NATIONAL RESEARCH UNIVERSITY
HIGHER SCHOOL OF ECONOMICS

Elena B. Pokryshevskaya,
Evgeny A. Antipov

ROBUST REGRESSION DISCONTINUITY ESTIMATES OF THE CAUSAL EFFECT OF THE TRIPADVISOR'S BUBBLE RATING ON HOTEL POPULARITY

BASIC RESEARCH PROGRAM

WORKING PAPERS

SERIES: MANAGEMENT

WP BRP 63/MAN/2020

Elena B. Pokryshevskaya[1],

Evgeny A. Antipov[2]

ROBUST REGRESSION DISCONTINUITY ESTIMATES OF THE CAUSAL EFFECT OF THE TRIPADVISOR'S BUBBLE RATING ON HOTEL POPULARITY[3]

In this paper we use detailed data on 4,599 hotels located in Rome collected from TripAdvisor, the world's largest travel platform, to examine the causal effects of bubble ratings (detailed to half-bubbles) on hotel popularity measured with the number of people viewing the hotel's page. By using a regression discontinuity design, we find that bubble presentation of ratings does not create any significant jumps at cutoffs. This result is different from those obtained in previous studies of similarly designed rating systems from other industries. Another finding is that web users tend to shortlist hotels with the bubble rating of at least 3. Despite that, there is no strong evidence of review manipulation around the 2.75 cutoff to make a transition from the 2.5-bubble rating to the 3-bubble rating. Potential uses of the number of views as a proxy of demand in hospitality and tourism research are outlined.

JEL Classification: L83, M31

Keywords: regression discontinuity, ratings, sales, booking, hotel reviews, TripAdvisor

---

[1] HSE University. Saint-Petersburg School of Economics and Management. Associate Professor. E-mail: epokryshevskaya@hse.ru

[2] HSE University. Saint-Petersburg School of Economics and Management. Associate Professor. E-mail: eantipov@hse.ru

[3] This Working Paper is an output of a research project implemented within the framework of the Academic Fund Program at HSE University in 2019 — 2020 (grant №19-01-070) and within the framework of the Russian Academic Excellence Project «5-100».Any opinions or claims contained in this Working Paper do not necessarily reflect the views of HSE.

1. Introduction

According to an Independent Phocuswright study of 14,991 global respondents prepared for TripAdvisor in April 2015 the majority of customers said TripAdvisor bubble ratings based on travelers' feedback were important when choosing an accommodation, restaurant, or attraction: 83% believed it is significant when choosing an accommodation, 70% when choosing a restaurant, and 58% when deciding what attractions to do[5]. The effect of consumer reviews on sales has been widely studied and the importance of the word-of-mouth has been shown for many industries (Floyd et al., 2014), including the hospitality industry (Phillips et al., 2017; Torres et al., 2015). However, in fact existing studies have merely shown the usefulness of ratings as predictors of hotel performance because user ratings are inevitably correlated with unobserved hotel quality and word-of-mouth sentiments, which causes omitted variable bias. As a result, existing estimates of the sensitivity of hotel performance to changes in user ratings themselves cannot be claimed to be causal effects of ratings that would ideally be obtained in an experiment where otherwise identical hotels were randomly assigned slightly different ratings. Not accidentally, the meta-analysis of Yang et al. (2018) has shown that the link between electronic word of mouth (eWOM) and hotel performance was significantly lower in panel data studies which are known to capture at least the time-invariant portion of unobserved heterogeneity of hotels.

There have been only a few studies employing rigorous quasiexperimental causal analysis techniques to identify distorting effects of systems reporting rounded ratings (usually depicted using stars, bubbles, or other graphical images) and none of them was related to hotels. Anderson and Magruder (2012) employed a quasi-experimental design (regression discontinuity) to estimate the effect of average Yelp.com ratings on restaurant reservations in San Francisco. They found that a half-star increase in rating results in a 19 percentage points increase in the probability of selling out during prime dining times, and this effect is even larger for restaurants that have external accreditation. Their results provide evidence on the importance of aggregate consumer reviews in informing restaurant quality. In line with Anderson and Magruder (2012), Luca (2016) used Yelp.com ratings and data from the Washington State of Revenue to associate average ratings with restaurant revenue. Luca (2016) found that an increase of one-star in the average rating in Yelp leads to an increase in revenue of about 5-9 percent for restaurants. Other studies corroborate these findings for many other products. For example, Chevalier and Mayzlin (2006) found a qualitatively similar result for book sales on Amazon and Barnes and Noble, and Zhu and Zhang (2010) found that online reviews are more influential for less popular games and games whose players have greater Internet experience. However, in contrast, Duan et al. (Duan et al., 2008) did not find a significant influence of the average rating on movies' box office revenues. Even more interestingly, according to a recent regression discontinuity study of online consumer reviews the star presentation can create negative, rather than positive, jumps at cutoffs (Wang et al., 2019). The authors provide the following reasoning. Consumers restrict their attention to a star category resulting in the "best" sellers in a lower star category being better off than the "worst" sellers in a higher star category. The incentive for review manipulation is strongly reduced, which in the long run will create trust and confidence for the review system as well as the sellers. For those sellers that are just below the cutoffs, simply crossing over the cutoffs would not lead to higher sales, unless they substantially improve their service quality to attract consumers.

[5] https://www.TripAdvisor/TripAdvisorInsights/w810

In this study we seek to answer whether an increase in a hotel's bubble rating at TripAdvisor *causes* an increase in the hotel's popularity. We believe that the effect of user ratings on hotel popularity should be weaker or, in many cases, even non-existent compared to restaurants (Anderson and Magruder, 2012; Luca and Zervas, 2016), because of the role of an alternative "official" categorization of hotels into "five-star hotels", "four-star hotels", etc. As a result an (almost) exogeneous small difference in average user ratings giving one hotel a 0.5 user rating's bubble advantage should not lead to a substantial increase in one hotel's popularity compared to the other's if they both have comparable underlying quality of reviews, locations and belong to the same "official" class.

Despite non-disclosure of actual booking data, data from TripAdvisor provides settings where causal effects of changes in bubble ratings on hotel popularity can be estimated:

- TripAdvisor reports ratings rounded to the nearest half-bubble. This allows utilizing these rounding rules and comparing hotels in the vicinity of each threshold. Because of this rounding, two nearly identically rated products could have different displayed ratings if they lie on opposing sides of a rounding threshold.
- TripAdvisor provides the number of people viewing each hotel's page at the current moment, which can serve as a good proxy for the hotel's popularity in the absence of the number of daily bookings.

2. Data

Data on 4,599 hotels located in Rome, Italy were collected in December 2019. They represent all hotels found on TripAdvisor in that region except for a small number of cases for which some of the key information was missing. The choice was driven by the fact that Rome is one of the most popular destinations and has one of the largest number of hotels among all cities in the world according to TripAdvisor. The causal inference technique employed in our study favors larger samples to provide enough instances in the vicinity of thresholds used to assign hotels with bubble ratings.

Dependent variable *views* is the number of people viewing the hotel's page at the time of data collection, i.e. its popularity at a given point in time. Such search data has limitations which are common to the data from search platforms (Brynjolfsson et al., 2010): we observe page views instead of actual bookings, and not every click results in a booking (Koulayev, 2010). Following Smith and Brynjolfsson (Smith and Brynjolfsson, 2001) and Koulayev (2010) we assume that a click is a revealed preference action. Smith and Brynjolfsson (2002) analyzed data from a book comparison website and had both click and sales data. They show that the click-to-buy ratio is rather homogeneous across merchants. Therefore, we assume that even if no booking is eventually made, the desire to gather more information about a particular hotel is indicative of preferences for that hotel.

Although the actual number of bookings is not available, the proxy used in our study ensures that all variables involved in our study are measured for the same point in time without any lag. In addition, the data collection was designed in such a way that differences in the time of data collection across hotels were negligible. In a pilot study conducted in May 2019 we revealed high internal consistency among 3 repeated measurements of the number of views for a sample of 582 hotels located in Rome (Cronbach's alpha=0.91). This preliminary result justifies the sufficiency of cross-sectional data (one-time measurement), which is important considering that TripAdvisor's policy discourages automated data collection. As zero number of views prevails in the sample, we also use an alternative binary dependent variable *view_binary* that equals 1 if there were any views and 0 otherwise.

The running variable *bubble_rating* varies from 1 to 5 with increments of 0.5. The running variable, underlying the rounded bubble rating, is *score*, which was computed as the exact average rating based on the number of 5-,4-,3-,2- and 1-bubble ratings. We identified only nine cases out of 4,599 where the bubble rating did not match the rounded score exactly, most likely due to some minor lags in the server-side recalculation of bubble ratings after recent reviews have been posted. For such hotels the assignment of bubble rating was especially close to being randomized at the time when the number of viewers was recorded. For these cases we adjusted the score so that it deviates from the threshold just by 0.01 point in the direction pointed out by the bubble rating (e.g., if bubble rating was 4, *score*=4.25 was changed to *score_adjusted*=4.24). The resulting adjustments were negligible, while having allowed us not to remove these valuable observations clearly lying in the vicinity of the thresholds and to use the sharp regression discontinuity design.

Our dataset also contains a rich set of covariates, which were used to test the similarity of the treatment and the control groups around cutoffs. These control variables were chosen to comprehensively reflect all information available to users as part of search filters and/or observable in search results before they choose pages of which hotels to visit for more details:

- *category_hotel*: 1 if the property belongs to the "Hotels in Rome" category
- *category_inn*: 1 if the property belongs to the "B&Bs / Inns in Rome" category
- *category_specialty*: 1 if the property belongs to the "Specialty Lodging in Rome" category.
- *class*: factor variable with 6 levels depending on the hotel's star category (no stars, 1 star, 2 stars, 3 stars, 4 stars, 5 stars), a third-party indicator disclosed by TripAdvisor and based on the property's available facilities, staff, and amenities
- *class_4_5*: 1 if class is "4 stars" or "5 stars", 0 – otherwise
- *class_3_4_5*: 1 if class is "3 stars", "4 stars" or "5 stars", 0 – otherwise
- *n_reviews*: numeric variable reflecting the total number of reviews.
- *location_grade*: TripAdvisor's measure of location quality (from 0 to 100) based on the number of restaurants and attractions located in the walking distance of the hotel
- *discount*: 1 if the standard price was struck through on the website and a price reduction was available, 0 - otherwise
- *discount_perc*: discount size (%)
- *price_curr_min*: minimum current price (RUB)
- *price_min:* minimum standard price for accommodation of 2 adults (RUB)
- *price_max*: maximum standard price for accommodation of 2 adults (RUB)
- *award_travellers_choice*: 1 if the hotel has the Travellers Choice Award 2019, 0 –otherwise
- *award_greenleaders*: 1 if the hotel has any Greenleaders Awards, 0 – otherwise
- *award_cert_excellence*: 1 if the hotel has the Certificate of Excellence, 0 – otherwise
- *photos*: number of photos on the hotel's page

3. Methods

We employ sharp regression discontinuity design (RDD) - a rigorous method for causal analysis, applications of which have received a lot of attention as being close to those from truly randomized experiments (Hill et al., 2017), but are still rare in marketing literature because of the lack of sources of exogeneous variation. Well-executed RDD studies can result in treatment effect estimates close to those from randomized studies that are generally not feasible in the context of online ratings. The regression discontinuity (RD) approach with application to the estimation of the impact of online ratings on demand is based on the idea that we can compare products that are characterized by similar continuous scores but different bubble ratings to isolate the impact of the bubble rating. For instance, two hotels with average ratings of 3.74 and 3.75 will be presented as 3.5- and 4-bubble hotels, respectively. Such hotels near the rounding threshold are likely to be, on average, almost identical except for their rounded bubble ratings. Hence, the causal impact of the bubble ratings can be obtained by comparing demand for hotels marginally above and marginally below the rounding threshold.

In order to detect whether there is evidence of self-selection (manipulation) around a cutoff where a bubble rating changes by a half-bubble we employ a manipulation testing procedure using the local second-order polynomial density estimator proposed in Cattaneo, Jansson and Ma (2019), a robust data-driven density test of falsification (McCrary, 2008). The method uses local quadratic approximation to construct the density point estimator. If the density below cutoff were considerably different from the one above, it would indicate that hotels have a possibility to manipulate their scores. For robustness check, we used two bandwidths: the data-driven optimal bandwidth automatically selected for each cutoff and the bandwidth of 0.1, which we find to be managerially reasonable as it ensures that samples used to estimate the density at each cutoff are of sufficient size and are very close to the threshold. Other parameters were set to default levels recommended by the procedure's developers: triangular kernel (ensures that observations closer to the cutoff are weighted substantially heavier than others), unrestricted density estimation, and jackknife robust variance-covariance matrix estimation (Cattaneo et al., 2019).

Regression discontinuity estimates of treatment effects were obtained for each cutoff of the *score_adjusted* using the optimal automatically selected bandwidth (Calonico et al., 2020) as well as the fixed bandwidth of 0.1. The estimation was conducted with a triangular kernel using a local-linear (order p=1) estimator with a local-quadratic (order q=2) bias-correction estimate. Robust standard error estimators were computed using 3 nearest-neighbors (Calonico et al., 2015). P-values associated with robust local-polynomial RD estimates are reported. Since the validity of the regression discontinuity design relies on those that were just barely treated (just above the cutoff) being the same as those who were just barely not treated (just below the cutoff), we examined the differences not only in the dependent variables, but also in all available covariates as well. Although some variables may differ for the two groups merely based on random chance, most of these covariates should be the same for the RDD analysis to be trustworthy.

4. Results

While 4.1% of hotels with the bubble rating of 3 or higher had non-zero views, only 0.3% of hotels with a lower rating had any views. Most views are thus concentrated in hotels with bubble ratings of 3 or higher, which is itself an interesting observation, implying that visitors avoid considering hotels rated lower than 3 bubbles out of five almost at all. The average number of views (mean of variable *views*) as well as the probability of being viewed (mean of variable *views_binary*) increases as the bubble rating goes up until rating 4.5 and then decreases for those having the bubble rating of 5. This decrease is possibly because good ratings are trustworthy only when they come along with a high number of reviews (Gavilan et al., 2018). In addition, hotels rated with 5 bubbles have somewhat higher prices, lower discounts, higher proportion of specialty lodging as opposed to traditional hotels or B&Bs/Inns, as well as a lower proportion of hotels with the certificate of excellence and those officially rated as belonging to the 3-5 star categories. A possible explanation for a smaller number of views/reviews for properties having 5 bubble ratings is that they are relatively more often special lodgings or B&Bs/Inns that can accommodate a relatively small number of people, which is why they are not available for booking on many dates unlike larger hotels thus generating a smaller number of views. Therefore, it is especially important to check covariate balance between treatment and control groups around the 4.75 cutoff.

While summary statistics by bubble rating (Table 1) indicate that the popularity of hotels having different bubble ratings tends to agree with their quality, our focus is on whether a half-bubble increase in consumer rating itself impacts the number (or probability) of views. To answer this question we will use regression discontinuity analysis procedures to check whether there are any discontinuities in the relationship between *score_adjusted* and *views*, as well as *score_adjusted* and *views_binary* that would indicate benefits or losses associated with getting a half-star higher *bubble_rating* while keeping other things equal.

Table 1. Summary of variables by bubble rating

		Bubble rating								
		1	**1.5**	**2**	**2.5**	**3**	**3.5**	**4**	**4.5**	**5**
Group size	**N**	65	38	105	169	312	538	992	1467	913
views	**Mean**	0.000	0.000	0.029	0.000	0.026	0.061	0.348	0.408	0.301
views_binary	**Mean**	0.000	0.000	0.010	0.000	0.006	0.011	0.060	0.054	0.026
score_adjusted	**Mean**	1.000	1.527	2.003	2.507	3.027	3.512	4.016	4.507	4.905
category_hotel	**Mean**	0.092	0.105	0.229	0.249	0.237	0.312	0.293	0.148	0.053
category_inn	**Mean**	0.323	0.526	0.476	0.527	0.522	0.524	0.550	0.707	0.702
category_specialty	**Mean**	0.585	0.368	0.295	0.225	0.240	0.164	0.156	0.145	0.245
class_4_5	**Mean**	0.000	0.026	0.057	0.053	0.074	0.147	0.161	0.132	0.047
class_3_4_5	**Mean**	0.000	0.079	0.152	0.160	0.221	0.349	0.391	0.367	0.143
n_reviews	**Mean**	1.7	17.3	43.9	79.7	72.9	175.2	249.1	213.4	106.0
location_grade	**Mean**	92.951	92.611	93.763	95.038	91.976	92.071	93.867	95.341	92.177
discount	**Mean**	0.000	0.026	0.038	0.071	0.071	0.158	0.146	0.166	0.108
discount_perc	**Mean**	0.000	1.037	0.368	1.360	1.505	2.781	2.160	2.473	1.368
price_curr_min	**Mean**	3801	2329	3290	3207	3697	4047	4569	5434	6001
price_min	**Mean**	4238	3800	3808	3907	3975	4129	4804	5459	5821
price_max	**Mean**	11083	8337	10511	11819	10185	10954	12638	13677	13223
award_travellers_choice	**Mean**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.011
award_greenleaders	**Mean**	0.000	0.000	0.000	0.000	0.003	0.002	0.016	0.015	0.010
award_cert_excellence	**Mean**	0.000	0.000	0.000	0.000	0.000	0.000	0.128	0.241	0.189
photos	**Mean**	7.1	20.0	40.4	53.4	57.1	111.7	157.2	138.4	77.2

74% of Roman hotels have a rating of 4 bubbles or higher. Such skewness of customer evaluations is typical of many platforms with online reviews (Hu et al., 2009). Given this skewness and the fact that hotels rated lower than 3 bubbles received almost no attention at all, we will concentrate on cutoff values of ratings around which sufficient variation in hotel popularity is available: 3.25, 3.75, 4.25, and 4.75. The density test also included 2.75 threshold associated with the transition from 2.5 to 3 bubbles to check for the manipulation around this threshold.

No systematic discontinuities in the density around thresholds have been detected by manipulation tests (Table 2): none of the differences in density estimates at the cutoff is statistically significant (all p-values>0.05), which supports the validity of the regression discontinuity design. It is not surprising as the site uses an algorithm to look at any "content integrity issues, animal welfare policies, and/or fraudulent activity" connected to the listing[6].

Table 2. RD manipulation test using local polynomial density estimation (robust p-values)

Cutoff	p-value for data-driven optimal bandwidth (bandwidth varies with cutoff)	p-value for constant bandwidth=0.1
2.75	0.0859	0.8912
3.25	0.411	0.0838
3.75	0.1078	0.8173
4.25	0.6514	0.2944
4.75	0.7921	0.7079

Similarity of treatment units to control units near the cutoffs has been checked and almost no significant differences in observable characteristics just above and just below the cutoffs have been found with a few exceptions. There is a significant spike in the probability of receiving the certificate of excellence when transitioning from 3.5 to 4 bubbles (by 0.078, or 7.8 percentage points, $p<0.05$) and the decrease of this probability when transitioning from 4.5 to 5 points (by 0.337, or 33.7 percentage points, $p<0.05$). While the sharp increase from 3.5 to 4 makes sense, as 4 bubble points is known to be the minimum eligibility rating used by TripAdvisor, a possible explanation for the negative effect of having 5 bubbles instead of 4.5 is as follows. While TripAdvisor does not disclose the exact algorithm for awarding hotels with this designation, it is known to be given to approximately 10% of all businesses on TripAdvisor that have consistently achieved great reviews over the past year. Most likely, the platform accounts not only for the average rating, but also for the number of reviews so that there is more trust in hotels with more reviews (Molinillo et al., 2016). We do not have the number of reviews over the last year at our disposal, but according to regression discontinuity analysis there is some weak evidence (under some bandwidth and estimation methods) that those whose rating is 4.75 or slightly higher have a significantly (at the 10% level) smaller number of reviews (and, thus, lower trust) than those whose rating is just below 4.75. More importantly, for the 4.75 cutoff there is strong evidence of a significantly lower proportion of hotels officially classified as 3-5 star hotels in the treatment group compared to the control group ($p<0.05$), which indirectly implies that TripAdvisor accounts for various reputational factors beyond the average consumer rating when designating organizations with their certificates.

[6] https://www.TripAdvisor/TripAdvisorInsights/w604

Table 3. Robust bias-corrected regression discontinuity estimates of the bubble rating's treatment effect for various cutoffs and dependent variables

Dependent variable	Cutoff	Data-driven optimal bandwidth (varies across dependent variables and cutoffs)		Constant bandwidth=0.1	
		Point Estimate	Robust p-value	Point Estimate	Robust p-value
views	3.25	0.131	0.163	0.007	0.809
views	3.75	-0.025	0.905	-1.231	0.342
views	4.25	-0.039	0.885	-0.189	0.579
views	4.75	-0.356	0.414	-0.174	0.756
views_binary	3.25	0.035	0.126	0.004	0.619
views_binary	3.75	0.008	0.743	-0.086	0.392
views_binary	4.25	0.014	0.711	0.018	0.710
views_binary	4.75	-0.027	0.527	-0.031	0.566

Significant jumps of neither the number of views (*views*) nor the probability of any views (*views_binary*) have been detected (Table 3). A plot illustrating the absence of significant jumps at one of the cutoffs is presented in Figure 1. It is clearly seen that in the vicinity of the 4.25 points cutoff the average number of views for the treatment and for the control groups almost mirror each other.

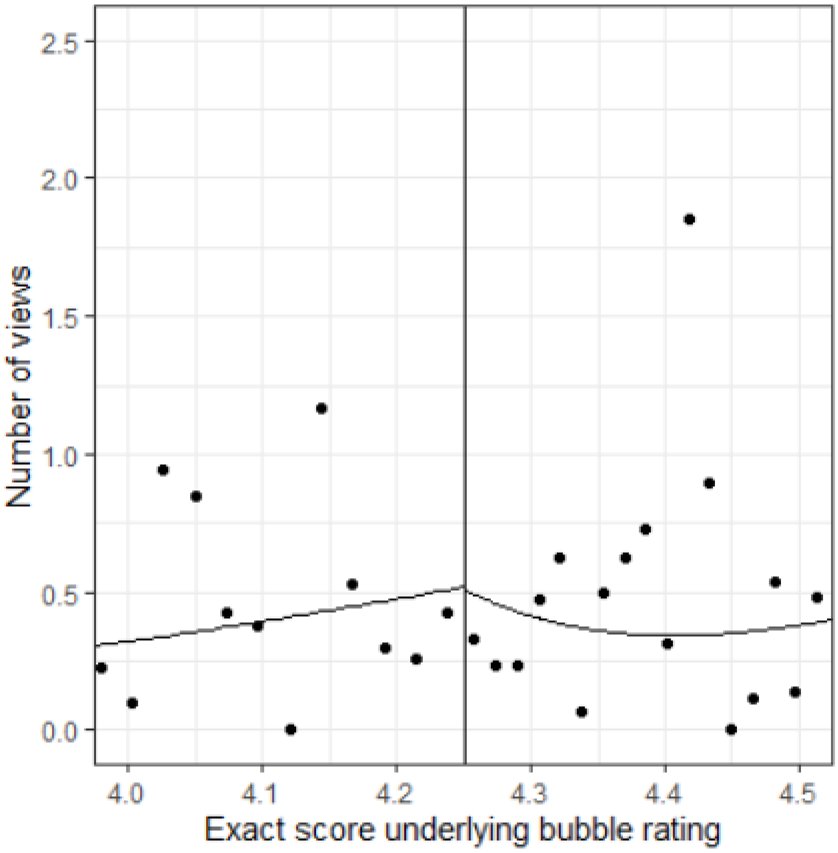

Figure 1. RDD plot (cutoff=4.25)

Even though it was previously shown that units just above and just below the cutoff are similar in most observable characteristics, we also conducted a series of robustness checks by rerunning the RD estimation with additional covariates (*class_3_4_5*, *n_reviews* and *price_curr_min* individually as well as in various combinations of 2 and 3 covariates). All estimates agreed with the insignificance of all treatment effects ($p>0.05$).

5. Conclusion

The number of people viewing a particular hotel is an indicator of the property's popularity, which is publicly available from TripAdvisor. Out study is the first to investigate the distribution of this indicator among hotels with various bubble ratings, as well as to figure out the impact of these bubble ratings themselves on hotel popularity. It turned out that Roman hotels with ratings lower than 3 receive almost no attention from TripAdvisor users, which agrees with previous findings that web users trust bad ratings more than good ratings and tend to shortlist hotels with better ratings (Gavilan et al., 2018). This is most likely because of the availability of a search filter, where users can limit search results to hotels with the bubble rating of not lower than some number of bubbles. At the same time, only very weak evidence of review manipulation has been detected using the density test. More specifically, there are no unusual statistically significant jumps (at the 5% or a lower significance level) in the density of scores to the right of the 2.75 threshold that would be expected if hotels tried to barely reach the 3 bubbles rating by creating fake review.

We have found no evidence that the rounding of bubble ratings to the nearest half-bubble creates discontinuities in the relationship between demand and bubble ratings such as those found by Anderson and Magruder (2012) and Luca (2016), who reported a significant effects of Yelp star ratings on restaurant bookings. The absence of a significant discontinuity implies that hotels of essentially the same quality (as measured by the exact score from 1 to 5 underlying the rounded bubble rating) but with a half-bubble difference attract the same level of attention at TripAdvisor. This implies that TripAdvisor's bubble rating system does not bias hotel quality signals, at least in the case of hotels with satisfactory to excellent ratings hotels located in Rome.

There are several plausible explanations of the difference between our results and those obtained in previous studies. First, Anderson and Magruder (2012) and Luca (2016) modeled restaurant bookings, while we used the number of people viewing a hotel as a measure of its popularity. The possibility that bubble ratings still influence the probability of actual booking cannot be ruled out. Data availability is a major barrier preventing an empirical test of this explanation, and even a small-scale study of the relationship between the number of views and actual bookings would shed light on the strength of this association and its heterogeneity across hotels. The second explanation is the existence of cross-category differences. Among other things, hotels are differentiated from restaurants by the availability of clearly specified prices for the user-specified type of accommodation (as opposed to just a price category available for restaurants at Yelp) and the traditionally large role of third-party star ratings as indicators of the hotel's class (Guillet and Law, 2010), which do not exist in most other markets. The third potential explanation is the difference in the way that Yelp and TripAdvisor disclose information about service providers: the platforms may differ by the sufficiency of information provided: the more comprehensive the amount of information disclosed by the comparison website, the smaller should be the role of rounded ratings as signals of quality. Further comparative research is needed on what features of review websites are responsible for mitigating such biases, but in the case of TripAdvisor it can possibly be the quality of summary information available before users click at the hotel's link, including hotel class (star rating), information about awards, some of the key amenities and, more importantly, a comprehensive set of more than 10 filters, each containing multiple checkboxes to select from in order to shortlist the hotels.

The main limitation of our study that can also explain differences between our study and the extant research is that it uses a proxy of demand for hotels, but considering that customers visit hotel web pages after reviewing search results, where a summary of hotel prices, class and amenities as well as awards is present, views are undoubtedly an important element of the sales funnel. The availability of such demand proxy makes empirical research with a hotel demand measure as the dependent variable possible, taking into account that reliable and up-to-date information on actual bookings is not available for most academic researchers. Being a publicly available performance metric, the number of views can equally be used in further research on the generalizability of results presented in this paper, as well as in more general studies of the determinants of hotel popularity in different regions of the world. Causal analysis techniques such as fuzzy regression discontinuity design and propensity score matching can be potentially useful for inferring some of the related effects, such as the effects of having "Traveler's choice" and "Certificate of Excellence" awards on hotel popularity.

6. References

1. Anderson, M., Magruder, J., 2012. Learning from the crowd: Regression discontinuity estimates of the effects of an online review database. Econ. J. 122, 957–989.
2. Brynjolfsson, E., Dick, A.A., Smith, M.D., 2010. A nearly perfect market? QME 8, 1–33.
3. Calonico, S., Cattaneo, M.D., Farrell, M.H., 2020. Optimal bandwidth choice for robust

bias-corrected inference in regression discontinuity designs. Econom. J. Forthcomin.
4. Calonico, S., Cattaneo, M.D., Titiunik, R., 2015. rdrobust: An r package for robust nonparametric inference in regression-discontinuity designs. R J. 7, 38–51.
5. Cattaneo, M.D., Jansson, M., Ma, X., 2019. Simple local polynomial density estimators. J. Am. Stat. Assoc. 1–7.
6. Chevalier, J., Mayzlin, D., 2006. The effect of word of mouth on sales: Online book reviews. J. Mark. Res. 43, 345–354.
7. Duan, W., Gu, B., Whinston, A.B., 2008. Do online reviews matter?—An empirical investigation of panel data. Decis. Support Syst. 45, 1007–1016.
8. Floyd, K., Freling, R., Alhoqail, S., Cho, H.Y., Freling, T., 2014. How online product reviews affect retail sales: A meta-analysis. J. Retail. 90, 217–232.
9. Gavilan, D., Avello, M., Martinez-Navarro, G., 2018. The influence of online ratings and reviews on hotel booking consideration. Tour. Manag. 66, 53–61.
10. Guillet, B.D., Law, R., 2010. Analyzing hotel star ratings on third-party distribution websites. Int. J. Contemp. Hosp. Manag.
11. Hill, R.C., Fomby, T.B., Escanciano, J.C., Hillebrand, E., Jeliazkov, I., 2017. Regression Discontinuity Designs: Theory and Applications. Emerald Group Publishing.
12. Hu, N., Zhang, J., Pavlou, P.A., 2009. Overcoming the J-shaped distribution of product reviews. Commun. ACM 52, 144–147.
13. Koulayev, S., 2010. Estimating demand in online search markets, with application to hotel bookings. Fed. Reserv. Bank Bost. Work. Pap.
14. Luca, M., 2016. Reviews, reputation, and revenue: The case of Yelp. com. (March 15, 2016). Harvard Bus. Sch. NOM Unit Work. Pap.
15. Luca, M., Zervas, G., 2016. Fake it till you make it: Reputation, competition, and Yelp review fraud. Manage. Sci. 62, 3412–3427.
16. McCrary, J., 2008. Manipulation of the running variable in the regression discontinuity design: A density test. J. Econom. 142, 698–714.
17. Molinillo, S., Ximénez-de-Sandoval, J.L., Fernández-Morales, A., Coca-Stefaniak, A., 2016. Hotel assessment through social media: The case of TripAdvisor. Tour. Manag. Stud. 12, 15–24.
18. Phillips, P., Barnes, S., Zigan, K., Schegg, R., 2017. Understanding the impact of online reviews on hotel performance: an empirical analysis. J. Travel Res. 56, 235–249.
19. Smith, M.D., Brynjolfsson, E., 2001. Consumer decision-making at an Internet shopbot: Brand still matters. J. Ind. Econ. 49, 541–558.
20. Torres, E.N., Singh, D., Robertson-Ring, A., 2015. Consumer reviews and the creation of booking transaction value: Lessons from the hotel industry. Int. J. Hosp. Manag. 50, 77–83.
21. Wang, W., Li, F., Yi, Z., 2019. Scores vs. stars: A regression discontinuity study of online consumer reviews. Inf. Manag. 56, 418–428. https://doi.org/10.1016/j.im.2018.08.002
22. Yang, Y., Park, S., Hu, X., 2018. Electronic word of mouth and hotel performance: A meta-analysis. Tour. Manag. 67, 248–260.
23. Zhu, F., Zhang, X., 2010. Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics. J. Mark. 74, 133–148. https://doi.org/10.1509/jmkg.74.2.133

Elena B. Pokryshevskaya
HSE University,
Kantemirovskaya St. 3, Saint-Petersburg, Russia, 194100
E-mail: epokryshevskaya@hse.ru