firms. The point-estimates on revenue are substantially larger than those on the number of orders: I find increases as high as 15.8% in weekly revenue for high-quality firms and drops of up to 9.2% for low-quality firms. These estimates are much noisier than in Panel A, and though always negative, the estimates on low-quality firms are statistically insignificant in Column (1) and Column (5). Nevertheless, the effects on high-quality firms and the difference in effects between high- and low-quality firms are consistently significant at the one-percent level across all specifications.

To conclude, the estimates in Table 4 complement the graphical analysis and support Proposition 1. The weak average effect masks considerable heterogeneity in firm performance following the entry of new competitors to the platform. The positive effects of entry are generated exclusively by the high-quality businesses. In contrast, low-quality firms are negatively affected by firm entry into the platform. While the highly-rated firms seem to benefit from the increased market size and the low-rated firms suffer from a decrease in market share, it would be incorrect to equate the magnitudes of these two forces with the reduced firm estimates by quality level. The reason is that both high and low types are affected by both forces, though to different degrees. The heterogeneous effects for each firm type capture different mixtures of the market-size and -share effects. In order to estimate each effect in complete isolation, I must impose additional structure on the model, which I

²² Since the estimates using continuous intensity are hard to interpret and are generally similar to the results when using the binary treatment definition, in the following tables, I will focus attention on the binary- and sharp binary-treatment definitions.

will discuss in more detail in Section 6.

5.3 Robustness of the Main Results

In this section, I briefly describe the robustness checks for the main results. A more complete description, as well as additional robustness tests, can be found in Appendix D.2. While the exact magnitudes of the treatment effects oscillate across different specifications, the main results are consistently robust to all of the alternative specifications discussed below.

Market Definition: Geographic area The main specification treats the city as the relevant market. However, this definition might be viewed as either too narrow (in large cities) or too broad (in clusters of small cities). To address these concerns, I reconstruct the data twice: first by narrowing the geographical definition of markets using the 5-digit zip code as the relevant market, and second by broadening market definition using the county as the relevant market. I find that the main effects are robust to the specific geographical definition of the relevant market. The results are presented in Table A3.

Market Definition: Food category The main specification considered only geographic boundaries in the market definition. One concern may be that other dimensions, such as food category, should be taken into account when defining relevant competitors. For instance, pizza restaurants in San Francisco might be competing with other pizza places in the city but not with sushi or Mexican restaurants. To address this concern, I reconstruct the data defining markets by city-category combinations. Potential pitfalls of this analysis are discussed in Appendix D.2. Nevertheless, the main results, presented in Table A4, are robust to the alternative market definition.

Treatment Definition So far, treatment intensity was defined as the change in the share of businesses on the platform. This definition, however, does not capture the change relative to the initial share in the market. I estimate equation 5 using:

$$TI = \frac{\# \text{ restaurants on YTP Post} - \# \text{ restaurants on YTP Pre}}{\# \text{ restaurants on YTP Pre}}$$

which is the percentage change in the share of restaurants on the platform as the relevant treatment intensity. I find that the main results, presented in Table A4, are robust to the alternative definition of treatment.

Outliers There are substantial differences between markets in the sample: Some are small towns with only a few businesses and others are huge metropolitan areas. One potential concern is that high-leverage outliers drive the results. To address this concern, I perform two sets of tests: First, I estimate the main specification while excluding the 5% largest and smallest cities. Second, I generate p-values using randomization inference tests, which

provide more robust and accurate inferences in the presence of high-leverage observations and when interacting the treatment assignment with unit characteristics (Young, 2016).²³ The main results, presented in Table A5, are robust to both of these tests.

Unbalanced Observables As mentioned in Section 4.2, there are significant differences in city characteristics across treatment and control cities. Treated cities are, on average, larger, have more restaurants, and have a higher share of restaurants on YTP. While this finding does not violate the identifying assumptions, I conduct several tests to verify that initial differences in market characteristics are not driving the results. First, I perform inverse probability weighting to account for the different probabilities of selection into treatment based on observables (Hirano et al., 2003) Second, I conduct a more demanding test, which takes advantage of the fact that treatment intensity is a continuous variable. In particular, I estimate the propensity score to receive treatment and then assign a binary treatment indicator within each propensity score bin, effectively changing the threshold for assignments into treatment as a function of the propensity score (a more detailed discussion appears in Appendix D.2). The main estimates, presented in Table A6, are robust to both of these tests.

5.4 Firm Response

In this section, I examine how firms respond to changes in market structure. I explore responses on three dimensions: pricing, investment in product quality, and advertisement.

Price Response I begin by studying how restaurants readjust prices in response to market expansion. Since restaurants sell multiple items, constructing relevant prices is not a trivial task. First, I can only estimate price changes for frequently ordered items. Second, I observe price paid rather than menu prices, which embeds noise in the price-response data, as dish modifications (e.g., “add chicken” or “make large”) are not always documented. To address these issues, first, I restrict attention to the six most ordered items in each restaurant, and second, I develop an algorithm to separate true price changes from dish modifications. The specific details are discussed in Appendix C.

Table 5 presents the results. Columns (1) and (2) present the estimates of equation

²³ Robustness inference provides a test for exact hypotheses using the random assignment of treatment instead of the (asymptotic) distribution of the error term. Intuitively, the estimation procedure iteratively reassigns units into treatment and control and estimates the treatment effect at each iteration. Then, p-values are calculated from the location of the true estimates in the distribution of estimates from potential treatment allocation. Since randomization inference is robust to small sample sizes, this approach is mostly recommended for analyzing experiments (Athey and Imbens, 2017). Nevertheless, Young (2016) shows that this methodology also performs better in settings with high-leverage observations or when interacting the treatment with unit characteristics.

4, with price as the dependent variable and item fixed effects. None of the coefficients on the average effects are statistically significant at the five-percent level. Moreover, the point estimates are generally economically small and have inconsistent signs across specifications. For instance, Column (1) finds that entry leads to an average increase of 4.4% in prices (this effect is significant at the ten-percent level), while Column (2) suggests a statistically insignificant decrease of 1.4% in average prices. Columns (3)–(6) present estimates of equation 5 using the different definitions of treatment assignment and high-quality. While the effect on prices in highly-rated restaurants seem to be generally positive and the effect on prices in low-rated restaurants are generally more negative, none of the coefficients on the heterogeneous effects are statistically significant and all are economically small.

To conclude, I do not find evidence that the entry of new businesses affected the prices of incumbent restaurants. The null effect may be an artifact of the specific setting and data limitations: First, restaurants’ prices are notoriously sticky and are often given as an example of businesses with high “menu cost” (Hobijn et al., 2006, Bils and Klenow, 2004, Zbaracki et al., 2004). Empirically, I find little price variation over time, as less than 20% of *all* food items in the sample have one or more price changes in a period of two years. Second, interviews conducted with YTP employees and delivery services suggest that online and offline prices rarely differ. Anecdotal evidence²⁴ as well as the relatively low weekly revenue from orders suggest that delivery and takeout constitute a relatively small portion of firms’ profit function and thus are not given much weight in setting prices. Finally, the data limitations discussed above may be restricting this study’s ability to detect price changes even if they are present in the data—when attempting to identify true price changes, I make quite a few assumptions and drop a significant portion of available data.

Investment in Quality The main analysis shows that entry increases revenue for high-quality sellers and decreases revenue for low-quality sellers. These findings directly imply that entry increases the *return to higher ratings*. Figure 7 demonstrates the intuition graphically. Panel 7a presents the treatment effect by rating decile. The change in revenue is monotonically increasing in rating decile and is negative for low-rated restaurants versus positive for high-rated ones. Panel 7b shows the relationship between the weekly number of orders and rating decile before and after the integration. As we can see, there is a strong positive relationship between sales and ratings prior to integration. Following integration, the trend line becomes even steeper, suggesting that moving up rankings increases sales by more than what would happen in the term prior to integration. As the return to high

²⁴ For instance, a [Morgan Stanley report](#) finds that online food delivery comprises only 6% of the total restaurant market.

quality grows, so do the incentives to invest in quality. Accordingly, if restaurants can (at least partially) affect their ratings, then we would expect to see subsequent increases in firms' ratings in treated markets.

To test this intuition, I estimate the effect of entry on the flow of incumbents' weekly ratings. Table 6 presents the estimation results of equations 4 and 5, with weekly ratings as the dependent variable. Columns (1) and (2) present the average effect and columns (3)–(6) present the heterogeneous effects by high- and low-rated restaurants. I find a small but statistically significant effect of integration on subsequent ratings, ranging between 0.6%–0.8%. The effect seems to be mostly driven by high-quality firms, though the differences between firm types are not statistically significant and the point estimates are sometimes positive. The positive effect on high-rated firms is consistent across specifications and exceeds 1% for the sharpest treatment and rating definitions.

To get a better understanding of the size of these effects, note that an increase of 1% in ratings will move the median business up by about 3 percentiles in the ratings distribution.²⁵ Using the results from Table 4, a back of the envelope calculation suggests that, for the median restaurant, a 1% increase in ratings percentile is associated with a little less than 1% increase in weekly revenue. The relatively small magnitude is unsurprising: Ratings on Yelp are a combination of reviews for delivery services as well as for the brick and mortar restaurant. It is unclear whether restaurants have much room to improve ratings, and even if so, whether this effort will have a meaningful impact on their ratings. Furthermore, reviews are relatively rare and thus extremely noisy; for the sample of incumbent restaurants, the median number of reviews per week is zero and even the 75th percentile is only one weekly review.

A potential concern is whether the estimated effects reflect true changes in quality or merely changes in rating behavior. First, results may be driven by rating inflation (Horton and Golden, 2018, Nosko and Tadelis, 2018). This concern is mitigated by the test conducted to support the identifying assumption: I test for changes in ratings trends for *non-YTP* businesses, and find null (and slightly negative) effects of entry into YTP on subsequent ratings of non-YTP businesses (Table A2). Nevertheless, to address concerns of differential rating inflation across markets, Appendix D.3 and Table A11 present a placebo specification in which integration is counterfactually coded at the middle of the pre-treatment period. The placebo test yields null results, suggesting that the effects are not driven by differential trends in rating inflation.

²⁵ In equilibrium, however, all firms will invest in quality, resulting in a red queen's race: all firms increase investment just to stay at the same rating percentile.

The second concern is that selection into specific services is correlated with rating behavior (Kovács and Sharkey, 2014, Fradkin et al., 2018). For instance, if users who use delivery services also tend to rate more leniently, then increases in online ordering might mechanically drive up the ratings of restaurants. To alleviate this concern, I tests for differential changes in raters’ leniency, which I define as the average rating across all reviews posted by the user (Table A11). I find no significant differences between raters’ leniency in treated or control markets following the integration.

Advertising The previous section examined how changes in market size affect firms’ incentives to invest in quality. I now study an alternative way for firms to attract consumers. Specifically, Yelp offers multiple services to improve firms’ appeal, with the two most common services being profile enhancement, which is a bundle designed to increase the attractiveness and conversion rate of the Yelp business page, and targeted ads. A more detailed description of the services Yelp offers to businesses is presented in Appendix C. In my analysis here, I use data on the total weekly revenue collected from businesses by Yelp, which I refer to simply as advertising expenditure.²⁶ Since treatment is administered by city, I restrict attention to campaigns purchased at the local-level, excluding national- or franchise-level campaigns.

Table 7 presents the results. Columns (1) and (2) present the average effect of market expansion on firms’ advertisement spending. I find that increased entry leads to a 2.7%–3.5% *decrease* in total firms’ spending on advertisements. Columns (3)–(6) decompose the effect by firm quality and find that the drop is concentrated among high-quality firms, which spend 3.4%–8.1% less on advertising following integration. In contrast, I do not find any consistent or statistically significant changes in spending among low-quality firms. The same patterns hold when examining the fixed and variable revenue separately, though the results on variable revenue are stronger both in terms of magnitude and statistical significance. Note that, in the period before integration, high-quality firms spent as much as 45% more on advertising than low-quality firms (consistent with Armstrong et al., 2009). Thus, though entry decreases the advertising gaps, high-quality firms still advertise more than low-quality firms in the same market.²⁷

This finding suggests that firm advertising and platform size expansion acts as substitutes; The increased sales generated by the growth in market size crowd-out the investment in advertising.²⁸ One possible mechanism is increasing marginal costs of production or, in

²⁶ Decomposition by revenue source can be found in Table A12.

²⁷ This finding does *not* have a casual interpretation, since rating and advertising behavior are not randomly assigned. For example, it is unclear whether high-rating leads to higher spending or whether more advertising leads to better reviews.

²⁸ This result is consistent with Hollenbeck et al. (2019) who find that demand generated by higher ratings substitutes advertising expenditure.

the extreme case, capacity constraints. Since I do not find any significant effect on prices, if per unit costs are increasing, then the net profit per unit is decreasing in the number of units, leading to a decrease in the marginal benefit of selling an additional unit. This consideration implies a decrease in the returns to advertising for high-quality firms. Intuitively, if firms have capacity constraints (or infinite marginal cost) then, once they reach capacity, there are zero returns to additional advertising.

5.5 Mechanisms

The Importance of Differentiation While the main analysis focuses on vertical differentiation (quality), horizontal differentiation, or the similarity between incumbents and entrants, may be an important determinant of incumbents’ outcomes. Specifically, I study how the overlap in the food category of entrants and incumbents affects incumbents’ performance.²⁹ The direction of the effect is theoretically ambiguous and relates to the forces described above: On the one hand, close alternatives compete more fiercely with incumbents, strengthening the negative *market share* effect (Hotelling, 1929, Barney, 1991). On the other hand, in a world with heterogeneous consumers, close alternatives can attract consumers who are more interested in the specific food category, increasing the positive *market size* effect.

I examine whether incumbents perform better or worse when a larger share of added restaurants are in the same food category. In particular, I regress a version of equation 4 in which I restrict attention in the analysis to cities that were affected by the partnership, and define treatment as the share of restaurants of the same food category out of the total number of entrants. For instance, if 100 new businesses joined YTP and 30 of the entrants were pizza places, then an incumbent pizza place will receive a treatment of 30%.

The results are presented in Table 8. In the odd columns, treatment is an indicator for whether any businesses of the same food category joined YTP, whereas the even columns use the continuous share of restaurants of the same food category out of the total number of entrants. The estimated effects of similarity on incumbents’ performance are negative in all specifications and are as low as -2.9% and -6.4% on the percentage change in the number of orders and revenue, respectively. The estimates, however, are only marginally statistically significant. Taken together, these results suggest that differentiation helps maintain incumbents’ competitive advantage and mitigates the deleterious effects of entry.

Is YTP Cannibalizing Revenue from Other Platforms? Table 4 shows that high-quality restaurants experienced increases in both weekly sales and revenue following

²⁹ I focus on the 21 largest food categories, which consists of about 87% of all observations. See discussion in “Market Definition: food category” in Appendix D.2 for details.

the addition of new restaurants to the platform. While it is clear that integration positively affected the revenue of high-quality restaurants on YTP, the effect of integration on the total revenue of high-quality restaurants remains unclear. In particular, the new consumers making orders on YTP might be substituting away from other forms of interaction with the restaurant, such as other delivery services.

It is hard to fully resolve these concerns since I do not have data on total firm revenue nor do I see consumers' transactions on other platforms. Nevertheless, note that it is difficult to ascribe the differential impacts by firm quality to cannibalization of other channels. In particular, such explanation would only work if platform expansion causes more cannibalization for high-quality firms compared to low-quality firms, which seems unlikely. In addition, I take advantage of the detailed search data on Yelp to generate evidence suggesting that the increase in orders from YTP is not cannibalizing other sources of revenue. Specifically, I examine whether users substitute away from ordering on other platforms. If incumbents offered delivery through other platforms, it may well be the case that, as YTP becomes more attractive, consumers leave the old platforms and switch to YTP. In this case, the total number of orders remains the same for a given restaurant, but the number of orders on YTP increases. This concern can be stated as follows: Consumers already know which restaurant to order from, and are merely selecting the channel to do so. I argue that if this situation is the case, we can expect to see changes in search patterns: We would expect consumers to search less and enter the order menu more quickly.

Table 9 presents the results of the estimating equation 4 at the city level, with search metrics as the outcomes. The sample includes only sessions in which a user ordered from a restaurant for the first time. Details regarding the construction of the sample, as well as limitations of this approach, are described in Appendix C. Panel A presents the effect of treatment on the average number of searches (search queries entered) prior to ordering. I find a weak *positive* effect on the number of searches, suggesting that integration caused users to search more intensively. Panel B presents the effect of treatment on the number of business pages viewed by users. The effects are economically small and insignificant. Panel C presents the total time spent on the platform. I find a significant and substantial *increase* in the time spent on the platform, ranging from 35% to 50%. Panel D presents the effect on the Levenshtein distance between the first search query and the name of the restaurant eventually chosen. The Levenshtein distance is defined as the number of character changes needed to move from the query to restaurant name, i.e., lower numbers imply a narrower search; if consumers already know which restaurant they want to order from, then we can expect the distance between search queries and the selected restaurants to decrease following the partnership. I find that average Levenshtein distance decreases by 3.3% to

5%, suggesting that users do use more specific queries. This result, however, is misleading; even after the reduction, the median distance in the sample is 12, which is so large that it is unlikely that consumers have a good idea of what they are looking for when beginning the search.³⁰ Moreover, almost 80% of the queries include a generic search term such as “Delivery,” “Chinese Takeout,” or “Thai food.” To examine whether integration affects the use of these generic search queries, I estimate a linear probability model with an indicator for a generic search as the outcome. The results are presented in Panel E; I find no evidence of reduction in generic searches in treated markets following integration.

Taken together, these results suggest that search intensity is not decreasing following integration and even seems to be slightly increasing. This observation is consistent with users ordering from restaurants for the first time, as opposed to simply changing the delivery platform, suggesting that the increase in revenue on YTP represents an increase in total revenue as well.

Are Consumers Responding to Firm Quality or to the Ordering of Search Results? The model argues that the main effects are driven by consumers’ selection of higher quality. A potential alternative explanation for the results is that the order of search results is affecting consumers’ choices. In particular, if Yelp ratings are highly correlated with search results’ sequence, then entry may be mechanically decreasing sales of low-quality business by reducing their salience in the search results and vice versa for high-quality types.

I first explore the importance of Yelp ratings in determining the order of search results.³¹ I find that while high-rated business are usually ranked slightly lower (we expect the relationship to be negative since lower ranks appear first), the correlation between Yelp rating and search result rank is smaller than -0.1 . Moreover, on average, improving the star-rating from two to five stars improves rankings by less than three ranks. In comparison, the mean rank in the YTP data is 18, and the difference between the average low- and high-quality businesses is about one star (3.5 and 4.5). Taken together, these results imply only a weak relationship between ratings and search results’ orders.

Second, I formally test whether controlling for search results’ orders changes the main results. To this end, I construct an index of the average weekly search-result ranks of a business, including all searches in which the business appears. As expected, the mean ranking for low-quality businesses is slightly lower than high-quality businesses, 15.5 and

³⁰ To get a sense of the magnitudes, the Levenshtein distance between the search query “Pizza” and the fictitious restaurant “Oren’s Pizza” is only 7. More alarming, the distance between “Chinese food” and the same restaurant is only 10!

³¹ Yelp’s search results ordering is proprietary, and I was not able to learn about the specific characteristics determining the orders or the relative weights given to each attribute. The observations I make are based solely on analysis of the relationship between firms’ characteristics and their relative ranks in search results.

18.5, respectively. I then re-estimate the main specification, flexibly controlling for the average weekly ranking.³² The results are presented in Table 10. I find that the main results are robust to the inclusion of average search results’ ranks. All coefficients are significant at the one-percent level and are slightly more negative compared to the main specification, which implies a stronger effect on low-quality types and (slightly) weaker effects for high-quality types. I thus conclude that firms’ quality is important even when controlling for the ordering of search results.

6 Structural Estimation

The reduced form results imply that entry has a positive impact on the sales and revenue of high-quality firms and a negative impact on low-quality firms. The main constraint of the reduced form analysis, however, is that these results are only relevant to a small segment of potential firm entry onto the platform. In particular, the median city in the sample has less than 5% of restaurants in the city on YTP even after integration. It is unclear if the results carry over to situations in which a larger percentage of the population of restaurants already participates in the market. For example, we can expect the market-size effect to have decreasing returns, i.e., the magnitude of additional increases in market size decreases when a large number of businesses are already on the platform. To study how entry affects firm performance as market participation grows, I impose additional structure on the model and the data-generating process. The structural model allows me to perform out-of-sample predictions for the full schedule of potential entrants.

6.1 Preliminaries

Setup I begin with the second stage of the model: Conditional on using the platform, consumers’ decisions follow the standard discrete choice model (McFadden et al., 1973, Berry et al., 1995). Specifically, I assume that the indirect latent utility of consumer i from buying product j in market t is:

$$U_{ijt} = \beta_i X_{jt} - \alpha_i P_j + \delta_t + \epsilon_{ijt} \tag{6}$$

Where X_{jt} are observed product characteristics, and P_j is the restaurant’s price range as captured by the Yelp Dollar Rating. δ_t are the combined statistical areas’ (CSAs’) fixed effects, and ϵ_{ijt} represents the random horizontal utility shock, assumed to be distributed

³² Formally, I estimate equation 5 with a third-order polynomial of the weekly rank index.

i.i.d. extreme value type 1. The random coefficients, β_i , are consumer-taste parameters for different product characteristics, and α_i captures the disutility from price. I allow α_i to vary across consumers according to:

$$\alpha_i = \alpha + \Pi D_i + \sigma_v v_i, \quad v_i \sim N(0, 1) \quad (7)$$

Where D_i represents observed consumer characteristics, Π is a matrix of coefficients that measures how taste characteristics vary with demographics, v_i are unobserved shock to preferences assumed to be normally distributed, and σ_v is the variance. Consumer characteristics, D_i , include gender, age, and income level. Restaurant characteristics include restaurant’s food category, Yelp rating, and mean income at restaurant location. I also include the quantile of a restaurant’s rating in the ratings distribution for the market (*ranking*) and the ranking of restaurant’s rating out of the restaurants in the same food category for the market (*category ranking*).³³

Moving to the first stage of the model, I assume that consumers’ entry costs are normally distributed and allow the distribution to vary by (potential) consumers’ income.³⁴

$$H(c) \sim N(\mu_i, \sigma), \text{ with } \mu_i = \mu + \gamma D_i \quad (8)$$

Empirically, potential market size is defined as the number of households in the city. The number of users interested in the platform is defined as the number of unique users in a given market searching for variations of the words “delivery,” “takeout,” or “pickup,” or using any of the YTP filters.³⁵ The share of consumers on the platform is given by the quintet of

³³ A few additional notes on the functional form of the utility function: 1) The utility function does not include business fixed effects since there are over 30,000 unique businesses in the final sample and only a handful of observations per business. In addition, I observe little-to-no variation over time in the covariates of interest. I use fixed effects by CSA, since this is the set from which I draw competitors in the simulation. 2) The formulation in 6 is more flexible than described in Section 2. The latter only allows for vertical and horizontal differentiation in consumers’ taste. The random coefficients specification also allows for a mixture of the two. 3) To address the concern that firms readjust prices in response to unobserved demand shocks, I instrument for dollar rating using BLP instruments (Berry et al., 1995). These instruments are likely to work well in this setting: First, as discussed in Section 5.4, I rarely see any changes in item prices, even after large shocks to the market. Second, entry of new firms is plausibly exogenous in this setting, and thus entering firms’ characteristics satisfy the instrumental variable exclusion restriction.

³⁴ First, there is no reason ex ante to assume that entry cost will follow the normal distribution, especially given the fact that we expect cost to be strictly positive. Nevertheless, I experimented with several alternative distributions, including log-normal, gamma, and exponential distributions, and the normal distribution seems to best fit the data. Second, income-levels are indicated by three levels (see Appendix C for details). Formally, I model the mean in CSA j as: $\mu_j = f_{j1} * \mu_1 + f_{j2} * \mu_2 + f_{j3} * \mu_3$, where f_{jn} and μ_n are the share of the population that has income in the n-th bin, and estimated the mean for the n-th group, respectively.

³⁵ For a discussion of the potential limitations of this approach, see Appendix C.

the above. Finally, I normalize the utility from the outside option, ω , to zero. A detailed description of the data used and the construction of variables for the structural model can be found in Appendix C.

Estimation Estimation begins at the second stage and follows the methods developed in Berry et al. (1995) and Nevo (2000). For this part, the market share of the outside good is measured as the share of consumers who do not order on the platform out of the number of consumers *searching on the platform*. I use the contraction mapping theorem and GMM estimation developed in Berry et al. (1995) to estimate the linear and non-linear parameters of the utility function, $\theta = \{\alpha, \beta, \Pi, \sigma\}$.

Given the distribution of ϵ , once we identified the parameters of the utility function, the expected utility from the set of restaurants available in the markets is given by:

$$E[\max_j u_{ijt}] = \int_{D_i} \log(\sum \exp\{u_{ijt}\}) \quad (9)$$

Since ω is normalized to 0, consumers only use the platform when $E[\max_j u_{ijt}] > c_i$, which happens with probability $\Phi(E[\max_j u_{ijt}])$. To identify the parameters of the entry-cost distribution $\theta_2 = \{\mu, \gamma, \sigma\}$, I use the estimated utilities and the empirical share of users searching on the platform, S_{jt} , to derive the minimum distance estimator:

$$\theta_2 = \arg \min_x \left\{ (\Phi(E[\max_j u_{ijt}]; x) - S_{jt})' \hat{\Lambda} (\Phi(E[\max_j u_{ijt}]; x) - S_{jt}) \right\} \quad (10)$$

Where $\Phi(\cdot)$ is the cdf of the normal distribution and $\hat{\Lambda}$ is the (empirical) efficient weighting matrix.

The fundamentals of the models are then used to simulate markets in which the percentage of firms on the platform out of total number of restaurants in the market increases from 1% up to 100%. The main outcomes of interest are consumer welfare, total market size, and firm performance under different market conditions. Since ratings play a crucial role in determining firms' outcomes, the results are presented by rating quantiles. The simulation algorithm and estimation details are described further in Appendix E.

6.2 Results

Model Parameters Panel A of Table 11 presents the estimates of the utility function parameters. Column (1) does not use an instrument for price ratings and presents the estimates from a simple logit model without user heterogeneities in the utility function. Column (2) presents the estimates from a logit model without user heterogeneities, but

instruments for dollar rating. Column (3) introduces a random coefficient on the dollar rating, and Column (4) adds consumer demographics to the random coefficient on the dollar rating.

First, when instrumenting for dollar ratings, higher dollar ratings, or higher average restaurant prices, lead to lower utility. Note that while the mean effect in Column (4) is positive, the mean effect in the population (taking into account the impact of demographics on the price coefficient and integrating over their respective prevalence in the population) is approximately -4.5. Second, both relative ranking and absolute rating affect consumer choices. As expected, ratings have a significant positive effect on consumers' utility. Ranking within category appears to have a stronger negative effect than general ranking. Conversely, ranking quantile (mostly) has an unexpected positive sign—though this finding should be interpreted cautiously, since ranking, ranking within category, and rating have strong correlations. When removing absolute rating or ranking within category from the model, the effect of total ranking becomes negative (not reported). All specifications also includes CSA fixed effects and food category dummies, which are not reported for brevity. Finally, as presented in Column (4), the estimates of the dollar rating coefficient shifters are extremely noisy, and none are statistically significant. For this reason, the results from Column (3) serve as the main specification, and the analysis using the estimates from the full model is presented in appendices A and E.

Panel B presents the estimated parameters of the entry cost distribution function, θ_2 in equation 10. In Columns (3) and (4), I allow the mean of the entry cost distribution to differ by income level and present the weighted average across demographics. The estimate mean, μ , and standard error, σ , are approximately 0.2 and 0.09, respectively.

Simulations Table 12 presents the simulation results from the random coefficient model without demographics.³⁶ In the baseline model only 5% of the firms in the city are available on the platform. Columns present the relative change in outcome when moving from 5% to 10%, 20%, and so forth. The first row presents the estimated change in welfare when the percentage of firms on the platform increases. Welfare is monotonically increasing in the percentage of firms on the platform, but at a decreasing rate. For instance, increasing the percent of firms available on the platform from 5% to 20% increases welfare by more than 80%. However, to get a similar increase starting at 20%, the percentage of firms on the platform has to grow to about 80%, almost four times the change, in order to have the same impact on welfare. The pattern of decreasing returns is presented graphically in

³⁶ The equivalents of Table 12 and Figure 9 from the simulation with demographics are presented in Table A13 and Figure A4 and described in Appendix E.

Figure 8. The same calculation can be used to derive the actual welfare change from YTP’s partnership with Grubhub. For the median treated city in the sample, the percentage of firms on the platform increased from 3.5% to 5.5%. These magnitudes imply an average welfare increase of 32.15%.

The second row presents the changes in market size as a function of the fraction of firms on the platform. Again, market size is monotonically increasing in the fraction of firms, but at a decreasing rate. This result is unsurprising, as the share of consumers depends on a concave transformation (through the distribution of entry cost) on the change in welfare. For the median treated city in the sample, this simulation result implies an increase of 27.5% in market size. For comparison, the reduced form estimates presented in Table 2 find an increase of approximately 35% in the number of unique users and weekly orders. The similar magnitudes suggest that the model does a fairly good job in fitting the data, and is, in fact, underestimating the impact of market size growth.

The second part of Table 12 presents the main results, namely, the change in firms’ performance by rating quantile. A more convenient way to understand the results is graphically, using Figure 9. The horizontal axis details the percentage of firms participating in the platform (1% to 100%), and the vertical axis present standardized sales. Each gray dot represents average sales for rating quantile over 1,000 simulations, the lines are the smoothing splines for each rating quantile, and the stars mark the maximum point of each smoothed trend-line. The table and figure reveal several results: First, though not imposed by the estimation algorithm, for relatively low participation rates, the simulation results are consistent with the reduced form analysis: For the median treated city in the sample, firms in the lowest rating quartile lose about 4.8% in sales and firms in the highest quartile gain 5%. In comparison, the simulation predicts a loss of 5.1% for firms in the lower-rating quantile and a gain of 5.2% for firms in the highest quantile. Second, for every percentage of firms on the platform, higher-rated firms perform strictly better. Graphically, the curves never cross: the highest-rated firms always sell more than the second highest, who sell more than the third highest, and so forth.

Third, except for firms in the lowest-rating quantile, for which sales are strictly decreasing in percentage of firms on the platform, there is a non-monotonic relationship between the percentage of firms on the platform and sales. In particular, initially, sales grow with the percentage of firms on the platform. As the percentage of firms continues to grow, however, the trend changes and the effect of market competition starts to dominate. Thus, sales start declining for all types of firms when the percentage of firms on the platform exceeds 50%. At 100% participation, for example, sales are lower for all firms as compared to 5%, except for the highest-rated ones. This finding is consistent with the concavity of the market-size

effect: as the percentage of firms grows, additional firms are less successful in attracting more consumers into the market but continue to compete with the incumbents. Finally, the bliss points, the points at which sales are maximized, are monotonically increasing in relation to firms' rating quantile. The bliss point is zero for lowest-rated firms, and increases to 3%, 11%, 14%, and 46% for firms in the 2nd, 3rd, 4th, and 5th quantiles, i.e., the highest-rated firms generate the highest sales when 46% of the firms participate in the platform.

7 Conclusion

To conclude, this paper studies an important yet unanswered question regarding the impact of new-firm entry on incumbent firms in two-sided markets. Collaborating with Yelp Transactions Platform, I investigate how entrants bring new value to the platform and expand total market size while—in parallel—threatening the performance and market share of incumbent firms. Using a difference-in-differences research design, I study the relative magnitude of these forces empirically and find strong evidence of network externalities, i.e., that supply-side growth led to increased demand and total platform usage. In addition, the entry of new firms benefited incumbent firms, but the positive effects are concentrated solely around high-quality firms; low-quality firms, in contrast, experienced a reduction in both sales and revenue following entry. I also find evidence suggesting that firms respond to the changes in the competitive environment by increasing their investments in quality and by adjusting their advertising behavior. Finally, using a simple structural model, I extrapolate the results to additional market settings and find that, in general, higher-quality firms prefer larger markets with more firms.

The main contribution of the paper is to highlight the importance of firm interactions and competition *within* the platform. The results described in this study have important implications to firms operating in platform settings, to platform strategy, and to competition policy. First, I find that the effects of entry on incumbent firms are positive on average, but critically depend on firm characteristics (i.e., quality) as well as platform maturity and size. Incorporating these considerations into the analysis yields important insights regarding how firms operate in a platform setting: For example, the main findings of this paper suggest that the platform environment and firm identity will mediate the benefits to incumbents from setting barriers to entry on platforms, with different platform characteristics and firm qualities dictating the direction and magnitude of such strategies. These new considerations should be taken into account when determining incumbents' responses towards the threat of entry in platform markets. Consequently, future research would benefit from unpacking the question of whether firms internalize the positive spillovers generated by potential entrants

and whether they differentially invest in deterring or even promoting the entry of new firms.

One limitation of the described analysis is that I only observe performance on one platform. While I present suggestive evidence that increased revenue on the platform is not strictly cannibalizing other sources of revenue, the model and empirical analysis generally abstain from addressing cross-platform competition and substitution patterns. A natural extension is to expand the analysis into additional datasets which include both multiple platforms and the larger market setting.

Second, expanding the number of suppliers on the platform naturally benefits the platform and is usually a central concern to platform managers. This paper contributes to the existing body of knowledge by demonstrating an additional channel by which expansion of the supply side may benefit the platform. Entry increases variety, creates additional value to consumers, and increases total transactions volume. At the same time, new firm entry raises the average quality on the platform: First, the volume high-quality sales grows while that of low-quality sales diminishes. Together, those imply that the average quality purchased by consumers is also increasing. Second, as the return to high quality increases, firms increase subsequent investments in quality. Finally, the simulations performed within this study further suggest that larger platforms are more attractive to high-quality firms, which subsequently signifies that the quality of the average entrant will improve as the platform grows. The current model studies only the short-term equilibrium and does not assess entry decisions, since entry into the market is quasi-random in our empirical setting. Nevertheless, these three mechanisms together suggest that entry does not just increase the size of the platform but also makes the average quality of the platform's firms better.

Lastly, policy-makers and regulators frequently direct attention to competition between platforms and the importance of restricting firms. This paper contributes to our understanding of the benefit and pitfalls of excessive regulation of platform markets. While addressing general impact on the market via the existence of multiple platforms is beyond the scope of this paper, the main results raise several important considerations that need to be incorporated into the decision-making process. In particular, restricting platform growth curtails the benefits detailed above—such as loss of positive network effects—and reduces competitiveness and quality on the platform. These negative implications must to be carefully examined as to not harm the businesses operating on the platform, the platform itself, and even the consumers that the regulators seek to protect.

References

- Aghion, P., Blundell, R., Griffith, R., Howitt, P., and Prantl, S. (2009). The effects of entry on incumbent innovation and productivity. *The Review of Economics and Statistics*, 91(1):20–32.
- Armstrong, M. (2006). Competition in two-sided markets. *The RAND Journal of Economics*, 37(3):668–691.
- Armstrong, M., Vickers, J., and Zhou, J. (2009). Prominence and consumer search. *The RAND Journal of Economics*, 40(2):209–233.
- Athey, S. and Imbens, G. W. (2017). The econometrics of randomized experiments. *Handbook of Economic Field Experiments*, 1:73–140.
- Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of management*, 17(1):99–120.
- Bass, F. M. (1969). A new product growth for model consumer durables. *Management science*, 15(5):215–227.
- Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, pages 841–890.
- Berry, S. T. (1992). Estimation of a model of entry in the airline industry. *Econometrica: Journal of the Econometric Society*, pages 889–917.
- Bils, M. and Klenow, P. J. (2004). Some evidence on the importance of sticky prices. *Journal of Political Economy*, 112(5):947–985.
- Bresnahan, T. F. and Reiss, P. C. (1991). Entry and competition in concentrated markets. *Journal of political economy*, 99(5):977–1009.
- Burdett, K. and Judd, K. L. (1983). Equilibrium price dispersion. *Econometrica: Journal of the Econometric Society*, pages 955–969.
- Caillaud, B. and Jullien, B. (2003). Chicken & egg: Competition among intermediation service providers. *RAND journal of Economics*, pages 309–328.
- Cao, G., Jin, G. Z., Weng, X., and Zhou, L.-A. (2018). Market expanding or market stealing? competition with network effects in bike-sharing. *Working paper*, National Bureau of Economic Research.
- Cennamo, C., Gu, Y., and Zhou, F. (2016). Value co-creation and capture in the creative industry: The u.s. home video game market. *Working paper*.

- Chevalier, J. A. and Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of marketing research*, 43(3):345–354.
- Corts, K. S. and Lederman, M. (2009). Software exclusivity and the scope of indirect network effects in the us home video game market. *international Journal of industrial Organization*, 27(2):121–136.
- Cullen, Z. and Farronato, C. (2019). Outsourcing tasks online: Matching supply and demand on peer-to-peer internet platforms. *Working paper*.
- Cusumano, M. A., Gawer, A., and Yoffie, D. B. (2019). *The business of platforms*. HarperCollins-Business London.
- Dixit, A. K. and Stiglitz, J. E. (1977). Monopolistic competition and optimum product diversity. *The American economic review*, 67(3):297–308.
- Dubé, J.-P. H., Hitsch, G. J., and Chintagunta, P. K. (2010). Tipping and concentration in markets with indirect network effects. *Marketing Science*, 29(2):216–249.
- Ellison, G. and Ellison, S. F. (2011). Strategic entry deterrence and the behavior of pharmaceutical incumbents prior to patent expiration. *American Economic Journal: Microeconomics*, 3(1):1–36.
- Farrell, J. and Saloner, G. (1985). Standardization, compatibility and innovation. *The RAND Journal of Economics*, 16:70–83.
- Farronato, C. and Fradkin, A. (2018). The welfare effects of peer entry in the accommodation market: The case of airbnb. Technical report, National Bureau of Economic Research.
- Fosfuri, A. and Giarratana, M. S. (2009). Masters of war: Rivals’ product innovation and new advertising in mature product markets. *Management Science*, 55(2):181–191.
- Fradkin, A., Grewal, E., and Holtz, D. (2018). The determinants of online review informativeness: Evidence from field experiments on airbnb.
- Fudenberg, D. and Tirole, J. (1984). The fat-cat effect, the puppy-dog ploy, and the lean and hungry look. *American Economic Review*, 74(2):361–366.
- Gawer, A. and Henderson, R. (2007). Platform owner entry and innovation in complementary markets: Evidence from intel. *Journal of Economics & Management Strategy*, 16(1):1–34.
- Goolsbee, A. and Syverson, C. (2008). How do incumbents respond to the threat of entry? evidence from the major airlines. *The Quarterly Journal of Economics*, 123(4):1611–1633.
- Hagiu, A. and Wright, J. (2015). Multi-sided platforms. *International Journal of Industrial Organization*, 43:162–174.

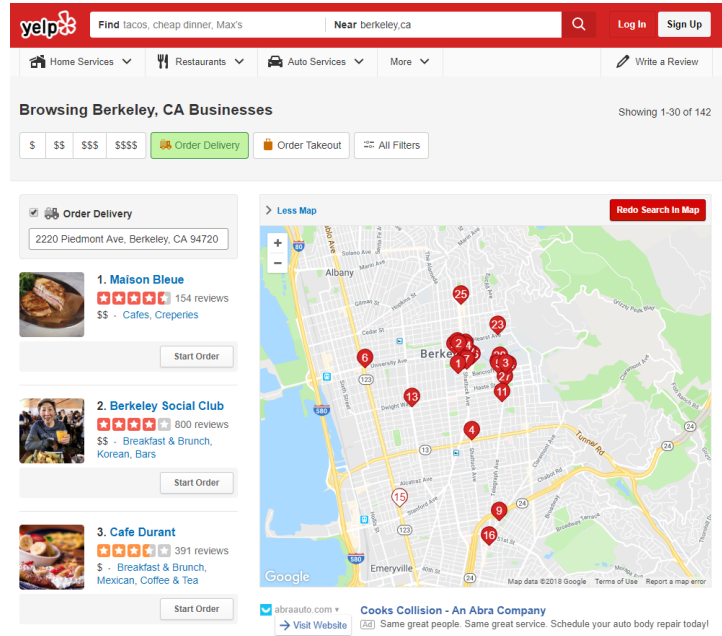
- Hauser, J. R. and Shugan, S. M. (2008). Defensive marketing strategies. *Marketing Science*, 27(1):88–110.
- Hirano, K., Imbens, G. W., and Ridder, G. (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*, 71(4):1161–1189.
- Hobijn, B., Ravenna, F., and Tambalotti, A. (2006). Menu costs at work: restaurant prices and the introduction of the euro. *The Quarterly Journal of Economics*, 121(3):1103–1131.
- Hollenbeck, B., Moorthy, S., and Proserpio, D. (2019). Advertising strategy in the presence of reviews: An empirical analysis. *Forthcoming Marketing Science*.
- Hortaçsu, A. and Syverson, C. (2007). Cementing relationships: Vertical integration, foreclosure, productivity, and prices. *Journal of Political Economy*, 115(2):250–301.
- Horton, J. and Golden, J. (2018). Reputation inflation: Evidence from an online labor market. *New York University, Working Paper*, 1.
- Hotelling, H. (1929). Stability in competition. *The Economic Journal*, 39(153):41–57.
- Illanes, G. and Moshary, S. (2018). Market structure and competition: Evidence from a natural experiment in liquor licensure. *Working paper*.
- Jia, P. (2008). What happens when wal-mart comes to town: An empirical analysis of the discount retailing industry. *Econometrica*, 76(6):1263–1316.
- Jin, G. Z. and Rysman, M. (2015). Platform pricing at sports card conventions. *The Journal of Industrial Economics*, 63(4):704–735.
- Kadiyali, V. (1996). Entry, its deterrence, and its accommodation: A study of the us photographic film industry. *The Rand Journal of Economics*, pages 452–478.
- Kaiser, U. and Wright, J. (2006). Price structure in two-sided markets: Evidence from the magazine industry. *International Journal of Industrial Organization*, 24(1):1–28.
- Katz, M. L. and Shapiro, C. (1985). Network externalities, competition, and compatibility. *American economic review*, 75(3):424–440.
- Kovács, B. and Sharkey, A. J. (2014). The paradox of publicity: How awards can negatively affect the evaluation of quality. *Administrative Science Quarterly*, 59(1):1–33.
- Lee, R. S. (2013). Vertical integration and exclusivity in platform and two-sided markets. *American Economic Review*, 103(7):2960–3000.

- Li, Z. and Agarwal, A. (2016). Platform integration and demand spillovers in complementary markets: Evidence from facebook’s integration of instagram. *Management Science*, 63(10):3438–3458.
- Luca, M. (2016). Reviews, reputation, and revenue: The case of yelp. com. *Com (March 15, 2016). Harvard Business School NOM Unit Working Paper*, (12-016).
- Mahajan, V., Sharma, S., and Buzzell, R. D. (1993). Assessing the impact of competitive entry on market expansion and incumbent sales. *Journal of marketing*, 57(3):39–52.
- Marshall, A. (1890). Principles of economics.
- McFadden, D. et al. (1973). Conditional logit analysis of qualitative choice behavior.
- Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *econometrica*, 71(6):1695–1725.
- Murphy, K. M., Shleifer, A., and Vishny, R. W. (1989). Industrialization and the big push. *Journal of political economy*, 97(5):1003–1026.
- Nair, H., Chintagunta, P., and Dubé, J.-P. (2004). Empirical analysis of indirect network effects in the market for personal digital assistants. *Quantitative Marketing and Economics*, 2(1):23–58.
- Nevo, A. (2000). Mergers with differentiated products: The case of the ready-to-eat cereal industry. *The RAND Journal of Economics*, pages 395–421.
- Nickell, S. J. (1996). Competition and corporate performance. *Journal of Political Economy*, 104(4):724–746.
- Nosko, C. and Tadelis, S. (2018). The limits of reputation in platform markets: An empirical analysis and field experiment. Technical report, National Bureau of Economic Research.
- Ohashi, H. (2003). The role of network effects in the us vcr market, 1978–1986. *Journal of Economics & Management Strategy*, 12(4):447–494.
- Pavcnik, N. (2002). Trade liberalization, exit, and productivity improvements: Evidence from chilean plants. *The Review of Economic Studies*, 69(1):245–276.
- Porter, M. E. (1989). How competitive forces shape strategy. In *Readings in strategic management*, pages 133–143. Springer.
- Porter, M. E. (1997). Competitive strategy. *Measuring business excellence*, 1(2):12–17.
- Rochet, J.-C. and Tirole, J. (2006). Two-sided markets: a progress report. *The RAND journal of economics*, 37(3):645–667.

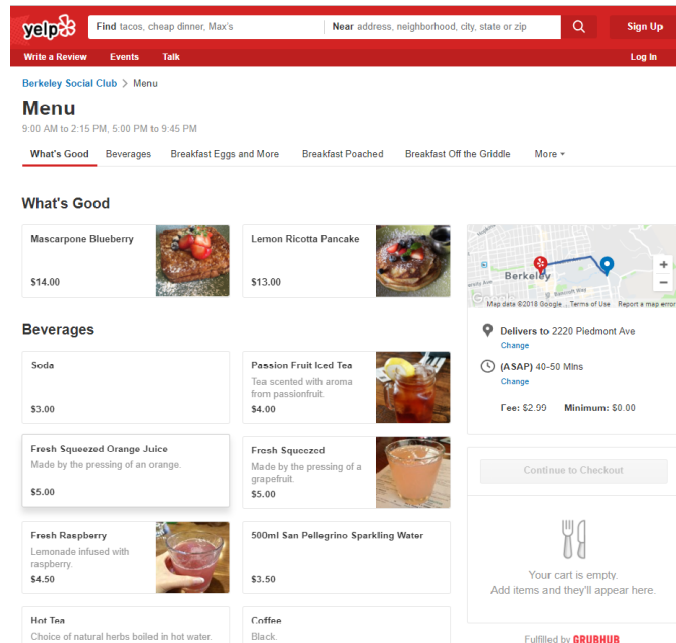
- Rysman, M. (2004). Competition between networks: A study of the market for yellow pages. *The Review of Economic Studies*, 71(2):483–512.
- Samuelson, P. A. (1951). *Economics-an introductory analysis*. McGraw-Hill Book Company, Inc.; New York; London; Toronto.
- Seamans, R. and Zhu, F. (2013). Responses to entry in multi-sided markets: The impact of craigslist on local newspapers. *Management Science*, 60(2):476–493.
- Seamans, R. C. (2013). Threat of entry, asymmetric information, and pricing. *Strategic Management Journal*, 34(4):426–444.
- Shen, Q. and Xiao, P. (2014). Mcdonald’s and kfc in china: Competitors or companions? *Marketing Science*, 33(2):287–307.
- Springel, K. (2018). Network externality and subsidy structure in two-sided markets: Evidence from electric vehicle incentives. *Working Paper*.
- Stigler, G. J. (1961). The economics of information. *Journal of political economy*, 69(3):213–225.
- Stremersch, S., Tellis, G. J., Hans Franses, P., and Bincken, J. L. (2007). Indirect network effects in new product growth. *Journal of Marketing*, 71(3):52–74.
- Syverson, C. (2004). Market structure and productivity: A concrete example. *Journal of Political Economy*, 112(6):1181–1222.
- Tadelis, S. (2016). Reputation and feedback systems in online platform markets. *Annual Review of Economics*, 8:321–340.
- Young, A. (2016). Channeling fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results. *London School of Economics, Working Paper*.
- Zbaracki, M. J., Ritson, M., Levy, D., Dutta, S., and Bergen, M. (2004). Managerial and customer costs of price adjustment: direct evidence from industrial markets. *Review of Economics and statistics*, 86(2):514–533.
- Zhu, F. and Iansiti, M. (2012). Entry into platform-based markets. *Strategic Management Journal*, 33(1):88–106.
- Zhu, F. and Liu, Q. (2018). Competing with complementors: An empirical look at amazon. com. *Strategic Management Journal*, 39(10):2618–2642.

Figures

Figure 1: Yelp Transactions Platform



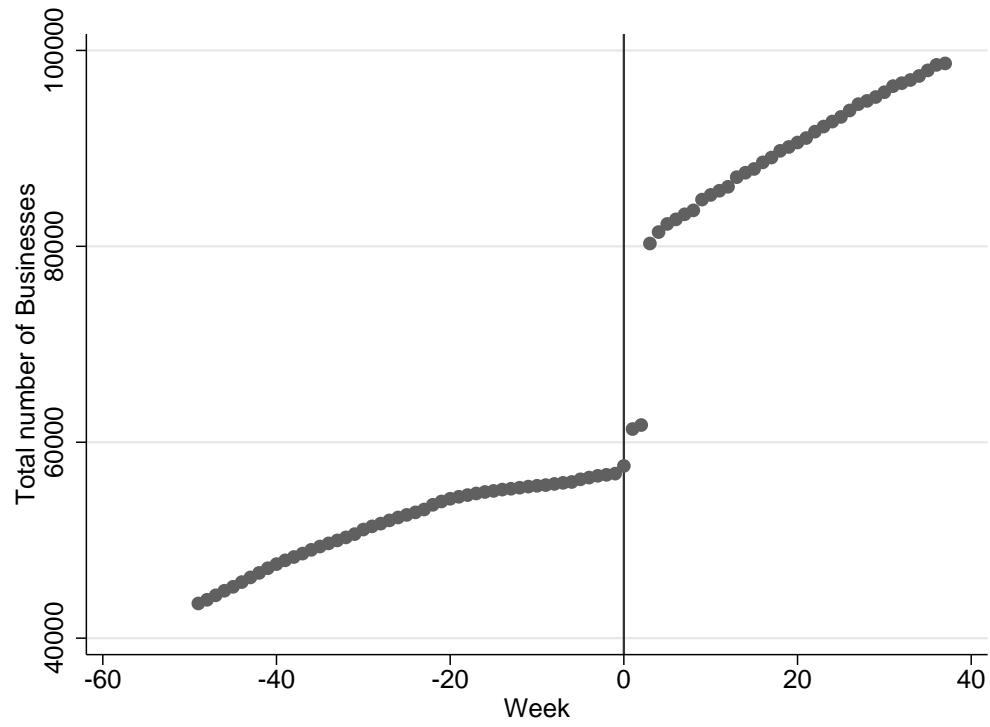
(a) Search on Yelp Transactions Platform



(b) Ordering on Yelp Transactions Platform

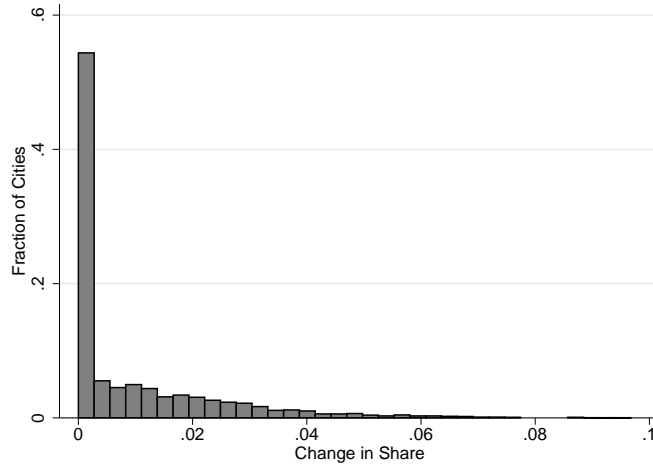
Note: Panel A presents the search results on YTP around Haas School of Business. Panel B presents a menu for a restaurant affiliated with Grubhub (as indicated in the bottom-right corner).

Figure 2: Number of Businesses on YTP over time

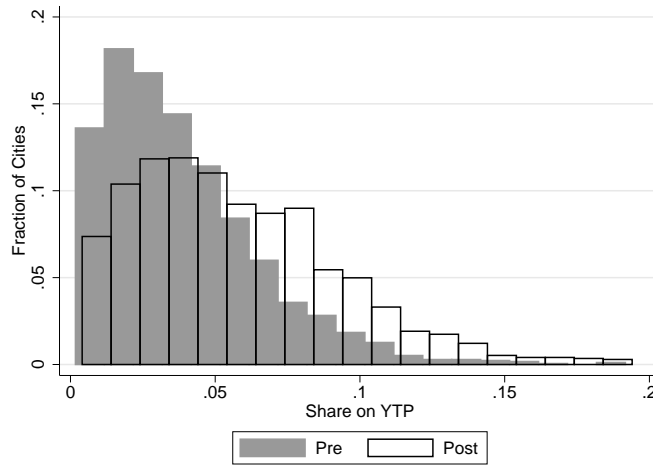


Note: The figure presents the development of the total number of restaurants available on YTP over time. The week of implementation is normalized to zero.

Figure 3: Variation in Treatment Intensity by City



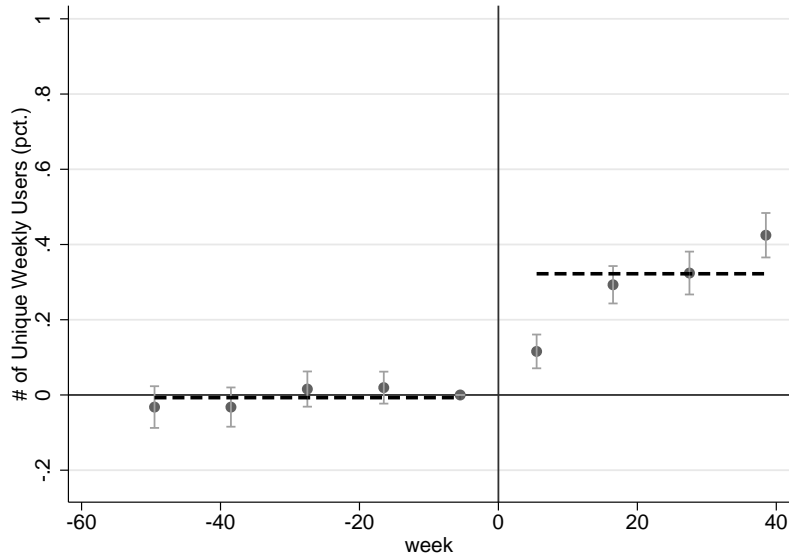
(a) Change in the Share of Businesses on YTP by City



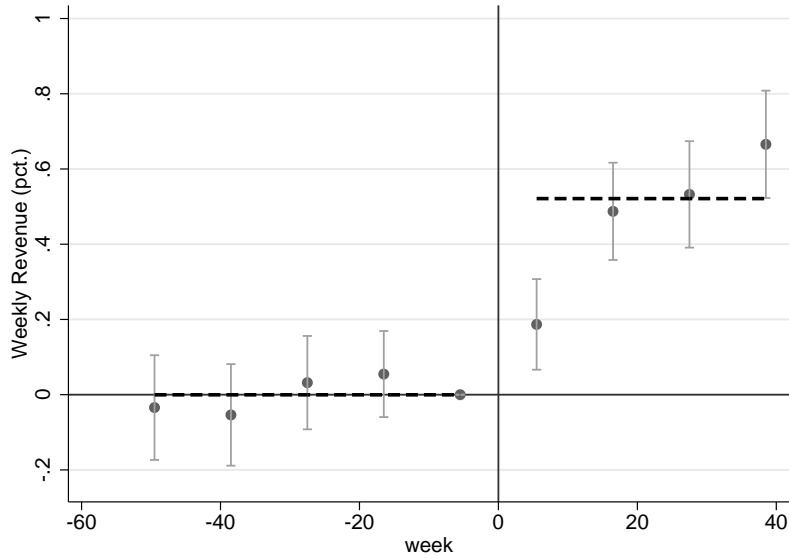
(b) Share of Businesses on YTP by City
(Conditional on Change)

Note: The figure presents the change in the percentage of restaurants available on YTP by city. Panel A presents the distribution of the percentage change by city. Panel B presents the distributions of the percentage of restaurants available on YTP before and after the partnership, for cities that were affected by the partnership.

Figure 4: Market-Level Changes (Percentages)



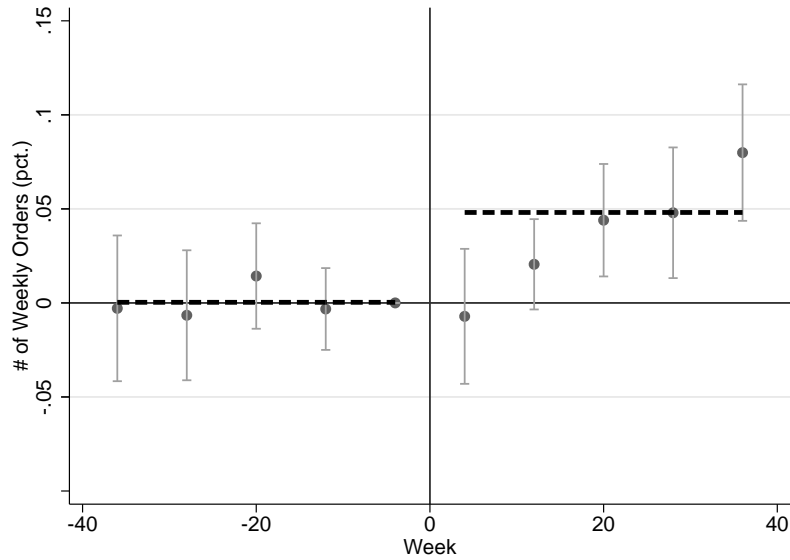
(a) Weekly Number of Unique Users



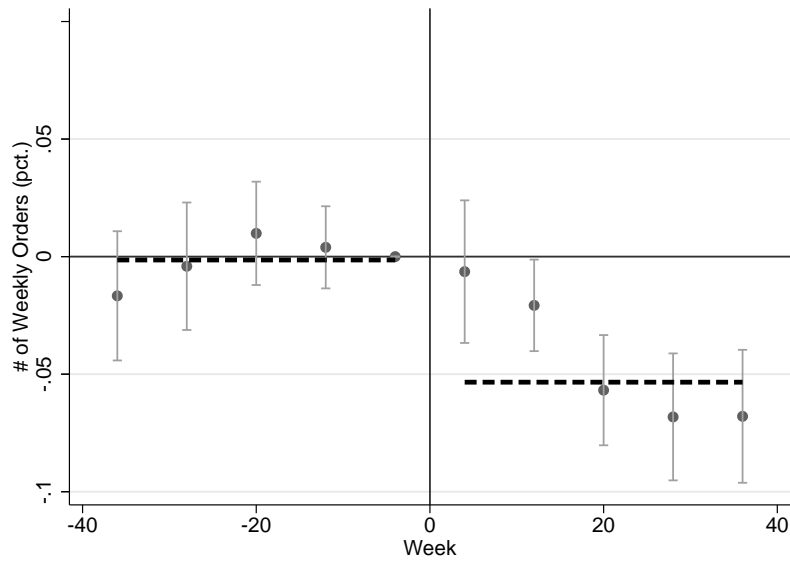
(b) Weekly Revenue

Note: This figure presents event-time estimates from a version of equation 4. The dependent variables are the inverse hyperbolic sine transformation and should be interpreted as percentage changes. The unit of observation is city-week, including both incumbent and newly added businesses. The dots represent point estimates from regressing the dependent variable on a treatment indicator interacted with nine-week bins, and city and week-state fixed effects. The treatment indicator compares cities that experienced almost no change in the percentage of businesses available on the platform to cities that experienced meaningful changes. The coefficient in the first period prior to implementation is normalized to zero. The vertical bars represent 95% confidence intervals, where standard errors are clustered at the city level.

Figure 5: Weekly Number of Orders Per Business (Percentages)



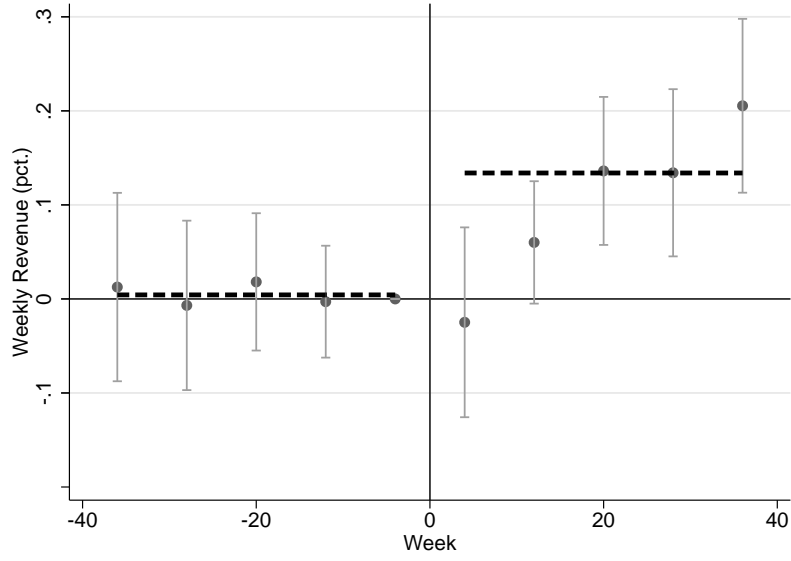
(a) High-rated Firms



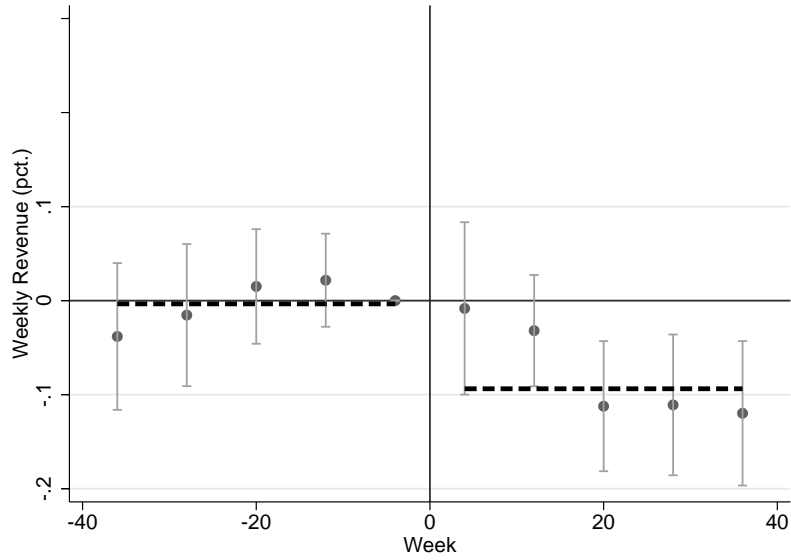
(b) Low-rated Firms

Note: This figure presents event-time estimates from a version of equation 5. The dependent variables are the inverse hyperbolic sine transformation and should be interpreted as percentage changes. The unit of observation is business-week. Panel A (B) includes only businesses with rating above the 75th percentile (below the 25th percentile) in the city. The dots represent point estimates from regressing the dependent variable on a treatment indicator interacted with nine-week bins, and business and week-state fixed effects. The treatment indicator compares cities that experienced almost no change in the percentage of businesses available on the platform to cities that experienced meaningful changes. The coefficient in the first period prior to implementation is normalized to zero. The vertical bar represent 95% confidence intervals, where standard errors are clustered at the city level.

Figure 6: Weekly Revenue Per Business (Percentages)



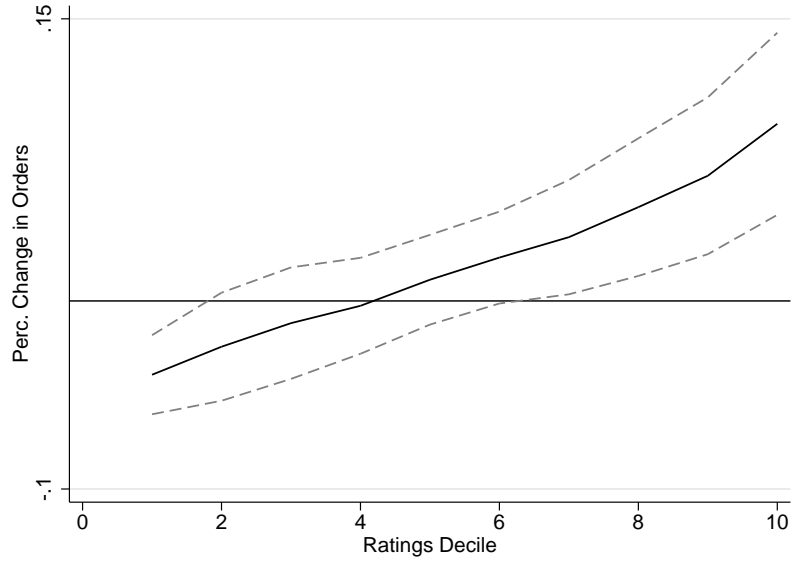
(a) High-rated Firms



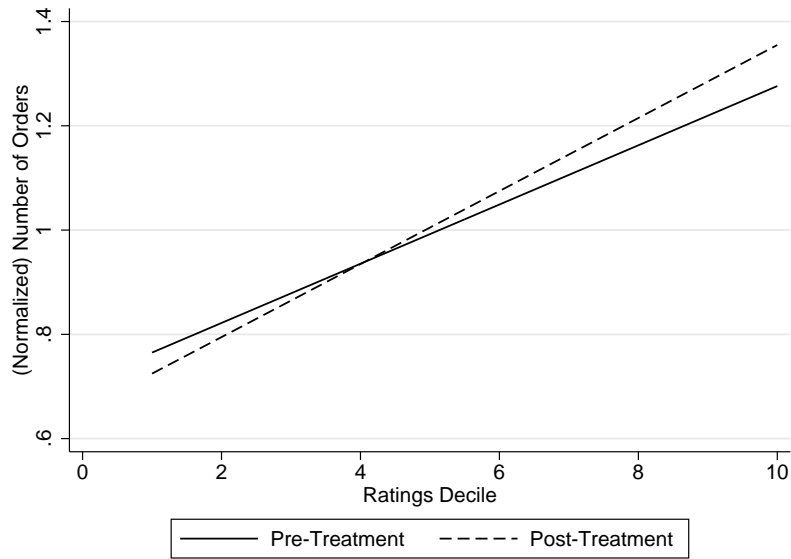
(b) Low-rated Firms

Note: This figure presents event-time estimates from a version of equation 5. The dependent variables are the inverse hyperbolic sine transformation and should be interpreted as percentage changes. The unit of observation is business-week. Panel A (B) includes only businesses with ratings above the 75th percentile (below the 25th percentile) in the city. The dots represent point estimates from regressing the dependent variable on a treatment indicator interacted with nine-week bins, and business and week-state fixed effects. The treatment indicator compares cities that experienced almost no change in the percentage of businesses available on the platform to cities that experienced meaningful changes. The coefficient in the first period prior to implementation is normalized to zero. The vertical bar represent 95% confidence intervals, where standard errors are clustered at the city level.

Figure 7: Return to Ratings



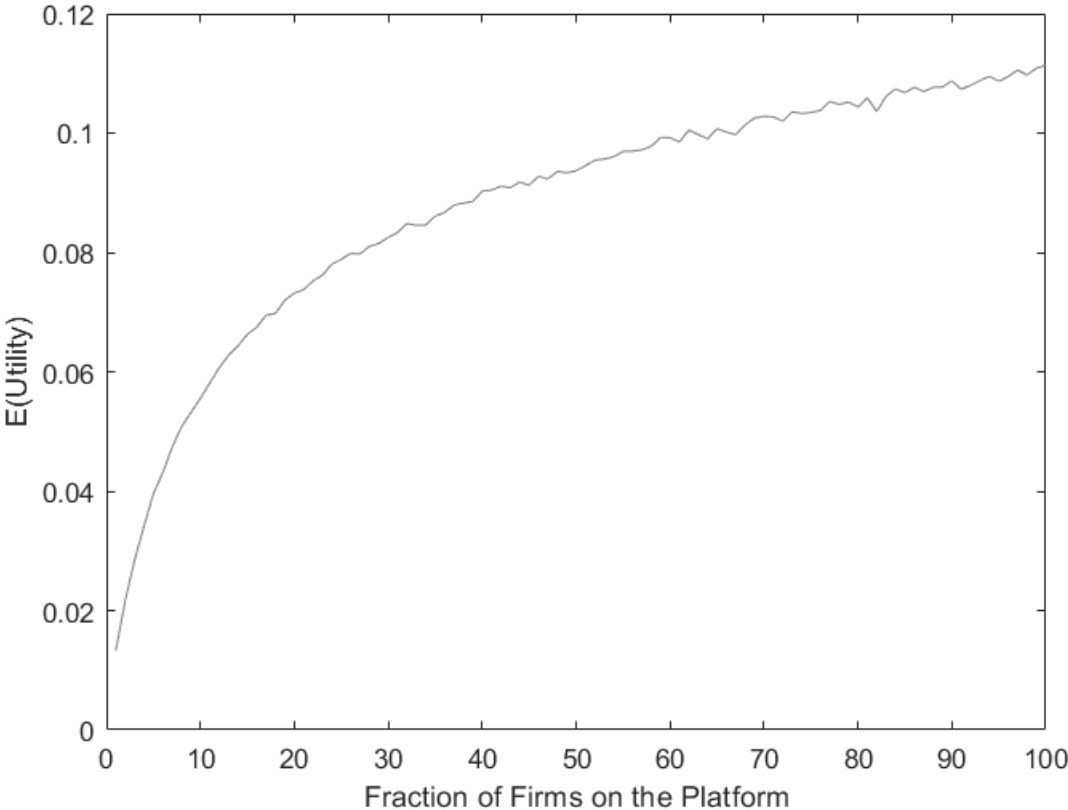
(a) Effect by Rating Decile



(b) Return to Ratings

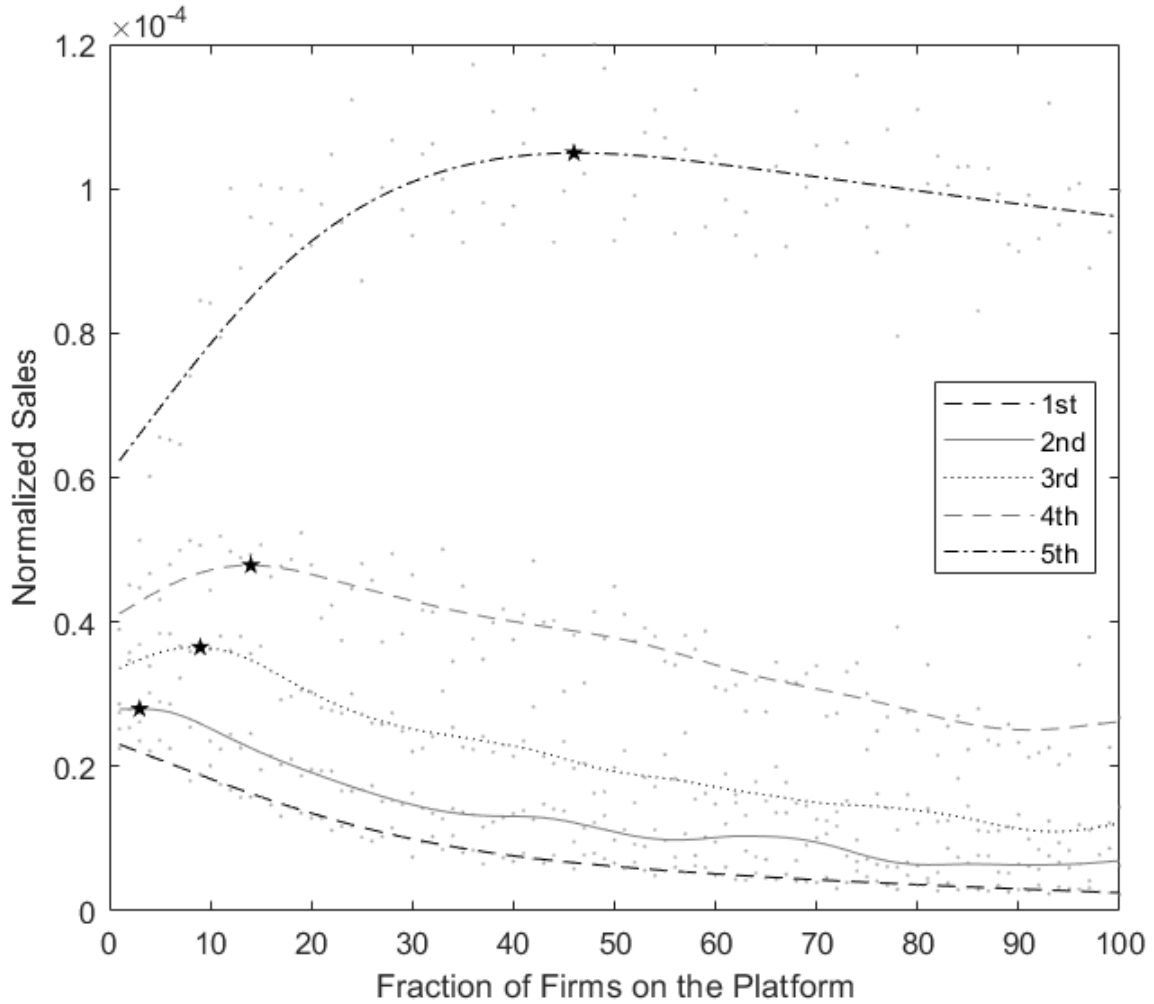
Note: This figure presents the change in correlation between rating decile within city and the percentage change in weekly revenue. The unit of observation is business-week. The solid line in Panel A presents the point estimates of triple interactions between treatment status (sharp treatment definition), an indicator for post integration, and rating decile dummies. The regression includes business and week-state fixed effects. The dashed lines represent 95% confidence intervals, where standard errors are clustered at the city level. Panel B presents the relations between rating deciles and revenue before and after implementation, in treated cities.

Figure 8: Expected Utility by Fraction of Firms on the Platform



Note: This figure present the change in expect utility as the share of firms on the platform grows. It depicts the average results of 500,000 simulated markets. The parameters used to simulate the data are presented in column 3 of Table 12. The simulation algorithm is described in appendix E.

Figure 9: Simulations of Firms' Sales by Rating Quantile



Note: This figure present the change in number of sales by rating quantile as the share of firms on the platform grows. The parameters used to simulate the data are presented in column 3 of Table 12. The simulation algorithm is described in appendix E. Each gray dot represent the average over one thousand simulations, and the dashed and solid line are the smoothing spline by rating quantile. The stars mark the maximum of each smoothing spline.

Tables

Table 1: Descriptive Statistics

	Mean	SD	Min	Max
A. Businesses (N=56,493)				
Ratings	3.63	0.61	1	5
Dollar Ratings	1.61	0.51	1	4
Weeks on YTP	94.8	23.8	1	114
Fraction on YTP (pre)	0.051	0.026	0	0.4
Share Change (Business-Level)	0.020	0.013	0	0.2
B. Cities (N=3,965)				
Total Businesses	424.3	1448.9	11	42180
Total on YTP (pre)	14.2	85.9	1	3459
Fraction on YTP (pre)	0.032	0.028	0	0.4
Share Change (City-Level)	0.012	0.018	0	0.2

Note: Panels A and B of this table report the characteristics at the business and city levels, respectively.

Table 2: Market-Level Analysis

	(1)	(2)	(3)
Panel A: Weekly Unique Users			
Treat*Post	0.364*** (0.015)	0.482*** (0.020)	10.271*** (0.634)
Panel B: Weekly Orders			
Treat*Post	0.367*** (0.015)	0.486*** (0.020)	10.343*** (0.639)
Panel C: Weekly Revenue			
Treat*Post	0.587*** (0.032)	0.762*** (0.042)	17.226*** (1.202)
Observations	327993	226157	327993
# of Clusters	3964	2781	3964
Treatment Def.	Median	25<>75	Change

Note: This table reports regression coefficients from nine separate regressions, three per panel. An observation is city-week, including both incumbent and newly added businesses. The dependent variables are the inverse hyperbolic sine transformation of the outcomes indicated in sub-headings and should be interpreted as percentage changes. Treatment status definitions are indicated below the table and are described further in the text. All regressions include city and week-state fixed effects. Standard errors are in parentheses and are clustered at the city level.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 3: Average Effects on Incumbent Firms

	(1)	(2)	(3)
Panel A: Weekly Orders			
Treat*Post	0.007 (0.005)	0.004 (0.007)	0.388* (0.161)
Panel B: Weekly Revenue			
Treat*Post	0.042*** (0.014)	0.044** (0.018)	2.068*** (0.416)
Observations	4409516	2623347	4409516
# of Clusters	3964	2781	3964
Treatment Def.	Median	25<>75	Change

Note: This table reports regression coefficients from six separate regressions, three per panel. An observation is business-week. The dependent variables are the per-business inverse hyperbolic sine transformation of weekly number of orders (Panel A) and weekly-revenue (Panel B), and should be interpreted as percentage changes. Treatment status definitions are indicated below the table and are described further in the text. All regressions include business and week-state fixed effects. Standard errors are in parentheses and are clustered at the city level.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 4: Heterogeneous Effects on Incumbent Firms

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Weekly Number of Orders						
Treat*Post	0.036*** (0.007)	0.050*** (0.010)	0.042*** (0.009)	0.065*** (0.012)	1.106*** (0.202)	1.432*** (0.269)
Treat*Post*Low	-0.062*** (0.010)	-0.098*** (0.013)	-0.084*** (0.012)	-0.119*** (0.015)	-1.654*** (0.286)	-2.342*** (0.377)
$\beta_1 + \beta_2$	-0.026	-0.048	-0.042	-0.054	-0.548	-0.910
Pvalue	0.001	0.000	0.000	0.000	0.015	0.000
Panel B: Weekly Revenue						
Treat*Post	0.098*** (0.017)	0.119*** (0.025)	0.121*** (0.022)	0.158*** (0.031)	3.793*** (0.508)	4.059*** (0.665)
Treat*Post*Low	-0.121*** (0.026)	-0.201*** (0.033)	-0.173*** (0.032)	-0.250*** (0.040)	-4.016*** (0.679)	-5.639*** (0.893)
Observations	4409516	2173244	2623347	1321619	4409516	2173244
# of Clusters	3964	3875	2781	2725	3964	3875
$\beta_1 + \beta_2$	-0.022	-0.082	-0.052	-0.092	-0.223	-1.580
Pvalue	0.279	0.000	0.049	0.000	0.691	0.015
Treatment Def.	Median	Median	25<>75	25<>75	Change	Change
Quality Def.	Median	25<>75	Median	25<>75	Median	25<>75

Note: This table reports regression coefficients from twelve separate regressions, six per panel. An observation is business-week. The dependent variables are the per-business inverse hyperbolic sine transformation of weekly number of orders (Panel A) and weekly-revenue (Panel B), and should be interpreted as percentage changes. The sum of the coefficient is presented below each panel along with the corresponding P value. The interaction between post and quality-level indicators is omitted for brevity. Treatment status and quality definitions are indicated below the table and are described further in the text. All regressions include business and week-state fixed effects. Standard errors are in parentheses and are clustered at the city level.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 5: Effect of Entry on Incumbents' Prices

	(1)	(2)	(3)	(4)	(5)	(6)
Treat*Post	0.044*	-0.014	0.052	0.024	-0.007	-0.002
	(0.025)	(0.027)	(0.034)	(0.040)	(0.033)	(0.048)
Treat*Post*Low			-0.028	-0.033	-0.025	-0.074
			(0.056)	(0.085)	(0.054)	(0.083)
Observations	5690994	2803318	5690994	2445849	2803318	1271390
# of Clusters	1703	1009	1703	1356	1009	804
$\beta_1 + \beta_2$			0.024	-0.008	-0.032	-0.076
Pvalue			0.523	0.906	0.464	0.277
Treatment Def.	Median	25<>75	Median	Median	25<>75	25<>75
Quality Def.			Median	25<>75	Median	25<>75

Note: This table reports regression coefficients from six separate regressions. An observation is item-business-week. The sample includes only prices for the most popular items, and excludes item modifications. See text for details. The dependent variable is the inverse hyperbolic sine transformation of item price and should be interpreted as percentage changes. The sum of the coefficient is presented below each panel along with the corresponding P value. The interaction between post and quality-level indicators is omitted for brevity. Treatment status and quality definitions are indicated below the table and are described further in the text. All regressions include item-business and week-state fixed effects. Standard errors are in parentheses and are clustered at the city level.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 6: Effect of Entry on Incumbents' Subsequent Yelp Ratings

	(1)	(2)	(3)	(4)	(5)	(6)
Treat*Post	0.006** (0.003)	0.008* (0.004)	0.007** (0.003)	0.007* (0.004)	0.009** (0.004)	0.012** (0.005)
Treat*Post*Low			-0.005 (0.006)	-0.005 (0.008)	-0.004 (0.007)	0.003 (0.010)
Observations	1449437	850538	1449437	687184	850538	408278
# of Clusters	3562	2420	3562	3380	2420	2314
$\beta_1 + \beta_2$			0.003	0.001	0.005	0.015
Pvalue			0.598	0.864	0.502	0.078
Treatment Def.	Median	25<>75	Median	Median	25<>75	25<>75
Quality Def.			Median	25<>75	Median	25<>75

Note: This table reports regression coefficients from six separate regressions. An observation is business-week. The dependent variable is the inverse hyperbolic sine transformation of the average rating received in a given week and should be interpreted as percentage changes. The sum of the coefficient is presented below each panel along with the corresponding p-value. The interaction between post and quality-level indicators is omitted for brevity. Treatment status and quality definitions are indicated below the table and are described further in the text. All regressions include business and week-state fixed effects. Standard errors are in parentheses and are clustered at the city level.
 * significant at 10%; ** significant at 5%; *** significant at 1%

Table 7: Effect of Entry on Advertising Purchases

	(1)	(2)	(3)	(4)	(5)	(6)
Treat*Post	-0.027* (0.014)	-0.035** (0.018)	-0.034* (0.019)	-0.034 (0.027)	-0.047** (0.024)	-0.081** (0.035)
Treat*Post*Low			0.011 (0.026)	0.050 (0.033)	0.016 (0.033)	0.096** (0.042)
Observations	4409516	2623347	4409516	2173244	2623347	1321619
# of Clusters	3964	2781	3964	3875	2781	2725
$\beta_1 + \beta_2$			-0.023	0.016	-0.031	0.015
Pvalue			0.250	0.419	0.208	0.540
Treatment Def.	Median	25<>75	Median	Median	25<>75	25<>75
Quality Def.			Median	25<>75	Median	25<>75

Note: This table reports regression coefficients from six separate regressions. An observation is business-week. The dependent variable is the inverse hyperbolic sine transformation of total advertising expenditure on Yelp and should be interpreted as percentage changes. The sum of the coefficient is presented below each panel along with the corresponding P value. The interaction between post and quality-level indicators is omitted for brevity. Treatment status and quality definitions are indicated below the table and are described further in the text. All regressions include business and week-state fixed effects. Standard errors are in parentheses and are clustered at the city level.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 8: Heterogeneity by Differentiation Between Incumbents and Entrants

	Num. of Orders (Prc.)		Revenue (Prc.)	
	(1)	(2)	(3)	(4)
Share Same Type*Post	-0.010* (0.006)	-0.029** (0.015)	-0.002 (0.014)	-0.064* (0.037)
Observations	3223311	3223311	3223311	3223311
# of Clusters	1917	1917	1917	1917
Similarity Definition	Positive Change	Continuous	Positive Change	Continuous

Note: This table reports regression coefficients from four separate regressions. An observation is business-week. The sample includes only cities that received above-median change in the share of businesses on YTP. The dependent variables are the inverse hyperbolic sine transformation of weekly number of orders and weekly revenue, and should be interpreted as percentage changes. Coefficients represent the interaction between the measure of similarity, and a dummy for post implementation and treatment status. In columns (1) and (2), the measure is an indicator for whether any business in the same category were added, whereas in columns (3) and (4), the measure is the share of business of the same food category as the incumbent out of the total number of added business. All regressions include business and week-state fixed effects. Standard errors are in parentheses and are clustered at the city level.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 9: Search Behavior

	(1)	(2)	(3)
Panel A: Number of Searches			
Treat*Post	0.011 (0.006)	0.024** (0.008)	0.374* (0.162)
Panel B: Number of Views			
Treat*Post	-0.014 (0.009)	-0.006 (0.012)	0.040 (0.236)
Panel C: Session Duration			
Treat*Post	0.355*** (0.023)	0.503*** (0.028)	8.016*** (0.782)
Panel D: Levenshtein Distance			
Treat*Post	-0.033*** (0.012)	-0.049*** (0.017)	-0.905** (0.354)
Panel E: Generic Queries			
Treat*Post	0.000 (0.003)	-0.000 (0.004)	-0.041 (0.084)
Observations	159912	94398	159912
# of Clusters	3729	2548	3729
Treatment Def.	Median	25<>75	Change

Note: This table reports regression coefficients from fifteen separate regressions, three per panel. An observation is business-week. The sample includes only Yelp sessions that ended in an order on YTP. The dependent variables are the per-business inverse hyperbolic sine transformations, and should be interpreted as percentage changes. Outcomes are indicated in the sub-headers and described further in the text. The interaction between post and quality-level indicators is omitted for brevity. Treatment status definitions are indicated below the table and are described further in the text. All regressions include business and week-state fixed effects. Standard errors are in parentheses and are clustered at the city level.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 10: Controlling for Ordering of Search Results

	(1)	(2)	(3)	(4)
	Orders	Orders	Revenue	Revenue
Treat*Post	0.049*** (0.010)	0.064*** (0.013)	0.112*** (0.026)	0.150*** (0.031)
Treat*Post*Low	-0.103*** (0.013)	-0.126*** (0.016)	-0.215*** (0.034)	-0.271*** (0.041)
Observations	2098250	1274303	2098250	1274303
# of Clusters	3830	2685	3830	2685
$\beta_1 + \beta_2$	-0.054	-0.063	-0.102	-0.120
Pvalue	0.000	0.000	0.000	0.000
Treatment Def.	Median	25<>75	Median	25<>75
Quality Def.	25<>75	25<>75	25<>75	25<>75

Note: This table reports regression coefficients from four separate regressions. An observation is business-week. The dependent variables are the per-business inverse hyperbolic sine transformation of weekly number of orders and weekly-revenue, and should be interpreted as percentage changes. The sum of the coefficient is presented below each panel along with the corresponding p-value. The interaction between post and quality-level indicators is omitted for brevity. Treatment status and quality definitions are indicated below the table and are described further in the text. All regressions include business and week-state fixed effects, as well as the average search results rank for a business in a given week. Standard errors are in parentheses and are clustered at the city level.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 11: Model Estimates

	(1)	(2)	(3)	(4)
Panel A: Utility Function				
<u>Mean Effects</u>				
Dollar Rating	-0.015 (0.011)	-1.554*** (0.215)	-2.393*** (0.292)	0.648 (.874)
Rating	0.224*** (0.016)	0.265*** (0.018)	0.289*** (0.020)	0.311*** (0.041)
Ranking (quantile)	0.020*** (0.007)	0.048*** (0.009)	0.029** (0.013)	-0.043 (0.029)
Ranking in Category	-0.089*** (0.001)	-0.092*** (0.001)	-0.080*** (0.004)	-0.051** (0.022)
Average Income in Zip-Code	-0.081* (0.043)	-0.809*** (0.111)	-0.598*** (0.149)	0.216 (0.219)
<u>Random Coefficient (Dollar Rating)</u>				
σ_v			1.361*** (0.302)	1.772 (2.130)
Income \$50,000-\$100,000				-1.032 (1.694)
Income above \$100,000				2.724 (5.483)
Female				-2.104 (5.619)
Age 25 to 44				-4.138 (18.72)
Age above 45				-3.991 (18.58)
Food Category FE	Yes	Yes	Yes	Yes
Instrument	No	Yes	Yes	Yes
Panel B: Entry Costs				
Average μ	.071 *** (0.001)	.205*** (0.006)	.197*** (0.10)	.203*** (0.009)
σ	0.010 *** (0.003)	0.091*** (0.003)	.093*** (0.005)	0.098*** (0.005)
Heterogeneous μ	No	No	Yes	Yes
Observations	124061	124061	124061	124061
# of markets	5500	5500	5500	5500

Note: This table reports four separate estimates of versions of the model described in Section 6. Panel A presents the parameters of the utility function, and Panel B presents the parameters of the entry cost distribution. Columns (1) and (2) present the model without user heterogeneity. Columns (3) and (4) allow for the coefficient of dollar rating to vary by unobserved- and unobserved-and-observed user characteristics, respectively. Columns (3) and (4) also allow the distribution of entry costs to vary by household income-levels. For these columns, only the weighted average μ over income groups is presented. All regressions include CSA fixed effects. Robust standard errors are in parentheses.

* significant at 10%; ** significant at 5%; *** significant at 1%

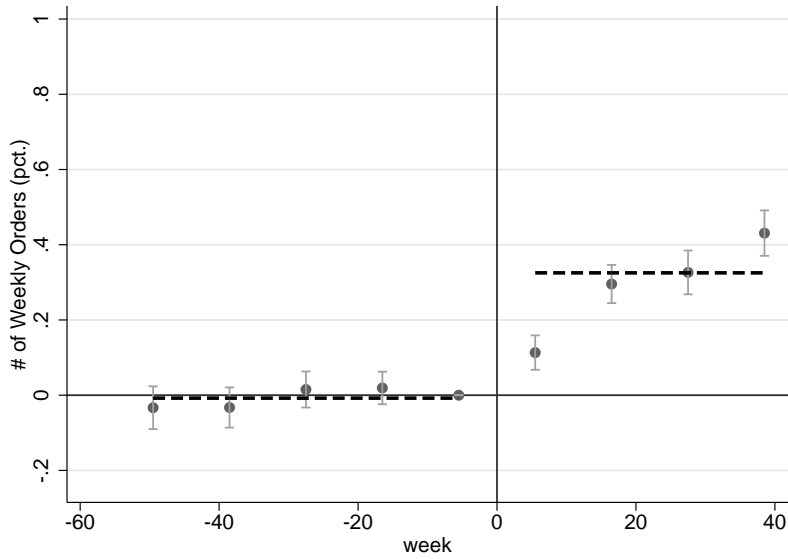
Table 12: Simulations

	10%	20%	40%	60%	80%	100%
Δ Welfare	0.40	0.85	1.28	1.51	1.64	1.81
Δ Market Size	0.41	0.94	1.58	1.98	2.19	2.58
<u>Δ Sales By Rating Quantile</u>						
1st Quantile	-0.13	-0.36	-0.64	-0.76	-0.83	-0.88
2nd Quantile	-0.12	-0.34	-0.55	-0.64	-0.79	-0.75
3rd Quantile	0.02	-0.17	-0.37	-0.53	-0.62	-0.67
4th Quantile	0.10	0.03	-0.12	-0.26	-0.38	-0.38
5th Quantile	0.12	0.32	0.47	0.46	0.40	0.36

Note: This table reports the average results of 500,000 simulated markets. The parameters used to simulate the data are presented in Column (3) of Table 12. The simulation algorithm is described in appendix E. The tables presents the percentage change in outcomes from the baseline. In the baseline, 5% of firms in the market are on the platform. Outcomes are indicated in row names and the subheading. Column headers indicate the simulated share of firms on the platform.

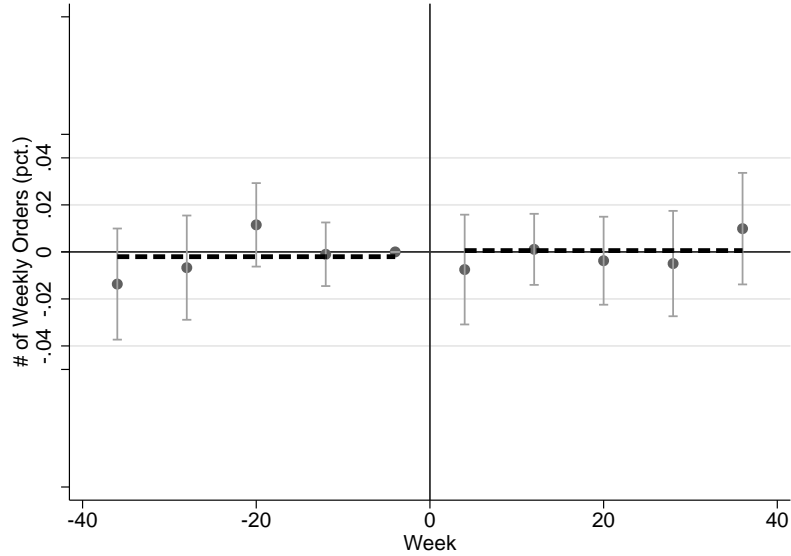
A Appendix Figures and Tables

Figure A1: Percentage Change in Market-Level Number of Orders

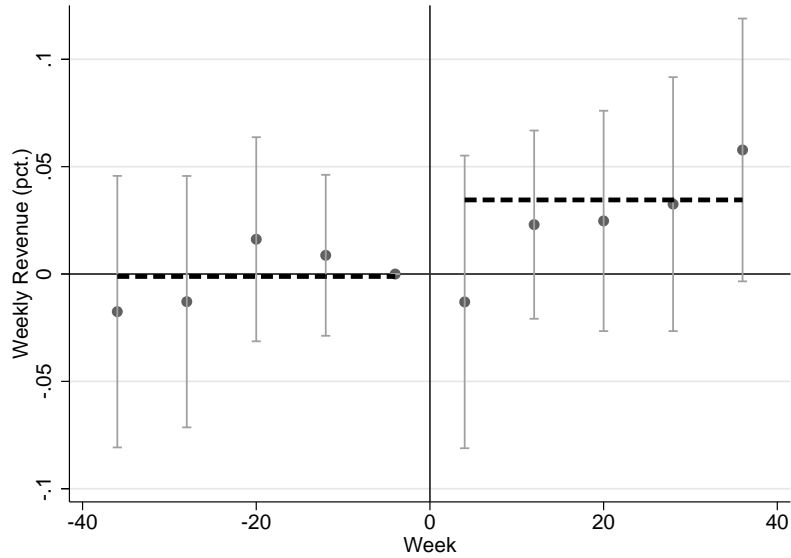


Note: This figure presents event-time estimates from a version of equation 4. The dependent variables are the inverse hyperbolic sine transformation, and should be interpreted as percentage changes. The unit of observation is city-week, including both incumbent and newly added businesses. The dots represent point estimates from regressing the dependent variable on a treatment indicator interacted with 9-weeks bins, and city and week-state fixed effects. The treatment indicator compares cities that experienced almost no change in the percentage of businesses available on the platform to cities that experienced meaningful changes. The coefficient in the first period prior to implementation is normalized to zero. The vertical bar represent 95% confidence intervals, where standard errors are clustered at the city level.

Figure A2: Impact of Entry on Incumbent Firms



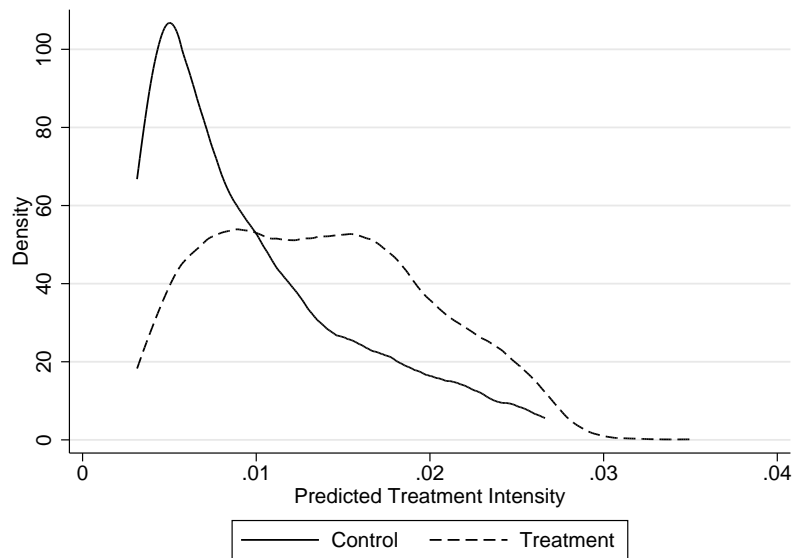
(a) Change in Weekly Number of Orders (Percentage)



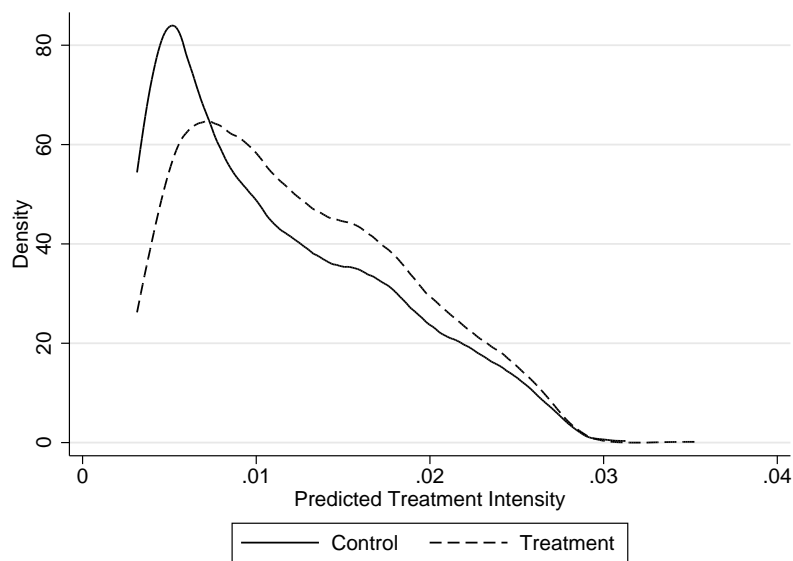
(b) Change in Weekly Revenue (Percentage)

Note: This figure presents event-time estimates from a version of equation 4. The dependent variables are the inverse hyperbolic sine transformations of weekly number of orders (Panel A) and weekly revenue (Panel B), and should be interpreted as percentage changes. The unit of observation is city-week. The dots represent point estimates from regressing the dependent variable on a treatment indicator interacted with 9-weeks bins, and city and week-state fixed effects. The treatment indicator compares cities that experienced almost no change in the percentage of businesses available on the platform to cities that experienced meaningful changes. The coefficient in the first period prior to implementation is normalized to zero. The vertical bar represent 95% confidence intervals, where standard errors are clustered at the city level.

Figure A3: Predicted Treatment Intensity by Treatment Assignment



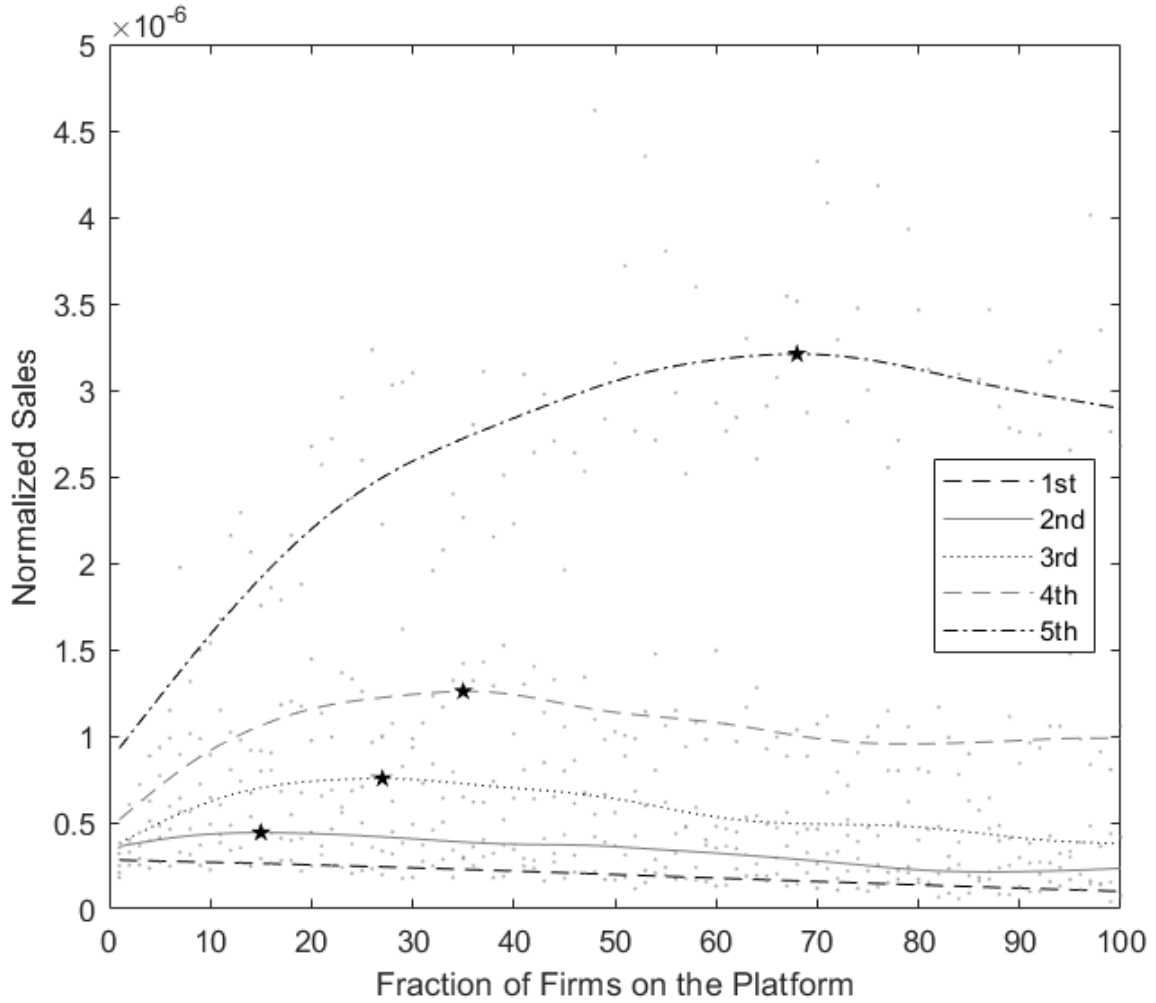
(a) Original Distributions



(b) Distribution Using Within Bin Assignment

Note: The figure presents the distribution of propensity score by treatment intensity. Propensity score are estimated on the continuous change in share of restaurants on YTP. Treatment is an indicator for above median treatment intensity. Panel A presents the distributions of propensity scores in the original data by treatment assignment. Panel B presents the distributions of propensity scores when treatment status is assigned by propensity score bins.

Figure A4: Simulations of Firms' Sales by Rating Quantile- Including Demographics



Note: This figure present the change in number of sales by rating quantile as the share of firms on the platform grows. The parameters used to simulate the data are presented in column 4 of Table 12. The simulation algorithm is described in appendix E. Each gray dot represent the average over one thousand simulations, and the dashed and solid line are the smoothing spline by rating quantile. The stars mark the maximum of each smoothing spline.

Table A1: Placebo Trends

	(1)	(2)	(3)
Panel A: Weekly Orders			
Treat*Post	-0.008 (0.007)	-0.003 (0.009)	-0.003 (0.013)
Treat*Post*Low		-0.010 (0.012)	0.007 (0.015)
$\beta_1 + \beta_2$		-0.013	0.004
Pvalue		0.179	0.686
Panel B: Weekly Revenue			
Treat*Post	-0.020 (0.018)	-0.017 (0.023)	-0.022 (0.033)
Treat*Post*Low		-0.010 (0.031)	0.047 (0.040)
Observations	1477208	1477208	740706
# of Clusters	2729	2729	2625
$\beta_1 + \beta_2$		-0.027	0.025
Pvalue		0.301	0.363
Treatment Def.	25<>75	25<>75	25<>75
Quality Def.		Median	25<>75

Note: This table reports regression coefficients from 9 separate regressions, 3 per panel. An observation is business-week. The dependent variables are the per-business inverse hyperbolic sine transformation of weekly number of orders (Panel A) and weekly-revenue (Panel B), and should be interpreted as percentage changes. Post is counterfactually set to the middle of the pre-treatment period. The sum of the coefficient is presented below each panel along with the corresponding Pvalue. The interaction between post and quality level indicators is omitted for brevity. Treatment status and quality definitions are indicated below the table and are described further in the text. All regressions include business and week-state fixed effects. Standard errors are in parentheses and are clustered at the city level.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table A2: Placebo Outcomes

	(1)	(2)	(3)	(4)	(5)
Panel A: Weekly Number of New Business					
Treat*Post	0.003 (0.003)	0.001 (0.004)	0.077 (0.084)		
Observations	321980	221935	321980		
# of Clusters	3788	2611	3788		
Panel B: Weekly Number of Review					
Treat*Post	-0.001 (0.001)	-0.001 (0.001)	0.025 (0.034)	0.000 (0.001)	-0.000 (0.002)
Treat*Post*Low				-0.002 (0.003)	-0.000 (0.003)
Observations	31833645	17587953	31833645	17587953	10751909
# of Clusters	3964	2781	3964	2781	2047
$\beta_1 + \beta_2$				-0.002	-0.001
Pvalue				0.365	0.689
Panel C: Average Weekly Rating					
Treat*Post	-0.000 (0.001)	-0.002 (0.001)	0.021 (0.028)	-0.001 (0.002)	-0.000 (0.002)
Treat*Post*Low				-0.003 (0.002)	-0.004 (0.003)
Observations	10236710	5685666	10236710	5685666	3444919
# of Clusters	3964	2781	3964	2781	2047
$\beta_1 + \beta_2$				-0.003	-0.005
Pvalue				0.049	0.009
Treatment Def.	Median	25<>75	Change	25<>75	25<>75
Quality Def.				Median	25<>75

Note: This table reports regression coefficients from 13 separate regressions, 3 in Panel A and 5 in Panels B and C. An observation is business-city in Panel A, and business-week in Panels B and C. The sample includes only non-YTP affiliated businesses and users. The dependent variables are inverse hyperbolic sine transformations and should be interpreted as percentage changes. Outcomes are indicated in the sub-headers and described further in the text. The sum of the coefficient is presented below each panel along with the corresponding p-value. The interaction between post and quality level indicators is omitted for brevity. Treatment status and quality definitions are indicated below the table and are described further in the text. All regressions include business and week-state fixed effects. Standard errors are in parentheses and are clustered at the city level.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table A3: Sensitivity to Geographical Market Definition

	By 5-Digits Zip Code				By County			
	(1) Orders	(2) Orders	(3) Revenue	(4) Revenue	(5) Orders	(6) Orders	(7) Revenue	(8) Revenue
Treat*Post	0.049*** (0.008)	0.066*** (0.011)	0.136*** (0.019)	0.180*** (0.028)	0.048* (0.019)	0.049** (0.018)	0.094* (0.045)	0.087 (0.047)
Treat*Post*Low	-0.076*** (0.011)	-0.108*** (0.014)	-0.158*** (0.026)	-0.234*** (0.034)	-0.100*** (0.023)	-0.111*** (0.022)	-0.184** (0.057)	-0.202*** (0.058)
Observations	2268248	1283178	2268248	1283178	1613929	1273658	1613929	1273658
# of Clusters	6666	4197	6666	4197	1260	921	1260	921
$\beta_1 + \beta_2$	-0.027	-0.042	-0.022	-0.054	-0.052	-0.062	-0.090	-0.114
Pvalue	0.000	0.000	0.232	0.024	0.000	0.000	0.015	0.004
Treatment Def.	Median	25<>75	Median	25<>75	Median	25<>75	Median	25<>75
Quality Def.	25<>75	25<>75	25<>75	25<>75	25<>75	25<>75	25<>75	25<>75

Note: This table reports regression coefficients from 8 separate regressions. An observation is business-week. Geographic market definition indicated in sub-headings and described further in the text. The dependent variables are the per-business inverse hyperbolic sine transformation of weekly number of orders and weekly-revenue, and should be interpreted as percentage changes. The sum of the coefficient is presented below each panel along with the corresponding Pvalue. The interaction between post and quality level indicators is omitted for brevity. Treatment status and quality definitions are indicated below the table and are described further in the text. All regressions include business and week-state fixed effects. Standard errors are in parentheses and are clustered at the city level.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table A4: Sensitivity to Alternative Market and Treatment Intensity Definitions

	By Market X Food Category				Alternative Treatment Definition			
	(1) Orders	(2) Orders	(3) Revenue	(4) Revenue	(5) Orders	(6) Orders	(7) Revenue	(8) Revenue
Treat*Post	0.030** (0.010)	0.027* (0.012)	0.067** (0.024)	0.072* (0.028)	0.047*** (0.010)	0.054*** (0.013)	0.115*** (0.022)	0.137*** (0.033)
Treat*Post*Low	-0.073*** (0.012)	-0.075*** (0.015)	-0.137*** (0.030)	-0.161*** (0.035)	-0.069*** (0.018)	-0.100*** (0.016)	-0.149*** (0.036)	-0.221*** (0.043)
Observations	1804041	1084294	1804041	1084294	2173127	781414	2173127	781414
# of Clusters	8540	6708	8540	6708	3863	2686	3863	2686
$\beta_1 + \beta_2$	-0.043	-0.047	-0.070	-0.089	-0.021	-0.046	-0.034	-0.085
Pvalue	0.000	0.000	0.000	0.000	0.071	0.000	0.153	0.003
Treatment Def.	Median	25<>75	Median	25<>75	Median	25<>75	Median	25<>75
Quality Def.	25<>75	25<>75	25<>75	25<>75	25<>75	25<>75	25<>75	25<>75

Note: This table reports regression coefficients from 8 separate regressions. An observation is business-week. Market and treatment intensity definitions indicated in sub-headings and described further in the text. The dependent variables are the per-business inverse hyperbolic sine transformation of weekly number of orders and weekly-revenue, and should be interpreted as percentage changes. The sum of the coefficient is presented below each panel along with the corresponding Pvalue. The interaction between post and quality level indicators is omitted for brevity. Treatment status and quality definitions are indicated below the table and are described further in the text. All regressions include business and week-state fixed effects. Standard errors are in parentheses and are clustered at the city level.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table A5: Sensitivity to Outliers

	Dropping Top & Bottom 5% of Cities				Randomization Inference			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Orders	Orders	Revenue	Revenue	Orders	Orders	Revenue	Revenue
Treat*Post	0.032** (0.011)	0.033* (0.016)	0.082** (0.028)	0.090* (0.040)	0.050*** [0.000]	0.065*** [0.000]	0.119*** [0.000]	0.158*** [0.000]
Treat*Post*Low	-0.058*** (0.015)	-0.060** (0.021)	-0.131*** (0.039)	-0.134* (0.056)	-0.098*** [0.000]	-0.119*** [0.000]	-0.201*** [0.000]	-0.250*** [0.000]
Observations	832690	423002	832690	423002	2173124	1321540	2173124	1321540
# of Clusters	2483	1422	2483	1422	3862	2714	3862	2714
$\beta_1 + \beta_2$	-0.026	-0.027	-0.049	-0.044				
Pvalue	0.016	0.082	0.089	0.302				
Treatment Def.	Median	25<>75	Median	25<>75	Median	25<>75	Median	25<>75
Quality Def.	25<>75	25<>75	25<>75	25<>75	25<>75	25<>75	25<>75	25<>75

Note: This table reports regression coefficients from 8 separate regressions. An observation is business-week. Columns 1 through 4 exclude the outliers cities. The dependent variables are the per-business inverse hyperbolic sine transformation of weekly number of orders and weekly-revenue, and should be interpreted as percentage changes. The sum of the coefficient is presented below each panel along with the corresponding Pvalue. The interaction between post and quality level indicators is omitted for brevity. Treatment status and quality definitions are indicated below the table and are described further in the text. All regressions include business and week-state fixed effects. In columns 1 through 4 standard errors are in parentheses and are clustered at the city level. In columns 4 through 8 randomization inference p-values based on 2000 draws are reported in square brackets.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table A6: Sensitivity to Initial Differences

	Propensity Score Weighting				Blocked Treatment Assignment			
	(1) Orders	(2) Orders	(3) Revenue	(4) Revenue	(5) Orders	(6) Orders	(7) Revenue	(8) Revenue
Treat*Post	0.044** (0.014)	0.066*** (0.016)	0.105** (0.034)	0.152*** (0.039)	0.033*** (0.009)	0.063*** (0.014)	0.090*** (0.021)	0.167*** (0.033)
Treat*Post*Low	-0.067*** (0.018)	-0.089*** (0.020)	-0.154*** (0.046)	-0.186*** (0.052)	-0.053*** (0.014)	-0.101*** (0.021)	-0.132*** (0.030)	-0.246*** (0.045)
Observations	814388	501376	814388	501376	2173127	1011399	2173127	1011399
# of Clusters	3412	2467	3412	2467	3863	2764	3863	2764
$\beta_1 + \beta_2$	-0.023	-0.023	-0.049	-0.035	-0.020	-0.037	-0.042	-0.079
Pvalue	0.045	0.070	0.111	0.333	0.029	0.003	0.042	0.006
Treatment Def.	Median	25<>75	Median	25<>75	Median	25<>75	Median	25<>75
Quality Def.	25<>75	25<>75	25<>75	25<>75	25<>75	25<>75	25<>75	25<>75

Note: This table reports regression coefficients from 8 separate regressions. An observation is business-week. In columns 1 through 4 observations are weighting by the inverse probability score, and in columns 5 through 8 treatment status is assigned by propensity score bins. See text for additional details. The dependent variables are the per-business inverse hyperbolic sine transformation of weekly number of orders and weekly-revenue, and should be interpreted as percentage changes. The sum of the coefficient is presented below each panel along with the corresponding Pvalue. The interaction between post and quality level indicators is omitted for brevity. Treatment status and quality definitions are indicated below the table and are described further in the text. All regressions include business and week-state fixed effects. Standard errors are in parentheses and are clustered at the city level.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table A7: Sensitivity to Partnership Date Definition

	Partnership on March 19th				Excluding Intermediate Period			
	(1) Orders	(2) Orders	(3) Revenue	(4) Revenue	(5) Orders	(6) Orders	(7) Revenue	(8) Revenue
Treat*Post	0.064*** (0.010)	0.078*** (0.012)	0.150*** (0.025)	0.190*** (0.031)	0.058*** (0.011)	0.074*** (0.013)	0.138*** (0.027)	0.184*** (0.032)
Treat*Post*Low	-0.114*** (0.013)	-0.136*** (0.016)	-0.231*** (0.033)	-0.287*** (0.040)	-0.112*** (0.014)	-0.133*** (0.016)	-0.228*** (0.035)	-0.287*** (0.041)
Observations	2068197	1257612	2068197	1257612	2040166	1231415	2040166	1231415
# of Clusters	3862	2714	3862	2714	3832	2723	3832	2723
$\beta_1 + \beta_2$	-0.050	-0.058	-0.080	-0.096	-0.054	-0.059	-0.091	-0.103
Pvalue	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Treatment Def.	Median	25<>75	Median	25<>75	Median	25<>75	Median	25<>75
Quality Def.	25<>75	25<>75	25<>75	25<>75	25<>75	25<>75	25<>75	25<>75

Note: This table reports regression coefficients from 8 separate regressions. An observation is business-week. In columns 5 through 8 observation between February and March 19th are excluded from the analysis. The dependent variables are the per-business inverse hyperbolic sine transformation of weekly number of orders and weekly-revenue, and should be interpreted as percentage changes. The sum of the coefficient is presented below each panel along with the corresponding Pvalue. The interaction between post and quality level indicators is omitted for brevity. Treatment status and quality definitions are indicated below the table and are described further in the text. In columns 1 through 4 post is indicator for before and after March 19th. All regressions include business and week-state fixed effects. Standard errors are in parentheses and are clustered at the city level.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table A8: Sensitivity to Business Attrition

	Excluding (Eventual) Exiters				Exit as Zero Sales & Revenue			
	(1) Orders	(2) Orders	(3) Revenue	(4) Revenue	(5) Orders	(6) Orders	(7) Revenue	(8) Revenue
Treat*Post	0.055*** (0.010)	0.074*** (0.013)	0.125*** (0.026)	0.171*** (0.032)	0.051*** (0.011)	0.067*** (0.013)	0.134*** (0.027)	0.177*** (0.034)
Treat*Post*Low	-0.099*** (0.014)	-0.123*** (0.016)	-0.194*** (0.035)	-0.246*** (0.042)	-0.096*** (0.013)	-0.113*** (0.016)	-0.201*** (0.034)	-0.239*** (0.042)
Observations	1914696	1116849	1914696	1116849	2282647	1387335	2282647	1387335
# of Clusters	3582	2509	3582	2509	3872	2722	3872	2722
$\beta_1 + \beta_2$	-0.044	-0.049	-0.069	-0.075	-0.044	-0.046	-0.067	-0.062
Pvalue	0.000	0.000	0.003	0.009	0.000	0.000	0.003	0.020
Treatment Def.	Median	25<>75	Median	25<>75	Median	25<>75	Median	25<>75
Quality Def.	25<>75	25<>75	25<>75	25<>75	25<>75	25<>75	25<>75	25<>75

Note: This table reports regression coefficients from 8 separate regressions. An observation is business-week. Column 1 through 4 excluded all firms that exit YTP during the analysis, in columns 5 through 8 sales and revenue of existing firms are coded as zero instead of missing. The dependent variables are the per-business inverse hyperbolic sine transformation of weekly number of orders and weekly-revenue, and should be interpreted as percentage changes. The sum of the coefficient is presented below each panel along with the corresponding Pvalue. The interaction between post and quality level indicators is omitted for brevity. Treatment status and quality definitions are indicated below the table and are described further in the text. All regressions include business and week-state fixed effects. Standard errors are in parentheses and are clustered at the city level.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table A9: Sensitivity to City-Level Time Trends

	(1)	(2)	(3)	(4)
	Orders	Orders	Revenue	Revenue
Treat*Post	0.060*** (0.010)	0.067*** (0.012)	0.139*** (0.025)	0.165*** (0.033)
Treat*Post*Low	-0.102*** (0.014)	-0.126*** (0.016)	-0.217*** (0.035)	-0.274*** (0.043)
Observations	2173244	1321619	2173244	1321619
# of Clusters	3875	2725	3875	2725
$\beta_1 + \beta_2$	-0.042	-0.059	-0.078	-0.109
Pvalue	0.000	0.000	0.001	0.000
Treatment Def.	Median	25<>75	Median	25<>75
Quality Def.	25<>75	25<>75	25<>75	25<>75

Note: This table reports regression coefficients from 4 separate regressions. An observation is business-week. The dependent variables are the per-business inverse hyperbolic sine transformation of weekly number of orders and weekly-revenue, and should be interpreted as percentage changes. The sum of the coefficient is presented below each panel along with the corresponding Pvalue. The interaction between post and quality level indicators is omitted for brevity. Treatment status and quality definitions are indicated below the table and are described further in the text. All regressions include business and week-state fixed effects. In addition, the specification allows for city-level time trends in establishment outcome. Standard errors are in parentheses and are clustered at the city level.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table A10: Effect of Entry on Users' Type

	Weekly Orders	Weekly Expenditure	Dollar Ratings	Variety (Restaurants)	Variety (Categories)
Treat*Post	-0.004 (0.005)	-0.004 (0.006)	0.002 (0.004)	-0.012*** (0.004)	-0.003 (0.004)
# of Clusters	2842	2843	2810	2843	2843

Note: This table reports regression coefficients from 6 separate regressions. An observation is one user. The sample includes only users in cities that were effected by the partnership and only includes behavior in the period after integration. The dependent variables are the inverse hyperbolic sine transformation of the outcomes indicated in sub-headings, and should be interpreted as percentage changes. Coefficients represent the interaction between a dummy for user joined YTP post implementation and treatment status is an indicator for whether the city experienced an above median change in the percentage of businesses on YTP. All regressions include city fixed effects. Number of observations are not reported in order to protect proprietary data. Standard errors are in parentheses and are clustered at the city level. * significant at 10%; ** significant at 5%; *** significant at 1%

Table A11: Effect of Entry on Incumbents' Subsequent Ratings- Robustness Checks

	Placebo Test				Type of Raters			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treat*Post	0.003 (0.008)	-0.001 (0.010)	0.008 (0.009)	-0.007 (0.015)	0.001 (0.003)	0.001 (0.003)	0.003 (0.003)	0.002 (0.004)
Treat*Post*Low			-0.059** (0.026)	-0.027 (0.020)			-0.008 (0.008)	-0.004 (0.010)
Observations	834964	488933	393194	233178	1464204	859456	695502	413378
# of Clusters	3356	2243	3110	2095	3827	2655	3680	2569
$\beta_1 + \beta_2$			-0.051	-0.034			-0.005	-0.002
Pvalue			0.050	0.109			0.539	0.810
Treatment Def.	Median	25<>75	Median	25<>75	Median	25<>75	Median	25<>75
Quality Def.			25<>75	25<>75			25<>75	25<>75

Note: This table reports regression coefficients from 8 separate regressions. In columns 1 through 4 an observation is business-week, in columns 5-8 an observation is user who rated a YTP restaurant and. The outcome is described further in the text. The dependent variable is the inverse hyperbolic sine transformation of the average rating received in a given week (columns 1-4) and the average rating given by user (column 5-8), and should be interpreted as percentage changes. The sum of the coefficient is presented below each panel along with the corresponding Pvalue. The interaction between post and quality level indicators is omitted for brevity. Treatment status and quality definitions are indicated below the table and are described further in the text. All regressions include business and week-state fixed effects. Standard errors are in parentheses and are clustered at the city level.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table A12: Effect of Entry on Advertising Purchases By Expense Type

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Fixed Revenue						
Treat*Post	-0.015 (0.009)	-0.024* (0.012)	-0.021 (0.012)	-0.031 (0.017)	-0.028 (0.015)	-0.052* (0.022)
Treat*Post*Low			0.010 (0.018)	0.041 (0.022)	0.003 (0.022)	0.055* (0.028)
$\beta_1 + \beta_2$			-0.011	0.010	-0.025	0.003
Pvalue			0.422	0.455	0.113	0.836
Panel B: Variable Revenue						
Treat*Post	-0.031** (0.013)	-0.040** (0.017)	-0.037** (0.018)	-0.034 (0.027)	-0.054** (0.023)	-0.083** (0.034)
Treat*Post*Low			0.009 (0.025)	0.045 (0.032)	0.023 (0.032)	0.097** (0.041)
Observations	4409516	2623347	4409516	2173244	2623347	1321619
# of Clusters	3964	2781	3964	3875	2781	2725
$\beta_1 + \beta_2$			-0.028	0.011	-0.031	0.014
Pvalue			0.135	0.553	0.189	0.537
Treatment Def.	Median	25<>75	Median	Median	25<>75	25<>75
Quality Def.			Median	25<>75	Median	25<>75

Note: This table reports regression coefficients from 12 separate regressions, 6 per panel. An observation is business-week. The dependent variable is the inverse hyperbolic sine transformation of total advertising expenditure on Yelp by expense type and should be interpreted as percentage changes. The sum of the coefficient is presented below each panel along with the corresponding Pvalue. The interaction between post and quality level indicators is omitted for brevity. Treatment status and quality definitions are indicated below the table and are described further in the text. All regressions include business and week-state fixed effects. Standard errors are in parentheses and are clustered at the city level.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table A13: Simulations Using Demographic Data

	10%	20%	40%	60%	80%	100%
Δ Welfare	0.49	1.19	1.88	2.34	2.57	2.77
Δ Market Size	0.62	1.70	2.95	3.88	4.38	4.79
<u>Δ Sales- By Rating Quantile</u>						
1st Quantile	-0.03	-0.08	-0.21	-0.36	-0.50	-0.63
2nd Quantile	0.08	0.09	-0.07	-0.20	-0.44	-0.41
3rd Quantile	0.25	0.48	0.40	0.07	-0.05	-0.24
4th Quantile	0.29	0.63	0.74	0.51	0.34	0.32
5th Quantile	0.29	0.79	1.31	1.59	1.54	1.36

Note: This table reports the average results of 500,000 simulated markets. The parameters used to simulate the data are presented in column 4 of Table 12. The simulation algorithm is described in appendix E. The tables presents the percentage change in outcomes from the baseline. In the baseline, 5% of firms in the market are on the platform. Outcomes are indicated in row names and subheading. Column headers indicate the simulated share of firms on the platform.

B Proofs

Proof of Lemma 1 In this proof (and only this proof) I assume that the idiosyncratic error is distributed according to the extreme value type 1 distribution, which is the distribution used on the structural model. Also, without loss of generality, I present the proof using expected quality distribution. Assuming a symmetric equilibrium, high-quality firms and low-quality firms solve:

$$\begin{aligned}
\max_{p_j} \Pi_j^h &= H(E[\max(u)] - \omega) * \frac{e^{q_h - p_j}}{1 + e^{q_h - p_j} + \frac{n-1}{2}e^{q_h - p_h} + \frac{n}{2}e^{-p_l}}(p_j - r) \\
&= \frac{2(r - p_h + 1)e^{p_h + p_l + p_j} + n(r - p_h + 1)e^{p_h + p_j} - (n-1)(-r + p_h - 1)e^{p_l + q_h + p_j} + 2e^{p_h + p_l + q_h}}{(2e^{p_h + p_l + q_h} + 2e^{p_h + p_l + p_j} + ne^{p_h + p_j} + (n-2)e^{p_l + q_h + p_j})^2} \\
&= 2(r - p_h + 1)e^{p_h + p_l + p_h} + n(r - p_h + 1)e^{p_h + p_h} - (n-1)(-r + p_h - 1)e^{p_l + q_h + p_h} + 2e^{p_h + p_l + q_h} \\
&= 2Ae^{p_h + p_l} + [(n-1)A - 2]e^{p_l + q_h} + nAe^{p_h} \tag{11}
\end{aligned}$$

$$\begin{aligned}
\max_{p_j} \Pi_j^l &= H(E[\max(u)] - \omega) * \frac{e^{-p_j}}{1 + e^{-p_j} + \frac{n-1}{2}e^{-p_l} + \frac{n}{2}e^{q_h - p_h}}(p_j - r) \\
&= \frac{2(-r + p_j - 1)e^{p_h + p_l + p_j} + (n-1)(-c + p_j - 1)e^{p_h + p_j} + n(-r + p_j - 1)e^{p_l + q_h + p_j} - 2e^{p_h + p_l}}{(2e^{p_h + p_l + p_j} + 2e^{p_h + p_l} + (n-2)e^{p_h + p_j} + ne^{p_l + q_h + p_j})^2} \\
&= 2(-r + p_j - 1)e^{p_h + p_l + p_l} + (n-1)(-r + p_j - 1)e^{p_h + p_l} + n(-r + p_j - 1)e^{p_l + q_h + p_l} - 2e^{p_h + p_l} \\
&= 2Be^{p_h + p_l} + [(n-1)B - 2]e^{p_h} + nBe^{p_l + q_h} \tag{12}
\end{aligned}$$

Where $A = (p_h - r - 1)$ and $B = (p_l - r - 1)$.

Combine the 11 and 12 to get:

$$\begin{aligned}
[A((n-1)B - 2) - nAB]e^{p_h} &= -[nBA - B((n-1)A - 2)]e^{p_l + q_h} \\
e^{p_h} &= \frac{B(A+2)}{A(B+2)}e^{p_l + q_h} \\
e^{p_h} &= \delta e^{p_l + q_h} \tag{13}
\end{aligned}$$

Substitute e^{p_h} with $\delta e^{p_l + q_h}$ into 12 to get:

$$e^{p_l} = -\left[\frac{n-1}{2} - \frac{1}{B} + \frac{n}{2\delta}\right]$$

Since $\delta = e^{p_h - p_l - q_h} > 0$, $\frac{n}{2\delta} > 0$, and $\frac{n-1}{2} > 0$, it must be that $B > 0 \implies p_l > r + 1$. Similarly, one can show that $p_h > r + 1$ by plugging 13 into 11.

Now, assume towards contradiction that $p_h = p_l$. This implies that $A = B$ and $\delta = \frac{B(A+2)}{A(B+2)} = 1$. But then by Equation 13 $e^{p_h} = e^{p_h + q_h} \implies q_h = 0$, a contradiction.

Now, assume toward contradiction the $B > A$. Then $\delta = \frac{B(A+2)}{A(B+2)} > 1$. However, noting that $B > A \implies p_l > p_h$ together with Equation 13: $\ln(A) = p_h - p_l - q_h$ implies that $0 > p_h - p_l - q_h = \ln(A) \implies A < 1$, a contradiction. Therefore $p_l < p_h$ implying that $A < 1 \implies \ln(A) < 0 \implies p_h - p_l - q_h < 0 \implies q_h - p_h > -p_l$.

Proof of Lemma 2 Inspecting equations 1 and 2, the first element does not depend of firm's quality. In contrast, lemma 1 implies that the second term, the per-unit profit, is larger for high-quality firms compared to low-quality firms. The third term captures the number of sales: The first element is equivalent for high-quality and low-quality, and the second and third elements are larger for high-types since $p_h - q_h < p_l$. Thus, high-quality firms sell more, charge higher prices, and generate higher profits.

Proof of Proposition 1 The following proof shows the effect of adding a high-quality firm. The proof when adding a low-quality firm is similar. Also, I assume here that N is sufficiently large. It is simpler to study the derivative of the natural logarithm, $\frac{\partial \ln(\pi_i)}{\partial N}$:

$$\begin{aligned} \ln(\pi_h) = & \ln(H(E_\epsilon[\max(u_{ij})] > \omega + c_i)) + \ln(p_h - r) + \\ & N_h \ln(1 - \tilde{G}(0)) + N_l \ln(1 - \tilde{G}(\bar{U}_l - \bar{U}_h)) + \ln(1 - G(\omega - \bar{U}_h)) \end{aligned} \quad (14)$$

$$\begin{aligned} \ln(\pi_l) = & \ln(H(E_\epsilon[\max(u_{ij})] > \omega + c_i)) + \ln(p_l - r) + \\ & N_h \ln(1 - \tilde{G}(\bar{U}_h - \bar{U}_l)) + N_l \ln(1 - \tilde{G}(0)) + \ln(1 - G(\omega - \bar{U}_l)) \end{aligned} \quad (15)$$

Where $\bar{U}_j \equiv q_j - p_j$ for $j \in \{h, l\}$. By the envelope theorem, we can ignore the impact the N_h has on p_h , p_l , $U_h(p_h)$, and $U_l(p_l)$, which means that we can ignore the second, fourth, and fifth terms in equations 14 and 15.³⁷ The third term in equation 15 is more negative than the third term in equation 14, and both are negative:³⁸

$$\ln(1 - \tilde{G}(\bar{U}_h - \bar{U}_l)) < \ln(1 - \tilde{G}(0)) < 0 \text{ since } \bar{U}_h > \bar{U}_l$$

³⁷ This depends on the assumption that N is large is quite large, so that the change in N_h is comparably small. Alternatively, we can reconstruct the model to have a continuum of firms with a measure μ and explore the impact of an infinitesimal increase in μ . For this formulation, however, we would have to make a direct assumption that $u'_\mu > 0$, at least in expectation, i.e., that consumers have some love-for-variety (which could be motivated by the model above).

³⁸ The same would be true if we were considering an increase in N_l :

$$\ln(1 - \tilde{G}(0)) < \ln(1 - \tilde{G}(\bar{U}_l - \bar{U}_h)) < 0 \text{ since } \bar{U}_h > \bar{U}_l$$

Since the effect of the first term $\frac{\partial \ln(H(E_\epsilon[\max(u_{ij}) > \omega + c_i])}{\partial N_h}$ is the same in on both types, this proved the second part of proposition 1.

Finally, we have to show that the first term is increasing in N_h . Since the both $\ln(\cdot)$ and $H(\cdot)$ are monotonically increasing transformations, and $\omega + c_i$ are independent of N_h , we only have to show that $E_\epsilon[\max(u_{ij})]$ is increasing in N_h . This follows from the fact that the maximum is increasing in the number of draws.³⁹ Together, this implies that the total effects on $\ln(\pi_h)$ and $\ln(\pi_l)$ are ambiguous, which established the first part of the proposition.

C Data Appendix

Sample selection and main results While Yelp keeps data on when restaurants join and exit YTP, I found multiple cases where transactions were made prior to a business ‘entering’ the platform or after the business ‘exited’ the platform. In cases on inconsistent data, I always code entry as the earliest date of the two and exit the the later date. For this reason, I leave a margin for 8 weeks at the beginning and end of my sample to separate between businesses with zero sales to businesses that have left the platform. Weeks in which I do not observe any transactions are coded as zero orders and zero revenue. Sales and revenue are coded as missing for the week before entry or after exit.⁴⁰ The final data used in the analysis consist of 88 weeks, from March 2017 to December 2018. I limit the analysis to cities in which there are ten or more businesses on the standard Yelp platform, since in very small places treatment intensities are extremely large mechanically. I excluded businesses that are marked by Yelp as bogus, spammy, or that are removed from users’ search results. 310 businesses in 291 cities are dropped from the analysis which amount to less than 30,000 observations in my data. The final sample consists of 3,956 cities. For the main part for the analysis, I use only the incumbent businesses, which joined YTP prior to the partnership with Grubhub; there are a total of 56,493 incumbent businesses and over 4 million business-week observations.

The Yelp system does not store historic businesses’ star-ratings. To calculate businesses’ rating on the eve of integration, I take the mean over all preceding review. I exclude reviews

³⁹ A simple, yet not very elegant proof, denote $\max(x_1, \dots, x_m)$ by a and the pdf of a as $g(a)$, then:

$$\begin{aligned} E[\max(x_1, \dots, x_m, x_{m+1})] &= \int \{P(x_{m+1} > a)E[x_{m+1}|x_{m+1} > a] + (1 - P(x_{m+1} > a))a\}g(a)da \\ &= \int \tilde{a}g(a)da > \int ag(a)da = E(\max(x_1, \dots, x_m)) \end{aligned}$$

⁴⁰ Appendix D.2 tests the sensitivity of the results to precise definitions.

that are marked by Yelp as untrustworthy, or are removed from business' page. To test whether this is a good approximation, I use the same method to calculate the current Yelp rating and find that there is a correlation of over 0.95. Restaurant categories are based on Yelp's classification. For the robustness checks presented in appendix D.2, I include only the top 21 most prevalent food categories (out of 244), which include 87% of all observations in the sample. Generally, Yelp collects little demographic information on its users. Users are encouraged to enter their gender and date of birth, but I found the fields to be mostly missing in my sample. Thus, I do not use individual-level characteristics in my analysis. For each user, however, I do observe the full history of transactions on YTP, which is used to differentiate between new and repeating consumers.

Prices As mentioned in Section 5.4 price data is extremely problematic since the data only includes prices for ordered items, as opposed to menu prices, and not all dish modifications are recorded. To address this issue, I attempt to identify 'true' prices using an algorithm developed for Luca & Reshef (in writing). The algorithm takes several step: (1) drop all item for which the name suggests possible modifications (e.g., 'customize', 'create' etc.) as well as items that were discounted or orders when a coupon was used, (2) Include only prices that appear 3 or more times, but no less than 10% of total price observations for the item, (3) calculate the inter quantile range (in terms of time) for each price level and excluded observations that are above (below) 1.5 times the upper (lower) bound, (4) find the first and last occurrence of each item-price, and (5) use only prices that do not overlap. This algorithm is extremely restrictive and ultimately discards more than half of the food items in the sample.

Search Orders data and search data are handled by different parts of the organization, and more importantly, are stored in different data clusters. Consequently, joining the two datasets is not a trivial task. To identify the search sessions which lead to an order, I develop an algorithm that matches each order with the most recent search session conducted by the user prior to finalizing the order. The algorithm has several disadvantages: First, it will not be able to match an order to a search session if the user was not signed in during the search process. Second, when a user performs multiple searches on the same day, the algorithm only picks up the last session. This might be an issue if users use multiple search sessions to choose a restaurant to order from. Though these issues create additional noise and reduce statistical power, they are unlikely to bias the results in any particular direction.

Advertising Some additional institutional details that are omitted from the text for brevity: Profile enhancement is a bundle designed to increase the attractiveness of Yelp business page. It includes several upgrades such as slideshow and videos, access to Yelp account management support and data analysis tools, and removal of ads purchased by

competitors the Yelp business page. Advertisements on Yelp appear in three places: (1) Ads appear at a premium location, above the organic Yelp search results, for keywords that are related to the business. (2) Ads appear on competitors' businesses pages (only if they are not advertising on Yelp as well) (3) On the Yelp mobile app. Payment is based, for the most part, on cost per click (CPC). Bidding for ads on Yelp is done using Yelp's auto-bidding algorithm; the only choice the client has is the period advertisement budget, which is usually exhausted. For the most cases, profile enhancement and targeted ads are purchased in one package. Some, more sophisticated, businesses purchase a custom or a la carte services. This type is both very rare and not well documented, so I exclude it from my analysis. Yelp uses a complicated algorithm to determine the payment per ads which includes user clicks, traffic, and actions on the business page. The Yelp auto-bidding algorithm cannot be manually overridden by the client. The algorithm determines key words to bid on, as well as bid amounts. I use data on business-level weekly revenue collected by Yelp to estimate the effect of entry on advertising behavior. I restrict attention to campaigns purchased at the local-level, excluding national- or franchise-level campaigns.

Placebo tests To test the validity of the research design, I consider three outcomes: First, the weekly flow for new businesses that are classified by Yelp as either 'food' or 'restaurants'. Second, the weekly flow of new ratings per business, and the average rating given. I include only businesses that are classified as either 'food' or 'restaurants' and exclude businesses that are marked as bogus, spammy, or that are removed from users' search. Importantly, businesses participating in YTP are also excluded from the analysis. I also exclude review that are marked as untrustworthy, were removed by Yelp, were given by paid users, or that are given by consumers that use YTP. To test alternative explanations for the increase in ratings, I test whether selection of more lenient reviewer is driving the results.⁴¹ To test for reviewer leniency, I construct the leave out average of all reviews given by a user. Weekly ranking in search order results is the average weekly rank across all search results in which the business appears. Businesses that are marked as bogus, spammy, or that are removed from users' search, and review that are marked as untrustworthy, were removed by Yelp, or were given by paid users, are excluded. Naturally, user who only gave one review in total, are excluded as well.

Data For Structural Model Demographic information on population, gender, age and income comes from the American Community Survey 5-Year Data 2017. Total population in each county is divided by 3, to approximate the number of households. I experimented

⁴¹ I also test a specification in which the implementation date is countfactually set at the middle of the pre-period. The results are reported in appendix [D.3](#).

with several factors ranging from 2 to 4, and it appears the main qualitative results are robust to such changes. Age is binned into 10-24, 25-44, and the omitted category is 45 and above. Adjusted gross income is binned by annual household income of (in thousands): below \$50, \$50-\$100, and \$100-\$200. These data are available at the county level and are merged with the city-level data by name and state name combination. Multiple matches are determined randomly. Data are merged with the main dataset by 5-digit zip codes. CSA data are merged by city-state name and county-state fips code (when available) combination. Beside, the county-level income data, I also use data on income at the zip-code level is used as part of businesses attributes. These data come from the IRS Individual Income Tax ZIP Code Data 2016. I use a continuous measure of the 6 Adjusted Gross Revenue bins presented above.

In order to reduce computation burden and since many businesses do not any orders at a given week, I aggregate the data by city and 20 weeks bins. The data are trimmed to include only 4 bins for each city, two in the period before integration and two afterwards. The outcome variable is the total number of order per business during each 20 weeks period. Since the Dollar Ratings for a given business rarely change over time, I use the same rating for a given business in all four markets. Rating of businesses are calculated by the rating at the mid-point of each 20 weeks period. Ranking and ranking within category are defined by the relative ranking of the businesses in a city-period or city-period-type combination. Total ranking in the city take values of 1 to 5 by the ranking quantile in the city. The reason to use the ranking quantile as opposed to absolute rank is that, for the simulations, the number of businesses increases and so does the range of ranks. Mechanically, at very large size the rank number can increase by hundreds and even thousands of percents and renders all other covariates meaningless. For the decomposition by rating level, I use rating quantiles across all cities in the sample. I aggregate food categories to 10 main categories.⁴² To improve simulation performance, the final sample excludes markets with less than 5 businesses (less than 5% of the sample) or more than 250 (less than 1% of cities). Finally, since the discrete choice models I use do not allow for zero market shares, I correct for zero market share by adding one order to that business in that city-period time.

To get the share of users of the platform out of the total number of potential buyer, I use the search data. I define user interested in the platform as the number of unique user IDs in a given city-period combination who have search for variations of the words “delivery”, “takeout” or “Pickup”, or use any of the YTP filters. There are some limitations to this

⁴² The ten categories are: Asian, pizza, Mexican & Latin, European, Arab-Indian, meat & seafood, American, coffee & pastry, sandwiches, and other.

definition: First, while I am interested solely in food orders, this algorithm will also pick up individuals who are interested in, for instance, flower delivery. Second, if users conduct search without logging in, the system will document them as new user. Both of these types of issues are likely to result in overestimation of the number of users on the platforms. Accordingly, the mean share of platform users across markets is 25%, which seems excessive. Nevertheless, though this is an inaccurate measure, it is unlikely to be correlated with any of the mechanisms of interest and thus I would not change the qualitative nature of the results.

D Robustness Checks Appendix

This Appendix discusses the robustness checks conducted for the main and subsequent results. All relevant figures and tables can be found in Appendix A. Subsection D.1 presents tests for the validity of the research design described in Section 4.2. Subsection D.2 presents the robustness checks for Section 5.3, robustness of the main results. Finally, subsection D.3 presents the robustness checks for the investment in quality results presented in Section 5.4.

D.1 Validity of the Research Design

This section presents additional tests to support the parallel trend assumption, which is the key identifying assumption of the difference-in-differences research design. First, I test for differences in trends between treated and control cities in the period before the partnership with Grubhub, and second, I test for difference in trends on outcomes that are unlikely to be affected by the institutional change. The research design is presented in Section 4.2.

Pre-trends The first suggestive evidence of parallel trends absent of treatment is to examine whether the main outcomes of interest trend similarly in the prior to the Grubhub partnership. Graphic evidence are presented and discussed in Section 5. In Table A1, I present a formal placebo test in which I counterfactually set the integration date to the middle of the pre-treatment period. I do not find any significant effects of the placebo on the average effect (column 1) or when examining the effect on high- or low-quality firms separately (columns 2 and 3). This results suggests that the main results are not driven by initial differences in trends between treated and control cities.

Placebo on non-YTP outcomes A second potential concern is the break in trends is driven by other unobserved changes at the city level that are unrelated to the Grubhub partnership. If that is indeed the case, then we can expect to find significant differences in other city-level outcomes, not directly related to food ordering. I conduct several placebo

tests to examine whether the partnership is correlated with outcomes of *non-YTP* businesses, such as the number of businesses on Yelp, the restaurant average weekly ratings, and the number of new weekly reviews per business. Table A2 presents the results. Panel A presents the results of estimating equation 4 the percentage change in the number of new restaurants on Yelp as the outcome variable. I do not find any significant effects, which suggests that, in general, the food industry is growing similarly in treated cities and control cities. Due to the large number of observations, Panels B and C present results at the monthly level. Again, columns 1 through 3 do not find any significant effects of treatment on the average weekly rating, or the number of review per business for *non-YTP restaurants*. Similarly, columns 4 and 5 present null effects of the partnership by quality levels.⁴³ Taken together, the null findings suggest that the results are not driven by unobserved changes in the city, the restaurant industry, or Yelp usage.

D.2 Robustness of the Main Results

Market definition-geographic area Table A3 presents the estimation results using alternative geographical definitions for the relevant market. Note that the number of observations decreases since not all observations include zip-code and county data. Columns 1-4 and 5-8 present the results when using the 5-digit code and county as the relevant markets, respectively. The first two columns in each group show the effect on number of orders and the last two columns in each group show the effect on weekly revenue. Qualitatively, the results are similar to the main estimation results: Entry leads to more sales and higher revenue for high-quality restaurants, and vice versa for low quality businesses. The point-estimates of the effects vary across specifications and market definition. This is especially true for when estimating the effect on revenue, using the county as the relevant market due to the low number of clusters and the noisiness of weekly revenue. Nevertheless, estimates are centered around the main results, do not change signs, and are generally statistically significant. For instance, columns 2 and 6 estimate the effect of entry on weekly orders using the sharp definitions for treatment and rating. They find a treatment effect of 4.9% to 6.6% for high types and -6.2% to -4.2% for low types. In comparison, using the city as the relevant market, I find an effect of 6.5% for high types and -5.4% for low types (column 2 in Table 4).

Market definition-food category An alternative definition treats each food category (pizza, Chinese, Mexican etc.) and city combination as a separate market. There are a few important disadvantages to this definition: First, it imposes strong restrictions on consumers'

⁴³ In contrast, in section 5.4 I find that treatment does effect the average and high-quality weekly ratings of YTP restaurants in treated cities compared to untreated cities.

decision-making process. In particular, this market definition implicitly assumes that consumers first decide which type of food they want to eat and only then choose the particular restaurant. This assumption is violated if, for instance, consumers search all restaurants in their area and choose the highest rated one. Second, even if restaurants only compete within food category, there are likely to be positive spillovers across categories, i.e., the SUTVA assumption is unlikely to hold for the market size effect. To see why this is the case, consider the implications of a market expansion in only one food category; the definition above implies that this will have no spillovers to other food types, and that consumers in that city will keep purchasing other food types from alternative channels. This assumption is extremely restrictive and is likely to fail. Finally, the borders of “food categories” are not clearly defined. There are over 244 unique food categories in the data, and some, such as Japanese food and sushi, are clearly not mutually exclusive and are likely to be decent substitutes. To address this issue without taking a stand on food “similarities”, I restrict my analysis to the 21 most prevalent food types (food categories with more than 50,000 observations), which constitute 87% of all observations.⁴⁴ I then reconstruct the entire dataset using the city-category market definition. I reapply the same rules as the main analysis, including redefining treatment intensities and high and low quality firms.

The results are presented in the first part of Table A4. First of all, note that the number of observations decrease since we restrict attention to a subset of restaurants, and since, as in the main analysis, very small markets are excluded from the analysis. In contrast, the number of clusters substantially increase because now each city is separated into several markets. The main qualitative results are robust to the change in market definition; I find statistically significant positive effect on weekly sales and revenue for high-quality firms, and vice versa for low-quality firms. The effects on low firms are about the same magnitude for low-quality firms, averaging around -4.5% and -8% for sales and revenue, respectively. The effects on high firms are smaller than the main specification. For instance, the estimated effect on high-quality restaurants’ weekly number of orders is only 2.7% – 3% compared to 5% – 6.5% in the main analysis. This result is consistent with arguments regarding positive spillovers across “markets” which introduce attenuation bias to the estimated coefficient.

Treatment definition The second part of Table A4 presents the results using the alternative market definition. This treatment definition is substantially noisier, especially due to places that had only a handful of restaurants on YTP prior to the integration. For instance, by that definition, some markets have a treatment intensity of over 200%.

⁴⁴ These categories are (in the order of importance): pizza, Chinese, sandwiches, Mexican, traditional American, Japanese, Indpak, hotdogs, Thai, Mediterranean, Italian, breakfast and brunch, seafood, cafes, burgers, new American, barbecue, delis, Asian fusion, and Vietnamese.

Nevertheless, the estimates of the effect of entry on weekly sales and revenue are robust to the alternative definition of treatment intensity; all coefficients have the same size and similar magnitudes as the main analysis, though the revenue estimates are substantially noisier. For example, column 6 suggests that, using the sharp definitions of both treatment and ratings, entry increases weekly sales of high-quality restaurants by 5.4% and decreases weekly sales by 4.6% compared to 6.5% and -5.4% when using the standard treatment definition.

Outliers To test whether the main results are driven by outliers, I perform two separate robustness checks. The results are presented in Table A5. First, I exclude the cities that are in the top (bottom) 5% in terms of the number of businesses on YTP prior to integration. This exercise turns out to be quite restrictive: To begin with, many small towns had only a handful of businesses prior to integration, so there is substantial mass at 5%. Additionally, the largest cities, with the most incumbent businesses, naturally contribute the most observations to the analysis. Thus, excluding outliers reduces both the number of clusters and the number of observations sententially. As columns 1-4 show, the qualitative results are similar to the main specification and are statistically significant. The estimated size of the effects diminishes in comparison to the main analysis, but we cannot reject the null hypothesis that the estimates are the same as the main specification.

Second, I estimate p-values using randomization inference instead of a traditional sampling-based approach. Randomization inference performs better in settings with concentration of leverage, the degree to which individual observations of right-hand side variables take extreme values and are influential, in a few observations (Young, 2016). Due to the large number of observations and the time it takes to run these specification I perform only 1000 iterations for each specifications. All of the p-values on the coefficients of interest are zero, i.e., the estimated effects were larger than all of the 1000 randomized treatment effects. Taking these results together, I conclude that it is unlikely that the estimates are driven solely by outliers.

Initial differences between treatment and control Though there are similar trends in treated and control markets prior to integration, there are substantial differences between markets. Specifically, treated markets tend to have more restaurants, have more restaurants on YTP, and have a larger share of restaurants on YTP.

To address the concern that these initial differences are driving the results, I perform two robustness checks: Firstly, I use inverse propensity score weighting to correct for the bias (Hirano et al., 2003). I estimate the propensity score using a third-order polynomial logit model with the total number of businesses, total number of businesses on YTP pre-integration, and the share of businesses on YTP pre-integration as predictors. Though there is generally common support on the full interval, I trim propensity above and below the 90th

and 10th percentile to correct for differences in mass. The results are presented in Table A6 columns 1 through 4. Though the point estimates are slightly lower, the estimated effect of market expansion on high-quality firms remains positive and statistically significant. The differences between the effects of high- compared to low- quality firms are both economically and statistically significant. The total effects on low-quality firms are both smaller in magnitude and noisier than the main specification

Secondly, since treatment and control looks very different on observables, I conduct an additional analysis, which takes advantage of the fact that treatment intensity is continuous. First, I run a linear probability model of treatment intensity (change in share of businesses on YTP) on a third-order polynomial with the total number of businesses, total number of businesses on YTP pre-integration, and the share of businesses on YTP pre-integration as predictors. Figure A3a presents the distribution of *predicted* treatment intensity by treatment status, where treatment is defined by the median (actual) treatment intensity. City characteristics clearly predict treatment intensity, though there is substantial overlap between the two groups. Secondly, I divide markets into 20 bins based on predicted treatment intensity, with an equal number of markets within each bin. I then assign markets into treatment and control based on their relative treatment intensity in their respective bin. Intuitively, bins with higher predicted intensity will tend to have higher thresholds to be included as treatment. Accordingly, some markets with very high predicted intensity might be coded as control even though their true intensity is relatively high, and vice versa for low intensity markets. Figure A3b presents the distribution of predicted treatment intensity by treatment status, where treatment is defined by the median *within bin* treatment intensity. It is clear the the distribution of treated and control markets is much more similar than in the original sample.

The estimation results, using the new treatment definition, are presented in Table A6 columns 5-8. The point estimates on all coefficients are similar in magnitude to the main specification and are statistically significant. Taken together, these results suggest that initial differences in market characteristics are not driving the main effects.

Additional robustness tests and results As discussed in Section 3 the beginning of the partnership between YTP and Grubhub is coded as February 19th, one month before system integration was completed. To test whether the main results are sensitive to this definition, I conduct two additional robustness checks: First, I define the integration date as March the 19th, and second, I completely exclude the period between February to March 19th. The results are presented in Table A7. In general, the main estimates remain statistically significant and are larger in magnitude when using either of these alternative definitions.

Another potential issue is attrition. Approximately 15% of the incumbent business (9,103) leave YTP before the end of period. In the main analysis, they are used as part of the analysis up to their last week before they leave the platform, then they are excluded from analysis. Since attrition is non-random, one potential concern is that using the data this way is biasing the main estimates. To address this concern I show that the main results are robust to two alternative specifications: First, excluding any establishment that ultimately leaves the platform from the analysis all together, and second, coding leaving businesses' weekly number of orders and weekly revenue as zero after existing the platform. The results are presented in Table A8. In general, I find that all of the main estimates remain statistically significant at 1 percent level and have similar magnitude in both of these alternative specifications.

In addition, I test whether new consumers joining the platform after market expansion are different than existing consumers. In particular, I examine for difference in behavior on the platform between new and old consumers in cities that were affected by the partnership. Table A10 presents the results.⁴⁵ In general, I find null difference in weekly number of order, expenditure, and the type and variety of restaurants ordered from. The only significant difference is that experienced users order, one average, for a slightly larger variety of restaurants. While statistically significant, the effect is relatively small, 1.2%.

D.3 Investment in Quality- Alternative Explanations

In section 5.4. I find that entry increased subsequent investments in quality, as measured by Yelp Ratings. This section addresses the alternative explanations to explain this result.

Rating inflation To address the concern that results are driven by rating inflation (Horton and Golden, 2018, Nosko and Tadelis, 2018), columns 1-4 in Table A11 presents a placebo specification in which integration is counterfactually coded at the middle of the pre-treatment period. Columns 1 and 2 find null average effects which here are both statistically and economically insignificant. Column 3-4 decompose the average effects by quality type; there are null effects for high-rated firms, and marginally significant *negative* effects on low-type. Taken together these results suggest that differential trends in rating inflation are not driving the increases in ratings following integration.

User selection The second concern is that selection into specific services is correlated with rating behavior (Kovács and Sharkey, 2014). For instance, if users who use delivery

⁴⁵ Since individual users are the unit of observation, the exact number of observations is omitted to protect proprietary Yelp data. Generally, the results presented in the table represent the behavior of millions of users. Also, the analysis excludes users that made only one order on the platform.

services also tend to rate more leniently, then increases in online ordering might mechanically drive up the ratings of restaurants. To alleviate this concern, I test for differential changes in raters' leniency. The results are presented in Panel B of Table A11. For each review rating, I calculate the average rating given by the user through her activity on Yelp. I do not find any evidence that raters' leniency changes following integration: The average effects and effects by quality-type on leniency are statistically and economically insignificant under all specifications.

E Structural Model

Simulations After estimating the fundamentals of the model, I simulate firm performance under different market conditions. I divide all firms into five bins, based on their rating quantile across the whole sample. The simulation goes as follows:

1. I choose the share of firms, P , on the platform out of the total number of restaurants in the city, from 1% to 100%.
2. I then draw one restaurant, X , from the relevant rating quantile.
3. Multiplying P by the total number of businesses in restaurant X 's city, gives the number of restaurants in the simulated market, N_s . I then draw $N_s - 1$ additional restaurants from the same CSA. These restaurants consist X 's simulated market. Note that the restaurants in X 's simulated market are drawn from *all* the restaurant in that CSA, not necessarily from the same rating quantile.
4. I recalculate the ranking and ranking within category for all restaurants in the simulated market, since these depend on the specific set of competitors in the market.⁴⁶
5. I then calculate firm X 's market share and the total expected utility from the market. When using demographics shifters on the random coefficient, market shares have to be integrated over the population. In practice, I only take into account the random coefficient on income, and since it discrete I simply take the weighted average over the three income levels.
6. I then repeat step 2-5 a thousand times, choosing different X firms and simulating different markets.

⁴⁶ Since I find null effects on firms' pricing behavior, simulation abstain from changes in pricing.

In total, I simulate 500,000 markets (5 rating quantiles, 100 share level for each, 1000 iterations).

Adding Demographic Shifters In Section 6, the main specification uses does not include demographic shifters of the random coefficient. The reason for ignoring the impact of observable demographics on the disutility from price level is that all demographics have a statistically insignificant effect on the dollar rating coefficient. In this section, I show that the main qualitative results are robust to the inclusion of demographics, but generally perform worse in fitting the data. Note that I only include income as a shifter, and average over the other demographic characteristics.

Simulation results are presented in Table A13. The first two rows show that the estimated market size effect is even stronger in the full model, both welfare and market size increase more rapidly when including demographic. Similarly, firms at all rating quantiles and all market sizes perform better compared to the main specification. These effects are also noticeable in graph A4. In particular, the bliss points are about 20 percentage points higher than in the main specification: 68, 35, 27, and 15 for the fifth, fourth, third and second rating quantiles, respectively. Nevertheless, the main qualitative results hold— high-quality firms perform strictly better, market size in increasing at a decreasing rate, and sales evolve non-monotonically with sellers participation.

Nevertheless, this set of simulations does poorly in fitting the data. First, the simulated entry cost distribution does a worse job fitting the empirical data, compared to the main specification (not shown). Second, the simulation implies the market size increase of over 45%, while the reduce form result suggest an increase of only 36%. Third, the simulation predicts a loss of only 1% of sales for the lowest rating firms (compared to 4.8% in the reduced form results), and an increase of over 13% percent for the highest rated firms (compared to over 13% in the reduced form results). Together, this results imply that using the point-estimates of the effect of observable demographics on the disutility from price does not do a good job fitting the data, and that those coefficient are better interpreted as zeros.