

The Financial Consequences of Online Review Aggregators: Evidence from Yelp Ratings and SBA Loans

Ruidi Huang^a

^a Southern Methodist University, Dallas, Texas 75275

Contact: ruidih@smu.edu,  <https://orcid.org/0000-0002-6061-5889> (RH)

Received: October 2, 2020

Revised: January 22, 2022; January 2, 2023;
July 9, 2023

Accepted: August 22, 2023

Published Online in Articles in Advance:
March 25, 2024

<https://doi.org/10.1287/mnsc.2020.03003>

Copyright: © 2024 INFORMS

Abstract. This paper demonstrates the financial and real consequences of online review aggregators. Exploiting a regression discontinuity design that overcomes the endogenous relationship between Yelp reviews and Small Business Administration loan outcomes, I show that higher coarse Yelp ratings lead to improved loan terms and performance. Specifically, a one-half-star increase in Yelp ratings corresponds to a 25-basis-point decrease in loan spread and 6% lower collateral requirements. The effects are more pronounced when banks have less borrower information. Higher Yelp ratings also contribute to increased consumer demand and the likelihood of future business openings. These findings indicate that online review aggregators influence both consumer choices and banks' financing decisions.

History: Accepted by Gustavo Manso, finance.

Supplemental Material: The online appendix and data files are available at <https://doi.org/10.1287/mnsc.2020.03003>.

Keywords: online review • credit rating • small business loan • bank lending

1. Introduction

Small businesses are the backbone of the U.S. economy, representing more than 95% of all companies and employing more than 50% of the private sector workforce. Because of limited access to public equity and debt markets, small businesses mainly rely on bank credit, especially term loans, to finance their growth. One major obstacle in small business lending is that banks have scarce reliable information about borrowers because small businesses are not subject to strict regulatory filing requirements. Thus, the tools lenders usually use might provide little to no benefit.

The proliferation of online review aggregators offers lenders alternative and less traditional views about potential borrowers when making lending decisions. These platforms aggregate ratings generated by individuals to an easily interpretable discrete number, allowing such information to quickly reach consumers as well as lenders. In this paper, I use Yelp, a leading platform, and exploit a regression discontinuity design to show that coarse ratings presented by online review aggregators have a real impact on small businesses. I find that a one-half-star increase in Yelp ratings leads to better loan terms and performance. In addition, higher Yelp ratings contribute to increases in consumer demand and the likelihood of subsequent business openings. These results suggest that online review aggregators influence both customers' consumption choices and banks' financing decisions.

Ex ante, it is not clear how information from online review aggregators impacts firm financing. On the one hand, lenders might regard the information produced by online review aggregators, especially the coarse ratings, as idiosyncratic, subjective, often biased by too few reviews, and difficult to benchmark against (i.e., highly variable across locations and industry). Thus, they may disregard reviews or apply some rule of thumb based on a threshold star level such that slight changes in reviews (e.g., 4 to 4.5) do not affect loan outcomes.

On the other hand, lenders prefer borrowers with a good credit history and solid fundamentals (Diamond 1991). Previous research show that online reviews have a positive impact on restaurant bookings (Anderson and Magruder 2012) and revenue (Luca 2016). Because of social learning (e.g., Chevalier and Mayzlin 2006, Cai et al. 2009, Moretti 2011), businesses with higher ratings are more likely to attract more consumers, which translates to higher revenue and cash flows. Borrowers with higher ratings, as indicated by the information provided by online review aggregators, therefore, have more financial slack to fulfill their loan commitments (Allen 1984). Lenders exploiting good prospects of highly rated businesses may make lending decisions that coincide with coarse online ratings.¹

However, one major challenge to studying the effect of coarse online ratings on loan outcomes is the potentially endogenous relationship between the two. Many

confounding factors that could also influence loan outcomes, such as good management, the quality of business plans, neighborhood characteristics, and taste-based discrimination, might coexist with online ratings. To address this identification issue, I exploit the coarse nature of ratings presented by a widely used online review aggregator, Yelp. Yelp aggregates all reviews and presents an overall rating of a business, shown as a number of stars. Because of the institutional design, Yelp ratings range from one to five stars in one-half-star increments, which means each overall rating is rounded to the nearest one-half star using predetermined rounding thresholds. For example, a business with an overall rating of 3.75 is rounded up to 4 stars, but a business with an overall rating of 3.74 is rounded down to 3.5 stars. I implement a regression discontinuity design (RDD) around the rounding thresholds. I compare the loan outcomes of businesses rounded up to those rounded down. The key identifying assumption is that the influence of potential confounding factors on loan outcomes is not expected to change discontinuously when a Yelp rating passes an arbitrary rounding threshold.

For my empirical analysis, I leverage a sample of loans made through the U.S. Small Business Administration (SBA) matched to Yelp ratings. Using the RDD tool, I first examine loan terms, specifically loan spreads and collateral. I document that a one-half-star increase in Yelp ratings leads to significantly lower financing costs. The average business just above the rounding thresholds enjoys a loan spread 25 basis points (bps) lower with two percentage points fewer collateral requirements than the average business just below the threshold. The effect is also economically substantial, corresponding to 9% cheaper loan pricing and 6% lower required collateral. I also examine loan performance in terms of default and the write-off amount upon default. I find that a one-half-star increase in Yelp ratings reduces the default rate by 45% compared with the mean, which is both statistically and economically significant. Overall, the results suggest that coarse Yelp ratings have a positive impact on loan outcomes. I note that the sample is selected on firms that make it into SBA approval data; consequently, the results do not speak to firms that do not apply for an SBA loan or to those denied an SBA loan.

Next, I explore whether lenders are more likely to rely on coarse Yelp ratings when they are less informed or have lower incentives to be informed about the borrower. First, I exploit a source of heterogeneity in the cost of retrieving information. My findings show that coarse Yelp ratings have a stronger impact on businesses located farther away from the lending banks. Banks need to devote extra monitoring costs when they are farther away. The ready availability of Yelp ratings helps alleviate this problem. Second, other sources of information could potentially dilute the effects of Yelp

ratings. I focus on small businesses that take out subsequent loans from the SBA, meaning the lenders have already had prior interactions with the borrowers. I find evidence that lenders rely less on Yelp ratings in subsequent lending decisions. However, the overall effects of Yelp ratings remain strong, indicating that Yelp is a good predictor of loan outcomes, even in repeated lending transactions. Overall, the cross-sectional analysis supports the notion that lenders refer to such ratings in their decision making.

An important dimension of information is the precision of the signal. Under the Bayesian learning theory, the more precise the signal is, the stronger the reaction to it will be. I identify cases in which ratings convey a more or less accurate signal by examining the number of reviews. I find that the results are stronger for businesses with a large number of Yelp reviews. In contrast, for businesses with fewer reviews, the point estimates are smaller and the standard errors are bigger, suggesting a noisier signal.

Next, I show the real effects of Yelp ratings. I focus on two outcomes, one from the consumer's perspective and one from the business'. I first study the changes in consumer demand. I proxy consumer demand using the number of reviews received by businesses and find that higher Yelp ratings lead to a higher number of reviews. I next measure subsequent location openings for companies already on Yelp to examine small businesses' investment decisions. I find that a one-half-star increase in Yelp ratings leads to an 8% increase in the likelihood of opening a new location. These results are in line with Iyer and Manso (2023), who assert that recommendations could have real effects.

My study contributes to several strands of inquiry. First, I add to a large body of work on resolving information asymmetry problems in bank lending. Earlier studies, both empirical and theoretical, focus on information derived from relationship lending, such as repeated transactions (e.g., Diamond 1991, Petersen and Rajan 1994, Berger and Udell 1995). More recent studies show that lending decisions are made with information beyond traditional sources, such as the borrower's physical appearance and claims of trustworthiness (e.g., Duarte et al. 2012, Iyer et al. 2016, Ravina 2019). I present evidence that banks use online reviews to address information asymmetry problems in the lending process. Because of technological advancements, I show that it has become easier for banks to incorporate increasingly nuanced information that has a significant impact on loan outcomes.²

My paper is also related to the literature on the feedback effects of reviews. Iyer and Manso (2023) show that, theoretically, recommenders not only inform consumers about the quality of a product, but also determine the quality of the product. If an influential recommender promotes one particular product, demand for that product is

likely to increase, which can prompt the firm to further improve it. Higher Yelp ratings lead to better bank financing, which translates to better products and services. The customers who were swayed by those ratings receive products and services that are consistent with their expectations. Such feedback effects give ratings generated by online review aggregators a self-fulfilling property, similar to that of credit ratings (Manso 2013).

Two related papers also study the effects of Yelp ratings. Anderson and Magruder (2012) document that Yelp ratings have a positive effect on restaurant bookings in San Francisco, and Luca (2016) argues that Yelp ratings lead to higher revenue for restaurants in Seattle. My paper complements those studies by showing that Yelp ratings significantly affect financing outcomes, that is, small business lending, in which banks have insufficient reliable knowledge about the borrowers' financial health. In addition, my sample encompasses all types of small businesses on Yelp across the United States. Furthermore, Anderson and Magruder (2012) and Luca (2016) illustrate that Yelp ratings resolve information asymmetry problems between businesses and customers. My analysis shows that Yelp ratings also resolve the information asymmetry problem between businesses (i.e., borrowers) and lenders. I also show that higher Yelp ratings are associated with far lower default rates, suggesting that using information from Yelp also benefits lenders.

Finally, this paper also adds to the studies about small business financing and entrepreneurship in general. I introduce a novel data source that covers more than 80 million distinct reviews of more than 1 million entrepreneurial firms. Furthermore, the existing studies mainly document the impact of small business loans on, for example, employment growth (Brown and Earle 2017), business growth (Hackney 2023), and local economic performance (Craig et al. 2005). Few studies examine the small business lending process. My paper takes a step back to understand how online review aggregators influence loan terms and outcomes.³

The remainder of the paper is organized as follows. Section 2 describes the identification strategy and the data. Section 3 focuses on the loan-level analysis. Section 4 examines the real consequences of Yelp ratings. Section 5 concludes.

2. Empirical Design, Institutional Details, and Data

This section begins by discussing the empirical strategy. Next, I describe the Yelp data and the data-collection process, followed by information about the SBA loan data and the institutional background of the SBA loan program. Finally, I present details on the final sample used in the analysis and present summary statistics.

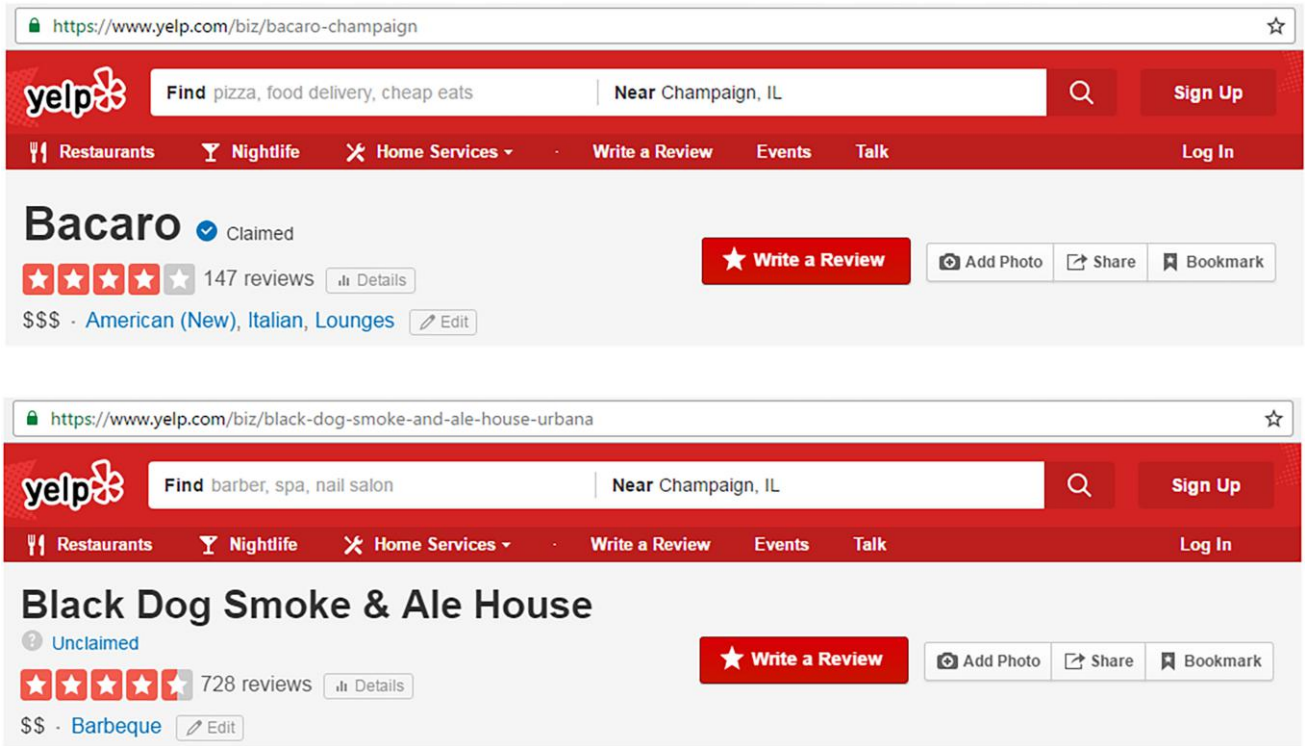
2.1. Empirical Strategy

Identifying the effects of coarse Yelp ratings on loan outcomes in general is challenging because unobservable factors, such as management quality, might be correlated with Yelp ratings. As a result, an ordinary least squares (OLS) regression of loan outcomes on Yelp ratings is subject to potential endogeneity problems.

To overcome this identification challenge, I exploit a design feature of the Yelp platform, following Anderson and Magruder (2012) and Luca (2016). When one looks up businesses on [Yelp.com](https://www.yelp.com), the website prominently displays the overall ratings of businesses on the top left corner under the business names. Figure 1 shows two examples of Yelp ratings. The overall ratings are presented as a number of stars, ranging from one to five. The star ratings increase in one-half-star increments, which yields nine possible rating categories: 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, and 5 stars. In the example shown in Figure 1, the business at the top has a Yelp rating of 4 stars, and the one at the bottom has a Yelp rating of 4.5 stars.

The star designations are average ratings calculated from all the individual reviews left by users. When Yelp aggregates those reviews, the platform uses a predetermined rule to round the average ratings to display them in one of the nine possible overall star ratings. Because of Yelp's institutional design, the rounding thresholds are set exactly at the midpoint of two neighboring star ratings. For example, in determining whether a business belongs to the 4- or 4.5-star group, Yelp uses the midpoint of 4.25. If the business has an average review below 4.25, then the overall rating is rounded down to 4 stars and shown as such on Yelp. Otherwise, the business's overall star rating is rounded up to 4.5 stars and displayed accordingly.

I implement an RDD strategy around the rounding thresholds to study their effects on loan outcomes. Because Yelp provides nine possible star ratings, there are eight rounding thresholds. I recenter the average ratings to their cutoff points and assign the businesses with average ratings just above and just below the rounding thresholds to the treatment and control groups, respectively. Assignment to the treatment is, thus, determined by whether the Yelp ratings are rounded up or down given the fixed rounding thresholds. Businesses in the two groups are similar in many relevant aspects. The average ratings may themselves be associated with loan outcomes, but this association is assumed to be smooth. As a result, I can interpret any discontinuity of the conditional distribution of loan outcomes as a function of Yelp ratings at the rounding thresholds as causal. The key identifying assumption is that, in the absence of a discontinuous jump in overall ratings around the rounding thresholds, no other discontinuous changes in the potential confounding factors directly affect the loan outcomes.⁴

Figure 1. (Color online) Examples of Yelp Ratings

Notes. The figures show examples of Yelp ratings for two sample businesses. Yelp displays ratings in the form of stars with half-star increments. The top business has a Yelp rating of 4 stars and the bottom business has a Yelp rating of 4.5 stars.

Throughout the paper, I adopt the parametric approach suggested by Lee and Lemieux (2010) and Roberts and Whited (2013) by utilizing all of the data, not just the data right around the cutoffs. I draw inferences by combining all of the data on both sides of the threshold and estimating a pooled regression of the outcome on the treatment indicator, the distance of the assignment variable to the cutoff, and the interaction of the two.⁵ To empirically carry out the identification strategy, I use the following regression:

$$\begin{aligned} Outcome_{i,t+} = & \beta_0 + \beta_1 \cdot I_{round\ up,i,t} + \beta_2 \cdot D_{i,t} \\ & + \beta_3 \cdot I_{round\ up,i,t} \times D_{i,t} + \beta_4 \cdot X + \phi_{i,t} \\ & + \tau_t + \eta_i + \theta_i + \epsilon_{i,t}, \end{aligned} \quad (1)$$

where $outcome_{i,t+}$ takes on different measures of loan outcomes, $I_{round\ up,i,t}$ is the treatment variable that equals one if the Yelp rating is rounded up and zero if rounded down, D is the distance between the unrounded average rating and the corresponding cutoff, $I_{round\ up,i,t} \times D_{i,t}$ allows for different slopes and intercepts on either side of the cutoff, X is a vector of business and loan characteristic controls (defined in Section 2.4), $\phi_{i,t}$ is cutoff fixed effects, τ_t represents year fixed effects, η_i is Yelp industry fixed effects, and θ_i is county fixed effects. Roberts and

Whited (2013) point out that using the entire support of the data introduces bias into the estimated treatment effect, so researchers face a trade-off between bias and variance when implementing an RDD. To address this concern, in Online Appendix B, I present robustness tests that use data right around the thresholds. I test three bandwidths around the cutoffs: 0.05, 0.10, and 0.15 points.

A potential concern with using Yelp data is that business owners are also aware of the rounding thresholds. They have an incentive to manipulate their ratings to appear to be one-half star higher in ratings, possibly to attract more customers or to window-dress their ratings before applying for financing. Anecdotally, Yelp has always maintained that it employs a team of engineers to combat fake reviews with complex and advanced computer algorithms. Additionally, federal judges have dismissed lawsuits against allegations of Yelp review manipulation. Both accounts afford me some confidence in ruling out manipulation by business owners. However, I formally test for manipulation by plotting the density of the running variable around the rounding threshold in Online Appendix C. I also carry out a McCrary (2008) density test to rule out the manipulation concern; results are detailed in Online Appendix D.

Another concern is potential selection bias in the sample, which is unlikely a representative sample of firms in the Yelp universe. On the one hand, only 11.88% of the sample firms have 2.5 stars or lower compared with 14.46% of the Yelp universe. The difference suggests that the Yelp universe has 22% more low-star firms relative to the sample. On the other hand, the sample has 5% more firms with three stars or higher compared with the universe, suggesting that low-star firms are comparatively less likely to be in the SBA data. In addition, the SBA sets a ceiling on the interest rate banks can charge. Tabulating the data used in the analysis, I find two to four times more firms with lower than or equal to 2.5 stars that are within one percentage point of the interest rate ceiling than with more than 2.5 stars. The low-star firms in my sample are the ones that survived and are possibly of similar quality with little variation among them. Consequently, the empirical limitation, that is, survivorship bias, prevents me from producing reliable estimates at low star levels. I discuss the direction and likely magnitude of the bias in the results section (i.e., Section 3).

2.2. Yelp Data

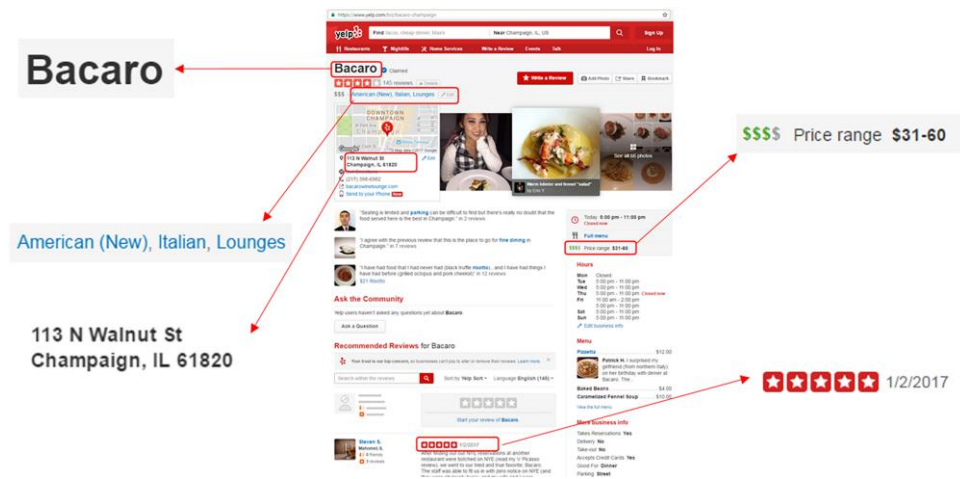
Founded in 2004, Yelp is a crowdsourced website on which people post their reviews of businesses. Each month, Yelp attracts 26 million unique visitors, on average, through the mobile application and 73 million unique visitors, on average, through the web. It contains more than 127 million reviews as of the first quarter of 2017. Anyone can sign up for a Yelp account free of charge to rate businesses and write reviews. The reviews consist of reviewer names, review dates, one-

to five-star ratings, and comments from the reviewers. Users can access Yelp through mobile devices and computers by directly searching the business names with search engines or on Yelp.com. Users can also look up businesses by a specific star rating, a particular location, a Yelp industry, or a price range among many other criteria.

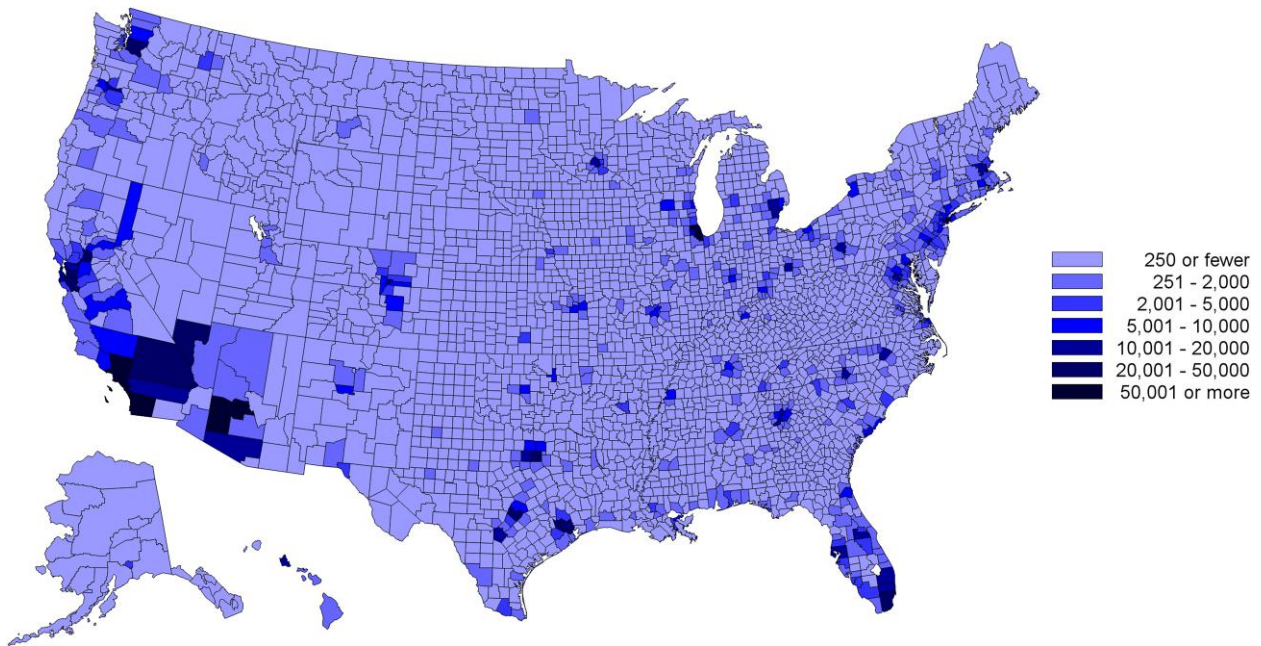
Figure 2 presents a sample business listed on Yelp to graphically illustrate the Yelp data. For each business on Yelp, I observe the business name (i.e., Bacaro, in the example), Yelp business category (i.e., American, Italian, Lounges), business address (i.e., 113 N Walnut St, Champaign, IL 61820), price range (i.e., \$\$\$ and \$31–60), and each reviewer rating along with the review date (i.e., 5 stars, 1/2/2017). The data are available on businesses located in the United States only. The data includes Yelp reviews starting in October 2004 (Yelp inception date) through the end of 2016. I require businesses to have more than five reviews at the time of data collection to be included in the sample for the empirical analysis.⁶

Figure 3 shows the Yelp data coverage at the county level. I assign a census-defined county code to each business based on the address. I count the number of businesses in each county. The counties with more businesses represented on Yelp are presented in darker color, and counties with fewer businesses on Yelp are in lighter color. The map shows that my data provide comprehensive coverage of most parts of the United States. Table 1, panel A, breaks down the Yelp data by year. Over time, Yelp has gained increasing popularity among reviewers. The average number of reviews per business increased from fewer than two in 2004 to

Figure 2. (Color online) Example of Yelp Data



Notes. This figure shows a sample business listed on Yelp. For each business, I observe the business name (i.e., Bacaro), Yelp business category (i.e., American, Italian, Lounges), address (i.e., 113 N Walnut St, Champaign, IL 61820), price range (i.e., \$\$\$ and \$31–60), each reviewer rating and review date (i.e., for example, 5 stars, 1/2/2017).

Figure 3. (Color online) Yelp Data Coverage

Notes. This map plots Yelp data coverage at the county level. Yelp data are available from 2004 to 2016. The number of businesses is aggregated to the county level to produce the figure. Regions range from darker (counties with higher numbers of businesses on Yelp) to lighter (counties with lower numbers of businesses on Yelp).

almost 13 in 2016. Table 1, panel B, divides the Yelp data into 22 Yelp-defined business categories. This panel shows the average Yelp rating, average number of reviews, and category weight. Of the 22 categories, restaurants account for around 20% and shopping accounts for about 10%. The average rating ranges from 3.49 to 4.49. The average number of reviews ranges from 5 to 120.

2.3. SBA Loan Data

The SBA 7(a) loan program caters to small businesses and offers a maximum loan amount of \$5 million. These loans provide the ideal environment to study the role of Yelp ratings because the covered businesses tend to have a single location and be family owned and operated. They are not subject to strict government filing requirements and do not have much publicly available information. The SBA 7(a) loans are given to small businesses unable to obtain credit elsewhere on reasonable terms. The SBA also requires the borrowers to have good prospects of repaying those loans and to be small among other aspects.

To obtain an SBA loan, borrowers apply to lenders directly. Depending on the lender type, the SBA performs different levels of review. Regular lenders are required to submit completed loan applications to the SBA, which makes the final credit decision. However, lenders in the preferred lenders program (PLP) can

decide on the outcome of the SBA loans without SBA review, making the SBA lending process comparable to that of commercial bank loans.⁷ The lending banks keep the loans on their books and also process, close, service, and liquidate the loans. The SBA sets a ceiling on the interest rate the banks can charge, which is normally a few percentage points over the prime rate. This limit makes the SBA loans no pricier than conventional small business loans. The application process for SBA loans is also timely. If the application is made with a regular lender, the decision process takes 5–10 business days. If the application is made with a preferred lender, the borrower could be approved within 24 hours.

One feature of the SBA 7(a) program is that the SBA provides loan guarantees to eligible businesses through the lenders. For loans of \$350,000 or more, borrowers pay a one-time, up-front guarantee fee between 0.25% and 3.75% of the guaranteed portion based on the loan amount and maturity.⁸ However, the SBA does not guarantee the entire loan amount. In the regular programs, SBA guarantees up to 85% for loans under \$150,000 and up to 75% for loans over \$150,000. SBA guarantees up to 50% on express loans. The loans do not always carry the maximum allowable guarantee because the borrowers try to reduce the guarantee fees paid to the SBA. As a design feature, lenders share the default risk with the SBA.

Table 1. Yelp Data Summary

Panel A. Yelp data by year			
Year	Number of businesses	Number of reviews	Number of reviews per business
2004	577	879	1.52
2005	17,729	48,764	2.75
2006	56,861	232,144	4.08
2007	124,862	681,431	5.46
2008	207,286	1,329,585	6.41
2009	286,580	2,053,547	7.17
2010	410,747	3,191,291	7.77
2011	561,496	4,841,449	8.62
2012	684,470	5,719,592	8.36
2013	847,984	7,694,181	9.07
2014	1,032,939	11,376,074	11.01
2015	1,239,229	15,425,396	12.45
2016	1,395,928	17,994,708	12.89
Panel B. Yelp data by category			
Yelp category	Average rating	Average number of reviews	Category weight, %
Restaurants	3.49	96.13	20.17
Shopping	3.77	15.91	10.30
Home services	3.91	11.72	8.77
Food	3.76	55.00	8.19
Health & medical	4.09	11.92	6.95
Beauty & spas	4.12	25.03	6.45
Local services	3.90	14.14	6.30
Automotive	3.72	19.12	5.78
Active life	4.23	20.71	4.07
Event planning & services	4.18	26.96	3.63
Nightlife	3.60	119.61	3.48
Professional services	4.15	8.26	2.93
Education	4.16	9.51	2.05
Hotels & travel	3.39	33.67	2.01
Financial services	3.54	6.65	1.99
Arts & entertainment	4.10	37.28	1.83
Pets	4.18	22.13	1.68
Real estate	3.63	8.42	1.59
Public services & government	3.53	19.01	0.77
Religious organizations	4.49	4.98	0.65
Local flavor	4.18	17.49	0.22
Mass media	3.53	8.78	0.17

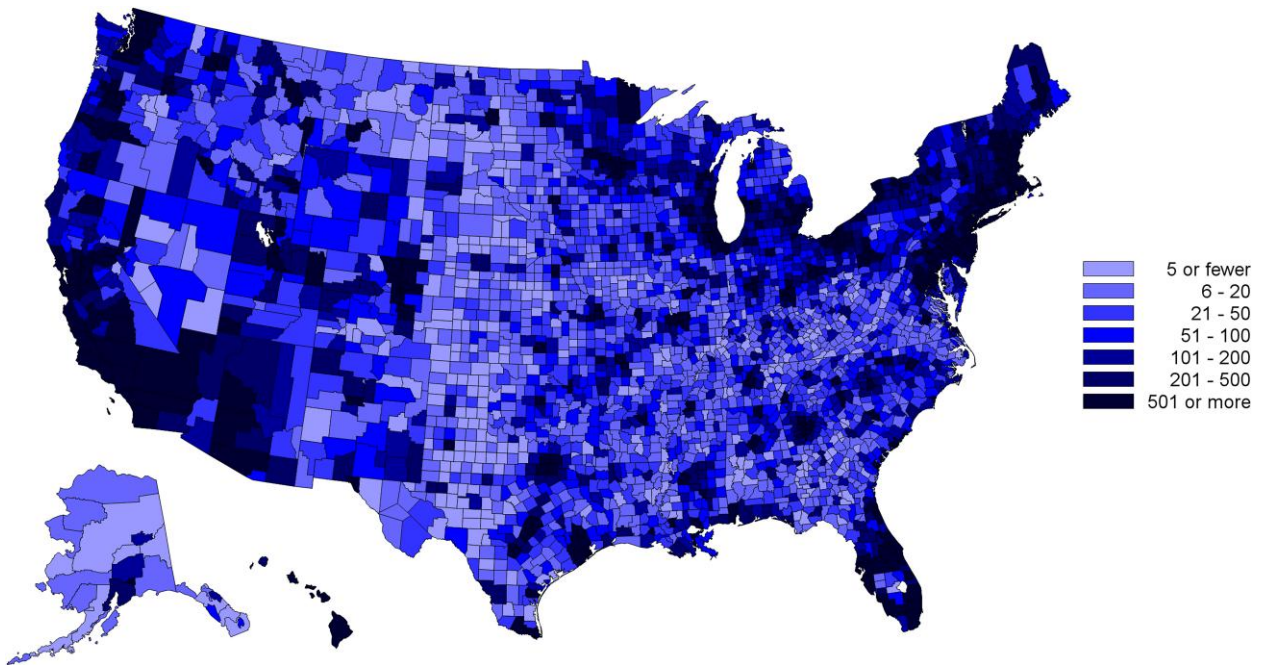
Notes. This table presents a summary of the Yelp data collected. Panel A presents the number of businesses, number of reviews, and the average number of reviews per business covered by Yelp in each year. Panel B shows a breakdown of the Yelp sample by the 22 Yelp business categories. The average rating, average number of reviews, and the category weight out of total Yelp businesses are reported for each category. Yelp data are available from 2004 to 2016.

I obtain the SBA loan data from the SBA through a Freedom of Information Act request. The data provide detailed information on each SBA loan, including business name, business address, bank name, bank address, loan-granting date, loan amount, interest rate, collateral, default status, and write-off amount. Figure 4 plots SBA loan coverage at the county level. For each business that participates in the SBA loan program, I assign a census-defined county code based on the address. I count the number of businesses in each county. Similar to the Yelp coverage map, the

counties with more (fewer) businesses in the SBA loan program are presented in darker (lighter) color. The map shows that businesses located throughout the United States take part in the SBA lending program.

2.4. Data Merges and Summary Statistics

To prepare the final sample for empirical analysis, I match the Yelp and SBA loan data in three steps using business and borrower names and addresses. First, I fuzzy match the standardized business names and addresses in both data sets. I manually check the

Figure 4. (Color online) SBA Loan Data Coverage

Notes. This map plots SBA loan data coverage at the county level. SBA loan data are available from 2004 to 2016. The number of SBA 7(a) loans is aggregated to the county level to produce the figure. Regions range from darker (counties with higher numbers of SBA loans) to lighter (counties with lower numbers of SBA loans).

matching results to identify matched pairs. Second, I geocode all of the addresses. I then fuzzy match standardized business names and require the geocoded addresses to be within 0.1 miles. Again, I manually check the matching results to identify additional accurate matches. Third, in the case of pairs with different names in the two data sets but that are within 0.1 miles or have similar addresses as measured by spelling distance, I search databases such as Reference USA, Yellow Pages, and county records to identify additional matches.

The matched Yelp–SBA data are at the business–loan level. For each originated loan, I merge on the firm’s Yelp rating from the previous month end given that SBA loans are approved within 15 days at most. As indicated in Equation (1), Yelp ratings are measured at the end of month t . The outcome variables, that is, $Outcome_{i,t+}$, are measured at different times in the future. Specifically, when studying loan terms, the outcomes are measured at month $t + 1$, and when studying loan performance measures, they are measured at the time when default takes place. Note that my Yelp data set ends in 2016, but performance information extends to 2021 because it takes some time to observe default.⁹

I next construct the variables of interest. The key outcome variables are *loan spread*, *collateral*, a *default indicator*, and *write-off amount*. *Loan spread* is the interest rate

charged on the loan (determined by the lending institution) minus the prime rate at the beginning of that month. *Collateral* is the amount required as collateral divided by the total loan amount. The *default indicator* is a dummy variable that equals one if the business defaults on the loan and zero otherwise. *Write-off amount* is the amount written off by the lender divided by the total loan amount.

I also include a set of variables to control for business and loan characteristics. *Lag log(number of reviews)* is the one-month lagged natural logarithm of the number of reviews. *Price ranges* are dummy variables based on the four Yelp price range categories displayed in the business profiles. For example, *price range (\$)* is a dummy variable that equals one if the business is in the one-dollar-sign price range category on Yelp and zero otherwise. *Loan amount* is the natural logarithm of total approved loan value. *Loan maturity* is the natural logarithm of loan term length in months.

Table 2, panel A, reports the number of businesses by star level. The sample covers approximately 20,000 businesses. Slightly more than 28% of businesses have four stars, whereas less than 6% of businesses have a two-star rating or below. Panel B of Table 2 presents the descriptive statistics, including the mean, standard deviation, P25, P50, and P75. Businesses on Yelp have an average rating of 3.70. One third of the businesses are in the one-dollar-sign price range, whereas 58% are in

Table 2. Summary Statistics

Panel A. Number of businesses by star level						
Star level	Number of businesses	Percentage			Cumulative percentage	
1	247	1.24			1.24	
1.5	268	1.35			2.59	
2	590	2.97			5.55	
2.5	1,259	6.33			11.88	
3	2,609	13.11			24.99	
3.5	4,236	21.29			46.28	
4	5,649	28.39			74.67	
4.5	3,028	15.22			89.89	
5	2,012	10.11			100.00	
Total	19,898	100.0				
Panel B. Descriptive statistics						
Variables	N	Mean	Standard deviation	P25	P50	P75
Average rating	19,898	3.7033	0.8210	3.2500	3.8159	4.2500
I _{round up}	19,898	0.4733	0.4993	0	0	1
Number of reviews	19,898	61.5497	131.4359	5	16	57
Lag log(number of reviews)	19,898	2.9551	1.4868	1.7918	2.7726	4.0254
Price range (\$)	19,898	0.3325	0.4711	0	0	1
Price range (\$\$)	19,898	0.5795	0.4937	0	1	1
Price range (\$\$\$)	19,898	0.0737	0.2612	0	0	0
Price range (\$\$\$\$)	19,898	0.0144	0.1190	0	0	0
Loan spread (%)	19,898	2.9389	1.3695	2	2.7500	3.5000
Collateral	19,898	0.3385	0.1514	0.2500	0.2500	0.5000
Default indicator	19,898	0.0374	0.1898	0	0	0
Write-off amount	19,898	0.0271	0.1449	0	0	0
Loan spread (%) (cutoff = 3.75)	4,898	2.9184	1.3431	2	2.7500	3.5000
Collateral (cutoff = 3.75)	4,898	0.3407	0.1496	0.2500	0.2500	0.5000
Default indicator (cutoff = 3.75)	4,898	0.0390	0.1936	0	0	0
Write-off amount (cutoff = 3.75)	4,898	0.0276	0.1449	0	0	0
Loan spread (%) (cutoff = 4.25)	4,997	2.8023	1.2581	2	2.7500	3.1500
Collateral (cutoff = 4.25)	4,997	0.3102	0.1530	0.1500	0.2500	0.5000
Default indicator (cutoff = 4.25)	4,997	0.0290	0.1679	0	0	0
Write-off amount (cutoff = 4.25)	4,997	0.0210	0.1278	0	0	0
Loan received indicator (× 1,000)	40,202,423	0.3897	19.7370	0	0	0
Number of reviews next month (full sample)	40,202,423	45.4522	105.9569	4	11	38
Number of reviews next month (cutoff = 3.75)	8,853,679	76.6362	143.6280	9	25	77
Number of reviews next month (cutoff = 4.25)	9,768,093	51.8753	119.7409	4	12	44
Subsequent business opening (full sample)	2,103,408	0.0113	0.1057	0	0	0
Subsequent business opening (cutoff = 3.75)	455,652	0.0086	0.0923	0	0	0
Subsequent business opening (cutoff = 4.25)	544,728	0.0119	0.1086	0	0	0

Notes. This table presents summary statistics for the sample. Panel A presents the number of businesses by star level, and panel B presents descriptive statistics. *Average rating* is the unrounded average monthly Yelp rating. *I_{round up}* is an indicator variable that equals one if the Yelp rating is rounded up and zero if rounded down. *Number of reviews* is the cumulative number of reviews at the end of each month. *Lag log(number of reviews)* is the one-month lagged natural logarithm of the number of reviews. *Price ranges* are dummy variables based on the four Yelp price range categories displayed in the business profile. For example, *price range (\$)* is a dummy variable that equals one if the business is in the \$ price range category on Yelp and zero otherwise. *Loan spread* is the interest rate charged on the loan that is determined by the lending institution minus the beginning-of-month prime rate. *Collateral* is the amount required as collateral divided by the total loan amount. *Default indicator* is a dummy variable that equals one if the business defaults on the loan and zero otherwise. *Write-off amount* is the amount written off by the lender divided by the total loan amount. *Loan received indicator* is a dummy variable that equals one if the business receives an SBA loan and zero otherwise. *Number of reviews next month* is the number of Yelp reviews in the following month. *Subsequent business opening* is a dummy variable that equals one if an existing business opens another location and zero otherwise.

the two-dollar-sign price range. The loans, on average, require 34% of the loan amount as collateral, command a 294-bps spread over the prime rate, and experience a

3.7% default rate.¹⁰ As seen in Online Appendix F, the average loan granted by the SBA is approximately \$385,000 with a maturity of 10 years (i.e., 124 months).

3. Loan-Level Analysis

In this section, I first examine the relationship between coarse Yelp ratings and loan outcomes, including loan terms and performance. I then further investigate the role of Yelp ratings in bank lending decisions by exploring the cross-sectional variation in the lending environment. Next, I explore the strength of the signal from Yelp ratings. Finally, I provide some suggestive evidence of loan probability.

3.1. Loan Terms

To examine whether coarse Yelp ratings affect loan terms, I consider two measures for loan terms following the banking literature: *loan spread* and *collateral*. I calculate the *loan spread* as the difference between the reported interest rate and the beginning-of-month prime rate. *Collateral* is the amount of collateral required for each loan divided by its respective total loan amount. Given that the SBA does not provide data on rejected loans, the analyses studying loan terms and loan performance are conditional on loan approval.

To carry out the analysis, I follow the RDD framework. I first plot the average loan spreads and collateral requirements in Figures 5 and 6, respectively, by rounding cutoffs. The graphs show discontinuous jumps in loan terms near the rounding thresholds. Taking the 3.75 cutoff as an example (i.e., panel (f)), the average firm to the left of the cutoff (just below) pays about 310 bps in loan spread compared with around 270 bps for the average firm just above the cutoff. Similarly, in terms of collateral, the average firm in the bin just above the cutoff pledges about five percentage points less in assets compared with the average firm in the bin just below the cutoff.

I formally test the relationship between Yelp ratings and loan terms by estimating Equation (1). I report the results in Table 3. The dependent variables in panels A and B are *loan spread* and *collateral*, respectively. In column (1), I utilize the full sample. When using *loan spread* as the dependent variable, the Yelp rating indicator has a statistically significant regression coefficient of about -0.25 ; thus, a half-star increase in Yelp ratings translates to a 25-bps reduction in *loan spread*. Given that the average loan spread is about 2.94% (294 bps), the finding is also economically significant, representing close to a 9% cost savings in business borrowing.¹¹ Using *collateral* as the dependent variable, I also document a negative and statistically significant result. More specifically, the businesses falling just above the rounding thresholds pledge 6% fewer assets as *collateral* compared with those just below the rounding thresholds.¹²

Columns (2) and (3) of Table 3 focus on businesses close to the 3.75 and 4.25 cutoffs, respectively. Among businesses at the 3.75 cutoff, those rounded up to 4 stars

are much more likely to have a lower *loan spread* compared with those rounded down to 3.5 stars. The regression coefficient represents an 11% decrease in *loan spread* compared with the unconditional mean. The results are similar when using *collateral* as the outcome variable. Together these results indicate that aggregated Yelp ratings lead to lower financing costs, suggesting that Yelp ratings influence lenders in their decision-making process.

As discussed in the empirical design section, my sample contains a lower percentage of low-star firms than the Yelp universe, indicating potential survivorship bias. This empirical limitation prevents me from finding any results at low star levels. If I were to assume that the effects of rounding up are the same at high versus low star levels, I would expect the point estimate to be larger in magnitude (i.e., more negative) than that reported in column (1) when using the full sample. To quantify the likely scale of the bias, the reduction in loan spread when pooling all star levels could be up to 5 bps higher in the absence of the survivorship bias.¹³

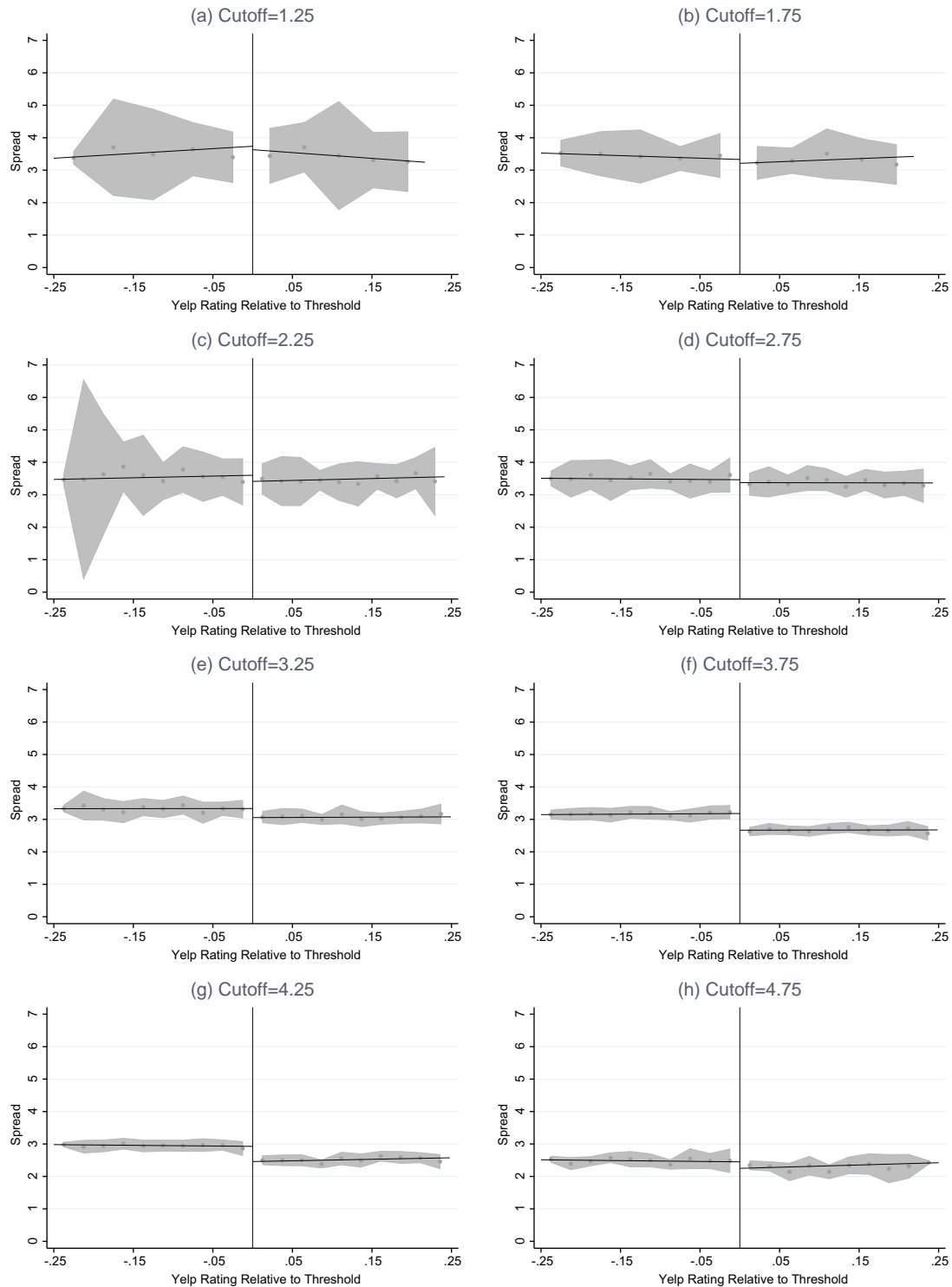
3.2. Loan Performance

I next focus on loan performance. I follow a similar empirical approach as that in the previous section, using two measures for loan performance. First, I look at *default*, which is an indicator variable that equals one if the borrower defaults on the loan and zero otherwise. Second, I use the outstanding loan amount that is written off by the lending institutions as the other measure for loan performance. I calculate the *write-off amount* as the percentage of the loan amount that is written off divided by the respective total loan amount.

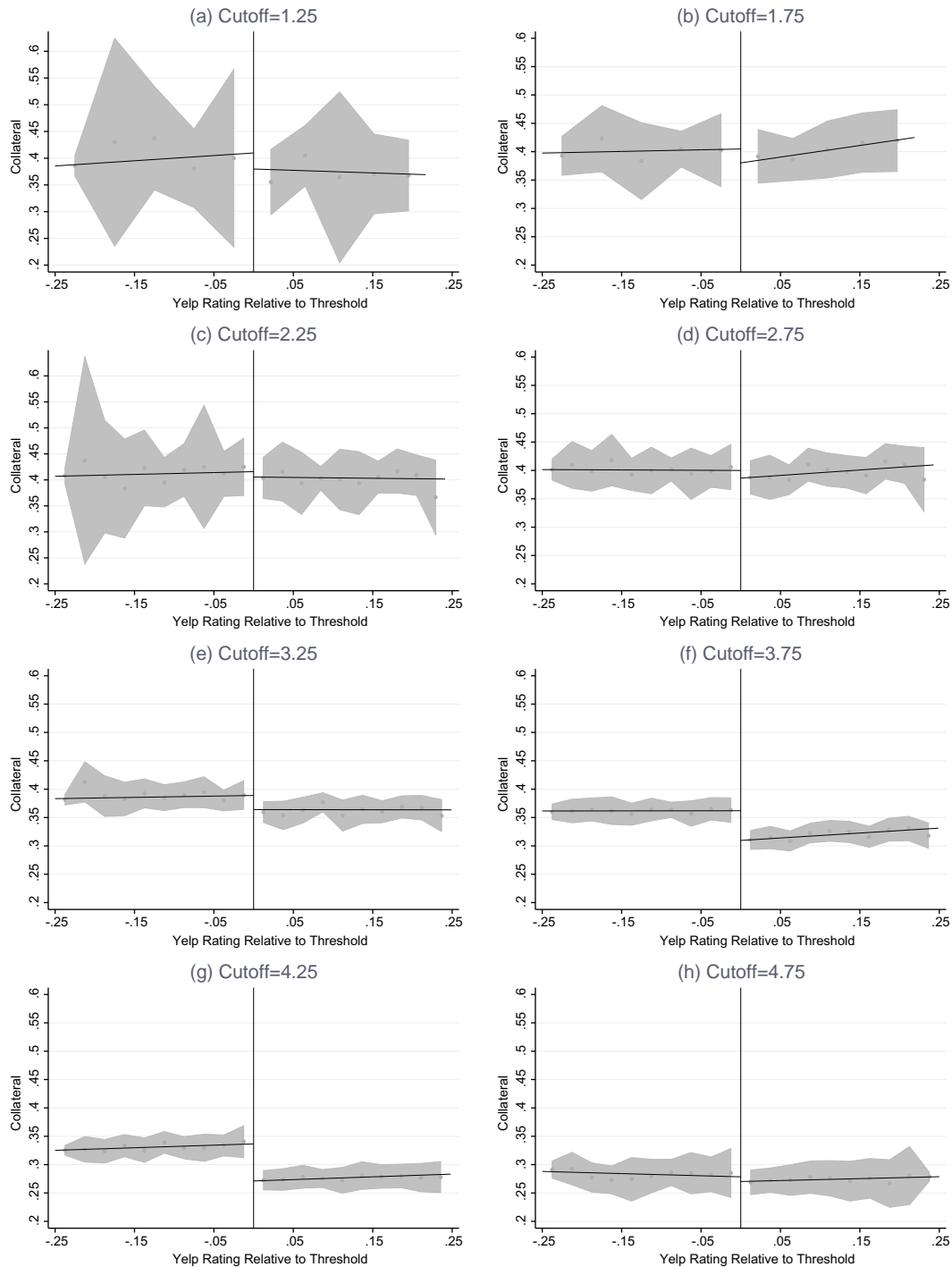
Again, I implement the RDD strategy. In Figures 7 and 8, I plot the loan *default indicator* and *write-off amount* on each side of the eight predetermined Yelp cutoffs separately. Businesses rounded up to the next Yelp rating categories exhibit lower default probabilities and write-off amounts compared with those rounded down. I next formally test this observation. The results are reported in Table 4. I use the *default indicator* as the dependent variable in panel A and the *write-off amount* as the dependent variable in panel B.

In column (1) of Table 4, I use the same specification as in the previous section to implement the full-sample parametric approach. I find that the regression coefficients on $I_{\text{round up}}$ are negative and statistically significant when using both the *default indicator* and *write-off amount* as the dependent variables. These findings indicate that coarse Yelp ratings have a causal relationship with loan performance, that is, higher Yelp ratings imply a greater likelihood that borrowers will repay the loans. These results are economically large. Compared with businesses that fall below the rounding

Figure 5. (Color online) Loan Spread



Notes. The figures plot the SBA loan spread as a function of Yelp ratings around each cutoff. Yelp rounds the average ratings of the businesses up and down to the nearest half point based on predetermined cutoffs. On a scale from 1 to 5, the ratings 1.25, 1.75, 2.25, 2.75, 3.25, 3.75, 4.25, and 4.75 are the cutoff points, and the graphs are displayed accordingly. Businesses with average ratings above (below) the cutoffs are assigned Yelp ratings that are rounded up (down) to the nearest half point. Yelp ratings are recentered around their respective cutoffs to zero. For every Yelp rating bin, the dots represent the average SBA loan spread in that bin, calculated as the average loan spread across all loans within the bin. The lines are first order polynomials fitted through the loan spreads on each side of the cutoff. The shaded regions represent the 95% confidence bounds.

Figure 6. (Color online) Loan Collateral

Notes. The figures plot the *collateral* required for SBA loans as a function of Yelp ratings around each cutoff. Yelp rounds the average ratings of the businesses up and down to the nearest half point based on predetermined cutoffs. On a scale from 1 to 5, the ratings 1.25, 1.75, 2.25, 2.75, 3.25, 3.75, 4.25, and 4.75 are the cutoff points, and the graphs are displayed accordingly. Businesses with average ratings above (below) the cutoffs are assigned Yelp ratings that are rounded up (down) to the nearest half point. Yelp ratings are recentered around their respective cutoffs to zero. For every Yelp rating bin, the dots represent the average collateral required in that bin, which is calculated as the collateral for each loan divided by the respective total loan amount averaged across all loans within the bin. The lines are first order polynomials fitted through the required loan collateral on each side of the cutoff. The shaded regions represent the 95% confidence bounds.

Table 3. Loan Terms

Variables	Full sample (1)	Cutoff = 3.75 (2)	Cutoff = 4.25 (3)
Panel A. <i>Loan spread</i>			
$I_{\text{round up}}$	−0.2511*** (0.0329)	−0.3163*** (0.0726)	−0.2972*** (0.0685)
<i>Lag log(number of reviews)</i>	0.0193*** (0.0074)	0.0552*** (0.0156)	0.0112 (0.0148)
<i>Price range</i> (\$)	−0.0223 (0.0803)	−0.1527 (0.2092)	−0.0420 (0.2046)
<i>Price range</i> (\$\$)	−0.0369 (0.0746)	−0.1637 (0.2055)	−0.0886 (0.1994)
<i>Price range</i> (\$\$\$)	−0.1124 (0.0789)	−0.2904 (0.2074)	−0.2124 (0.1917)
Loan characteristics control	✓	✓	✓
Cutoff fixed effects	✓		
Year fixed effects	✓	✓	✓
Yelp industry fixed effects	✓	✓	✓
County fixed effects	✓	✓	✓
Observations	19,665	4,662	4,783
Adjusted R^2	0.4299	0.4283	0.3621
Panel B. <i>Collateral</i>			
$I_{\text{round up}}$	−0.0208*** (0.0037)	−0.0249*** (0.0085)	−0.0235*** (0.0081)
<i>Lag log(number of reviews)</i>	0.0008 (0.0007)	0.0005 (0.0020)	0.0035* (0.0018)
<i>Price range</i> (\$)	−0.0011 (0.0074)	−0.0244 (0.0180)	0.0134 (0.0234)
<i>Price range</i> (\$\$)	−0.0007 (0.0067)	−0.0208 (0.0173)	0.0068 (0.0204)
<i>Price range</i> (\$\$\$)	−0.0090 (0.0067)	−0.0135 (0.0176)	−0.0015 (0.0235)
Loan characteristics control	✓	✓	✓
Cutoff fixed effects	✓		
Year fixed effects	✓	✓	✓
Yelp industry fixed effects	✓	✓	✓
County fixed effects	✓	✓	✓
Observations	19,665	4,662	4,783
Adjusted R^2	0.4531	0.4337	0.3806

Notes. This table reports OLS regression results for SBA loan terms and Yelp ratings. In panel A, the dependent variable is the *loan spread*, calculated as the interest rate charged on the loan that is determined by the lending institution minus the beginning of month prime rate. In panel B, the dependent variable is the amount required as collateral divided by total loan amount; $I_{\text{round up}}$ is an indicator variable that equals one if the Yelp rating is rounded up and zero if rounded down. Column (1) utilizes the full sample with all cutoffs. Columns (2) and (3) examine cases in which the cutoff equals 3.75 and 4.25, respectively. The cutoffs are defined in Figures 5 and 6. Loan characteristics control includes *loan amount* and *maturity*. The variables are defined in Table 2. Robust standard errors clustered by county are in parentheses.

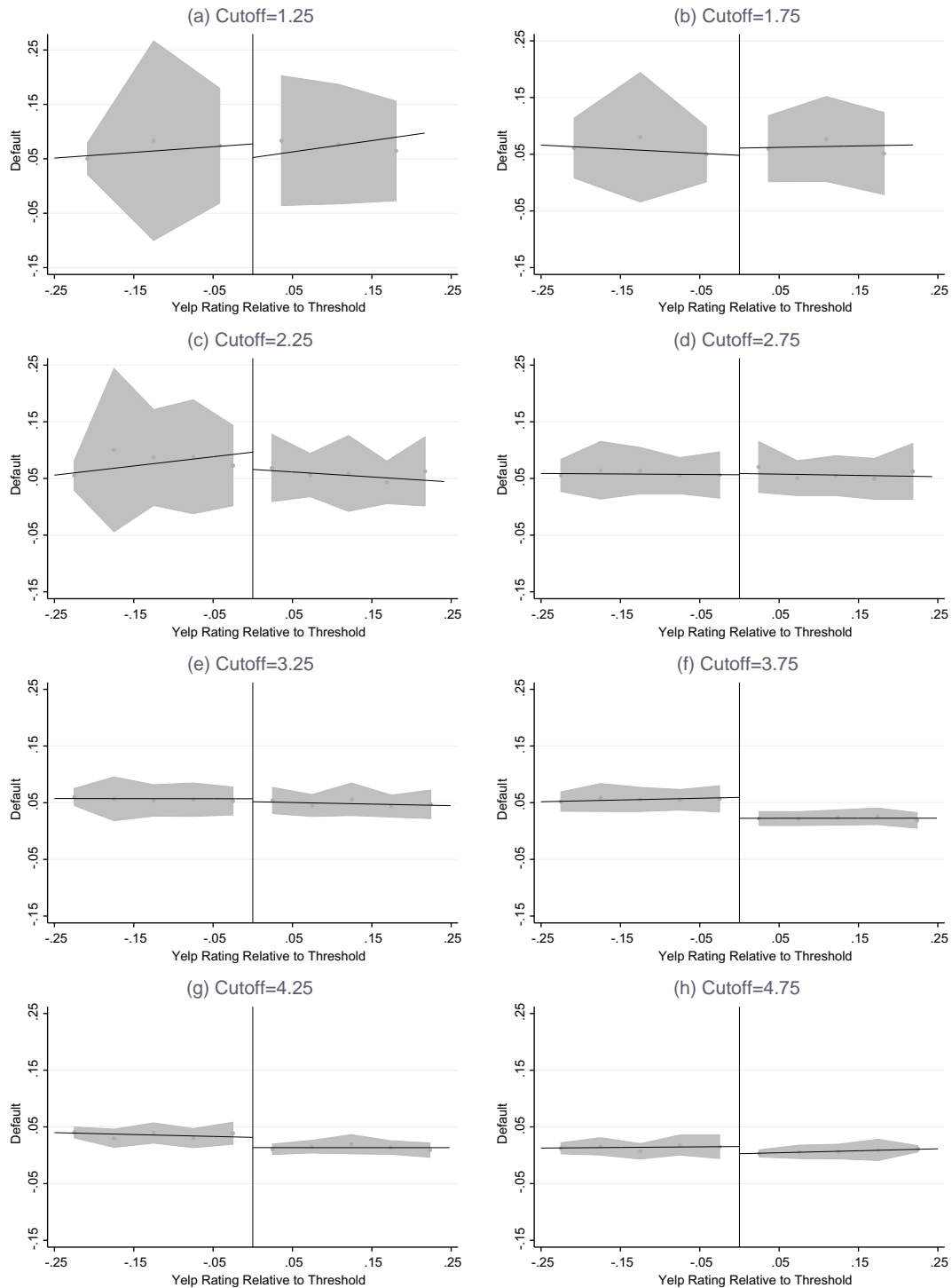
***, **, and * denote statistical significance at the 1%, 5% and 10% level, respectively.

thresholds, those above the thresholds are 45% less likely to default on SBA loans. When using the *write-off amount* as the outcome variable, I find that businesses whose ratings are rounded up compared with those rounded down have 76% lower loan write-offs relative to the sample mean.

In columns (2) and (3) of Table 4, I again restrict the sample to businesses near the 3.75 and 4.25 cutoffs, respectively. Column (2) shows not only a statistically significant relationship between loan performance and Yelp ratings, but also a greater economic meaning compared with the full sample. Businesses rounded up to 4 stars are 72% less likely to default, and the write-off

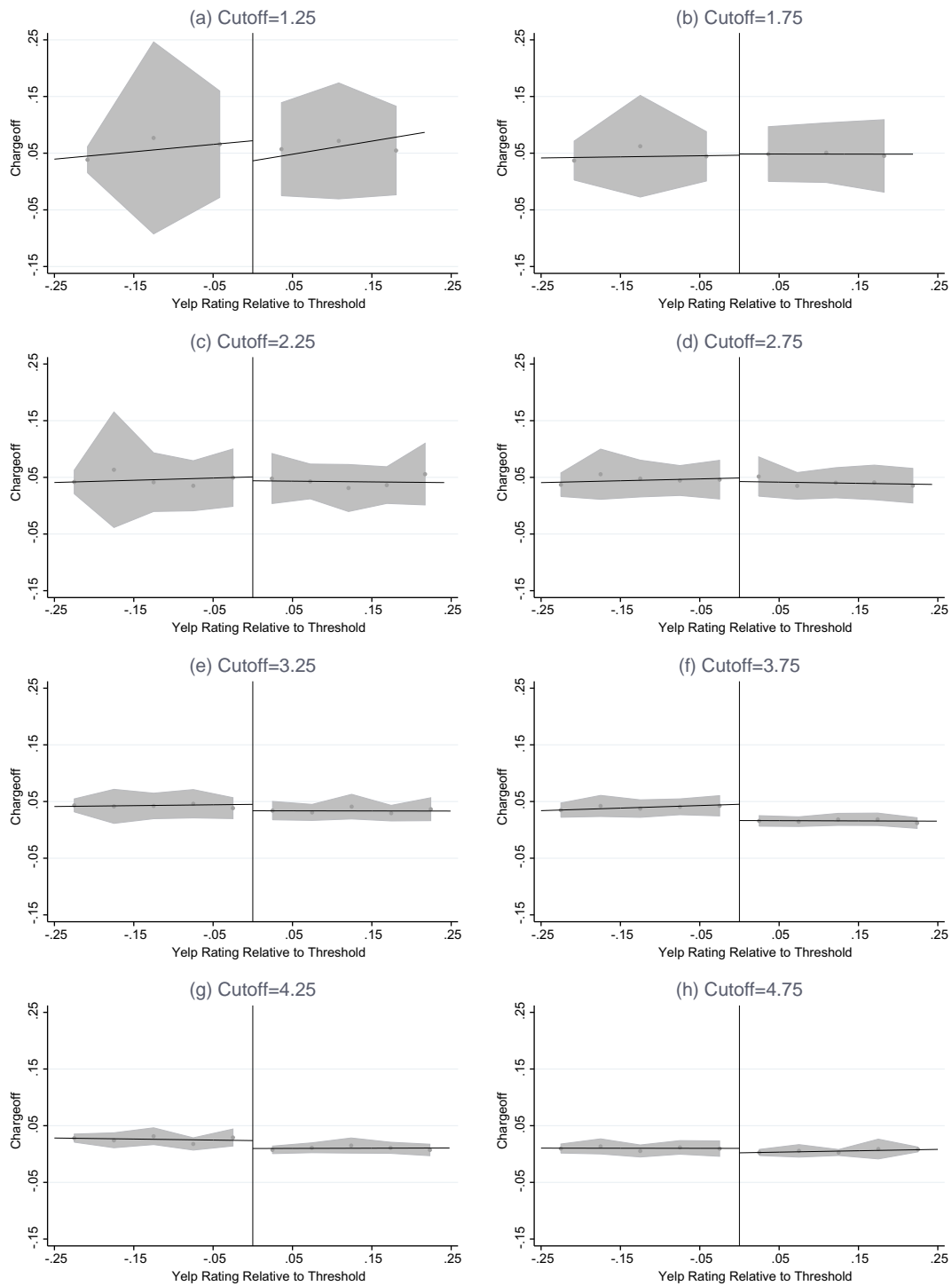
amounts are 1.3 times lower than those of businesses rounded down to 3.5 stars. Column (3) shows similar results for the 4.25 cutoff. Overall, these results indicate that a one-half-star higher Yelp rating leads to better SBA loan performance. This evidence suggests that using information from Yelp benefits both the borrowers and the lenders.¹⁴

I am interested in how much of the loan performance effect of Yelp ratings likely is attributable to better loan terms versus a revenue channel. I, therefore, rerun the loan performance tests controlling for loan pricing and present the results in Online Appendix K. When using either the *default indicator* or the *write-off amount* as the

Figure 7. (Color online) Loan Default Indicator

Notes. The figures plot the *default indicator* as a function of Yelp ratings around each cutoff. Yelp rounds the average ratings of the businesses up and down to the nearest half point based on predetermined cutoffs. On a scale from 1 to 5, the ratings 1.25, 1.75, 2.25, 2.75, 3.25, 3.75, 4.25, and 4.75 are the cutoff points, and the graphs are displayed accordingly. Businesses with average ratings above (below) the cutoffs are assigned Yelp ratings that are rounded up (down) to the nearest half point. Yelp ratings are recentered around their respective cutoffs to zero. For every Yelp rating bin, the dots represent the average probability of default in that bin, which is calculated as the number of loans default over the total number of loans within the bin. The lines are first order polynomials fitted through the probabilities of default on each side of the cutoff. The shaded regions represent the 95% confidence bounds.

Figure 8. (Color online) Loan Charge-off Amount



Notes. The figures plot the SBA loan *write-off* amount upon default as a function of Yelp ratings around each cutoff. Yelp rounds the average ratings of the businesses up and down to the nearest half point based on predetermined cutoffs. On a scale from 1 to 5, the ratings 1.25, 1.75, 2.25, 2.75, 3.25, 3.75, 4.25, and 4.75 are the cutoff points, and the graphs are displayed accordingly. Businesses with average ratings above (below) the cutoffs are assigned Yelp ratings that are rounded up (down) to the nearest half point. Yelp ratings are recentered around their respective cutoffs to zero. For every Yelp rating bin, the dots represent the average write-off amount in that bin, which is calculated as the write-off amount for each loan divided by the respective total loan amount averaged across all loans within the bin. The lines are first order polynomials fitted through the write-off amounts on each side of the cutoff. The shaded regions represent the 95% confidence bounds.

Table 4. Loan Performance

Variables	Full sample (1)	Cutoff = 3.75 (2)	Cutoff = 4.25 (3)
Panel A. <i>Default indicator</i>			
$I_{round\ up}$	−0.0205*** (0.0058)	−0.0353** (0.0141)	−0.0222** (0.0086)
<i>Lag log(number of reviews)</i>	−0.0007 (0.0012)	0.0004 (0.0026)	−0.0016 (0.0021)
<i>Price range (\$)</i>	0.0043 (0.0124)	−0.0118 (0.0406)	−0.0652** (0.0303)
<i>Price range (\$\$)</i>	0.0049 (0.0122)	−0.0095 (0.0406)	−0.0599* (0.0308)
<i>Price range (\$\$\$)</i>	−0.0060 (0.0129)	−0.0315 (0.0376)	−0.0538* (0.0300)
Loan characteristics control	✓	✓	✓
Cutoff fixed effects	✓		
Year fixed effects	✓	✓	✓
Yelp industry fixed effects	✓	✓	✓
County fixed effects	✓	✓	✓
Observations	19,665	4,662	4,783
Adjusted R^2	0.0389	0.0487	0.0455
Panel B. <i>Write-off amount</i>			
$I_{round\ up}$	−0.0168*** (0.0044)	−0.0279*** (0.0105)	−0.0168*** (0.0063)
<i>Lag log(number of reviews)</i>	−0.0003 (0.0009)	−0.0010 (0.0020)	−0.0003 (0.0015)
<i>Price range (\$)</i>	−0.0018 (0.0102)	−0.0220 (0.0330)	−0.0557** (0.0252)
<i>Price range (\$\$)</i>	−0.0011 (0.0100)	−0.0220 (0.0329)	−0.0508** (0.0255)
<i>Price range (\$\$\$)</i>	−0.0078 (0.0103)	−0.0360 (0.0305)	−0.0434* (0.0242)
Loan characteristics control	✓	✓	✓
Cutoff fixed effects	✓		
Year fixed effects	✓	✓	✓
Yelp industry fixed effects	✓	✓	✓
County fixed effects	✓	✓	✓
Observations	19,665	4,662	4,783
Adjusted R^2	0.0384	0.0497	0.0417

Notes. This table reports OLS regression results for SBA loan performance and Yelp ratings. In panel A, the dependent variable is a dummy variable that equals one if the business defaults on the loan and zero otherwise. In panel B, the dependent variable is the write-off amount by the lender divided by total loan amount. $I_{round\ up}$ is an indicator variable that equals one if the Yelp rating is rounded up and zero if rounded down. Column (1) utilizes the full sample with all cutoffs. Columns (2) and (3) examine cases in which the cutoff equals 3.75 and 4.25, respectively. The cutoffs are defined in Figures 7 and 8. Loan characteristics control includes *loan amount* and *maturity*. The variables are defined in Table 2. Robust standard errors clustered by county are in parentheses.

***, **, and * denote statistical significance at the 1%, 5% and 10% level, respectively.

dependent variable, the adjusted R^2 increases by around 4%, suggesting that loan pricing contributes to the treatment effect on loan performance in addition to the revenue channel. The coefficients on the main independent variable, $I_{round\ up}$, decrease by only about 10%. In sum, the results indicate that loan pricing plays a smaller role in determining loan performance than the revenue channel.

Taking the loan-level analyses together, I interpret the loan outcome results as consistent with the notion that aggregated Yelp ratings provide reliable information

about borrowers' future earnings potential. My evidence indicates that Yelp star ratings are a significant factor in the SBA lending process. Thus, Yelp ratings help to address costly information-acquisition problems when reliable information about the borrowers is scarce.

3.3. Cross-Sectional Analyses

To better understand the role of Yelp ratings in bank lending decisions, I explore whether lenders are more likely to rely on coarse Yelp ratings when they are less informed or have lower incentives to be informed. First,

Table 5. Cross-Sectional Analysis

Variables	Loan spread (1)	Collateral (2)	Default indicator (3)	Write-off amount (4)
Panel A. Distance to bank				
$I_{\text{round up}}$	−0.1340*** (0.0388)	−0.0130*** (0.0044)	−0.0126** (0.0063)	−0.0085* (0.0048)
<i>Far-from-lender indicator</i>	0.3983*** (0.0495)	0.0390*** (0.0064)	0.0089* (0.0049)	0.0084** (0.0039)
$I_{\text{round up}} \times$ <i>far-from-lender indicator</i>	−0.2346*** (0.0331)	−0.0161*** (0.0037)	−0.0152*** (0.0051)	−0.0159*** (0.0041)
Business characteristics control	✓	✓	✓	✓
Loan characteristics control	✓	✓	✓	✓
Cutoff fixed effects	✓	✓	✓	✓
Year fixed effects	✓	✓	✓	✓
Yelp industry fixed effects	✓	✓	✓	✓
County fixed effects	✓	✓	✓	✓
Observations	19,665	19,665	19,665	19,665
Adjusted R^2	0.4396	0.4615	0.0392	0.0390
Panel B. Repeated borrowing				
$I_{\text{round up}}$	−0.2697*** (0.0317)	−0.0221*** (0.0035)	−0.0229*** (0.0061)	−0.0185*** (0.0046)
<i>Previous SBA loan indicator</i>	−0.3066*** (0.0333)	−0.0093** (0.0040)	−0.0184*** (0.0048)	−0.0127*** (0.0041)
$I_{\text{round up}} \times$ <i>previous SBA loan indicator</i>	0.1791*** (0.0523)	0.0108** (0.0050)	0.0200*** (0.0074)	0.0146*** (0.0054)
Business characteristics control	✓	✓	✓	✓
Loan characteristics control	✓	✓	✓	✓
Cutoff fixed effects	✓	✓	✓	✓
Year fixed effects	✓	✓	✓	✓
Yelp industry fixed effects	✓	✓	✓	✓
County fixed effects	✓	✓	✓	✓
Observations	19,665	19,665	19,665	19,665
Adjusted R^2	0.4329	0.4532	0.0393	0.0387

Notes. This table reports OLS regression results for SBA loan outcomes and Yelp ratings in the cross-section; $I_{\text{round up}}$ is an indicator variable that equals one if the Yelp rating is rounded up and zero if rounded down. Panel A includes the *far-from-lender indicator*, which is a dummy variable that equals one if the distance between the firm and the lender is above the sample median and zero otherwise. Panel B includes the *previous SBA loan indicator*, which is a dummy variable that equals one if the borrower had a loan from the SBA previously and zero otherwise. Business characteristics control includes different *price ranges* and *lag log(number of reviews)*. Loan characteristics control includes *loan amount* and *maturity*. The variables are defined in Table 2. Robust standard errors clustered by county are in parentheses.

***, **, and * denote statistical significance at the 1%, 5% and 10% level, respectively.

I exploit a source of heterogeneity in the cost of retrieving information about the borrower. For businesses located farther away from the banks, banks incur extra communication, transportation, and monitoring costs to obtain the information they need to approve and oversee the loans (Degryse and Ongena 2005). In contrast, soft information is less costly to collect the closer the bank is to the borrower (Petersen and Rajan 2002), and some information can only be collected locally (Agarwal and Hauswald 2010). Thus, I expect that the information provided in Yelp ratings is less valuable if the lender is closer and has better soft information about the borrower. Conversely, Yelp ratings could help lenders obtain better knowledge about borrowers located farther away.¹⁵ Second, other sources of information could potentially dilute the effects of Yelp ratings. A long line of research establishes that lenders derive valuable information about the borrowers from

repeated interaction (e.g., Petersen and Rajan 1994, Berger and Udell 1995). For small businesses that are repeat participants in the SBA loan program, banks already have credit history information for these businesses as well as other borrower-specific information. Consequently, I expect banks to be less reliant on Yelp ratings in this setting.

I conduct empirical tests following the main analysis and use the full sample with all cutoffs. I again use *loan spread*, *collateral*, a *default indicator*, and the *write-off amount* as measures for loan outcomes. The results are presented in Table 5. In panel A, I examine the geographic distance between banks and borrowers. I calculate straight-line distances using businesses' physical addresses. I create a *far-from-lender indicator*, which equals one if the distance between the firm and the lender is above the sample median and zero otherwise. Panel B focuses on repeated lending relationships. I

create a *previous SBA loan indicator*, which equals one if the borrower has taken out a previous SBA loan and zero otherwise.¹⁶ I also construct interaction terms between these two indicator variables and the Yelp rating dummy, that is, $I_{\text{round up}}$, to capture the effects of distance and repeated interaction.

Using the RDD setting, I show that a one-half-star increase in Yelp ratings leads to lower loan spreads, collateral requirements, likelihood of default, and write-off amounts for businesses located farther from banks. For businesses with a previous SBA loan, the coefficients on the interaction terms are positive and significant, suggesting that lenders rely less on Yelp ratings in subsequent lending decisions compared with the first borrowing encounters. However, if I examine the overall effects of Yelp ratings in the presence of existing banking relationships, Yelp ratings are still good indicators of loan outcomes as the additional sources of information do not entirely counteract the Yelp effects.¹⁷

Overall, the results in Table 5 suggest that Yelp ratings are helpful when banks have little information about borrowers. Yelp ratings serve as good indicators of the future revenue potential of businesses, and they are readily available to lenders. The results provide suggestive evidence that banks reference Yelp ratings when making lending decisions. By exploiting situations in

which available information on borrowers varies greatly among the banks, I show that Yelp ratings become more effective when information is scarce, and they remain effective when additional soft information is present.¹⁸

3.4. Signal Precision

Another important dimension of information is the precision of the signal. Under Bayesian learning theory, the more precise a signal is, the stronger the reactions to it will be. Even though each review written on Yelp by itself is considered noisy at best, the aggregated rating from more reviews contains more information. In addition, because the number of reviews for each business is prominently displayed next to the aggregated ratings (see Figure 1 for two examples), customers are likely to apply more weight to ratings with more reviews than those with fewer reviews when making a decision using Yelp. Similarly, I expect lenders to react more strongly to a more precise signal as well. As a result, the effect of Yelp ratings on loan outcomes is stronger when the number of reviews is high.

I split the sample based on the number of reviews. I again utilize the RDD setting and use loan terms and performance as the dependent variables. The results are presented in Table 6. In panel A, I focus on the subsample of

Table 6. Signal Precision

Variables	Loan spread (1)	Collateral (2)	Default indicator (3)	Write-off amount (4)
Panel A. Number of reviews above 75th percentile				
$I_{\text{round up}}$	−0.3753*** (0.0531)	−0.0302*** (0.0060)	−0.0359*** (0.0068)	−0.0270*** (0.0054)
Business characteristics control	✓	✓	✓	✓
Loan characteristics control	✓	✓	✓	✓
Cutoff fixed effects	✓	✓	✓	✓
Year fixed effects	✓	✓	✓	✓
Yelp industry fixed effects	✓	✓	✓	✓
County fixed effects	✓	✓	✓	✓
Observations	5,160	5,160	5,160	5,160
Adjusted R^2	0.4379	0.4264	0.0450	0.0522
Panel B. Number of reviews below 25th percentile				
$I_{\text{round up}}$	−0.1014 (0.0893)	−0.0216** (0.0107)	−0.0081 (0.0164)	−0.0113 (0.0118)
Business characteristics control	✓	✓	✓	✓
Loan characteristics control	✓	✓	✓	✓
Cutoff fixed effects	✓	✓	✓	✓
Year fixed effects	✓	✓	✓	✓
Yelp industry fixed effects	✓	✓	✓	✓
County fixed effects	✓	✓	✓	✓
Observations	4,917	4,917	4,917	4,917
Adjusted R^2	0.4151	0.4925	0.0608	0.0600

Notes. This table reports OLS regression results for SBA loan outcomes and Yelp ratings using subsamples based on the number of reviews; $I_{\text{round up}}$ is an indicator variable that equals one if the Yelp rating is rounded up and zero if rounded down. Panel A focuses on businesses with a number of reviews above the sample 75th percentile. Panel B focuses on businesses with a number of reviews below the sample 25th percentile. Business characteristics control includes different *price ranges* and *lag log(number of reviews)*. Loan characteristics control includes *loan amount* and *maturity*. The variables are defined in Table 2. Robust standard errors clustered by county are in parentheses.

***, **, and * denote statistical significance at the 1%, 5% and 10% level, respectively.

businesses whose total number of reviews is above the 75th percentile. I find that the point estimates are higher for those businesses compared with the baseline results. In panel B, I use the subsample of businesses whose number of reviews is below the 25th percentile. In contrast, I find that the point estimates are smaller for those businesses and the standard errors are bigger compared with the baseline results.¹⁹

In sum, the findings show that the effects of Yelp ratings are stronger for businesses with a large number of Yelp reviews, suggesting that fewer total reviews offer a noisier signal. The results provide evidence that banks make decisions consistent with Bayesian learning, possibly through a combination of their own behavior and taking advantage of their customers' learning behavior. Overall, the documented effects in this section provide additional evidence that lenders use information from Yelp ratings when making loan decisions.

3.5. Loan Probability

The most appropriate method to assess whether lenders reference Yelp ratings in their lending decisions is to examine the probability of receiving approval for loans.²⁰ Ideally, I would use a sample of firms that applied for SBA loans and the outcome of those applications, that is, loan approval or denial. Unfortunately, the SBA only made information available about approved loans. Instead, I compare firms with SBA loans to the rest of the firms on the Yelp platform. Because I cannot observe whether businesses without SBA loans applied and were denied or never applied at all, the assumption is that the

demand for loans is similar for firms on either side of the rounding thresholds such that the results would not be driven by differentials in loan demand. Given those caveats, the results provided here should be considered suggestive.

Users can post Yelp ratings at any time, and Yelp updates the overall ratings after receiving each new rating. To strike a balance between keeping valuable information and maintaining a reasonable data analysis process, I calculate an average rating for each business at the end of every month. More specifically, at each month end, I average the ratings that the businesses have received since they first appeared on Yelp. These data provide a snapshot of every business's average rating for each month, effectively constructing a business-month level panel. My outcome variable, a *loan received indicator*, equals one if the business receives an SBA loan in the following month and zero otherwise.²¹

To formally test the relationship between having higher Yelp ratings and receiving SBA loans, I estimate a slightly modified version of Equation (1). More specifically, I regress the *loan received indicator* on the treatment dummy $I_{round\ up}$, the distance between unrounded average rating and the cutoff, the interaction of the two, $lag\ log(number\ of\ reviews)$, an indicator for whether the firm had received an SBA loan before, cutoff fixed effects, firm fixed effects, and county \times year fixed effects. In Table 7, column (1), the coefficient on $I_{round\ up}$ is highly statistically significant and positive at the 1% level, meaning that a one-half-star increase in Yelp ratings

Table 7. Loan Probability

Variables	Loan received indicator		
	(1)	(2)	(3)
$I_{round\ up}$	0.0910*** (0.0158)	0.0411** (0.0180)	0.1030*** (0.0153)
$I_{round\ up} \times far\ from\ nearest\ lender\ indicator$		0.0989*** (0.0161)	
$I_{round\ up} \times prior\ banking\ relationship\ indicator$			−0.7196*** (0.2061)
Firm level control	✓	✓	✓
Cutoff fixed effects	✓	✓	✓
Firm fixed effects	✓	✓	✓
County \times year fixed effects	✓	✓	✓
Observations	40,198,897	40,198,897	40,198,897
Adjusted R^2	0.0145	0.0145	0.0145

Notes. This table reports OLS regression results for SBA loan probability and Yelp ratings. The unit of observation is at the firm-month level. The dependent variable is a dummy variable that equals one if the business receives an SBA loan in the following month and zero otherwise and is scaled up by a factor of 1,000; $I_{round\ up}$ is an indicator variable that equals one if the Yelp rating is rounded up and zero if rounded down. The *far-from-the-nearest-lender indicator* equals one if the distance between the firm and the nearest SBA lender is above the sample median and zero otherwise. The *prior banking relationship indicator* equals one if the borrower took out an SBA loan previously and zero otherwise. Firm level control includes an indicator for whether the firm has received an SBA loan in the past and $lag\ log(number\ of\ reviews)$. The variables are defined in Table 2. Robust standard errors clustered by Yelp industry and county are in parentheses.

***, **, and * denote statistical significance at the 1%, 5% and 10% level, respectively.

leads to a significant increase in the probability of receiving loans. This effect is also economically significant, representing a 23% increase in the probability of getting SBA loans.

So far, the results provide suggestive evidence that Yelp ratings help resolve the information asymmetry problem between lenders and borrowers about the businesses' prospects. However, Yelp ratings may also impact consumer demand, which affects the creditworthiness of the business. I conduct two tests to control for the effect of Yelp ratings on consumer demand. In column (2) of Table 7, I explore the differential effects of distance to the closest SBA lender. I expect that the distance from the closest SBA lender is likely correlated with information asymmetry between the lender and business but is not correlated with the effect of Yelp rating on business demand. I include an indicator variable, *far-from-nearest-lender*, that equals one if the distance between the firm and the nearest SBA lender in the same county is above the sample median in the regression. I focus on the interaction between $I_{round\ up}$ and the *far-from-nearest-lender* indicator. I find that the effect of coarse Yelp ratings on the probability of receiving SBA loans is stronger if businesses are farther away from the nearest SBA lender.

In column (3) of Table 7, I explore whether having obtained SBA loans in the past makes Yelp ratings less effective in predicting loan probability. I expect that lenders rely less on Yelp ratings when they have extant information about the borrowers. The existence of prior lending relationship is also unlikely correlated with business demand. The interaction term between $I_{round\ up}$ and the *prior banking relationship indicator*, which equals one if the borrower had a loan from the SBA before, is negative. This finding is consistent with my prediction. Taken together, the results provide further evidence that Yelp ratings help resolve the information asymmetry problem between lenders and borrowers about businesses' prospects. The findings suggest that online review aggregators, such as Yelp, influence banks' financing decisions.

4. The Real Effects of Yelp Ratings

In this section, I explore the real effects of coarse Yelp ratings, focusing on outcomes from the consumer and business perspectives. I first study consumer demand as a result of Yelp ratings and then examine the effects of Yelp ratings on small businesses' investment decisions. More specifically, I study subsequent business openings.

4.1. Consumer Demand

In this section, I study the effects of Yelp ratings from a consumer perspective. The social learning mechanism predicts that businesses with higher Yelp ratings attract

a larger number of customers. I empirically test this hypothesis. In an ideal situation, I would observe the number of customers who visit the businesses. However, such data are not available. I, therefore, use the number of Yelp reviews for the businesses to proxy for the number of customers, assuming that the two variables are highly correlated. Using the number of reviews the businesses receive each month, I examine whether higher Yelp ratings lead to more reviews received, that is, higher consumer demand.

To perform the empirical test, I construct a business-month panel similar to the one used for the loan probability test. I again utilize the RDD setting. The dependent variable is the *number of Yelp reviews in the following month*. The key independent variable is $I_{round\ up}$, an indicator that equals one if the Yelp rating is rounded up and zero if rounded down. Given that the number of Yelp reviews is discrete in nature, I carry out Poisson regressions.

I report the results in Table 8. In column (1), I use the full sample with cutoff fixed effects. I find a positive and statistically significant coefficient on $I_{round\ up}$. This result is also economically meaningful. Businesses with one-half-star higher Yelp ratings receive 0.35% more reviews in the following month, equivalent to a 4% increase in number of reviews over one year. In columns (2) and (3), I focus on cutoffs 3.75 and 4.25, respectively. I continue to find a positive and significant relationship between consumer demand and Yelp ratings. When I compare businesses just above and just below the cutoffs, the former receives more reviews than the latter, indicating that higher customer demand is due to higher Yelp ratings.

Table 8. Consumer Demand

Variables	Full sample (1)	Cutoff = 3.75 (2)	Cutoff = 4.25 (3)
$I_{round\ up}$	0.0035*** (0.0002)	0.0007*** (0.0002)	0.0051*** (0.0004)
Firm level control	✓	✓	✓
Cutoff fixed effects	✓		
Firm fixed effects	✓	✓	✓
County × year fixed effects	✓	✓	✓
Observations	40,198,897	8,833,270	9,742,833

Notes. This table reports Poisson regression results for consumer demand and Yelp ratings. The unit of observation is at the firm-month level. The dependent variable is the *number of Yelp reviews in the following month*. $I_{round\ up}$ is an indicator variable that equals one if the Yelp rating is rounded up and zero if rounded down. Column (1) utilizes the full sample with all cutoffs. Columns (2) and (3) examine cases in which the cutoff equals 3.75 and 4.25, respectively. Firm level control includes *lag(number of reviews)* and second-order polynomials. The variables are defined in Table 2. Robust standard errors clustered by Yelp industry and county are in parentheses.

***, **, and * denote statistical significance at the 1%, 5% and 10% level, respectively.

4.2. Subsequent Business Opening

I next explore the relationship between investment decisions by small businesses and Yelp ratings. More specifically, I examine whether businesses with high Yelp ratings are more likely to open a second location. The purpose of this test is twofold. First, subsequent openings serve as an alternative indicator of whether Yelp ratings determine and convey business prospects. Second and more importantly, nobody knows more about a particular business than the owner. Owners only open subsequent stores when they believe their current operations are profitable and full of potential. Consequently, the prediction of Yelp ratings coincides with the businesses' actual behavior as measured by opening subsequent stores.

To carry out this test, I construct a business-year panel. I identify whether the businesses' subsequent locations share the same business name, Yelp industry, and county. I use the first review date of those subsequent locations on Yelp as the subsequent opening date. Given that an opening is a quasi-absorbing state (once it is open, it is open for some time), I start by examining businesses with one location and then drop them from the analysis once they open a new establishment. I create the *subsequent business opening* variable as an indicator variable that equals one if an existing business opens another location in the following year and zero otherwise. Yelp ratings are measured at the end of each year.

I follow a similar approach as the main analysis and report the results in Table 9. I use the full sample and

include cutoff bin fixed effects in column (1). I document that a one-half-star increase in Yelp ratings leads to a statistically significant higher probability of subsequent business openings. This result is also economically significant. The average business falling just above the rounding thresholds is 0.09 percentage points more likely to open a subsequent location compared with the average business falling just below the thresholds, representing an 8% increase over the unconditional mean. In columns (2) and (3), I again focus on the two rounding cutoffs that matter the most, namely, 3.75 and 4.25, respectively. Businesses rounded up to 4 stars are 0.18 percentage points more likely to open a new location compared with those rounded down to 3.5 stars. Similarly, compared with businesses rounded down to 4 stars, businesses rounded up to 4.5 stars have a 0.09-percentage-point higher probability of opening another location.²²

Considering the changes in consumer demand and subsequent business opening results together, I document that Yelp ratings are significant indicators of businesses' future performance and investment decisions. Furthermore, Yelp ratings are nontrivial from both the consumer and business perspectives. In sum, the results support the argument that recommendations have real effects, confirming the theoretical findings by Iyer and Manso (2023).

5. Concluding Remarks

In this paper, I investigate the financial and real consequences of online review aggregators using coarse Yelp ratings in the SBA 7(a) loan lending process. Aggregated Yelp ratings provide reliable information about the borrowers' future earnings potential. Lenders reference Yelp ratings when making lending decisions. Higher future revenue resulting from consumers' social learning behavior, combined with better loan terms, may lead to better loan performance.

Using an RDD empirical setting that exploits the Yelp rating rounding thresholds, I show that coarse Yelp ratings are good indicators of loan outcomes. More specifically, a one-half-star increase in Yelp ratings leads to 9% savings in loan pricing, a 6% reduction in required collateral, and a 45% lower default probability. In the cross-section, my findings are stronger when businesses are located farther from the banks. I study repeated loan transactions and show that the overall effects of Yelp ratings remain powerful in the presence of existing lending relationships. I also document the real effects of Yelp ratings. Higher Yelp ratings lead to a higher number of future Yelp reviews and an increase in the likelihood of opening new locations.

The identification strategy in the paper requires the use of coarse ratings, which exploits the unique design

Table 9. Subsequent Business Opening

Variables	Full sample (1)	Cutoff = 3.75 (2)	Cutoff = 4.25 (3)
$I_{\text{round up}}$	0.0009*** (0.0002)	0.0018*** (0.0004)	0.0009** (0.0004)
Firm level control	✓	✓	✓
Cutoff fixed effects	✓		
Firm fixed effects	✓	✓	✓
County × year fixed effects	✓	✓	✓
Observations	2,040,152	388,716	461,232
Adjusted R^2	0.1790	0.1642	0.1505

Notes. This table reports OLS regression results for subsequent business openings and Yelp ratings. The unit of observation is at the firm-year level. The dependent variable is a dummy variable that equals one if an existing business on Yelp opens another location under the same name in the same county and Yelp industry in the following year and zero otherwise; $I_{\text{round up}}$ is an indicator variable that equals one if the Yelp rating is rounded up and zero if rounded down. Column (1) utilizes the full sample with all cutoffs. Columns (2) and (3) examine cases in which the cutoff equals 3.75 and 4.25, respectively. Firm level control includes $\log(\text{number of reviews})$. The variables are defined in Table 2. Robust standard errors clustered by Yelp industry and county are in parentheses.

***, **, and * denote statistical significance at the 1%, 5% and 10% level, respectively.

feature of Yelp. However, it is a valid concern whether the results are only relevant to Yelp. A major competitor to Yelp ratings is Google reviews. However, Google rounds its star ratings to one decimal place and displays the numerical scores. As a result, Google reviews do not have the coarse nature that Yelp ratings do. I further discuss the generalizability of Yelp ratings in Online Appendix Q.

Overall, my results shed light on the importance of online review aggregators in the determination of financial and real outcomes. As technology advances rapidly, such resources only become more critical, providing easy access to this type of knowledge about a business. Future research could extend this analysis to other situations and incorporate other useful content generated by internet users.

Acknowledgments

The author thanks Gustavo Manso (the editor), an associate editor, two anonymous referees, Rajesh Aggarwal, Heitor Almeida, Allen Berger, Bo Bian, James Brau, Sudipto Dasgupta, David Denis, Nuri Ersahin, Umit Gurun, John Hackney, Gerald Hoberg, Jiekun Huang, Timothy Johnson, Charles Kahn, José Liberti, Chen Lin, James Linck, Baixiao (Tony) Liu, Ping Liu, Erik Mayer, George Pennacchi, Joshua Pollet, Tarun Ramadorai, Jay Ritter, Anthony Saunders, David Skeie, Philip Strahan, Scott Weisbenner, Yufeng Wu, Yuhai Xuan, and Mao Ye as well as seminar participants at Auburn University, Cheung Kong Graduate School of Business, Chinese University of Hong Kong, Florida State University, Fordham University, Hong Kong University, Lehigh University, Northeastern University, Oklahoma State University, Queens University, San Diego State University, Southern Methodist University, Temple University, University at Buffalo, University of Central Florida, University of Florida, University of Illinois at Urbana-Champaign, University of Pittsburgh, University of South Carolina, University of Sydney, University of Texas at Dallas, University of Utah, the 2018 London Business School Transatlantic Doctoral Conference, the 2018 Financial Intermediation Research Society annual meeting, and the 2018 Northern Finance Association annual meeting for helpful discussions and comments.

Endnotes

¹ The information submitted to lenders is questionable because of the lack of formal audits, rendering financial documents unreliable. In contrast, online reviews could be considered forward-looking performance measures, hence providing valuable information to lenders.

² Petersen and Rajan (2002) show that the lender–borrower distance has increased over time, primarily because of technological and communications advancements. I show that Yelp ratings could be part of the explanation by making information acquisition easier.

³ Brau and Osteryoung (2001) study the SBA sample of the National Survey of Small Business Finances and provide descriptive statistics of the borrowing firms and loan characteristics. Levine et al. (2020) study the impact of within-banking-organization communication on small business lending using Community Reinvestment Act data.

⁴ For the RDD framework to be valid, the treatment must satisfy the ignorability assumption, that is, it must be randomly assigned conditional on observables. In the RDD setting at hand, this assumption is satisfied trivially. More specifically, when the actual rating is above (below) the cutoff, the treatment takes the value of one (zero). Conditional on the actual rating, there is no variation left in the treatment dummy, making it impossible to be correlated with any other factor. Nevertheless, I plot the covariates, lag log(number of reviews) and price range, in Online Appendix A and show that they are indeed continuous around the cutoffs.

⁵ The interaction between the treatment indicator and the distance of the assignment variable to the cutoff allows for different slopes and intercepts on either side of the cutoff because there is no strong reason to constrain the functional form on both sides of the cutoff to be the same.

⁶ It is crucial to ensure that the ratings I calculate from my data match the actual ratings displayed on Yelp, especially the ones right around the rounding thresholds. To provide a validation test, I focus on ratings within 0.05 of the rounding thresholds. I assign a star rating to each business based on the imputed average rating following Yelp's rounding rule. For example, for a cutoff of 3.75, I focus on ratings from 3.70 to 3.79. I assign the ones below 3.75 a rating of 3.5 stars and the ones equal to or above 3.75 a rating of 4 stars. When comparing the imputed star ratings to the actual star ratings posted on Yelp at the time of data collection, I find fewer than 0.1% of the businesses display any difference in ratings. The numbers of businesses with false lower versus higher ratings are very similar.

⁷ The preferred lender advantage is clearly stated on the SBA website and also widely reported in the news media. Borrowers are encouraged to work with those lenders, primarily banks that frequently underwrite SBA loans.

⁸ For loans of \$350,000 or less, the guarantee fee is 0% regardless of loan maturity as of October 1, 2021.

⁹ In the final sample, some businesses appear more than once because they subsequently apply for additional SBA loans. As shown in Online Appendix E, the results are similar if I only study the first loan for each business in the sample.

¹⁰ In Online Appendix F, I compare the sample loan characteristics with the universe of SBA loans. Collateral, maturity, default indicator, and write-off amount are very similar between the sample and all SBA loans. The only difference is the loan amount. The loans in the sample are about \$23,000 smaller, on average, compared with all SBA loans. In addition, I report the average loan characteristics from all bank loans reported in the Federal Reserve Board's Survey of Terms of Business Lending.

¹¹ Assuming the small business owner's take-home pay is \$100,000 per year with a marginal tax rate of 20%, the interest expense savings for the average loan because of a half-star increase in Yelp ratings translates to roughly 0.5% of the owner's take-home pay.

¹² Gelman and Imbens (2019) suggest using estimators based on a linear or quadratic polynomial. I adopt linear functions on either side of the cutoffs throughout the analysis. Following Lee and Lemieux (2010), I show that the results are robust to the inclusion of polynomials up to the fourth order in Online Appendix G. In Online Appendix H, I include county \times year fixed effects and cluster the standard errors at the Yelp industry and county level. I also saturate the model with year–month fixed effects and two-way cluster on year–month and county. The results are consistent when using those alternative specifications.

¹³ More specifically, columns (2) and (3) of Table 3, panel A, use the 3.75 and 4.25 cutoffs, respectively, neither of which suffers from the potential survivorship bias. The point estimates suggest that the effect of a one-half-star increase in Yelp ratings is around a 30-basis-point reduction in loan spread for those two cutoffs. If the effect of rounding

up is the same across all star levels, then a one-half-star increase in Yelp ratings should lead to a 30-bps reduction in loan spread across the board. If this were the case, the reduction in loan spread could be as much as 5 bps higher than the 25-bps estimate currently presented in column (1). The same analogy could be applied to the collateral analysis.

¹⁴ In Online Appendix I, I show that the results are robust to adding employment and credit score as control variables, using data from Reference USA. In the same test, when studying loan performance, I further include loan terms, that is, loan spread and collateral, as control variables. This allows me to control for the effects of better financing terms on loan outcomes. In Online Appendix J, following Roberts and Whited (2013), I run placebo tests using different cut-offs to further test the validity of the RDD setting.

¹⁵ Similar arguments are made in the credit rating literature. Credit ratings are particularly helpful to less-informed lenders in the syndicated loan market (Sufi 2009). Furthermore, Sufi (2007) shows that lenders participating in syndicates are geographically closer to the borrowing firm and have previous relationships with the borrower when the borrower has no credit rating available to reduce the need for information gathering.

¹⁶ In Online Appendix L, I focus only on businesses with more than one SBA loan to explore the heterogeneity in lending banks. I define a same bank indicator, which equals one if the subsequent loan is taken out with the same bank and zero otherwise. I also include the interaction term of the same bank indicator and $I_{round\ up}$ to capture the effect of a same-bank lending relationship. I find similar results.

¹⁷ One potential concern with examining the distance to the lender and prior banking relationships is that these variables could vary by county population, capturing the effect of Yelp in urban versus rural areas. In Online Appendix M, I control for county-level population interacted with $I_{round\ up}$; the results are similar.

¹⁸ Ideally, the variation in information would be uncorrelated with the fundamentals of the business. Otherwise, the cross-sectional heterogeneity found in spreads and collateral could be explained by cross-sectional heterogeneity in how Yelp ratings affect the fundamentals of the firm. One potential concern with the statistically significant interaction terms in columns (3) and (4) of Table 5 is that the distance to the lender and prior banking relationship variables could be correlated with fundamentals (proxied by default), which could muddy the interpretation of the spread and collateral results. As a remedy, I examine business closings. In Online Appendix N, I show that a one-half-star increase in Yelp ratings is negatively associated with the probability of business closure. However, the interaction terms with distance to lender and prior banking relationship are neither statistically nor economically significantly related to business closing. These results support the assumption that the closest SBA lender and a prior banking relationship are not correlated with the effect of Yelp rating on business fundamentals.

¹⁹ In Online Appendix O, I run weighted OLS regressions using the number of reviews as the weights. In Online Appendix P, I split the sample at the sample median. I find similar results in those robustness tests, indicating that a higher number of reviews offers a more precise signal.

²⁰ Anecdotally, the SBA and Yelp have partnered to help small businesses succeed with online reviews (see <https://www.sba.gov/about-sba/sba-initiatives/sba-and-yelp-present-success-online-reviews>, last retrieved July 15, 2017). Similarly, many other major financial institutions, such as J.P. Morgan Chase, offer advice on their websites to educate small business owners about social media (see <https://www.chase.com/news/111416-social-media>, last retrieved July 15, 2017). Mainstream news outlets, such as Forbes and the BBC, have also published articles about social media and the ability to qualify for bank loans (see <http://www.bbc.com/news/business-37224847> and

<https://www.forbes.com/sites/chynes/2017/04/25/how-data-will-help-drive-universal-financial-access/#1466f1eb57e6>, last retrieved July 15, 2017).

²¹ This analysis focuses on PLP and express loans because the lender makes the lending decisions on those loans instead of the SBA. PLP and express loans comprise about 78% of the sample.

²² One potential concern with this test is franchises. Even within a county, the franchises could be owned and operated by different franchisees. Those cases could potentially contaminate the test. I retrieve a list of franchises from <https://www.franchising.com/franchises> (last retrieved June 1, 2020) and exclude them from the data. I find similar results. However, that list of franchises is unlikely to be exhaustive.

References

- Agarwal S, Hauswald R (2010) Distance and private information in lending. *Rev. Financial Stud.* 23(7):2757–2788.
- Allen F (1984) Reputation and product quality. *RAND J. Econom.* 15(3):311–327.
- Anderson M, Magruder J (2012) Learning from the crowd: Regression discontinuity estimates of the effects of an online review database. *Econom. J. (London)* 122(563):957–989.
- Berger AN, Udell GF (1995) Relationship lending and lines of credit in small firm finance. *J. Bus.* 68(3):351–381.
- Brau JC, Osteryoung JS (2001) An empirical examination of SBA guaranteed loans: Rates, collateral, agency costs, and the time to obtain the loan. *J. Entrepreneurial Finance* 6(1):1–23.
- Brown JD, Earle JS (2017) Finance and growth at the firm level: Evidence from SBA loans. *J. Finance* 72(3):1039–1080.
- Cai H, Chen Y, Fang H (2009) Observational learning: Evidence from a randomized natural field experiment. *Amer. Econom. Rev.* 99(3): 864–882.
- Chevalier JA, Mayzlin D (2006) The effect of word of mouth on sales: Online book reviews. *J. Marketing Res.* 43(3):345–354.
- Craig BR, Jackson WE, Thomson JB (2005) SBA-loan guarantees and local economic growth. Preprint, submitted May 3, <https://dx.doi.org/10.2139/ssrn.712622>.
- Degryse H, Ongena S (2005) Distance, lending relationships, and competition. *J. Finance* 60(1):231–266.
- Diamond DW (1991) Monitoring and reputation: The choice between bank loans and directly placed debt. *J. Political Econom.* 99(4): 689–721.
- Duarte J, Siegel S, Young L (2012) Trust and credit: The role of appearance in peer-to-peer lending. *Rev. Financial Stud.* 25(8): 2455–2484.
- Gelman A, Imbens G (2019) Why high-order polynomials should not be used in regression discontinuity designs. *J. Bus. Econom. Statist.* 37(3):447–456.
- Hackney J (2023) Small business lending in financial crises: The role of government guaranteed loans. *Rev. Finance* 27(1):247–287.
- Iyer G, Manso G (2023) Recommendations with feedback. *Rev. Financial Stud.* 36(2):501–533.
- Iyer R, Khwaja AI, Luttmer EF, Shue K (2016) Screening peers softly: Inferring the quality of small borrowers. *Management Sci.* 62(6): 1554–1577.
- Lee DS, Lemieux T (2010) Regression discontinuity designs in economics. *J. Econom. Literature* 48(2):281–355.
- Levine R, Lin C, Peng Q, Xie W (2020) Communication within banking organizations and small business lending. *Rev. Financial Stud.* 33(12):5750–5783.
- Luca M (2016) Reviews, reputation, and revenue: The case of Yelp. com. Harvard Business School Working Paper No. 12-016, Harvard University, Boston.
- Manso G (2013) Feedback effects of credit ratings. *J. Financial Econom.* 109(2):535–548.

- McCrary J (2008) Manipulation of the running variable in the regression discontinuity design: A density test. *J. Econometrics* 142(2): 698–714.
- Moretti E (2011) Social learning and peer effects in consumption: Evidence from movie sales. *Rev. Econom. Stud.* 78(1):356–393.
- Petersen MA, Rajan RG (1994) The benefits of lending relationships: Evidence from small business data. *J. Finance* 49(1):3–37.
- Petersen MA, Rajan RG (2002) Does distance still matter? The information revolution in small business lending. *J. Finance* 57(6): 2533–2570.
- Ravina E (2019) Love & loans: The effect of beauty and personal characteristics in credit markets. Preprint, submitted March 6, 2008, <https://dx.doi.org/10.2139/ssrn.1107307>.
- Roberts MR, Whited TM (2013) Endogeneity in empirical corporate finance. Constantinides GM, Harris M, Stulz RM, eds. *Handbook of the Economics of Finance*, vol. 2, Part A (Elsevier, Amsterdam), 493–572.
- Sufi A (2007) Information asymmetry and financing arrangements: Evidence from syndicated loans. *J. Finance* 62(2):629–668.
- Sufi A (2009) The real effects of debt certification: Evidence from the introduction of bank loan ratings. *Rev. Financial Stud.* 22(4):1659–1691.

Copyright 2025, by INFORMS, all rights reserved. Copyright of Management Science is the property of INFORMS: Institute for Operations Research and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.