Do the past experiences of others influence our decision of where to eat? Evidence from a natural experiment on Yelp

SFU SIMON FRASER UNIVERSITY

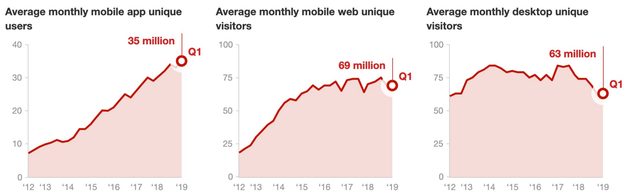
Jordan Hutchings, Simon Fraser University

Research Question

How do the star ratings left on a restaurant's Yelp page shape the future flow of customer traffic to that restaurant?

Motivation

The consumer is faced with large amounts of information asymmetry when making consumption decisions. Review platforms aid in filling in some of these information gaps, and often improve the consumer's decision making abilities. Think of the last time you were determining where to travel to, which movie to see, or where to eat. Did crowd-sourced rating information factor into your decision?

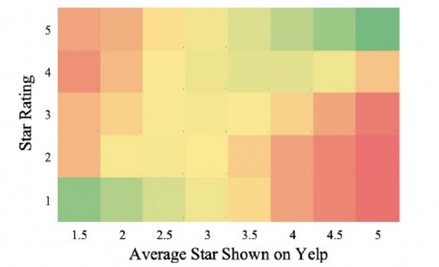


Source: https://www.yelp.ca/factsheet

The Simultaneity Problem

Restaurant quality is highly correlated with the Yelp rating it receives. The simultaneity problem is the largest obstacle when estimating the impact Yelp ratings have on the attendance to a given restaurant.

Figure 11: Probability of receiving a star rating given average star shown

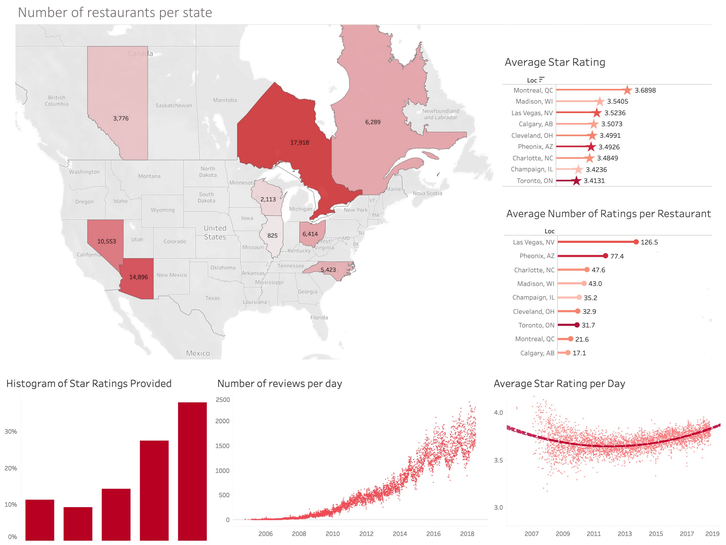


Each square is the conditional probability of receiving a star rating given the average star shown. Computed using the random sample of 2^{20} ratings from restaurants with over 25 ratings from Figure 10.

A good restaurant will most likely receive a good rating and have very little change in its attendance, similarly a bad restaurant receiving a bad rating causes very little change in the number of weekly customers.

Data

The Yelp Academic Dataset provides 2.9 million Yelp ratings for a sample of 58,290 restaurants located throughout North America collected since Yelp's conception in 2004.



Natural Experiment

Low Yelpers are users who exogenously rate restaurants poorly. Ratings provided by this subset of Yelp users provide exogenous variation in restaurant attendance as a result of the star rating provided.

Table 4: *Low_Yelper* Sample

Name	Average Stars	Review Count	Friends	Useful	Funny	Cool	Ratings in Sample
Justin	1.00	28	0	1	1	0	4
Kris	1.04	37	0	63	4	13	1
Peter	1.09	31	3	40	1	0	10
Joe	1.10	41	0	6	0	0	6
Jamie	1.10	35	5	38	0	0	5
Frederic	1.11	35	12	4	1	0	2
Lou	1.11	29	2	0	0	0	3
Benjamin	1.12	32	109	0	0	0	1
J	1.12	26	0	3	0	0	2
Steve	1.13	28	8	3	1	0	1
Victor	1.17	37	0	87	1	0	13
Ron	1.17	37	0	6	0	0	1
Ryan	1.18	42	0	0	0	0	21
M	1.19	63	43	47	42	11	11
Glen	1.19	25	7	0	0	0	2

Table 16: Rating Summary Statistics

	Nobs	Mean	Median	Stdev	Skewness
Panel A: Yelp Rounded					
Treated	318	3.50	3.5	0.68	-0.52
Restaurants	18782	3.66	3.5	0.63	-0.54
Panel B: True Average					
Treated	318	3.50	3.63	0.67	-0.62
Restaurants	18782	3.64	3.70	0.62	-0.59

Restaurant averages taken at the period immediately prior to Low_Yelper's rating, or at latest possible date for non-treated restaurants. Rounded panel uses true restaurant averages

Empirical Strategy

2SLS regression model:

$$Stars_{irt} = \delta_0 + \gamma_{FS} Low\_Yelper_{irt} + \delta_1 \boldsymbol{X}_{irt} + \lambda_t + \theta_r + \varepsilon_{\text{irt}} \quad (1)$$

$$\Delta Performance_{irt} = \beta_0 + \tau \widehat{Stars}_{irt} + \beta_1 \boldsymbol{X}_{irt} + \widehat{\lambda_t} + \theta_r + \epsilon_{irt} \quad (2)$$

$$\Delta Performance_{irt} = \sum_{i=t+1}^{t+\eta} Ratings\_Received_{ri} - \sum_{j=t-\eta}^{t} Ratings\_Received_{rj} - 1$$

Results & Discussion

Each incremental star rating is responsible for a 0.27 (13%) increase in the ratings received by the restaurant in the subsequent week.

Table 15: Dependent Variable Coefficient Estimates

	Main Model	Las Vegas	LBSN	High Comp.	Low Comp.	Chain & Franchise
Intercept	-4.31***	-5.77***	-15.97*	-2.745***	-3.9***	-1.57***
Star.hat	0.27***	0.39**	0.12**	0.58*	0.23***	0.19***
Rating Controls						
Text Length	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
Useful	0.01**	<0.01	<0.01	0.05	0.01	0.001***
Funny	0.05***	0.08***	0.02***	0.08**	0.04***	0.05***
Cool	-0.06	-0.08*	-0.02***	-0.14**	-0.04***	-0.04***
Restaurant Controls						
Average Star	0.25***	0.46***	-0.03	0.22***	0.22***	0.27***
Price Range 2	0.55***	0.88***	0.15***	0.9***	0.50***	0.53***
Price Range 3	1.05***	1.78***	0.16***	0.49***	1.16***	1.10***
Price Range 4	0.04	0.19***	0.05	-0.86***	0.09***	-0.17
Champaign, IL	0.43***		0.08*		0.43***	0.42***
Charlotte, NC	0.50***		0.14***		0.54***	0.51***
Cleveland, OH	0.29***		0.29***	-0.50	0.28***	0.29
Las Vegas, NV	1.65***		0.34***	0.78***	1.46***	1.73***
Madison, WI	0.39***		0.18***		0.40***	0.40***
Montréal, QC	0.13***		0.09*	-0.36***	0.40***	0.15***
Pittsburgh, PA	0.44***		0.11***		0.41***	0.46***
Toronto, ON	0.37***		0.14***	-0.12***	0.28***	0.38***
Observations	2,732,794	1,310,285	178,143	46,581	1,738,412	179,122
R^2	0.055	0.031	0.026	0.208	0.062	0.029
Adjusted R^2	0.031	0.227	0.025	0.205	0.062	0.028

*p<0.1; **p<0.05; ***p<0.01

All columns but LBSN column 3 are in terms of Yelp ratings received in subsequent week.
LBSN estimates are in terms of LBSN checkins.
Not shown: Time & restaurant fixed effects
Price Range 1 and Calgary, AB dropped from restaurant controls.

Conclusion

Yelp ratings are found to have a significant impact on the weekly number of restaurant attendees. Restaurants located in more restaurant-dense areas are more susceptible to Yelp ratings. Conversely, consumer flows to chains and franchises are less affected by Yelp ratings. This evidence supports the claim that crowd-sourced social rating platforms do affect consumer demand within the restaurant industry.