Getting On The Map: The Impact of Online Listings on Business Performance

Michael Luca
Getting On The Map: The Impact of Online Listings on Business Performance

Michael Luca
Harvard Business School

Working Paper 23-031

Copyright © 2022 by Michael Luca.
Working papers are in draft form. This working paper is distributed for purposes of comment and discussion only. It may not be reproduced without permission of the copyright holder. Copies of working papers are available from the author.

Funding for this research was provided in part by Harvard Business School.
Getting On The Map: 
The Impact of Online Listings on Business Performance

Abstract

We evaluate the extent to which small businesses maintain an online presence, looking at restaurant listings on a major online review platform. While the majority of restaurants have an online presence, we find that roughly 18 percent in our sample have no presence as of the end of 2017, despite being able to list their business for free. Using temporal variation in a restaurant’s online presence as well as a natural experiment arising from a data acquisition that led over a thousand businesses to be added to the platform, we find that maintaining an online presence is valuable for small businesses. Establishing an online presence leads to a 5% revenue increase. Using data on customer reviews post-add, we explore potential channels that could explain why restaurants continue to remain offline despite potential benefits from digital representation. Consistent with models of voluntary disclosure, we find that businesses that had stayed offline until the acquisition ended up having lower ratings, on average, than those that added themselves or were added by customers.
“A map is not the territory it represents” - Alfred Korzybski (1958)

1 Introduction

Online review platforms have the potential to facilitate the growth of small businesses by reducing search costs and enabling small businesses to more easily develop an online reputation. Yet, not all businesses maintain an online presence. In this paper, we explore data from a large review platform to answer two questions, focusing on the restaurant industry as a case study. First, how many businesses maintain an online presence? Second, what is the effect of being added to the platform?

Creating a new listing on platforms such as Yelp is often free, and takes little time or technical sophistication. Yet, a business might be deterred if they believe there is no benefit of maintaining an online presence. Moreover, businesses that are worried about attracting negative reviews might be especially likely to remain offline. Our empirical strategy enables us to measure the extent to which local businesses are not represented on online platforms and assess the performance implications of digital representation.

We focus on the online representation of restaurants and bars in the state of Texas via listings on the review platform Yelp as a case study. We rely on administrative tax data on the virtual universe of permanent establishments that serve alcohol in Texas at any point from 2007–2017, which we merge with proprietary data on the universe of Yelp listings during the same period and the date on which each listing was created. Our tax data allows us to observe not only the universe of local businesses, but also their performance in terms of monthly revenues in dollars.

Our first contribution is to examine the extent to which establishments in the real world are represented on online platforms. Examining the full sample of establishments that appear in the tax data reveals that 4,810, or 33.6 percent, never listed on Yelp. Of the 3,754 businesses open throughout the entire 11-year panel, the vast majority have listings. However, even out of this group, approximately 11.7 percent do not have a Yelp listing. The extent to which businesses maintain a presence on Yelp has also evolved over time. Three years after Yelp’s entry in Texas nearly 50 percent of establishments do not have an online listing, but this gap reduces to about 18% at the end of 2017. These findings indicate that the set of missing businesses without an online listing has reduced over time but remains sizeable.

This result is relevant for policymakers, who have formed initiatives such as a recent OECD program to encourage small businesses maintain a digital presence (OECD 2022). Our results highlight the unevenness of maintaining an online presence among small businesses.

Next, we consider the effect of being added to Yelp on business outcomes. Because we can observe when an establishment without a listing gets added to Yelp, we can tease apart the causal effect
of a new listing by relying on intra-business variation in online representation. We leverage a recent two-way fixed-effects estimator (Gardner 2021) to identify the effect of being added to Yelp, controlling non-parametrically for business and county-specific month fixed effects. These results suggest that, in aggregate, an establishment’s sales increase over 5% after being added to Yelp.

While significant, these effects may conflate the possibility that business owners are adding new listings along with contemporaneous business changes (e.g., a renovation project or change in management or prices), which might confound the interpretation of our main effect. Our baseline results are robust to performing a number of additional analyses to account for this possibility.

We note that a business can receive a new listing through three different channels: (a) being added by a customer or Yelp employee, (b) added by the owner and (c) added through a data acquisition by Yelp (a “bulk” addition). In particular, we isolate business listings added either by Yelp community members who are in the top 1% most active in terms of adding new places or en masse through data purchase agreements, to address potential endogeneity concerns. When we examine the effects of being added to Yelp for establishments that were added by users who are in the top 1% most active (“super adders”) revenues increase by almost 9% after a Yelp listing is created. When we consider the effects of bulk additions, where in a space of two days in July 2010, Yelp added 1,295 new listings to their database, revenues increase by more than 10%. In addition, we also perform other checks such as dropping establishments that receive a listing in the first three months after opening (to account for “new business” effects) and also check for pre-existing trends in revenues prior to listing. Overall, the robustness analyses help to bolster our baseline results - suggesting a positive effect, on average, of being added to the platform.

Our findings contribute to the literature on designing online platforms (Ellison and Ellison 2005; Chevalier and Mayzlin 2006; Dai et al. 2018; Reimers and Xie 2019). In practice, many platforms are incomplete, and there is often a gap between what is represented online and on the ground. It has been long understood in the field of cartography (Harley 1992) and recently in economics and management (Nagaraj 2016; Nagaraj and Stern 2020) that “the map is not the terrain.” Platform choices and data acquisitions are a common strategy to seed online platforms and to create a more complete map. This research suggests that creating listings for businesses can directly increase their sales.

Our work also connects to the literature on management practices (Bertrand and Schoar 2003; Bloom et al. 2012) documenting heterogeneity in management and marketing practices and situations in which changes to business practices can improve business performance. For example, many firms have uniform prices across geographies, which limit their profits and are suggestive of managerial mistakes at the geographic market level (DellaVigna and Gentzkow 2019). As a second example, research has found that left-digit bias can lead to kinks in demand at round numbers, resulting in popularity of $9 and .99 cent price endings (Anderson and Simester 2003). However, firms might not fully adjust prices for the extent to which demand responds to .99 cent pricing...
Our results suggest that creating an online presence might be valuable even for businesses that had not taken the time to do so. This highlights the potential for small, low-cost interventions to help improve business practices at the margin. Our results also suggest that being added to the platform was beneficial for businesses that had not added themselves, even though they could have. This highlights the potential for programs aimed at encouraging digital engagement among small businesses, such as the OECD Digital for SMEs Global Initiative, can be valuable and straightforward interventions.

Third, we contribute to the literature on disclosure decisions, which has explored the situations in which firms do and do not disclose information and the effects of disclosure on business performance (Bederson et al. 2018; Jin et al. 2021; Dai and Luca 2020). This literature has pointed to a strategic component of non-disclosure, which suggests that businesses with less favorable information would be less likely to create an online listing (Gao et al. 2015). Our results are consistent with this, as businesses that are added to Yelp through data acquisitions - as opposed to added by the business or users – end up having lower ratings than other businesses. Nonetheless, we find that even these establishments benefit from having a Yelp listing, on average.

In section 2, we describe the data and empirical context. In section 3, we discuss the identification strategy and results. Section 4 concludes.

2 Empirical Context and Data

2.1 Yelp Local Business Listings and Bulk Additions

Yelp is a major local business platform with over 96 million unique monthly visitors as of October 2019 (Yelp 2020). Yelp helps customers discover, rate and review different types of local businesses such as bars, restaurants, beauty salons etc. To receive ratings and reviews, a business must first have a listing on Yelp. A listing is a unique page dedicated to a given business on the Yelp platform that provides basic information (such as its address, name, contact information etc). Once a listing has been created, a business can then receive reviews and ratings and show up in search results on the Yelp platform. As of October 2019, Yelp had 4.9 million active listings and 205 million user reviews.

Although Yelp is a national platform, because the focus in the present study is on bars and restaurants in the state of Texas, it useful to understand Yelp’s history in that state. Yelp was founded in 2004 with a focus on San Francisco and the company began expanding in the state of Texas in April 2007 when it appointed local staff to manage the community in Austin and Houston (Payne 2020). Following this launch, Yelp users could not only add reviews on pre-existing listings in

---

1Users could add listings and reviews for businesses across the country prior to this date, but usage seems to have been low.
Texas, but they also were able to create a new listing for a business if one did not exist already. Business owners who wished to be listed on the site and had not yet been added by patrons could create a new listing for free by registering as a Yelp user. Accordingly, many listings for restaurants and bars were added to Yelp organically in the early years of operation, i.e. between 2007-2010. If a Texas restaurant or bar did not have a listing on Yelp by mid-2010, it was likely because users or business owners had not created a free listing for it. Although Yelp charges businesses for advertising, because creating and managing listings has always been free for business owners, cost is unlikely to be the reason for not listing.

In July 2010, the company decided to supplement its business listings in the state of Texas from third-party sources. In particular, the company turned to external data providers including local business data aggregators which provide data on business listings in the U.S. Such data are largely derived from digitizing local business directories (such as Yellow Pages). Yelp worked with one such data provider as a source of business listing data for listings it did not already capture in its database. Over the course of two days in the summer of 2010 (22nd and 23rd July), Yelp created 1,295 new listings for restaurants and bars in the state of Texas using information from this provider. Part of our analysis and research design focuses on this sample of “bulk” additions.

2.2 Data

To understand the extent to which local business establishments are represented on Yelp and to estimate the business implications of increasing representation, we need three different types of data: (a) a full census of establishments in a given region over time, (b) a measure of performance at the establishment level, and (c) data on all Yelp listings created over time, matched to the relevant establishment in the census. We rely on administrative data from Texas to obtain a census of establishments linked to performance (points a and b), and proprietary data from Yelp for listings activity (point c) as described below.

A. Texas Data on Tax Receipts:

We obtain a full census of business establishments from the Texas Comptroller of Public Accounts’ office. In Texas, any establishment that serves “distilled spirits, beer, ale and wine” is subject to “mixed beverage” taxes and their monthly tax receipts are a matter of public record. We use these data to construct an unbalanced panel of monthly tax receipts for every establishment that served alcoholic beverages in Texas between 2007 and 2017. These data allow us to not only measure performance in terms of assessed revenue, but they also implicitly provide a census of all active establishments that are subject to taxes. These data provide establishment-level revenue data

---

2 The name of the firm Yelp worked with has not been publicly disclosed, but firms providing similar data include InfoUSA and Localeze

3 These data can be obtained at https://data.texas.gov/Government-and-Taxes/Mixed-Beverage-Gross-Receipts/naix-2893
including the name of the establishment and its address, which we leverage to assess representation on Yelp. While the tax data do not provide a unique identifier for each establishment, in Appendix B we document our procedure to assign every establishment with a unique identifier. In all, we have 14,381 unique establishments in our sample.

One limitation of the data is that information is provided on revenue linked to alcohol (but not food) expenditures. Appendix D provides evidence that alcohol revenue constitutes a valid proxy for overall revenue, and shows clicks on Yelp’s website to closely track alcohol revenues. We are thus reassured that alcohol revenues are a reliable indicator of business popularity. Past work comparing non-alcohol revenue to alcohol-related revenues has shown that the two track each other quite closely (Goldfarb and Xiao 2019; Basuroy et al. 2020). Our use of mixed-beverage receipts also means that we do not capture revenue for any establishments that do not have a license to serve alcohol, even though they might be affected by whether or not they are represented on Yelp. We are unable to focus on such establishments, but believe that our results from places that do serve alcohol could generalize to all types of local businesses including those that are not authorized to serve alcohol.

B. Yelp Listings Data and Matching:

After acquiring the census of alcohol-serving establishments in Texas, we turn to assessing their representation on Yelp. We exploit proprietary access to Yelp’s data and algorithms provided through a research partnership with the firm to leverage a matching algorithm developed for internal use that matches an establishment’s name and geographic coordinates (latitude and longitude) with the corresponding Yelp listing (if one exists). We employ this algorithm to match Texas data on tax receipts with Yelp’s listings database.

For each match identified, we obtain a unique id for a listing, details about the Yelp page including average rating and number of reviews, and, crucially, a field that provides the date on which the Yelp page was created. By comparing the date when an establishment first started operations (or April 2007, whichever is later) and the date when it was added to Yelp’s database, we can estimate the lag in online representation for a given establishment, or how long it took for the business to be added to Yelp. After cleaning our data by removing establishments that appear on Yelp prior to appearing in the tax data and dropping establishments that report months of 0 revenues, we are left with Yelp listings for 9,571 of the 14,381 establishments in our sample.4

3 Empirical Results

Armed with data matching Yelp listings to tax records, we now turn to our analysis that descriptively explores online representation on Yelp and its implications for business performance.

4The accuracy of this estimate relies on our matching procedure, which we validate through manual checks as described in Appendix B
3.1 Evaluating Digital Representation

Panel A of Figure 1 examines how the coverage of establishments on Yelp changes over time beginning in January 2007. The series in blue represents the number of establishments filing taxes in each month while the series in red represents the number of active establishments with a Yelp listing. Given Yelp’s limited popularity and reliance on user-generated information to create new listings, Yelp’s coverage in the early months of operation was quite low – hovering around just 20 restaurant listings at the start of 2008. This gap diminishes secularly over time, and by the end of our sample, a large portion of the gap is reduced, partially owing to Yelp’s increasingly dominance in the market for online reviews. However, even as of Jan 2018, almost 20% of tax-paying establishments in our data do not have a Yelp listing. In an average month, between 2007 and 2018, roughly 30% of establishments do not have a listing on Yelp. Also note how the dramatic improvement in Yelp’s coverage due to the addition of the “bulk” listings that were added to Yelp after a data purchase is clearly visible in mid-2010. The proportion of establishments not on Yelp falls from 50% to 30%, representing the most drastic increase in the site’s coverage in the short space of two days. As one might anticipate, Yelp’s coverage in urban areas is relatively more complete at the outset; 59% of establishments in Austin have Yelp listings in 2008 (compared with a statewide average of 21%). However, at the end of our time period, Yelp coverage in Austin is identical to the state-wide average (82%).

Table 1 presents summary statistics delving deeper into this sample for 5 different cuts of the data: (1) establishments never listed on Yelp, (2) all those with Yelp listings, (3) listed establishments for which we have 3 months of data prior to being added Yelp and that were more than 3 months old at the time of their add, (4) all establishments whose listings were created by the top 1% of Yelp users, whom we call “superadders”, and (5) all establishments in the sample that were added in “bulk” on the two days in July 2010. As is clear from comparing the first 2 columns, places added to Yelp are more urban, are in business longer, and have over 1.5 times the amount of monthly revenue as compared to places that are never listed on Yelp. This suggests that online representation is more likely to be a challenge for more rural businesses of a smaller size, which potentially might be unaware of the growing importance of digital platforms in shaping consumer demand. Figure ?? provides a visual overview of this variation. Note the concentration of establishments in a few key urban areas in Texas (with Houston, Dallas, Austin and San Antonio metro areas being the main centers) and the wide variation in terms of whether or not places are represented on Yelp.

Panel B of Figure 1 focuses on a 5% sample of establishments in our data that were added to Yelp and provides another way to visualize our data. Here we show only those establishments which have a Yelp listing. Each establishment is represented by a single line, which is grey from the business’ inception until it is added to Yelp and turns green afterwards for the remainder of its lifetime. The length of each line indicates the number of months for which a given business was active. Establishments are arranged based on the first month in which they began to report
Getting On The Map: The Impact of Online Listings on Business Performance

taxes. Each dot indicates the date at which the indicated business was added to Yelp. Yellow dots represent businesses added in a bulk data purchase or by superadders, while blue dots represent all other businesses. Again, in this sample, the magnitude of the bulk import of listings to Yelp is evident in the number of businesses that are added on the same date in mid-2010. Aside from this single date, there is no apparent trend regarding when businesses are added to Yelp. Even businesses that have been open since the beginning of our panel are added to Yelp as businesses are added to Yelp at varying points in their lifetimes.

This figure provides the inspiration for our two-way fixed-effects estimation strategy. There is wide variation in when establishments were added to Yelp, so we can evaluate the effect of digital representation and its effect on revenues by comparing pre-Yelp outcomes with post-Yelp, while controlling for establishment fixed effects and time-trends.

3.2 The Performance Implications of Digital Representation

3.2.1 Research Design:

An empirical challenge in our setting is that our data features establishments that vary by treatment time and our panel is unbalanced in calendar time, whereas most applications of difference-in-differences involve settings with strongly balanced panels in calendar time, although they may not be balanced in time relative to treatment. Further, because Yelp’s importance is growing throughout the study period, its impact is likely to vary significantly, leading to treatment effect heterogeneity that has been shown to be problematic in the context of two-way fixed-effects estimators (Goodman-Bacon 2021).

A number of estimators have been developed to address this fundamental problem (De Chaisemartin and d’Haultfoeuille 2020; Callaway and Sant’Anna 2021; Sun and Abraham 2021; Borusyak et al. 2021; Gardner 2021). The approaches differ slightly in the formulation of the parallel trends assumptions, types of treatment effects possible to identify, and data or treatment conditions necessary for identification and estimation. Assessing the relative computational complexity of different estimators in our setting, which has a large number of treated and control units, we consider the two-stage imputation approach from Gardner (2021) to be the most suitable estimator for our setting. The key identifying assumption of this approach is that the counterfactual $Y_i(d,0)$ can be characterized by $Y_i(d,0) = \alpha_i + \gamma_t$, an additive function of a unit fixed effect and a time trend. Since we are mainly concerned about Yelp’s growing popularity which is likely to be additively separable from any establishment’s fixed effect, this assumption accords well with our setting. For the sake of comparison, we also estimate traditional two-way fixed effects regressions (presented in Appendix Table ??) both with and without the sample restrictions necessary for the two-stage imputation estimator. More detail on our empirical approach can be found in Appendix E.
3.2.2 Full-Sample Analysis:

Table 2 reports the results of our analysis on the full sample of establishments in our data. Column 1 of this table captures the aggregate effect on revenues of being added to Yelp. We want to ensure that we are not measuring a new business effect if businesses are added to Yelp as soon as they open and then see a natural increase in revenues due to age effects. To address this concern, we run the same analysis but limit our sample to establishments open for at least three months before being added to Yelp. We find that the effect of being added to Yelp on revenues is significant, ranging from 5%-10%, and the size of the effect is not attenuated when we exclude new establishments. Taken at face value, these results represent significant effects of mere digital inclusion on online platforms – comparable in magnitude to the effect of gaining one additional star (Luca 2016).

3.2.3 Super-Adders:

While these analyses are highly informative, it is possible that the results might be biased upwards if business owners are adding their establishments when they are on a positive trend or undergoing other business upgrades. We deal with this concern by examining two other samples.

First, we focus on businesses added by “super-adders,” a group of 56 individuals in the top 1% of all users in terms of the distribution of number of establishments added. In practice, each super-adder has added at least 10 establishments although a few have added many more. We focus on these users because they are unlikely to be connected to restaurant management and the timing of their adding activity is likely due to factors unrelated to restaurant revenue, such as a change in business ownership. We find that this group of Yelp users has added 747 businesses in our sample to the site. The businesses added by these “super-adders” are, on average younger than the mean establishment when they are added to Yelp. However, these are not brand new establishments, and these active users are not just those who are the fastest to add listings for new businesses.

Table 2 examines the revenue effect of being added to Yelp for this set of establishments (third column from the left). We find that the causal effect of being added to Yelp for this sample is about 8.9%, which is a larger effect than we find for our full sample. This result increases our confidence in the validity of our estimates as capturing the effect of having a presence on Yelp on business outcomes.

3.2.4 Bulk Adds:

Next, we focus on a set of listings added exogenously by Yelp via bulk data acquisition. As well as data adds by users and business owners, listings were acquired inorganically by Yelp through purchases from data providers. In particular, we identify the exogenous improvement in Yelp’s coverage through a partnership with an external data provider that automatically created new
listings for a set of 1,295 businesses over the course of two days in July 2010. This incident offers a unique opportunity to cleanly identify the effect of such “bulk” additions.

The last column of Table 2 reports estimates from the sample restricted to establishments acquired through the “bulk” add. We find that the overall effect of a bulk addition is larger than our main effects, with a 10% positive and significant increase in revenues. Both the super-adder and bulk adds exercise both offer significant support for our baseline conclusion.

3.2.5 Pre-Trends:

Finally, we also test for the possibility that some businesses were already on a positive trend (perhaps due to a change in ownership) that was then followed by the creation of a Yelp listing. To do so, we examine in the difference in trends between establishments yet to be added to Yelp and those already added. Figure 2 plots estimates for $\beta_t$ for these leads and lags for each of the samples in Table 2. Two points are clear from these figures. First, there is no significant trend in revenue before establishments are added to Yelp. The pre-trends are relatively flat and tightly estimated. Second, the rise in revenues is immediate and persistent for all samples, suggesting that adding a new listing unlocks a new source of customers and revenue to the restaurant that stays persistent over time.

3.3 Correlates of Missing Listings:

Thus far we have shown the absence of a listing can be potentially quite costly for establishments in terms of lost revenue. While we do not attempt to offer specific mechanisms in terms of why businesses are unable to take advantage of online platforms, in Figure 3 we compare different samples of establishments based on their Yelp listing status in terms of whether or not they are in urban locations (Panel A) and customer ratings (Panel B). Comparing the five samples we examined in Table 1, we find that establishments that are never added to Yelp are more likely to be located in rural areas than Yelp listings in aggregate and those added by other, non-bulk channels. Bulk adds seem to be similar to those never added, in that they are more rural. Establishments added by super-adders on the other hand, represent a more urban sample of establishments, perhaps because of the density of businesses in urban versus rural locations. Considering that our treatment effect for bulk adds was in the range of 10%, it seems that the more rural set of establishments that are not on Yelp could benefit significantly from an online listing given that they are quite comparable to the set of places added through bulk adds. In other words, it appears that the reason that more rural places are not listing on Yelp is likely due to information frictions or management practices, rather than that explanation that Yelp listings do not drive improved performance in more rural areas due to differences in customer usage of Yelp, for example.

Panel B performs a similar comparison of customer reviews on Yelp, excluding the sample of
establishments that are never added to Yelp since we do not observe ratings for them. As is clear from this figure, bulk adds receive lower ratings than establishments in the other three samples. This is also visible in Table 1, in which we see that the mean rating for bulk-added businesses is significantly lower than for other samples. However, even given that they may be of lower quality, as we find in our analysis, bulk adds benefit significantly from having a presence on Yelp.

Collectively, these results suggest that even though the establishments that intentionally remain off of online platforms are more likely to receive negative reviews once they are listed online, not having an online presence still may lead to lost revenue for these businesses.

4 Discussion

In this work, we investigate the role that online listing platforms play in driving business outcomes. First, we find that a sizeable proportion of active businesses—almost 20%—are absent from Yelp at the end of 2017, and over 30% of the businesses in our sample are never listed on Yelp. Next, we evaluate the effect of being added to Yelp on revenues. We leverage quasi-random variation in how businesses were added to the site and identify that establishments see significant and positive revenue increases after being listed on Yelp. Our results also show that firms who will have a worse reputation are less likely to create an online listing to begin with. However, we find that even these establishments benefit from the exposure of having a listing on Yelp.

This paper contributes to the literature on the role of digital platforms and broadens the scope of this work by looking beyond implications just for establishments that already have an online presence. Multisided platforms play an increasingly large and important role in organizing economic activity. A large body of work has highlighted the ways in which digital platforms are reducing search costs and reshaping markets (Bakos 1997; Smith et al. 1999; Ellison and Ellison 2005). Further, a robust literature has looked at how reputation systems within platforms can affect business outcomes (Zhu and Zhang 2010; Chevalier and Mayzlin 2006; Luca 2016). Growing acknowledgment of the importance of platform design points to the types of information that might help to create a vibrant reputation system (Teubner et al. 2017; Aguiar and Waldfogel 2018; Kim and Luca 2019). While prior research has focused mainly on the intensive margin (through enhanced ratings, reviews or rankings), our work highlights the importance of thinking about the extensive margin as well, by shaping a consumers’ consideration sets. Much as the “the map is not the terrain” (Korzybski 1958), platforms are unlikely to always be completely comprehensive, suggesting that this phenomenon is likely to generalize to a variety of online platforms beyond Yelp.

Our results also shed light on business decisions about whether to engage on platforms like Yelp. In principle, market forces could lead all businesses to post listings, as long as the information is verifiable and the costs of disclosure are sufficiently small (Viscusi 1978; Grossman and Hart 1980; Grossman 1981; Milgrom 1981). Under these conditions, all but the very worst firms would choose
to post - following the information unraveling hypothesis. A more recent literature, however, has shown that in practice, disclosure is often incomplete (Brown et al. 2012; Luca 2016; Jin et al. 2021). We find that even those establishments who do not receive positive ratings benefit from the exposure of having a listing on Yelp. Consistent with the emerging literature on behavioral firms and management practices, our findings show that there are businesses without an online presence that might benefit from creating one.

Our findings highlight several rich areas for future study. Consistent with Simester (2017), this work suggests that field experiments can be a rich tool by which to identify ways that firms can optimize marketing. We find that the revenue benefits of being added to Yelp are not depressed by low ratings, but we do not directly examine the mechanisms by which a presence on Yelp increases revenues. Understanding how this occurs and identifying interventions that might help businesses leverage online platforms to improve their outcomes constitutes is a fruitful topic for future exploration. Additionally, one limitation of this work is we do not observe the macro-level effects of increasing the coverage of a site like Yelp. That is, we do not identify the spillover effects on establishments not listed on Yelp and the net effect on consumer demand. In future work, we plan to further examine the implications of Yelp’s growing footprint on business outcomes.
Getting On The Map: The Impact of Online Listings on Business Performance

References


Getting On The Map: The Impact of Online Listings on Business Performance


Figure 1: Assessing the Coverage of Yelp Listings over Time


Note: Panel A plots the unique number of establishments reporting revenues in the administrative tax data (blue line) compared to the number of Yelp listings (red line) for the period Jan 2007 to Nov. 2017. In Panel B, we explore how establishments enter and exit our baseline sample and the timing of when they are added to Yelp. Each horizontal line represents one establishment and is drawn from the month in which they began operation (or Jan 2007 whichever is later) and the month in which they stopped operating (or Dec 2017 if they did not). The color of this line switches to green once an establishment is listed on Yelp and is grey otherwise. The month in which the establishment is added to Yelp is indicated with a blue dot. Establishments are arranged from bottom to top in the order of the month in which they first opened.
Figure 2: The Impact of being added to Yelp on Log Revenues

(a) Full Sample  
(b) >3 months Age at Yelp Add

(c) Added by Super-Adder  
(d) Bulk Adds

Note: This figure shows the effect of being added to Yelp on total revenues for different samples of the data.
Figure 3: Evaluating Yelp Listings

Panel A: Proportion of Establishments in Urban Areas

Panel B: Distribution of Yelp Ratings
### Table 1: Comparing Establishments by Yelp Status

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Added to Yelp</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Located in urban area</td>
<td>0.773</td>
<td>0.864</td>
<td>0.856</td>
<td>0.967</td>
<td>0.747</td>
</tr>
<tr>
<td></td>
<td>(0.419)</td>
<td>(0.342)</td>
<td>(0.351)</td>
<td>(0.180)</td>
<td>(0.435)</td>
</tr>
<tr>
<td>Months reporting taxes</td>
<td>38.30</td>
<td>82.63</td>
<td>101.7</td>
<td>94.86</td>
<td>98.41</td>
</tr>
<tr>
<td></td>
<td>(41.34)</td>
<td>(46.52)</td>
<td>(38.83)</td>
<td>(44.51)</td>
<td>(38.36)</td>
</tr>
<tr>
<td>Year opened</td>
<td>2007.6</td>
<td>2005.9</td>
<td>2002.8</td>
<td>2002.4</td>
<td>2004.1</td>
</tr>
<tr>
<td></td>
<td>(6.220)</td>
<td>(6.678)</td>
<td>(5.658)</td>
<td>(5.142)</td>
<td>(5.300)</td>
</tr>
<tr>
<td>Year closed</td>
<td>2012.5</td>
<td>2015.5</td>
<td>2015.4</td>
<td>2014.5</td>
<td>2015.2</td>
</tr>
<tr>
<td></td>
<td>(3.850)</td>
<td>(2.592)</td>
<td>(2.701)</td>
<td>(3.397)</td>
<td>(2.548)</td>
</tr>
<tr>
<td>Average monthly tax receipts</td>
<td>20853.5</td>
<td>30911.1</td>
<td>29843.8</td>
<td>37016.2</td>
<td>21470.7</td>
</tr>
<tr>
<td></td>
<td>(39199.8)</td>
<td>(39443.7)</td>
<td>(36149.1)</td>
<td>(41689.4)</td>
<td>(27274.5)</td>
</tr>
<tr>
<td>Added by Yelp user</td>
<td>0.616</td>
<td>0.668</td>
<td>1</td>
<td>0.00541</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.486)</td>
<td>(0.471)</td>
<td>(0)</td>
<td>(0.0734)</td>
<td></td>
</tr>
<tr>
<td>Yelp rating</td>
<td>3.033</td>
<td>2.902</td>
<td>3.173</td>
<td>2.088</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.174)</td>
<td>(1.247)</td>
<td>(0.819)</td>
<td>(1.673)</td>
<td></td>
</tr>
<tr>
<td>Proportion post-Yelp (months)</td>
<td>0.777</td>
<td>0.693</td>
<td>0.829</td>
<td>0.634</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.237)</td>
<td>(0.236)</td>
<td>(0.198)</td>
<td>(0.197)</td>
<td></td>
</tr>
<tr>
<td>Year added to Yelp</td>
<td>2010.5</td>
<td>2009.5</td>
<td>2007.9</td>
<td>2010</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.879)</td>
<td>(2.307)</td>
<td>(1.229)</td>
<td>(0)</td>
<td></td>
</tr>
<tr>
<td>Listings delay (months)</td>
<td>17.96</td>
<td>25.70</td>
<td>9.855</td>
<td>31.63</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(21.22)</td>
<td>(21.69)</td>
<td>(9.681)</td>
<td>(12.64)</td>
<td></td>
</tr>
</tbody>
</table>

*:*p<0.10; **:p<0.05; ***:p<0.01. Standard errors clustered at block-level shown in parentheses. Note: Summary statistics for the cross-section of establishments that (a) are never listed on Yelp, (b) are listed on Yelp, (c) are listed with a delay of over 3 months, (d) are added by the top 1% of Yelp users, and (e) are in the sample of bulk additions in July 2010.
Table 2: Does being added to Yelp affect sales?

<table>
<thead>
<tr>
<th></th>
<th>All listings: Log Revenues</th>
<th>3+ month delay: Log Revenues</th>
<th>Superadder added: Log Revenues</th>
<th>Bulk added: Log Revenues</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post x Yelp</td>
<td>0.0520*** (0.0108)</td>
<td>0.0520*** (0.0108)</td>
<td>0.0894*** (0.0225)</td>
<td>0.1044*** (0.0182)</td>
</tr>
<tr>
<td>Establishment FE</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>County X Month</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>N</td>
<td>269,512</td>
<td>268,984</td>
<td>81,586</td>
<td>105,483</td>
</tr>
</tbody>
</table>

*: p<0.10; **: p<0.05; ***: p<0.01. Standard errors clustered at block-level shown in parentheses.

Note: Estimates of the effect of a new Yelp listing on establishment revenues.