

Consumers, Experts, and Online Product Evaluations:
Evidence from the Brewing Industry

Grant D. Jacobsen*
University of Oregon

Published: Journal of Public Economics, 126 (2015) 114-123

Abstract

The growth of the Internet has led to a dramatic increase in the number of consumer or “user” product ratings, which are posted online by individuals who have consumed a good, and are available to other individuals as they make decisions about which products to purchase. These ratings have the potential to substantially improve the match between products and consumers, however the extent to which they do so likely depends on whether the ratings reflect actual consumer experiences. This paper evaluates one potential source of bias in consumer ratings: mimicry of the reviews of experts. Using a rich dataset on consumer product ratings from the brewing industry and a regression discontinuity empirical framework, I show that expert reviews influence consumer ratings. Consumer ratings fall in response to negative expert reviews and increase in response to positive expert reviews. The results are most pronounced for strongly negative or strongly positive expert reviews. This mimicry limits the extent to which information on product quality from actual consumer experiences diffuses to the population. I suggest that “nudges” could be implemented to limit the extent to which mimicry affects ratings.

Keywords: consumer ratings; user ratings; product ratings; electronic-word-of-mouth; eWOM, user-generated content; expert reviews; beer; brewing industry; nudges
JEL Codes: D83, D12, D71

*I thank Matt Lang, Bree Lang, Ben Hansen, and two anonymous reviewers for helpful comments and suggestions. Post: 1209 University of Oregon, 119 Hendricks Hall, Eugene, OR, 97403-1209, Tel: (541) 346-3419, Fax: (541) 346-2040, Email: gdjaco@uoregon.edu.

1 Introduction

The Internet has led to a massive increase in the number of consumer or “user” product ratings. These ratings, which have sometimes been referred to as “electronic word-of-mouth”, provide a complement or substitute to expert reviews, which have historically been of greater significance to consumers. Expert reviews and consumer ratings are both common, and these two types of evaluations are often presented alongside each other in web sites. Rotten Tomatoes, the popular web site for movie reviews, presents both a “Tomatometer”, which is based on critic reviews, and an Audience rating, which is based on user ratings. CNET, which provides ratings for electronic products, displays both an Editor’s review and an average user rating. Edmunds.com provides both an expert review and consumer ratings for cars. Amazon.com often provides summaries of expert book reviews from major news outlets before presenting the average consumer rating for the book. Additionally, for all products, Amazon presents the average rating for a product, as well as “the most helpful reviews”, which are displayed more prominently.

The availability of a new widely accessible source of information on product quality provided through consumer product ratings holds the promise of substantially improving the match between products and consumers. However, the value of consumer ratings likely depends on whether these ratings reflect actual consumer experiences. One potential source of bias is mimicry of the reviews of others. A substantial literature on individual behavior suggests that individual decision-making is influenced by information on the actions of others, and that certain types of individual responses can lead to negative welfare consequences. Most notably, as developed in Banerjee (1992), mimicry of the behavior of others can lead

to socially inefficient outcomes because “herd behavior”—in which individuals are influenced by the decisions that have previously been made by others due to its informational content—limits the diffusion of information to the rest of the population. Ottaviani and Sørensen (2001) show theoretically that the presence of experts can exacerbate herding problems and lead to worse outcomes, especially when experts do not have fully accurate information and when the population contains heterogeneous preferences.

While mimicry and herding are most often discussed in the context of financial markets, such behavior is applicable in many settings, and recent research has focused on its potential role in consumer product ratings. Muchnik et al. (2013) implement a field experiment with a social news aggregation web site and find that a randomly assigned positive initial rating for an item leads to large and lasting effects in the amount of positive ratings received by the item. They conclude that mimicry can lead to “ratings bubbles” which biases the opinion of “collective intelligence.”¹ Moe and Trusov (2011) find similar evidence of mimicry, using data on fragrance and beauty products to show that ratings behavior is significantly affected by previously posted ratings. Mimicry in ratings is consistent with the notion of conformity, in which individuals behave in patterns consistent with the social norm in order to protect their status within the community (Bernheim 1994).

For products in which both consumer and expert ratings are available, both the previous ratings of others and the opinions of the experts serve as potential sources of mimicry. Though the response of consumer ratings to expert reviews has not previously been examined, mimicry of expert reviews may be of relatively greater significance because individu-

¹Collective intelligence is a relatively recent term within the psychology literature that refers to the decision-making capabilities of groups (Wooley 2010).

als appear to be more responsive to the opinion of high-status individuals than low-status individuals. For example, Kumru and Vesterlund (2010) present experimental evidence that “low-status” individuals mimic “high-status” individuals in charitable giving. Whether mimicry is driven by the opinions of peers or the opinions of an expert, the extent to which it exists is likely to be socially harmful because mimicry limits the diffusion of information regarding actual consumer experiences.

In this paper, I provide the first evaluation of whether consumers mimic the reviews of experts. The analysis is based on data from the brewing industry, and I use a rich dataset on expert reviews and consumer ratings from the leading site for beer evaluations, Beer Advocate. Using a regression discontinuity empirical framework based on the timing of expert reviews, I find clear evidence that expert reviews influence subsequent consumer ratings. In particular, consumers mimic the reviews of experts. Positive expert reviews lead to an increase in subsequent consumer ratings and negatives reviews lead to a decrease. The response is most evident for strongly positive or strongly negative expert reviews. The evidence is most consistent with users changing their ratings in response to the expert review, as opposed to the expert review changing the selection of users who choose to rate the beer.

In addition to relating to the literature on mimicry and herd behavior, this research contributes to a rapidly growing literature on online product ratings which reflects their expanding role in modern marketplaces. While economists have been aware of the importance of information in consumer decisions for some time (Akerlof, 1970; Nelson, 1970), as well as the potential for the Internet to enhance the role of consumer ratings (Avery et al., 1999), recent research has focused on identifying empirical relationship. Consumer ratings

have been shown to influence demand in many settings, including menu inserts and choice of food items (Cai et al., 2009), Yelp reviews and choice of restaurant (Anderson and Magruder, 2012; Luca 2013), online movie reviews and box office receipts (Chintagunta et al., 2010) and Amazon.com reviews and book sales (Chevalier and Mayzlin, 2006). Other studies have shown that the demand for an item sold through eBay depends on the ratings that the item's seller has received from previous buyers (Cabral and Hortacsu, 2010; Lucking-Reiley et al., 2007; Jin and Kato, 2006; Melnik and Alm, 2002). Researchers have also examined whether firms actively manipulate consumer ratings and have found mixed evidence.²

As with the literature on consumer ratings, the literature on expert reviews has focused on consumer demand as the primary outcome of interest. Reinstein and Snyder (2005) find evidence that expert reviews influence consumer demand for movies.³ Both Higler et al. (2011) and Friberg and Grönqvist (2012) find evidence the expert reviews influence the demand for wine, with the former study using a field experiment and the latter study exploiting a quasi-experiment based on the timing of reviews. Berger et al. (2010) find that receiving a favorable book review from the *New York Times* is associated with increased sales. DellaVigna and Hermle (2014) provide one of the few studies focusing on potential bias in expert reviews, finding no evidence that expert movie reviews are biased when they are reported in a media outlet that is owned by the same company that produced the film.

The remainder of the paper proceeds as follows. In Section 2, I describe background information on the brewing industry and Beer Advocate, as well as the features of the rating-level dataset that provides the basis for the analysis. In Section 3, I evaluate the effect of

²Mayzlin et al. (2014) find evidence of such manipulation in Hotel ratings on TripAdvisor. Anderson and Magruder (2012) do not find evidence of rating manipulation in restaurant reviews on Yelp.

³Eliashberg and Shugan (1997) also examine the impact of expert reviews on box office performance, but do not find a significant relationship.

expert reviews on consumer ratings using a regression discontinuity framework. I conclude by describing some implications of the findings in Section 4.

2 Background and Data

I study consumer ratings in the context of the brewing industry, which is an appealing setting for three primary reasons. First, the brewing industry is substantial, accounting for \$99 billion in sales and 200 million barrels of beer in the U.S. in 2012 (Brewers' Association, 2013). Secondly, the brewing industry involves a large number of products. Beers vary in their ingredients, style, and brewer and each beer variety constitutes a different product. There are currently more than 2,500 breweries in the U.S., and breweries typically produce multiple varieties of beer.⁴ Third, unlike many other products, beers are not typically reviewed by “experts” as soon as they are produced. The delayed review of most beers enables the empirical framework that I employ, which is based on how ratings change immediately following the expert review, and which enables plausible identification of causal effects (Angrist and Pischke, 2010). One of the benefits of focusing on how user ratings change following an expert review is that it leads to estimates that are free from concerns related to reverse causality, in which user reviews influence expert reviews, as opposed to vice versa.

I examine reviews of beers reported on Beer Advocate (beeradvocate.com). Beer Advocate is the primary source for beer ratings, and receives more than 25 million page views and 2.5

⁴The number of breweries has increased dramatically in recent years due to the expansion of the craft brew sector. Relative to large commercial breweries, craft breweries tend to have low production volumes, a focus on local or regional markets, and an emphasis on more specialized varieties of ales (stout, porter, IPA, etc.). The volume of craft brew produced grew at a rate of 10 percent annually between 2007 and 2012 (DeMeter, 2013).

million unique visitors per month (Beer Advocate, 2013).⁵ Ratings are provided by the Beer Advocate founders, Jason and Todd Alström, and by site members. Member registration is free and, as of 2012, there were over 30,000 members contributing to the reviews on Beer Advocate. Beers receive an overall rating, as well as separate ratings for taste, appearance, aroma, and mouthfeel.⁶ The overall rating is chosen by the member, as opposed to being generated through a formula based on the other ratings. Ratings are scored on a 1 to 5 point scale. Over the time period in which the data was collected, Beer Advocate required that ratings must be accompanied by reviewer comments of at least 250 characters in length.

Figure 1 presents a screen shot of a beer's rating page on Beer Advocate. The page includes two summary scores for the beer, as well as a member review (more member reviews could be accessed by a site visitor by scrolling down). One summary score represents member ratings (the "BA Score") and one represents the founders ("The Bros"). In the context of this paper, "The Bros" score represents the "expert" review. The Bros' score is based on the reviews of Jason or Todd.⁷ The summary scores are a function of the underlying ratings but are scored on a 0 to 100 point scale that is structured similarly to the grades used in an academic setting.⁸ A beer appears in Beer Advocate when a member submits it through an online form. Members must record twenty reviews before they are eligible to add beers to the database. Members cannot enter multiple different ratings for the same beer.

I obtained data on Beer Advocate ratings from the Stanford Network Analysis Project.⁹

⁵The main alternative to Beer Advocate is RateBeer.com. Google search volume for Beer Advocate is about three times as large as search volume for Rate Beer (Google, 2013).

⁶In the analysis, I focus on the overall rating of a beer because ratings for taste, appearance, aroma, and mouthfeel are not displayed prominently at beeradvocate.com.

⁷Jason and Todd each have their own individual account, as well as a mutual account "BeerAdvocate".

⁸According to the key on beeradvocate.com, a score below 60 is "awful", 60 to 69 is "poor", 70 to 79 is "okay", 80-84 is "good", "85-89" is "very good", "90-94" is outstanding, "95-100" is world-class.

⁹The data were downloaded at <http://snap.stanford.edu/data/web-BeerAdvocate.html>.

Each observation in the data refers to a beer rating and includes the rating as well as reviewer, beer, and date identifiers. While the complete dataset includes more than 1.5 million observations recorded between the founding of Beer Advocate in 1996 and the end of the sample in January 2012,¹⁰ I restrict the data to the set of beers that the Bros have rated, and to ratings that were made within 4 years of the Bros' rating for most of the analysis.¹¹ The goal of dropping these observations is primarily to relax the assumptions required for causal identification,¹² though they also enhance computational tractability for some of the specifications with a large number of non-parametric controls. The final dataset consists of 218,095 observations. Much of the analysis is based on a preferred dataset of 49,715 observations that occurred within 6 months of the expert rating.

I generate several key variables. *Rated* is a binary variable that equals 1 if the rating was entered after the Bros' rating or 0 if before. *Review Order* indicates the order in which a rating was placed relative to all other ratings received by a beer. *Average User Score at Time of Rating* reports the average user score for a beer at the time the rating was posted. *User Experience* indicates the number of reviews previously entered by the Beer Advocate member entering the review. *Relative Days* reports the day of the observation relative to when the Bros' review was placed. For example, a value of 7 indicates the review was placed a week after the Bros' rating, and a value of 0 indicates the rating was made the same day as the Bros. If both Bros rated the beer, *Relative Days* is based on the time of the first

¹⁰While the Brother's ratings appear as early as 1996, ratings from regular members do not appear in the database until October 2000.

¹¹I also drop any beer for which there were no member reviews either preceding or following the Bros' rating. Additionally, I drop the ratings of the Bros such that the data only include member ratings (though I use the Bros' rating to generate several variables, as I describe in the next paragraph).

¹²In regression discontinuity studies, limiting the sample to observations that occur within a certain bandwidth of a primary threshold is a common and recommended procedure (Imbens and Lemieux, 2008).

rating. *Relative Months* is generated by collapsing *Relative Days* into 30 day bins. *Rating Difference* reports the difference between the Bros' rating and the mean user rating made prior to the Bros' rating. As with the relative time variables, if both Bros rated the beer, the *Rating Difference* is calculated based on the first review. *Strong Negative*, *Weak Negative*, *Weak Positive*, and *Strong Positive* are binary variables that represent ranges of *Rating Difference*. *Strong Negative* equals 1 if the Bros' rating was more than 0.5 points below the mean user rating, *Weak Negative* equals 1 if the Bros' rating was between 0.5 and 0 points below the mean user rating, *Weak Positive* equals 1 if the Bros' rating was equal between 0 and 0.5 points above the mean user rating, and *Strong Positive* equals 1 if the Bros' rating was more than 0.5 points above the mean user rating.

Summary statistics are reported in Table 1. Statistics are reported for samples consisting of ratings that occurred within six months of an expert review, ratings that occurred within forty-eight months of an expert review, all ratings from beers that received an expert review, and the full sample of all ratings.¹³ The means of the variables are generally similar across samples and especially similar with respect to *User Rating*.¹⁴ Focusing on the six-month sample, the average user rating for a beer is 3.8. About 140 ratings have typically already been placed for a beer at the time a rating is entered, and a user has typically entered about 350 previous ratings at the time of recording their rating. The Bros are slightly less generous than the average member, with a mean rating 0.1 points lower. A substantial share of ratings fall in each category of *Rating Difference*, though a strong positive rating

¹³It is not possible to compute the variables that are based on the expert rating for ratings of beers that were not rated by an expert, so for the full sample I do not report summary statistics for these variables.

¹⁴The primary difference is that the means of *Review Order*, *Relative Months*, and *Relative Days* are smaller when a narrower bandwidth is used. Due to the growth of Beer Advocate, expanding the bandwidth is more likely to lead to the inclusion of recent reviews. These reviews are likely to have occurred substantially after the Bros' rating because the Bros have been rating beers since the beginning of Beer Advocate.

is the least common. There are about 8,000 member profiles in the 6-month bandwidth sample, 15,000 member profiles in the 48-month bandwidth sample, 22,000 members in the sample of all rated beers, and 33,000 member profiles in the full sample. All samples limited to expert-rated beers have about 1,500 beer varieties.¹⁵

To help characterize the variation of primary interest in the analysis, I present a histogram of *Rating Difference* in Figure 2. The figure shows that the difference between the Bros' rating and the average user rating placed prior to the Bros' rating is approximately normally distributed with a mean near zero. Notably, there is little evidence of a bi-modal distribution in *Rating Difference*, which might be expected if the Bros' were primarily entering reviews that were meant to be "corrections" in cases when they viewed the Beer Advocate community as erroneously rating a beer too generously or too negatively. The takeaway from Figure 2 is that the Bros' rankings do not appear to systematically differ from the ratings previously entered by users and that there is meaningful variation across beers in the relative rating received from the Bros.

3 Analysis and Results

I examine the effect of an expert review on user ratings by evaluating how the ratings of Beer Advocate members change in the period following the Bros' review, and how this change relates to how the Bros rated the beer. I focus on how the Bros' rating compares to the average member rating that was entered prior to the Bros' review because the analysis focuses on changes in how members rate a beer. The Bros' rating is only likely to influence

¹⁵For the six-month bandwidth sample, 19 beers are omitted from the sample because they were not rated by a user within the six-month period around the expert review.

member ratings if it differs from the way members had previously been rated the beer.

By focusing on changes in ratings that occur discontinuously at the time of the Bros' review, I am able to limit the potential for bias in the estimates. In particular, in order for an omitted factor to lead to bias, it must change discontinuously at the time of the expert review. For example, within the empirical framework, an association between member ratings and the Bros' ratings cannot be explained simply by "correlated tastes," in which the Bros' and the typical member prefer the same beers, because this correlation in tastes should not change exactly at the time that a beer is reviewed by an expert. Concerns about reverse causality, in which member ratings influence expert reviews, are also not an issue because the findings are driven by ratings that are entered after the Bros' have already reviewed the beer.

In addition to employing a regression discontinuity framework, I also control for certain factors that are known to influence product ratings. Godes and Silva (2014) show that review order, the current review average, and the time at which a review is entered influence product ratings.¹⁶ I control explicitly for these factors in each regression.¹⁷ I also control for the individual who enters the rating, and the specific beer that is being scored.

In the remainder of this section, I first present trends in user ratings relative to the time of the Bros' rating. I then present a series of estimates of the effect of the Bros' rating, as well as a set of econometric robustness checks and a discussion of selection dynamics in user ratings. I conclude with an examination of the mechanism behind the results.

¹⁶Moe and Schweidel (2012) present evidence that infrequent posters are particularly likely to be swayed by the average user rating at the time the rating is posted. As I describe in section 3.5, I examine whether the influence of an expert review diminishes with user experience as well. I find little evidence that the influence of an expert review diminishes with user experience.

¹⁷Results are very similar when these factors are omitted, primarily because they do not typically change discontinuously at the time of an expert review.

3.1 Trends in Means and Residuals

Figure 3 presents the average user rating across *Relative Months*. I produce separate plots for beers that received strong negative, weak negative, weak positive, and strong positive ratings from the Bros. The plots indicate that a strong negative rating leads to a decline in subsequent user ratings and a strong positive rating leads to an increase in subsequent user ratings. These patterns provide some initial visual evidence that user ratings are influenced by the reviews of experts.

In order to isolate the variation in user ratings that cannot be explained by a number of potentially confounding factors, I regress *User Rating* on individual (e.g., member), beer, and month-of-sample fixed effects and then calculate the residuals for each observation. These residuals represent variation in user ratings that cannot be explained by the individual who performed the review, the beer that was reviewed, or the time of the review. I plot the residuals by *Relative Months* in Figure 4. I present separate plots for each category of *Rating Difference*. The results are similar to those presented in Figure 3 though the residuals are less noisy and do not appear to be trending systematically across *Relative Months*.¹⁸ These results provide further evidence that expert reviews influence user ratings.

¹⁸The trends in rating means and residuals presented in this subsection are not materially affected by the choice of bin size. I demonstrate this in Figure A.1 in the Online Appendix, which replicates the residual plot presented in the “Strong Negative” panel of Figure 4 using bin sizes of 1, 7, 30, and 365 days.

3.2 Estimates

I use a regression discontinuity framework to more formally evaluate the effect of the expert review. In particular, I estimate models based on specifications of the following form,

$$\begin{aligned} \text{User Rating}_i = & \alpha_j + \gamma_k + \omega_t + \beta_1 \text{Rated}_i + \beta_2 \text{Relative Months}_i + \\ & \beta_3 \text{Relative Months}_i \times \text{Rated}_i + \sum_k B_k X_{ik} + \epsilon_i, \end{aligned} \quad (1)$$

where i indexes ratings. The terms α_j , γ_k , and ω_t represent individual, beer, and month-of-sample fixed effects. The summation term includes linear controls for review order and the average user rating at the time that a review was entered. The coefficient of primary interest is β_1 , as it represents the change in the average user rating following the Bros' rating. I perform regressions based on equation 1 for each category of *Rating Difference*. In all regressions throughout the analysis, I cluster standard errors by member profile.¹⁹

I also estimate models for the full sample that are of the form,

$$\begin{aligned} \text{User Rating}_i = & \alpha_j + \gamma_k + \omega_t + \beta_1 \text{Rated}_i + \beta_2 \text{Rated}_i \times \text{Rating Difference}_i + \\ & \beta_3 \text{Relative Months}_i + \beta_4 \text{Relative Months}_i \times \text{Rated}_i + \sum_k B_k X_{ik} + \epsilon_i. \end{aligned} \quad (2)$$

The primary coefficient of interest is β_2 , which indicates how a one-point increase in the Bros' rating affects subsequent user ratings.

Estimates are presented in Table 2. Columns 1-4 present estimates based on equation 1. A strong negative expert rating is associated with a statistically significant decline of 0.08 points in the average user rating. A weak negative expert rating is associated with a

¹⁹Results are robust to alternative clustering choices, such as clustering by beer variety.

statistically significant decline of 0.05 points in the average user rating. A weak positive expert rating is not associated with a statistically significant change in the average user rating. A strong positive expert rating is associated with a statistically significant increase of 0.14 points in the average user rating. Column 5 presents estimates based on equation 2. The coefficient on *Rated* \times *Rating Difference* indicates that a 1 point increase in the Bros' rating is associated with a statistically significant 0.15 point increase in subsequent consumer ratings.

The most appropriate interpretation of the coefficients is that they indicate the extent by which a Bros' rating of each variety causes individual reviewers to change their own ratings. There are two other tempting alternative interpretations of the coefficients, both of which are likely to be incorrect. First, one might explain the relationship through a sample selection mechanism in which the Bros' rating induces certain types of members to subsequently rate the beer (potentially those who are more inclined to give a positive or negative review), however this type of response would not be reflected in the estimates because the models include individual fixed effects. I describe issues related to selection dynamics further in subsection 3.4. A second tempting explanation for the result is that the Bros can more accurately predict the quality of a beer, and therefore the change in user ratings is simply driven by a trend over time of a beer's rating toward its true quality. However, given the small bandwidth used in the estimates, and the clear discontinuities that are evident in Figures 3 and 4, this is also unlikely to explained the observed results.

3.3 Robustness Checks

To investigate the robustness of the results, I examine the impact of alternate bandwidths on the estimates, as well as the impact of modeling *Relative Months* using global polynomials of various orders. For the robustness checks, I focus on the coefficient on *Rated x Rating Difference* in the models based on the full sample, as this provides the most concise test of the effect of the Bros' rating. The results are presented in Table 3. Each panel presents results from a different sample bandwidth, and each column presents results from a different polynomial order. Across samples and specifications, the coefficient on *Rated x Rating Difference* is nearly unchanged from the coefficient reported in column 5 of Table 2.²⁰

In addition to estimating the above robustness checks, I conduct a placebo test in order to investigate potentially overlooked sources of bias in the way the data are compiled or analyzed. In particular, I compile and analyze the data such that the member of Beer Advocate with the most ratings is actually the Bros. When all reviewers are considered, including the Bros and member-contributors, the most prolific reviewer is "NorthYorkSammy", who has 5,800 beer reviews.²¹ I report these placebo results in Table 4 in a manner that mirrors the robustness checks presented in Table 2. At small bandwidths, the coefficient on *Rated x Rating Difference* in the placebo tests is small in magnitude and statistically insignificant. At higher bandwidths, the coefficient becomes statistically significant, but remains small in magnitude. Overall, the small magnitude of the coefficients and inconsistent statistical significance in the placebo tests support the validity of the empirical design.²²

²⁰I also estimate effects by bandwidth for each category of *Rating Difference*. These estimates are presented in Figure A.2 in the Online Appendix and show that the estimates for each category are stable across bandwidths.

²¹For context, North York Sammy has rated roughly 58 beers per month since his first review.

²²One explanation for the slight, yet statistically-significant, coefficient is that NorthYorkSammy has

3.4 Selection Dynamics

One potential concern with the analysis is that the results are caused by selection dynamics. For example, if the Bros' review induces certain users to rate a beer, then observed changes in user ratings could be caused by selection effects as opposed to or in addition to individuals actually changing their rating based on the Bros' review. Such a selection-based response would not change the overall impact of the Bros' review on subsequent user ratings entered into Beer Advocate, but it would change the interpretation of the mechanism by which the response occurs.

I initially examine the possibility of selection using a of *Relative Months*, as presented in Figure 5. The histogram shows that reviews are most likely to occur around the time of the Bros' rating. The first explanation that comes to mind for this spike is that the Bros' review invokes ratings from Beer Advocate members, but this is unlikely to be the case because the spike in ratings occurs both before and after the Bros' review. Note that the spike in ratings is nearly as high in the 30-days *prior* to the Bros' review, as the 30-days following the Bros' review. Additionally, I further investigate whether the Bros' review invokes user ratings by examining whether the spike in ratings is more prominent when the Bros give a positive rating, which would be expected if selection were present because a positive Bros' ratings should lead to a larger increase in demand. I find little evidence of such a pattern (see Figure A.3 in the Online Appendix).

Instead, ratings likely surge around the time of the Bros' review because the ability of the Bros to acquire and rate a beer indicates a period beer when a beer was highly likely to be in production and available outside of its local market, and thus also accessible to Beer

achieved "quasi-expert" status among Beer Advocate members due to his large number of ratings.

Advocate members for their review. Beers come in and out of production either because a Brewer chooses to create or discontinue them altogether, or because the Beer is a seasonal beer that may be more likely to be brewed or bought during certain months of the year.²³ This latter explanation also explains that general seasonality observed in the number of user ratings across *Relative Months*.²⁴

While Figure 5 does not provide strong evidence that the Bros' review induces user ratings, the empirical framework should eliminate any selection effects that do exist because the individual fixed effects control for the member who is rating the beer.²⁵ Additionally, as one would expect based on Figures 3 and 4, the empirical results are all nearly unchanged by dropping observations from the months immediately surrounding the Bros' rating. In sum, the evidence indicates that the observed change in user ratings is almost certainly caused by individuals changing their ratings as opposed to the expert review inducing a systematically different selection of users to rate the beer.²⁶

²³Certain dark beers, such as imperial stouts, tend to be more sought after in the winter, as well as more likely to be brewed and distributed. Certain light beers, such as saisons or wheat beers, are more common in the summer. Some breweries distribute the same set of seasonal beers repeatedly across years. For example, Deschutes Brewery in Oregon produces the Jubel Ale in winter, the Twilight Ale in the summer, and the Red Chair Ale in spring.

²⁴If a beer is a seasonal beer and the Bros' rate the beer during peak season, it will contribute to a 12-month pattern across *Relative Months* regardless of the calendar month that corresponds to the beer's peak season.

²⁵Results are very similar regardless of whether individual effects are included, indicating that the response of Beer Advocate members to the Bros' rating occurs exclusively through a change in the ratings chosen by members (relative to what they would have chosen absent the Bros' review), as opposed to a change in the sample of members rating the beer.

²⁶There is also little evidence that the Bros' reviews are induced by user ratings. For example, if an indicator variable is included in the regressions that equals 1 for the three observations entered for a beer variety immediately prior to a Bros' rating, the coefficient for the indicator variable is generally insignificant. The implication of this result is that anomalous user reviews do not appear to prompt the Bros' to enter their review.

3.5 Mechanism and Application to Other Categories

There are several mechanisms that may explain why expert reviews influence consumer ratings. First, consumers may be following a “herding” instinct whereby they suppress their own opinions and mimic the opinions of experts for psychological or social reasons, such as a general preference for conformity. Secondly, expert reviews may change the manner in which individuals experience a beer—perhaps by highlighting a key characteristic of the beer—and thereby influence the subsequent rating entered by the consumer. This mechanism could be called “persuasion.” Finally, reviewers may update their rating after reading an expert review because they believe that the expert opinion is more accurate than their own. This might be called a “learning” mechanism. These different mechanisms likely have different welfare implications, with the herding mechanism likely to be the most socially damaging.

In order to examine the mechanism, I evaluate how the effect of an expert review varies depending on user experience. The assumption behind this part of the analysis is that as users become more experienced—in regards to both exposure to beer varieties and to entering ratings—the effect of an expert review should dissipate if the mechanism is persuasion or learning because users are likely to develop more clearly defined tastes and more confidence in their own opinions as they gain experience. If the mechanism is herding behavior, on the other hand, then there should be no interaction between user experience and the effect of an expert review.

I present estimates based on a specification that is identical to equation 1 with the exception that it includes an interaction of *Rated* and *User Experience* (scaled by 10 to ease

exposition) in the first four columns of Table 5.²⁷ I limit the sample to ratings that were amongst the first 500 ratings entered by a user, as the effect of experience, if it exists, is likely to reach a saturation point as more reviews are entered.²⁸ The coefficient on the new interaction term is small and statistically insignificant across samples. I also present results that are based on a modified version of equation 2 that includes a new set of interaction terms, most notably *Rated x Rating Difference x User Experience / 10*. Here again, the coefficient on the key interaction term is small and statistically insignificant. While it is difficult in this setting to identify the mechanism behind the results with complete certainty, the implication from the results in Table 5 is that the mechanism driving the results is herding as opposed to learning or persuasion.²⁹

It is possible that mimicry would fail to exist or could be driven by different mechanisms in other settings. In particular, for goods or services where quality is easier to observe, consumers may feel less desire to conform. For example, a consumer is likely to enter a negative rating if they discover that their hotel room is unsanitary, regardless of any previous reviews to the contrary. Similar logic might apply in the case of ratings of software or smart phone applications, where “buggy” programs can often be easily and unambiguously detected. The results in my setting are most likely to generalize to settings where there are less objective measures of quality, such as ratings for restaurants, cars, or apparel.

²⁷To ease interpretation of the results, I center *User Experience* by subtracting the mean from each observation, such that the coefficient on *Rated* can be interpreted as the effect of an expert review for a user with an average amount of experience.

²⁸Results are robust to limiting the sample even further (the interaction term is consistently insignificant across any choice of threshold below 500).

²⁹Across models, estimates of the effect of an expert review (for an individual with an average level of experience) are noisier than than the estimates reported in Table 2, but in all cases the confidence interval includes the point estimate from Table 2.

4 Conclusion

The Internet has created an environment in which individuals can easily provide publicly accessible product evaluations. By increasing the availability of information, these ratings hold the promise of improving the match between consumers and products. However, the value of these ratings depends on whether the ratings accurately reflect consumer experiences. In this paper, I evaluate one potential source of bias, which is mimicry of the reviews of experts. I find clear evidence that consumer ratings are swayed by the reviews of experts, increasing their ratings in response to positive expert reviews and decreasing them in response to negative expert reviews.

It is likely that the welfare effects of mimicry are negative. If the typical user who submits a rating has similar preferences as the typical person who views a rating, then overtime an item's rating should converge toward its true quality. However, if distortions exist through mimicry of experts, then an item's rating will be a function of both the expert's review and the "true" quality of the item. If an expert's preferences do not exactly match the typical visitor's preference for every item, then mimicry will limit the diffusion of valuable information to the population. The assumption that the typical user who submits a rating has similar preferences as the typical individual viewing the rating seems unlikely to be strongly violated, and the assumption that the preferences of experts are not identical to the typical visitor across all items seems to be very weak. However, if these assumptions were violated, the effects of mimicry are unclear and could potentially be positive.

Assuming that the effects of mimicry are negative, one implication of mimicry in consumer ratings is that "nudges" that increase the likelihood that consumers submit their rat-

ings without seeing the reviews of experts or other consumers may be a successful strategy for improving that extent to which product ratings reflect actual consumer experiences.³⁰ On Beer Advocate, the “review” button that members click to rate a beer is currently located immediately adjacent to two prominent boxes that summarize the opinions of the Bros and other Beer Advocate members (see Figure 1). It would be easy to modify the site to hide these reviews from those submitting new ratings. For example, members could be required to navigate to a general “Submit Ratings” page and then select the beer of interest from a drop-down menu or search bar in order to submit their rating.

The idea of using nudges to minimize distortions in ratings could be easily applied to other settings. For example, Amazon.com currently allows members to submit products ratings either by clicking a button on the “Your Orders” page that does not include any information on previous product ratings or by clicking a button on the product’s page that is located directly next to a frequency chart of all previous consumer ratings. The latter option seems much more likely to induce bias from previous reviews and this option could be eliminated in order to limit the potential influence of mimicry. Nudges that increase the accuracy of product ratings would likely be appealing to organizations that provide consumer ratings because more accurate ratings would make their web sites more appealing to consumers. Additionally, policymakers have demonstrated an interest in minimizing distortions in online consumer ratings. New York’s Attorney General fined numerous companies for illegal and deceptive business practices for writing fake online reviews on web sites including Yelp, Google, and CitySearch (Streitfeld, 2013).

³⁰“Nudges” refer to actions that influence behavior using “choice architecture” that manipulates the context of a decision (Thaler and Sunstein 2008).

Advances in mobile technology have allowed consumers to have nearly continuous access to online product ratings, and these ratings are likely to play an increasingly significant role in the markets for many different varieties of goods and services, whether for day-to-day products, like beer, or for more expensive products, such as cars, legal counsel, and medical services. Despite the prevalence of these ratings, they have been the focus of relatively little research. Future research that further evaluates how consumer ratings are formed, how they influence consumer behavior, and how they influence market equilibrium should be a priority for economists and other researchers.

References

- Akerlof, G. A. (1970). The Market for 'Lemons': Quality Uncertainty and the Market Mechanism. *Quarterly Journal of Economics* 84(3), pp. 488-500.
- Anderson, M. L. and J. Magruder (2013). Learning from the Crowd: Regression Discontinuity Estimates of the Effects of an Online Review Database. *Economic Journal* 122(563), pp. 957-989.
- Angrist, J. D. and J. Pischke (2010). The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con out of Econometrics. *Journal of Economic Perspectives* 24(2), pp. 3-30.
- Avery, C., P. Resnick and R. Zeckhauser (1999). The Market for Evaluations. *The American Economic Review* 89(3), pp. 564-584.
- Banerjee, A. V. (1992). A Simple Model of Herd Behavior. *Quarterly Journal of Economics* 107(3), pp. 797-8173.
- Beer Advocate (2013). About Beer Advocate. Online at beeradvocate.com/about. Accessed September 3, 2013.

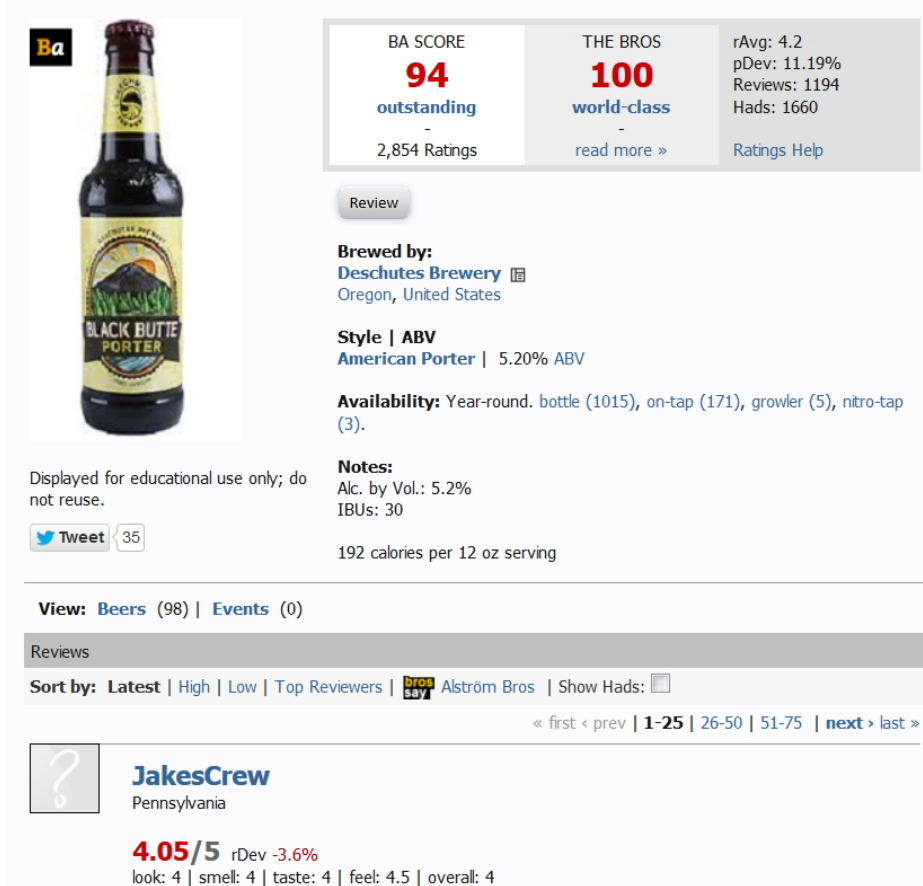
- Bernheim, B. D. (1994). A Theory of Conformity *Journal of Political Economy* 102(5), pp. 841-877.
- Berger, J., A. T. Sorensen, and S. J. Rasmussen (2010). Positive Effects of Negative Publicity: When Negative Reviews Increase Sales. *Marketing Science* 29(5), pp. 815-27.
- Brewer's Association (2013). Craft Brewing Facts. Online at brewersassociation.org. Accessed September 3, 2013.
- Cabral, L. and A. Hortacsu (2010). The Dynamics of Seller Reputation: Evidence from eBay. *Journal of Industrial Economics* 58(1), pp. 54-78.
- Cai, H., Y. Chen, and H. Fang (2009). Observational Learning: Evidence from a Randomized Natural Field Experiment. *American Economic Review* 99(3), pp. 864-882.
- Chevalier, J. A. and D. Mayzlin (2006). The Effect of Word of Mouth on Sales: Online Book Reviews. *Journal of Marketing Research* 43(3), pp. 345-354.
- Chintagunta, P., S. Gopinath, and S. Venkataraman (2010). The Effects of Online User Reviews on Movie Box Office Performance: Accounting for Sequential Rollout and Aggregation Across Local Markets. *Marketing Science* 29(5), pp. 944-957.
- DellaVigna, S. and J. Hermle (2014). Does Media Concentration Lead to Biased Coverage? Evidence from Movie Reviews. Working Paper.
- Demeter Group (2013). State of the Craft Beer Industry 2013. Online at demetergroup.net. Accessed September 3, 2013.
- Eliashberg, J. and S. M. Shugan (1997). Film Critics: Influencers or Predictors? *Journal of Marketing* 61(2), pp. 68-78.
- Friberg, R. and E. Grönqvist (2012). Do Expert Reviews Affect the Demand for Wine? *American Economic Journal: Applied Economics* 4(1), pp. 192-211.
- Godes, D. and J. C. Silva (2012). Sequential and Temporal Dynamics of Online Opinion. *Marketing Science* 31(3), pp. 448-473.

- Google (2013). Google Trends. Online at google.com/trends. Accessed September 3, 2013.
- Hilger, J., G. Rafert, and S. Villas-Boas. (2011). Expert Opinion and the Demand for Experience goods: An Experimental Approach in the Retail Wine Market. *Review of Economics and Statistics* 93(4), pp. 1289-1296.
- Imbens, G. W. and T. Lemieux (2008). Regression Discontinuity Designs: A Guide to Practice. *Journal of Econometrics* 142(2), pp. 615-635.
- Jin, G. Z. and A. Kato (2006). Price, Quality, and Reputation: Evidence from an Online Field Experiment. *RAND Journal of Economics* 37(4), pp. 983-1005.
- Kumru, C. S. and L. Vesterlund (2010). The Effect of Status on Charitable Giving. *Journal of Public Economic Theory*, 12(4), pp. 709-735.
- Luca, M. (2013). Reviews, Reputation, and Revenue: the Case of Yelp.com. Working Paper.
- Lucking-Reiley, D., D. Bryan, N. Prasad, and D. Reeves (2007). Pennies from eBay: The Determinants of Price in Online Auctions. *Journal of Industrial Economics* 55(2), pp. 223-233.
- Mayzlin, D., Y. Dover, and J. A. Chevalier (2012). Promotional Reviews: An Empirical Investigation of Online Review Manipulation. *American Economic Review* 104(8), pp. 2421-55.
- Melnik, M. I. and J. Alm (2003). Does a Seller's eCommerce Reputation Matter? Evidence from eBay auctions. *Journal of Industrial Economics* 50(3), pp. 337-349.
- Moe, W. M. and D. A. Schweidel (2012). Online Product Opinions: Incidence, Evaluation, and Evolution. *Marketing Science* 31(3), pp. 372-386.
- Moe, W. M. and M. Trusov (2011). The Value of Social Dynamics in Online Product Ratings Forums. *Journal of Marketing Research* 48(3), pp. 444-456.
- Muchnik, L., S. Aral, and S. J. Taylor (2013). Social Influence Bias: A Randomized Experiment. *Science* 341, pp. 647-651.


- Ottaviani, M. and P. Sørensen (2001). Information aggregation in debate: who should speak first? *Journal of Public Economics* 81, pp. 393-421.
- Nelson, P. (1970). Information and Consumer Behavior. *Journal of Political Economy* 78(2), pp. 311-329.
- Reinstein, D. A. and C. M. Snyder (2005). The influence of Expert Reviews on Consumer Demand for Experience Goods: A Case Study of Movie Critics. *Journal of Industrial Economics* 53(1), pp. 27-51.
- Streitfeld, D. (2013). Give Yourself 5 Stars? Online, It Might Cost You. *New York Times*. September 22, 2013, pp. A1.
- Thaler, R. H. and C. R. Sunstein (2008). *Nudge*. New Haven: Yale University Press.
- Wooley, A. W., C. F. Chabris, A. Pentland, N. Hasmi, and T. W. Malone (2013). Evidence for a Collective Intelligence Factor in the Performance of Human Groups. *Science* 330, pp. 686-687.

Figures and Tables

Black Butte Porter - Deschutes Brewery




Ba



BA SCORE 94 outstanding - 2,854 Ratings	THE BROS 100 world-class - read more »	rAvg: 4.2 pDev: 11.19% Reviews: 1194 Hads: 1660 Ratings Help
---	--	--

Review


Brewed by:
Deschutes Brewery 
Oregon, United States

Style | ABV
American Porter | 5.20% ABV

Availability: Year-round. bottle (1015), on-tap (171), growler (5), nitro-tap (3).


Notes:
Alc. by Vol.: 5.2%
IBUs: 30
192 calories per 12 oz serving

Displayed for educational use only; do not reuse.


 Tweet 35

View: Beers (98) | Events (0)

Reviews

Sort by: Latest | High | Low | Top Reviewers |  Alström Bros | Show Hads:

<< first < prev | **1-25** | 26-50 | 51-75 | next > last >>

 **JakesCrew**
Pennsylvania

4.05/5 rDev -3.6%
look: 4 | smel: 4 | taste: 4 | feel: 4.5 | overall: 4

Figure 1: **Example of a Beer's Rating Page on Beer Advocate.** This image was captured from beeradvocate.com on September 3, 2013.

Table 1: Summary Statistics

Variable	Sample		
	Bandwidth: 6 Months	Bandwidth: 48 Months	All Rated Beers
User Rating	3.79 (0.73)	3.81 (0.74)	3.83 (0.73)
Rated (by an expert at time of user rating; 1 = yes)	0.56 (0.50)	0.77 (0.42)	0.88 (0.33)
Year	2007.15 (3.22)	2006.52 (2.65)	2007.79 (2.51)
Review Order	134.06 (216.78)	190.53 (236.78)	437.35 (476.80)
Avg. User Score at Time of Rating	3.83 (0.44)	3.84 (0.44)	3.85 (0.41)
User Experience	354.71 (510.67)	276.55 (413.01)	252.92 (390.33)
Relative Days	12.36 (82.53)	443.67 (606.43)	1419.88 (1175.63)
Relative Months	-0.09 (2.75)	14.29 (20.21)	46.83 (39.19)
Rating Difference	-0.13 (0.61)	-0.10 (0.64)	-0.10 (0.68)
Strong Negative (Rating Difference < -0.5)	0.24 (0.43)	0.21 (0.41)	0.21 (0.41)
Weak Negative (-0.5 ≤ Rating Difference < 0)	0.28 (0.45)	0.29 (0.45)	0.28 (0.45)
Weak Positive (0 ≤ Rating Difference ≤ 0.5)	0.35 (0.48)	0.37 (0.48)	0.39 (0.49)
Strong Positive (0.5 < Rating Difference)	0.12 (0.32)	0.13 (0.33)	0.13 (0.33)
Num. of Individuals (i.e. members, users)	7,978	14,780	22,357
Num. of Beers	1,506	1,525	1,525
Observations	49,715	218,095	456,377

Notes: Each observation refers to a user rating. Means and standard deviations are reported in the top panel. The 6 and 48 bandwidth columns report statistics for ratings that occurred within 6 and 48 months of the Bros' rating, respectively. The "All Rated Beers" beers column reports statistics from the sample of all beers that were rated by the Bros. The "All Obs." column reports statistics for the whole sample. *User Experience* measures the number of ratings that the user had previously entered at the time the rating was recorded. *Relative Days* and *Relative Months* reports time relative to the Bros' rating. *Rating Difference* reports the difference between the Bros' rating and the average user rating that occurred prior to the Bros' rating. *Strong Negative*, *Weak Negative*, *Weak Positive*, and *Strong Positive* are binary categories that reflect the ranges of *Rating Difference* reported in the parentheses.

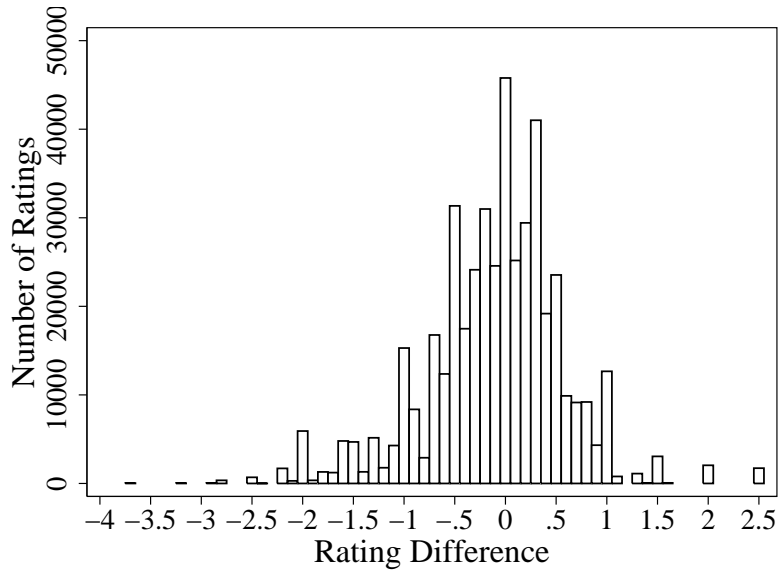


Figure 2: **Histogram of Rating Difference.** *Rating Difference* records the difference between the Bros' rating and the average user rating that was posted prior to the Bros' rating. The width of each bar is one-tenth of a point.

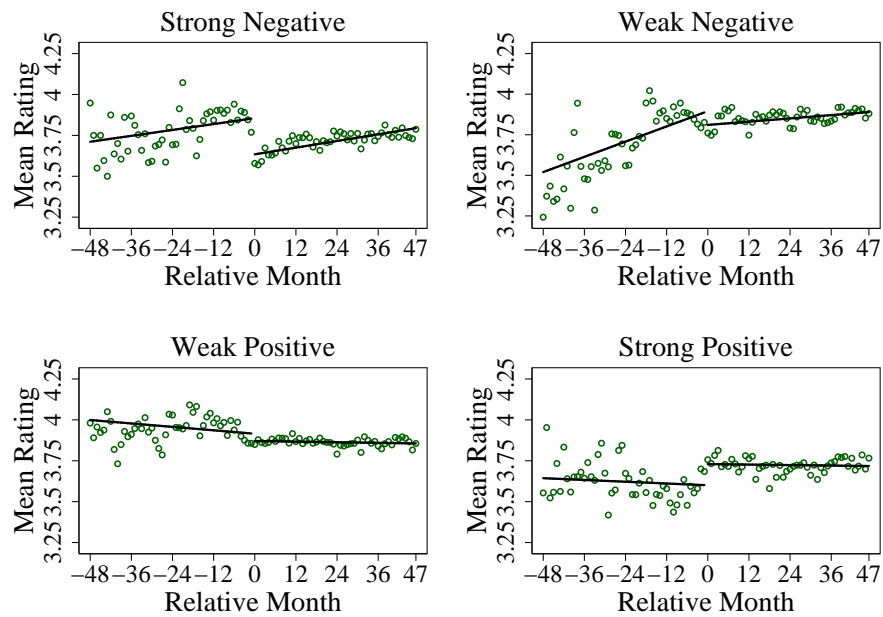


Figure 3: Average User Rating across Relative Months for Each Category of Expert Review. Each graph plots the mean user rating across *Relative Months*, which measures time relative to the expert review. Each panel presents results for a sample of beers that received an expert review of a certain type, as described in the panel titles. The fitted line represents the predicted values from a regression of *User Rating* on before and after linear *Relative Month* trends and *Rated*.

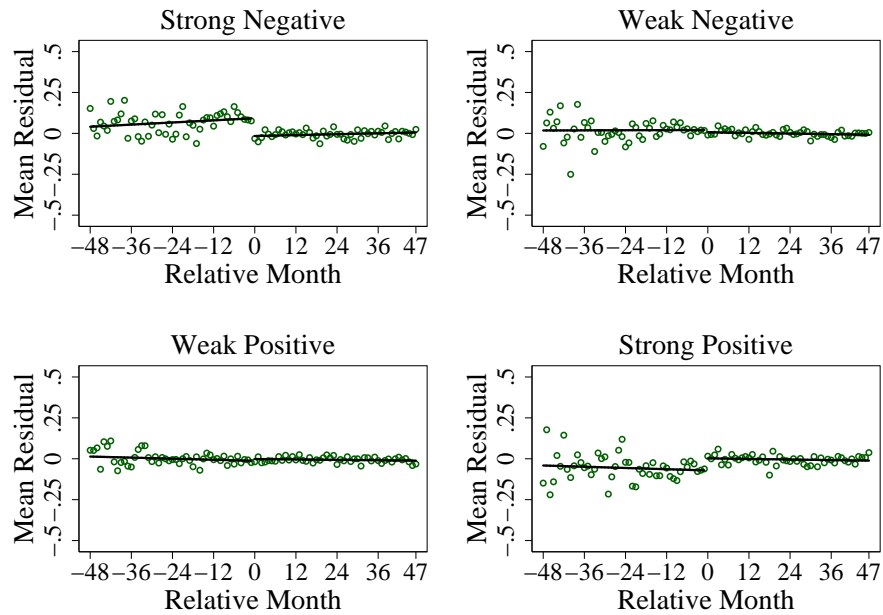


Figure 4: Average Residuals across Relative Months for Each Category of Expert Review. Each graph plots the mean residual from a regression of *User Rating* on beer fixed effects, month-of-sample fixed effects, and individual fixed effects across *Relative Months*, which measures time relative to the expert review. Each panel presents results for a sample of beers that received an expert review of a certain type, as described in the panel titles. The fitted line represents the predicted values from a regression of the residuals on before and after linear *Relative Month* trends and *Rated*.

Table 2: Estimates of the Effect of an Expert Review on User Ratings

Sample	Str. Neg. (1)	Wk. Neg. (2)	Wk. Pos. (3)	Str. Pos. (4)	Full (5)
Rated	-0.08** (0.03)	-0.05** (0.02)	0.03 (0.02)	0.14** (0.05)	0.00 (0.01)
Rated x Rating Difference					0.15** (0.01)
Individual Effects	Yes	Yes	Yes	Yes	Yes
Review Order	Yes	Yes	Yes	Yes	Yes
Avg. User Score at Time of Rating	Yes	Yes	Yes	Yes	Yes
Linear Relative Month Trends	Yes	Yes	Yes	Yes	Yes
Beer Effects	Yes	Yes	Yes	Yes	Yes
Month-of-Sample Effects	Yes	Yes	Yes	Yes	Yes
Observations	11,958	13,970	17,386	5,783	49,097

Notes: The unit of observation is a user rating. The dependent variable is the rating score. All observations are from ratings that occurred within a 6-month bandwidth of the Bros' rating. The first four columns report estimates from samples that are restricted to beers that received an expert review of a certain type, as described in the column headings. The fifth column includes all observations that occurred within the 6-month bandwidth. The yes/no items at the bottom of the table indicate the inclusion of various control variables. Standard errors are reported in parentheses and are clustered by member profile. One and two stars indicate 10% and 5% percent significance, respectively.

Table 3: Estimates of the Effect of an Expert Review by Relative Months Polynomial and Bandwidth

Global Polynomial Order	1st	2nd	3rd	4th
	(1)	(2)	(3)	(4)
<i>Bandwidth: 3 Months</i>				
Rated	-0.00	-0.00	0.00	0.00
	(0.01)	(0.01)	(0.02)	(0.02)
Rated x Rating Difference	0.15**	0.15**	0.15**	0.15**
	(0.02)	(0.02)	(0.02)	(0.02)
Observations	33,782	33,782	33,782	33,782
<i>Bandwidth: 6 Months</i>				
Rated	-0.00	-0.00	-0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Rated x Rating Difference	0.15**	0.15**	0.15**	0.15**
	(0.01)	(0.01)	(0.01)	(0.01)
Observations	49,097	49,097	49,097	49,097
<i>Bandwidth: 12 Months</i>				
Rated	-0.01	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Rated x Rating Difference	0.14**	0.14**	0.14**	0.14**
	(0.01)	(0.01)	(0.01)	(0.01)
Observations	78,373	78,373	78,373	78,373
<i>Bandwidth: 24 Months</i>				
Rated	-0.01*	-0.01	-0.00	-0.00
	(0.01)	(0.01)	(0.01)	(0.01)
Rated x Rating Difference	0.13**	0.13**	0.13**	0.13**
	(0.01)	(0.01)	(0.01)	(0.01)
Observations	129,376	129,376	129,376	129,376
<i>Bandwidth: 48 Months</i>				
Rated	-0.00	0.00	-0.00	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Rated x Rating Difference	0.12**	0.12**	0.12**	0.12**
	(0.01)	(0.01)	(0.01)	(0.01)
Observations	216,691	216,691	216,691	216,691

Notes: The unit of observation is a user rating. The dependent variable is the rating score. Each panel reports results from a sample that consists of ratings that occurred within the corresponding bandwidth of the Bros' rating. Each column reports results from specifications that model *Relative Month* trends using a global polynomial of the corresponding order. All models include individual, beer, and month-of-sample effects; and linear controls for *Review Order* and *Average User Score at Time of Rating*. Standard errors are reported in parentheses and are clustered by member profile. One and two stars indicate 10% and 5% percent significance, respectively.

Table 4: Placebo Estimates

Global Polynomial Order	1st	2nd	3rd	4th
	(1)	(2)	(3)	(4)
<i>Bandwidth: 3 Months</i>				
Rated	0.01 (0.02)	0.01 (0.02)	0.00 (0.02)	0.00 (0.02)
Rated x Rating Difference	0.02 (0.02)	0.02 (0.02)	0.02 (0.02)	0.02 (0.02)
Observations	31,405	31,405	31,405	31,405
<i>Bandwidth: 6 Months</i>				
Rated	-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)
Rated x Rating Difference	0.02** (0.01)	0.02** (0.01)	0.02** (0.01)	0.02** (0.01)
Observations	55,595	55,595	55,595	55,595
<i>Bandwidth: 12 Months</i>				
Rated	-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)
Rated x Rating Difference	0.02** (0.01)	0.02** (0.01)	0.02** (0.01)	0.02** (0.01)
Observations	101,751	101,751	101,751	101,751
<i>Bandwidth: 24 Months</i>				
Rated	-0.01* (0.01)	-0.01** (0.01)	-0.02** (0.01)	-0.02** (0.01)
Rated x Rating Difference	0.03** (0.01)	0.03** (0.01)	0.03** (0.01)	0.03** (0.01)
Observations	186,455	186,455	186,455	186,455
<i>Bandwidth: 48 Months</i>				
Rated	-0.01** (0.00)	-0.01** (0.00)	-0.01** (0.01)	-0.01** (0.01)
Rated x Rating Difference	0.02** (0.01)	0.02** (0.01)	0.02** (0.01)	0.02** (0.01)
Observations	326,726	326,726	326,726	326,726

Notes: The estimates correspond to those reported in Table 3 except the data are compiled and analyzed such that NorthYorkSammy is the Bros. NorthYorkSammy is the Beer Advocate member that has supplied the most ratings.

Table 5: Estimates of How the Effect of an Expert Review Interacts with User Experience

Sample	Str. Neg. (1)	Wk. Neg. (2)	Wk. Pos. (3)	Str. Pos. (4)	Full (5)
Rated	-0.056 (0.038)	-0.083** (0.027)	0.021 (0.025)	0.226** (0.057)	0.009 (0.016)
User Experience / 10	0.001 (0.001)	-0.002 (0.001)	0.000 (0.001)	0.001 (0.002)	0.000 (0.001)
Rated x User Experience / 10	-0.001 (0.001)	-0.002 (0.001)	-0.000 (0.001)	-0.003 (0.002)	-0.001 (0.001)
Rated x Rating Difference					0.155** (0.022)
Rating Difference x User Experience / 10					0.000 (0.001)
Rated x Rating Difference x User Experience / 10					0.000 (0.001)
Individual Effects	Yes	Yes	Yes	Yes	Yes
Review Order	Yes	Yes	Yes	Yes	Yes
Avg. User Score at Time of Rating	Yes	Yes	Yes	Yes	Yes
Beer Effects	Yes	Yes	Yes	Yes	Yes
Linear Relative Month Trends	Yes	Yes	Yes	Yes	Yes
Month-of-Sample Effects	Yes	Yes	Yes	Yes	Yes
Observations	9,240	10,558	13,561	4,622	37,981

Notes: The unit of observation is a user rating. The dependent variable is the rating score. All observations are from ratings that occurred within a 6-month bandwidth of the Bros' rating. The first four columns report estimates from samples that are restricted to beers that received an expert review of a certain type, as described in the column headings. To ease exposition, *User Experience* was centered at zero prior to the regression (by subtracting the mean from each observation). I exclude observations entered by users that have already made more than 500 reviews. The yes/no items at the bottom of the table indicate the inclusion of various control variables. Standard errors are reported in parentheses and are clustered by member profile. One and two stars indicate 10% and 5% percent significance, respectively.

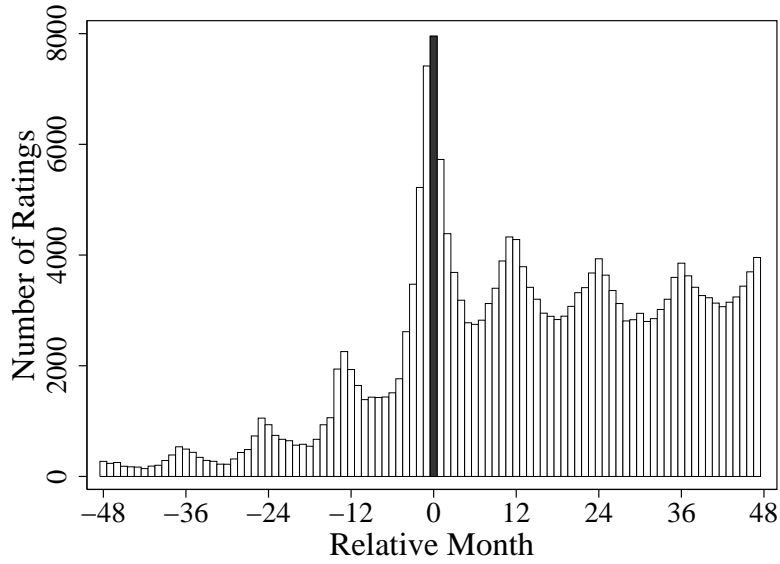


Figure 5: **Histogram of Relative Months.** *Relative Months* records the difference between the time when the user rating was entered and the time when the Bros' rating was entered (as measured in 30-day bins). The darkly shaded bar refers to the first 30 days following the Bros' rating.