MIS Quarterly

RESEARCH ARTICLE

ON SELF-SELECTION BIASES IN ONLINE PRODUCT REVIEWS

Nan Hu
School of Business, Stevens Institute of Technology, 1 Castle Point Terrace,
Hoboken, NJ 07030 U.S.A. {nhu4@stevens.edu}

Paul A. Pavlou
Fox School of Business, Temple University, 1801 N. Broad Street,
Philadelphia, PA 19122 U.S.A. {pavlou@temple.edu}

Jie Zhang
College of Business, University of Texas, Arlington, 701 S. Nedderman Drive,
Arlington, TX 76019 U.S.A. {jiezhang@uta.edu}

Appendix A

The J-Shaped Distribution of Online Product Reviews

A random sample of product information and their corresponding consumer reviews were collected from Amazon in 2005 using Amazon Web Service (AWS) for more than 77,000 books, DVDs, and Videos from Amazon (Table A1).

Table A1. Descriptive Statistics for Amazon's Data

Product Category	Number of Products	Number of Reviews	Mean of Reviews
Books	32,878	967,075	4.02
DVDs	17,978	2,034,552	4.19
Videos	28,983	1,248,992	3.99

Figure A1 shows the distribution of the average rating for all books, DVDs, and videos on Amazon.com.

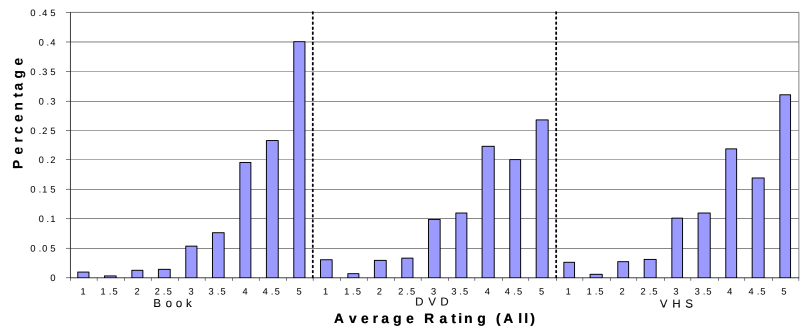

Figure A1. The Distribution of Online Product Reviews for All Books, DVDs, and Videos on Amazon.com

To verify that the J-shaped distribution does not vary over time, we split all Amazon's reviews into four equal groups (initial stage, early stage, late stage, final stage)[1] based on their posts. The J-shaped distribution persists (Figure A2).

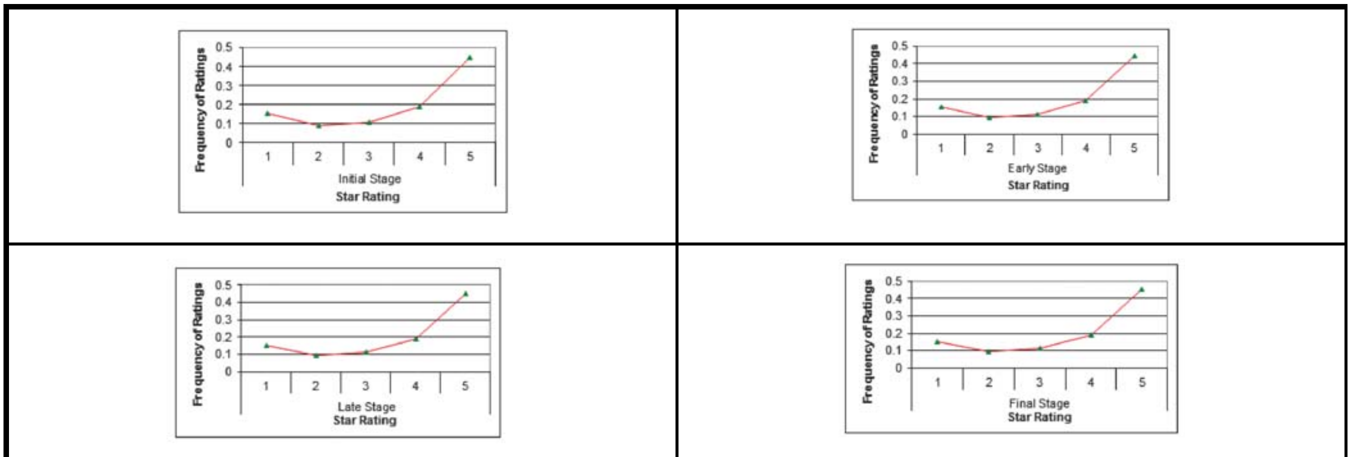

Figure A2. The Distribution of Online Product Reviews Over Time

Figure A3 shows the distribution of three randomly selected products in each of the three popular product categories with more than 2,000 reviews. The results show that these products also have a bimodal, asymmetric, left-skewed distribution, thus confirming that the observed J-shaped distribution is not due to the small number of product reviews.

[1]These four stages and their labels are proposed in a relative sense. Specifically, the initial stage reflects the earliest stage the product was first released; the final stage is the latest period. The J-shaped pattern still holds irrespective of periods and the absolute age of the reviews.

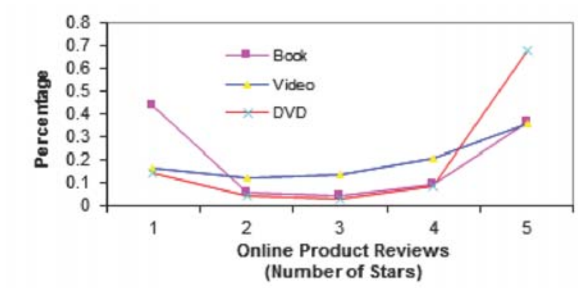

Figure A3. The Distribution of Online Product Reviews for Books, DVDs, and Videos with n > 2,000 Reviews

Figure A4 presents the distributions of online product reviews for freeware on Download.com with fewer than 20 reviews, which follow a bimodal, J-shaped distribution. This is a strong indication that the distribution is *not* normal. Also, to show that the J-shaped distribution applies to products with a different mean rating, Figure A4 also shows the distribution of products with a mean of 3.5-star (roughly in the middle) and 4-star (right hand side).[2]

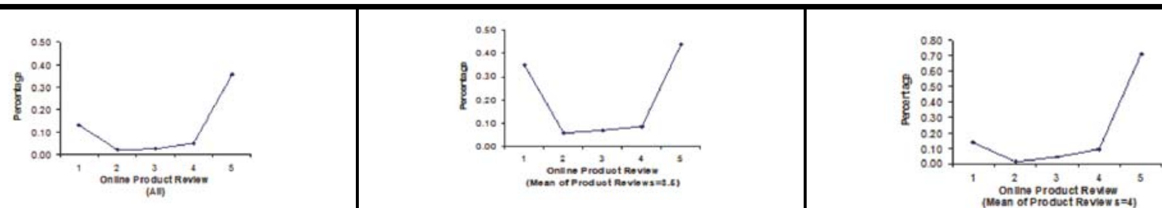

Figure A4. Distribution of Online Product Reviews of Freeware on Download.com (≤ 20 reviews)

Figure A5 shows the bimodal distribution of online reviews for products with a mean star rating of 3 and 4.[3]

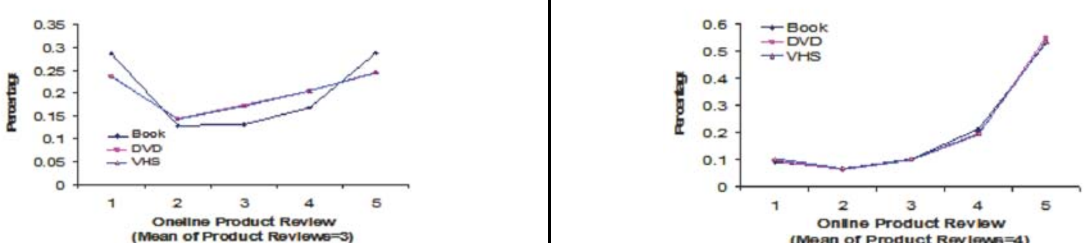

Figure A5. The Distribution of Books, DVDs, and Videos with Mean Rating of Three- or Four-Stars

---

[2]We calculated the mean of each product's online reviews based on all observations. Since the mean can be decimal number, such as 1.2 or 2.1, we used the following classification: If the mean of the online product reviews was between 1 and 1.5, we classified the product into a group with a mean = 1; if the mean was between 1.5 and 2, we classified the product into a group with a mean = 1.5, etc. We also tried to classify products with a mean rating around 1 (e.g., between 0.9 and 1.1), into a group with mean = 1, and the results were very similar.

[3]Bimodality is not due to a truncated distribution since consumers cannot write reviews higher than five or lower than one star. Graphic plots of the mean of product ratings other than 3.0 or 2.5 stars reveal that there are fewer consumers writing a review with a five-star rating than those writing a review with a four-star rating. In those cases, however, there are still fewer consumers writing a three-star review.

Figure A5 shows that as the mean rating increases, the U-shaped distribution becomes more left-skewed, thus turning into a J-shaped distribution. Besides the graphical inspections (Figures A1–A5), we formally tested whether the distribution of online product reviews is normal using the Kolmogorov-Snirnov test (Chakravarti et al. 1967), which examines if a sample comes from a normally distributed population.[4] This test at the individual product level showed that nearly all products do *not* follow a normal distribution.

To further test if the distribution of online product reviews for individual products is bimodal, the nonparametric DIP test (Hartigan and Hartigan 1985) was used. The DIP test is a measure of departure from unimodality; the DIP statistic for a unimodal distribution approaches zero, while the DIP statistic of a bimodal distribution approaches a positive constant.[5] We obtain the DIP statistics using ***R***.[6] The DIP test show that 90.17% of products have a distribution of online reviews that is neither unimodal nor normal. Virtually all products with a mean star rating between 1.5 and 4 stars do *not* follow a unimodal distribution. Even most products with a mean star rating around 5 stars do not follow a unimodal distribution.

To further establish the J-shaped distribution, we test the quadratic Equation A1 (Anderson 1998) using the following model:

$$f_{ij} = \alpha_{0j} + \alpha_{1j} s_{ij} + \alpha_{2j} s_{ij}^{\ 2} + \sum \beta_{mj} x_{mj} + \varepsilon_{ij} \tag{A1}$$

where f_{ij} is the number of product reviews with score i, for item j, $s_{ij} = i \in \{1, 2, 3, 4, 5]$ is review score for product j, x_{mj} are other variables that might influence the rating of item j, such as price, mean rating and product category, and ε_{ij} is an error term. The null hypothesis to accept the bimodal distribution is given by H_0: $\alpha_1 < 0$ and $\alpha_2 > 0$.

To account for potential differences in product characteristics and means, we ran a fixed effect model by regressing the number of product reviews on the star rating (number of stars). As a robustness check, we ran separate regressions for different groups composed of products from the same product category and with similar mean rating of product reviews, and we then estimated the mean coefficient across these categories. The results are qualitatively the same. The results when all products are pulled together show a significant negative $\alpha_1 = -40.54$ and a significant positive value $\alpha_2 = 9.09$. Therefore, the estimated quadratic curve of [A1] is symmetric in terms of the rating $s_i = 2.2$, which lies to the left of the median point of 3, implying that online product reviews for virtually all products collected from Amazon within the range of 1–5 star ratings have a J-shaped (left skewed bimodal) distribution. As an additional robustness check, we ran this J-shaped test on individual product level. Our results indicate that 83.1% of books, 82.2% of DVD, and 76.0% of VHS follow a J-shaped distribution.

---

[4] We also employed the Cramer-von Mises (Thode 2002) and the Anderson-Darling (Stephens 1974) tests with similar results.

[5] Besides DIP test, other nonparametric tests of unimodality are available, such as the excess mass test (Muller and Sawitzki 1991), and the Silverman (1981) test. The DIP and excess mass tests are equivalent in the one-dimensional case as the excess mass statistic is exactly twice the DIP statistic (Cheng and Hall 1998). However, the DIP test is simpler and more conservative (Cheng and Hall 1998; Henderson et al. 2000). Therefore, if the DIP test shows a large percentage of online product reviews to have a bimodal distribution, the other tests are likely to provide even more pronounced results.

[6] For more information on ***R***, see http://www.r-project.org.

Appendix B

Lab Experiment on Self-Selection Biases in Online Product Reviews

We asked 218 subjects to review four products (music CD, movie DVD, Access software, IS textbook) as well as their review, purchase intentions, and purchase importance on a 1–5 scale. These products were chosen to vary in terms of product category (music, movies, software, textbook), prior ownership, familiarity, importance, and price level ($10–$250). For each product, we assured that the subjects were familiar with each product. They were asked to hear all twelve 30-second clips of the music CD, and they were also asked to watch movie "Titanic" if they did not. Subjects used Access as part of a required class assignment, while for them the IS textbook was a required class textbook. Subjects were asked to rate the product and also report whether (1) they had already owned the product, (2) the importance of the product to them, and (3) their intention and passion to report a product review. The purpose is to compare the distribution of online product reviews from almost all respondents in the lab experiment with that of reviews on Amazon.com. Table B1 shows the number of respondents for each product, and descriptive statistics of their responses. The response rate of over 92% shows that the products were reviewed by almost all participants in our sample, and nonresponse bias tests showed that the nonrespondents did not differ from the respondents.

Table B1. Sample Characteristics

Product	Number of Subjects	Number of Reviews on Amazon	Prior Ownership (Percentage)	Intention to Review (Mean – STD)	Purchase Importance (Mean – STD)
Music CD	197	157	8%	2.03 (0.95)	2.23 (0.99)
Movie DVD	199	2107	35%	2.25 (1.00)	2.69 (1.16)
Software	203	10	66%	2.18 (0.97)	2.59 (1.06)
Textbook	201	13	83%	2.54 (1.03)	3.67 (1.17)

Amazon's and the experiment's mean ratings for each of the four products are quite different (Table B2). While the music CD and textbook are rated higher on Amazon ($p < .001$), the movie DVD is rated higher among the experiment's respondents; finally, the mean rating for the Access software is roughly the same between Amazon and the experiment.

Table B2. Differences in Mean Star Ratings of Product Reviews Survey Versus Amazon

Product	Field Study	Amazon	Equality Test (p-value)
Music CD	3.25	3.90	0.0000
Movie DVD	4.09	3.56	0.0000
Access Software	3.53	3.60	0.7931
IS Textbook	3.51	4.79	0.0000

Besides graphical differences (Figures 1 and 2), we specified a system of equations to isolate the self-selection biases:

$$Rating = \alpha_0 + \alpha_1 Ownership + \alpha_2 Intention + \alpha_3 Importance + \Sigma_{i=1}^{3} \alpha_i ProductDummy_i + \varepsilon \quad \text{[B1]}$$

$$Intention = \beta_0 + \beta_1 Rating + \beta_2 Rating^2 + \beta_3 Importance + \Sigma_{i=1}^{3} \beta_i ProductDummy_i + \eta \quad \text{[B2]}$$

where *Rating* = The respondent's star rating on a five-point Likert-type scale anchored between one star and five stars.
Ownership = Binary variable whether the resondent already owns the product.
Intention = The respondent's intention to write a product review at Amazon.com on a five-point scale.
Importance = The respondent's assessment of how salient the purchase is on a five-point scale.
ProductDummy = Represents the fixed effects due to potential differences across the three categories.

In Equation [B1], we used *Ownership* as a proxy for acquisition bias. Equation [B1] summarizes the predictors of the star rating. The utility theory suggests that prior ownership is expected to increase the mean rating. In fact, the mean rating of subjects who already owned the product was significantly higher than those who do not ($p < .05$) (Table B3). Purchase importance was controlled for its positive effect on the star rating since consumers who perceive the purchase to be important are more likely to be positively predisposed toward the product and to write a positive review.

In Equation [B1], intention to write a review was used as a proxy for underreporting bias. If consumers are more likely to write a review when they are either extremely satisfied or dissatisfied, *Intention* is positively correlated with *ExtremeStarRating* (a dummy variable =1 when consumers leave a one-star or five-star rating and 0 otherwise). Wald's test in Table B4 based on Equation [B3] supports this positive correlation ($\chi^2 = 40.98$, $p < .0001$). We estimate Equation [2] in which the subjects' intention to report a review on Amazon is determined by the subject's star rating. *Rating* and *Rating²* are included to account for a potential nonlinear effect of the respondent's star rating and her intentions to write a review.

Table B3. Differences in Star Ratings of Online Product Reviews Based on Prior Ownership

Product	Prior Ownership (mean)	No Ownership (mean)	Difference: Sign	Difference: Difference	t-value	p-value
Music CD	4.20	3.18	+	1.02	4.3600	<.0001
Movie DVD	4.26	4.00	+	0.26	2.2000	0.029
Access Software	3.64	3.30	+	0.34	2.8600	0.0047
IS Textbook	3.58	3.20	+	0.38	2.8700	0.0046

Table B4. Likelihood of Writing an Extreme Product Review

	Coefficient	Wald Chi-Square	p-value
Intercept	-3.68	105.24	<.0001
Intention	0.68	40.98	<.0001
DVD Dummy	1.31	20.74	<.0001
Software Dummy	-0.04	0.01	0.9081
Textbook Dummy	-1.13	7.90	0.0050

Since the *Intention* variable in Equation [B1] is a linear combination of other variables in Equation [B2], we adopted the limited-information maximum likelihood (LIML) estimation to simultaneously estimate the system of equations for acquisition bias and underreporting bias.[7] The results are reported in Table B5.[8]

$$ExtremeStarRating = \gamma_0 = \gamma_1 Intention + \sum_{i=2}^{4} \gamma_i ProductDummy_i + \varepsilon \qquad [B3]$$

[7]Asymptotically, 2SLS and LIML estimators have the same distribution (Anderson 2005). Even though it is easier to compute 2SLS, LIML was used because (1) the parameter estimation method of simultaneous equation models was based on ML that is commonly believed to yield superior estimators (Anderson 2005); (2) LIML takes into account the covariances of the error terms; (3) 2SLS estimator treats the components of β asymmetrically, which runs contrary to simultaneous equations (Anderson 2005, p. 9).

[8]As a robustness check, besides estimating the system of Equations 1 and 2, we also estimated these equations independently using both OLS and logistic regression. Furthermore, we estimated Equation 2 with the sample composed of all respondents, or those respondents who already owned the product before. All of these tests have qualitatively the same results.

Table B5. Acquisition Bias and Underreporting Bias Based on System Equations [1] and [2]

Acquisition Bias		Underreporting Bias	
Variable	**Parameter *α***	**Variable**	**Parameter *β***
Intercept	2.23***	Intercept	1.10***
Prior Ownership	0.13**	Rating	-0.19*
Intention to Write Review	0.05	Rating²	0.04**
Purchase Importance	0.28***	Purchase Importance	0.44***
CD_Dummy	0.28***	CD_Dummy	0.13*
DVD_Dummy	0.94***	DVD_Dummy	0.07
Software_Dummy	0.37***	Software_Dummy	0.11
N	800	N	800
Adjusted R²	28.9%	Adjusted R²	29.9%
DW	1.98	DW	2.07

***$p < 0.001$; **$p < 0.05$; *$p < 0.10$

For acquisition bias, after controlling for purchase importance, prior ownership ($\alpha_1 = 0.13$, $p < 0.05$) was positively linked to rating, and intention to report a review was not significant ($\alpha_3 = 0.05$, $p > 0.10$). For underreporting bias, the negative coefficient of *Rating* was marginally significant ($\beta_1 = -0.19$, $p < 0.10$), and the *Rating*² coefficient ($\beta_2 = 0.04$, $p < 0.05$) was positive and significant. That implies that consumers are more likely to write a review when they are either satisfied or dissatisfied, while they are the least likely to report a review when their star rating is moderate ($-\frac{\partial\beta_2}{2\partial\beta_1} = 2.4$ stars) according to Equation [B2]. These results reinforce that acquisition bias (reflected through prior ownership) and underreporting bias (reflected through higher intentions to write a product review) result in the observed J-shaped distribution.

Figure B1 shows that there is a stark contrast between the means of Amazon's online product reviews and those of the lab experiment reviews for the exact same four products (music, movie, software, textbook). Amazon's online product reviews resemble a J-shaped distribution, whereas the lab experimental data follow a unimodal, roughly normal distribution. Furthermore, while the majority of Amazon's online product reviews are extreme or polarized (one-star or five-star), the majority of the lab experiment's product reviews (over 90%) are moderate (two-star, three-star, or four-star). Finally, while Amazon's online product reviews are mostly positive (five-star), the lab experiment's results are balanced across all star ratings between one-star and five-stars.

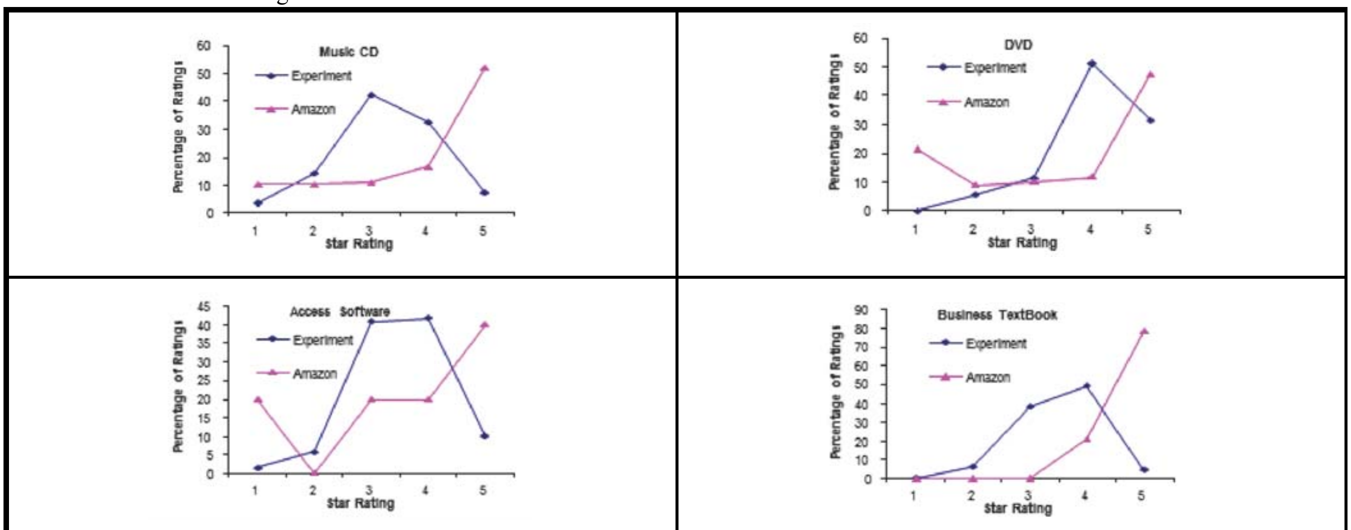

Figure B1. Comparison of Amazon's and the Lab Experiment's Distributions of Online Product Reviews

In addition to showing that the lab experiment's data follow a normal distribution (Figure B1), we isolated the two self-selection biases by plotting the distribution of the respondents with prior ownership (capturing acquisition bias) and respondents with high intentions (≥ 3) to write an online review on Amazon.com (capturing underreporting bias). As shown in Figure B2 for the movie DVDs (the other products follow a similar pattern and are omitted for brevity), prior ownership shifts the distribution toward higher ratings. Besides, selecting those respondents with high intentions to write a review largely omits the moderate star ratings, resulting in a distribution that resembles Amazon's observed J-shaped distribution. In sum, the combination of the two self-selection biases is shown to jointly shift a normal distribution of all respondents to a left-skewed bimodal distribution, which resembles Amazon's J-shaped distribution.

Figure B2. Distribution of Online Product Reviews from Amazon Versus Experimental Field Study Results

Appendix C

Comparison of Models with and Without Self-Selection Biases in Online Product Reviews

To compare the proposed "dual mode" model against the three competing models (mean average model, weighted mean average model, and extreme rating controlled model), we randomly selected another 10,000 books, DVDs, and videos from our Amazon sample starting in July of 2005, with the SAS random function. For each product, we collected its price, sales rank, and all online product reviews for several months on three-day intervals.[9] Therefore, since we have panel data from July 2005 to January 2006, we can compare which model has the highest power in terms of predicting future product sales using longitudinal secondary data. The models, whose predictive validity is shown in Table C1, are compared below.

Proposed Dual Mode Model

$$\ln(SalesRank_{i+1}) = \beta_0 + \beta_1 AvgRating_t + \beta_2 X_{Lt} + \beta_3 X_{Ut} + \beta_4 StdevRating_t \\ + \beta_5 \ln(SalesRank_t) + \beta_6 \ln(Price_t) + \beta_7 \ln(NumRev_t) \\ + \beta_8 Book\_Dummy + \beta_9 DVD\_Dummy + \varepsilon_t$$

Model 1 (Simple Mean)

$$\ln(SalesRank_{i+1}) = \alpha_{01} + \alpha_{11} Mean\_Rating_i + \alpha_{21} \ln(Sales\_Rank_i) + \alpha_{31} \ln(Price_i) + \alpha_{41} \ln(Num\_Rev_i) \\ + \alpha_{51} Book\_Dummy + \alpha_{61} DVD\_Dummy + \varepsilon_{i1}$$

Model 2a (Weighted Mean a*)

$$\ln(SalesRank_{i+1}) = \alpha_{02} + \alpha_{12} Weighted\_Mean\_Rating_{1i} + \alpha_{22} \ln(SalesRank_i) + \alpha_{32} \ln(Price_i) + \alpha_{42} \ln(Num\_Rev_i) \\ + \alpha_{52} Book\_Dummy + \alpha_{62} DVD\_Dummy + \varepsilon_{i2}$$

Model 2b (Weighted Mean b*)

$$\ln(SalesRank_{i+1}) = \alpha_{03} + \alpha_{13} Weighted\_Mean\_Rating_{2i} + \alpha_{23} \ln(SalesRank_i) + \alpha_{33} \ln(Amazon\_Price_i) \\ + \alpha_{43} \ln(Num\_Rev_i) + \alpha_{53} Book\_Dummy + \alpha_{63} DVD\_Dummy + \varepsilon_{i3}$$

Model 2c (Weighted Mean c*)

$$\ln(SalesRank_{i+1}) = \alpha_{04} + \alpha_{14} Weighted\_Mean\_Rating_{3i} + \alpha_{24} \ln(SalesRank_i) + \alpha_{34} \ln(Price_i) + \alpha_{44} \ln(Num\_Rev_i) \\ + \alpha_{54} Book\_Dummy + \alpha_{64} DVD\_Dummy + \varepsilon_{i4}$$

$*Weighted\_Mean\_Rating_{1i} = Avg((HelpfulReviews/TotalReviews) * ReviewRating)$
$*Weighted\_Mean\_Rating_{2i} = Sum((HelpfulReviews/TotalReviews) * ReviewRating)/Sum(HelpfulReviews/TotalReviews)$
$*Weighted\_Mean\_Rating_{3i} = Sum(HelpfulReviews * ReviewRating) / Sum(HelpfulReviews)$

[9]Ideally, we would like to collect data on a daily basis because price, sales, and online product reviews change on a daily basis. However, due to the instability of Amazon's web service, it often takes more than three days to collect a batch of data. Thus, to ensure that we get clean copy for each batch of data, we used a three-day instead of a one-day interval.

Model 3 (One-Star and Five-Star Model)

$$\ln(SalesRank_{i+1}) = \alpha_{05} + \alpha_{15}Percent_{1star} + \alpha_{25}Percent_{5star} + \alpha_{35}\ln(SalesRank_i) + \alpha_{45}\ln(Price_i) + \alpha_{55}\ln(Num\_Rev_i) + \alpha_{65}Book\_Dummy + \alpha_{75}DVD\_Dummy + \varepsilon_{i5}$$

As shown in Table C1, controlling for previous sales rank,[10] price, and the total number of product reviews, the model with self-selection controlled explains a substantial amount of the variance (R^2 adjusted = 77.29%) in future product sales, which is significantly higher than all other models ($p < .0001$). The mean of the online product reviews has a significant effect, explaining .16% of the variance in future product sales. The two dual modes X_L and X_U *also* have significant effects ($p < .001$),[11] explaining .64% and .33% of the variance in future product sales, respectively. Interestingly, the variance explained by X_L (lower mode) is almost twice as much as that explained by the upper mode X_U. These findings are consistent with the literature (e.g., Chevalier and Mayzlin 2006) that suggests that consumers pay more attention to negative reviews (which are generally captured by X_L) compared to positive reviews (which are captured by X_U). Finally, the STD is statistically significant (p<.05), explaining .034% of the variance. This is consistent with Clemons et al. (2004) who showed that the variance of online product reviews affects future sales.

Table C1. Model Comparisons

	Proposed Model X	Model 1	Model 2a	Model 2b	Model 2c	Model 3
Mean_Product Reviews	-0.0513***	-0.0351***	-0.0232**	-0.0272**	-0.0194*	
X_L	0.0573***					
X_U	-0.0685***					
STD†	0.0086*					
Percent(1-star reviews)						0.1088*
Percent(5-star reviews)						-0.0452
Ln (Current Sales Rank)	0.7114***	0.7206***	0.7218***	0.721***	0.7215***	0.7208***
ln(Price)	0.0780***	0.0999***	0.100***	0.0989**	0.0986***	0.099***
ln (# of Product Reviews)	-0.0869***	-0.0533***	-0.056***	-0.050***	-0.051***	-0.0528***
Book Dummy	0.311***	0.334***	0.333***	0.331***	0.390***	0.333***
DVD Dummy	-0.154***	-.1042***	-0.105***	-0.108***	-0.108***	-0.105***
Intercept	3.7890***	3.5234***	3.448***	3.482***	3.444***	3.3936***
Adjusted R^2	**77.29%**	**75.18%**	**75.17%**	**75.25%**	**75.24%**	**75.17%**
Difference in R^2		**2.11%**	**2.12%**	**2.04%**	**2.05%**	**2.12%**
F-Value		23.444***	23.555***	22.666***	22.778***	23.555***
N	7573	7573	7573	7573	7573	7573

†STD is weighted by helpvote/totalvote.
****p* < .001; ***p* < .01; **p* < .05; +*p* < .10. All *p*-values are two-sided.

We used the following equation for calculating the significance between two regression models:

$$F_{(kx-ki),(n-kx-ki)} = \frac{\left[R^2\left(Model\_X\right) - R^2\left(Model\_i\right)\right] / K_x - K_i}{\left[\left(1 - R^2\left(Model\_X\right)\right)\right] / \left(N - K_x - K_i\right)} \qquad [7]$$

[10]The time difference between $t + 1$ and t is 130 days. We also tested other time lag values (e.g., 100 days, 110 days), which yielded similar results.

[11]Following Aigner (1971), the variance explained was decomposed among the independent variables by multiplying the standardized regression coefficients by the correlation of the independent variables with the dependent variable.

where K_x is the number of independent variables in the proposed Model X
K_i is the number of independent variables in the competing Model I
N is the sample size

There are several criteria that can be used to choose among competing models, such as the Adjusted R^2, Akaike information criterion (AIC), Schwarz information criterion (SIC), Mallow's C_p criterion, and forecast χ^2 (chi-square). These criteria aim at minimizing the residual sum of squares, or increasing the adjusted R^2 value. The AIC imposes a harsher penalty than the R^2, while the SIC imposes an even harsher penalty than the AIC. However, as argued by Diebold and Kilian (2001), no criterion is necessarily superior. For simplicity, we evaluated the performance of the various competing models by comparing their adjusted R^2, which is the most widely used criterion for model comparison. In terms of other comparisons beyond the F-test (Equation 7), following Davidson and MacKinnon (1993, p. 456): "For linear regression models, with or without normal errors, there is of course no need to look at likelihood, W, and LR at all, since no information is gained from doing so over and above what is already contained in F." Therefore, we did not perform other comparisons for the nested models. For non-nested model comparison, we also used the Davidson-MacKinnon J test, which showed similar results.

The proposed self-selection controlled model explains at least 2% higher variance compared to the five competing models (which roughly explain about the same variance).[12] This difference in variance explained is statistically significant ($p < .0001$), as the F-tests in Table C1 attest. Besides the high F-values that denote that the 2% improvement in variance explained is statistically significant, from a practical standpoint, one may question this improvement. However, it is important to recognize that the great majority of the variance is explained by the control variables. Specifically, the current sales rank explains 52% of the variance,[13] the number of online product reviews explains 4.55%, the product dummies explain 7%, and price only explains .11%. Given these influential control variables, the variance explained by the new proposed independent variables (X_L, X_U, STD) is also substantial from a practical standpoint, attesting to the need for including these distributional parameters when predicting future product sales.

Table C2. Comparison between Proposed Modes and Percentage of Polarized Reviews

	Regression Model
Mean_Product Reviews	-0.0906*
X_L	0.0563***
X_U	-0.0688***
STD	0.0083*
% 1-star reviews	-0.1895$^{N/S}$ (p = .1739)
% 5-star reviews	0.0343 $^{N/S}$ (p = 0.684)
ln (Current Sales Rank)	0.71152***
ln(Price)	0.0778***
ln (# of Reviews)	-0.0866***
Book Dummy	0.318***
DVD Dummy	-0.152***
Intercept	3.95***
Adjusted R^2	77.29%
N	7573

[12]Interestingly, none of the three proposed weighted means of online product reviews is superior to the simple mean rating. Perhaps this is because consumers only observe the simple mean and do not otherwise process the number of useful reviews.

[13]We also ran the same regression models (Table C1) by omitting the current sales rank as a control variable, and the results were very similar (Model X outperformed all others by over 2%). However, since the variance explained in future sales is substantially lower (circa 35%) when omitting current sales rank, we only report the results with all control variables.

While the proposed prediction model with the proposed X_L and X_U modes is superior to the one with the polarized (one-star and five-star) reviews (Model 3), we still wanted to have a direct comparison of their joint impact. Therefore, we ran a regression model in which we included both X_L and X_U and also the percentage of one-star and five-star reviews (Table C2). The density mass in the bimodal distribution of online product reviews that obtains the X_L and X_U are different from the percentage of one-star and five-star reviews, allowing us to simultaneously include them in a regression model. As shown in Table C2, both the percentage of polarized (1-star and 5-stars) reviews become insignificant when X_L and X_U are included in the regression model. Accordingly, the inclusion of polarized reviews did not improve the variance explained (77.29%) in future product sales. These results attest to the superiority of the X_L and X_U parameters obtained by the DIP test versus the percentage of polarized reviews.

Appendix D

Proofs

Proof of Lemma 1:

With a window as described in Equation [6], the density function of reviews is derived as

$$f\left(q_i|\theta_i\right)=\begin{cases} 0 & \text{if } \underline{\delta}\le u_i-E\left(u_i|\theta_i\right)\le\bar{\delta} \\ \dfrac{\phi\left(\frac{q_i-q^e-\rho\sigma_q\left(\theta_i-\mu_\theta\right)/\sigma_\theta}{\sqrt{1-\rho^2}\sigma_q}\right)\Big/\sqrt{1-\rho^2}\sigma_q}{\Phi\left(\frac{\underline{\delta}+\left(q^e-q\right)}{\sqrt{1-\rho^2}\sigma_q}\right)+\left(1-\Phi\left(\frac{\bar{\delta}+\left(q^e-q\right)}{\sqrt{1-\rho^2}\sigma_q}\right)\right)} & \text{Otherwise} \end{cases} \tag{D1}$$

The expected rating from consumer i who purchases the product with $E(u_i|\theta_i)\ge 0$ that is, $\theta_i\ge\alpha(p, q^e)$ with $\alpha\left(p,q^e\right)=\frac{\sigma_\theta\left(p-q^e\right)+\rho\sigma_q\mu_\theta}{\sigma_\theta+\rho\sigma_q}$, is derived by

$$\begin{aligned}
&E(r_i\,|\,\theta_i\ge\alpha)=E(q_i\,|\,\theta_i\ge\alpha)=\int q_i f(q_i\,|\,\theta_i)dq_i \\
&=q+\rho\frac{\sigma_q}{\sigma_\theta}(\theta_i-\mu_\theta)+\Bigg(\int_{-\infty}^{\underline{\delta}+q^e+\rho\frac{\sigma_q}{\sigma_\theta}(\theta_i-\mu_\theta)}\frac{\frac{q_i}{\sqrt{1-\rho^2}\sigma_q}\phi\left(\frac{q_i-q-\rho\frac{\sigma_q}{\sigma_\theta}(\theta_i-\mu_\theta)}{\sqrt{1-\rho^2}\sigma_q}\right)}{\Phi\left(\frac{\underline{\delta}+q^e-q}{\sqrt{1-\rho^2}\sigma_q}\right)+\left(1-\Phi\left(\frac{\bar{\delta}+q^e-q}{\sqrt{1-\rho^2}\sigma_q}\right)\right)}dq_i+ \\
&\quad\int_{\bar{\delta}+q^e+\rho\frac{\sigma_q}{\sigma_\theta}(\theta_i-\mu_\theta)}^{+\infty}\frac{\frac{q_i}{\sqrt{1-\rho^2}\sigma_q}\phi\left(\frac{q_i-q-\rho\frac{\sigma_q}{\sigma_\theta}(\theta_i-\mu_\theta)}{\sqrt{1-\rho^2}\sigma_q}\right)}{\Phi\left(\frac{\underline{\delta}+q^e-q}{\sqrt{1-\rho^2}\sigma_q}\right)+\left(1-\Phi\left(\frac{\bar{\delta}+q^e-q}{\sqrt{1-\rho^2}\sigma_q}\right)\right)}dq_i\Bigg) \\
&=q+\rho\frac{\sigma_q}{\sigma_\theta}(\theta_i-\mu_\theta)+ \\
&\quad\sqrt{1-\rho^2}\sigma_q\Bigg(\int_{-\infty}^{\frac{\underline{\delta}+q^e-q}{\sqrt{1-\rho^2}\sigma_q}}\frac{\frac{q_i-q-\rho\frac{\sigma_q}{\sigma_\theta}(\theta_i-\mu_\theta)}{\sqrt{1-\rho^2}\sigma_q}\phi\left(\frac{q_i-q-\rho\frac{\sigma_q}{\sigma_\theta}(\theta_i-\mu_\theta)}{\sqrt{1-\rho^2}\sigma_q}\right)}{\Phi\left(\frac{\underline{\delta}+q^e-q}{\sqrt{1-\rho^2}\sigma_q}\right)+\left(1-\Phi\left(\frac{\bar{\delta}+q^e-q}{\sqrt{1-\rho^2}\sigma_q}\right)\right)}d\left(\frac{q_i-q-\rho\frac{\sigma_q}{\sigma_\theta}(\theta_i-\mu_\theta)}{\sqrt{1-\rho^2}\sigma_q}\right)+ \\
&\quad\int_{\frac{\bar{\delta}+q^e-q}{\sqrt{1-\rho^2}\sigma_q}}^{+\infty}\frac{\frac{q_i-q-\rho\frac{\sigma_q}{\sigma_\theta}(\theta_i-\mu_\theta)}{\sqrt{1-\rho^2}\sigma_q}\phi\left(\frac{q_i-q-\rho\frac{\sigma_q}{\sigma_\theta}(\theta_i-\mu_\theta)}{\sqrt{1-\rho^2}\sigma_q}\right)}{\Phi\left(\frac{\underline{\delta}+q^e-q}{\sqrt{1-\rho^2}\sigma_q}\right)+\left(1-\Phi\left(\frac{\bar{\delta}+q^e-q}{\sqrt{1-\rho^2}\sigma_q}\right)\right)}d\left(\frac{q_i-q-\rho\frac{\sigma_q}{\sigma_\theta}(\theta_i-\mu_\theta)}{\sqrt{1-\rho^2}\sigma_q}\right)\Bigg) \\
&=q+\rho\tfrac{\sigma_q}{\sigma_\theta}\left(\theta_i-\mu_\theta\right)+\sqrt{1-\rho^2}\sigma_q\Lambda_0\left(q^e\right)-p \qquad \text{where } \Lambda_0\left(q^e\right)=\frac{\phi\left(\frac{\bar{\delta}+\left(q^e-q\right)}{\sqrt{1-\rho^2}\sigma_q}\right)-\phi\left(\frac{\underline{\delta}+\left(q^e-q\right)}{\sqrt{1-\rho^2}\sigma_q}\right)}{\Phi\left(\frac{\underline{\delta}+\left(q^e-q\right)}{\sqrt{1-\rho^2}\sigma_q}\right)+\left(1-\Phi\left(\frac{\bar{\delta}+\left(q^e-q\right)}{\sqrt{1-\rho^2}\sigma_q}\right)\right)}
\end{aligned} \tag{D2}$$

Q.E.D.

Proof of Proposition 1:

Following [D2], we derive the expected review score from a consumer i with prior quality expectation q^e:

$$\begin{aligned} E(r_i) &= E\left(E(q_i - p)\right) \\ &= E\left(q + \rho \tfrac{\sigma_q}{\sigma_\theta}(\theta_i - \mu_\theta) + \sqrt{1-\rho^2}\,\sigma_q \Lambda_0\left(q^e\right) - p\right) \\ &= q + \rho\sigma_q \frac{\phi\left(\alpha\left(p,q^e\right)\right)}{1-\Phi\left(\alpha\left(p,q^e\right)\right)} + \sqrt{1-\rho^2}\,\sigma_q \Lambda_0\left(q^e\right) - p \end{aligned} \quad \text{[D3]}$$

If we let $\lambda(x) = \frac{\phi(x)}{1-\Phi(x)}$, then [D3] can be expressed as

$$E(r) = q + \sqrt{1-\rho^2}\,\sigma_q \Lambda_0\left(q^e\right) + \rho\sigma_q \lambda\left(\alpha\left(p,q^e\right)\right) - p \quad \text{[D4]}$$

$\rho\sigma_q\lambda\left(\alpha\left(p,q^e\right)\right)$ and $\sqrt{1-\rho^2}\,\sigma_q\Lambda_0\left(q^e\right)$ are the acquisition bias and the underreporting bias, respectively. If $\rho \neq 1$, getting rid of the underreporting bias requires $\Lambda_0 = 0$, which can only be achieved when $\underline{\delta}$ and $\overline{\delta}$ are symmetric in terms of the difference between the realized quality and prior, that is, $\overline{\delta} + \underline{\delta} = 2(q - q^e)$, if $\overline{\delta} = \underline{\delta} = 0$ is impossible due to the prevalence of underreporting bias.

The variance of the rating of consumer i who purchases the product, $E(\mu_i | \theta_i) \geq p$, is obtained by

$$\begin{aligned} &Var(r_i \mid \theta_i \geq \alpha) = Var(q_i \mid \theta_i \geq \alpha) = E(q_i - E(q_i \mid \theta_i \geq \frac{\sigma_\theta(p-q^e) + \rho\sigma_q\mu_\theta}{\sigma_\theta + \rho\sigma_q}))^2 \\ &= E(q_i - q - \rho\frac{\sigma_q}{\sigma_\theta}(\theta_i - \mu_\theta) - \sqrt{1-\rho^2}\,\sigma_q\Lambda_0)^2 \\ &= E(q_i - q - \rho\frac{\sigma_q}{\sigma_\theta}(\theta_i - \mu_\theta))^2 - (1-\rho^2)\sigma_q^2\Lambda_0^2 \\ &= \int_{-\infty}^{\underline{\delta}+q^e+\rho\frac{\sigma_q}{\sigma_\theta}(\theta_i-\mu_\theta)} \frac{\dfrac{(q_i - q - \rho\frac{\sigma_q}{\sigma_\theta}(\theta_i-\mu_\theta))^2}{\sqrt{1-\rho^2}\,\sigma_q}\phi(\dfrac{q_i - q - \rho\frac{\sigma_q}{\sigma_\theta}(\theta_i-\mu_\theta)}{\sqrt{1-\rho^2}\,\sigma_q})}{\Phi(\dfrac{\underline{\delta}+q^e-q}{\sqrt{1-\rho^2}\,\sigma_q}) + (1-\Phi(\dfrac{\overline{\delta}+q^e-q}{\sqrt{1-\rho^2}\,\sigma_q}))} dq_i + \\ &\int_{\overline{\delta}+q^e+\rho\frac{\sigma_q}{\sigma_\theta}(\theta_i-\mu_\theta)}^{+\infty} \frac{\dfrac{(q_i - q - \rho\frac{\sigma_q}{\sigma_\theta}(\theta_i-\mu_\theta))^2}{\sqrt{1-\rho^2}\,\sigma_q}\phi(\dfrac{q_i - q - \rho\frac{\sigma_q}{\sigma_\theta}(\theta_i-\mu_\theta)}{\sqrt{1-\rho^2}\,\sigma_q})}{\Phi(\dfrac{\underline{\delta}+q^e-q}{\sqrt{1-\rho^2}\,\sigma_q}) + (1-\Phi(\dfrac{\overline{\delta}+q^e-q}{\sqrt{1-\rho^2}\,\sigma_q}))} dq_i - (1-\rho^2)\sigma_q^2\Lambda_0^2 \\ &= (1-\rho^2)\sigma_q^2(1 + \Lambda_1 - \Lambda_0^2) \end{aligned} \quad \text{[D5]}$$

Where:

$$\Lambda_1 = \frac{\dfrac{\overline{\delta}+q^e-q}{\sqrt{1-\rho^2}\,\sigma_q}\phi(\dfrac{\overline{\delta}+q^e-q}{\sqrt{1-\rho^2}\,\sigma_q}) - \dfrac{\underline{\delta}+q^e-q}{\sqrt{1-\rho^2}\,\sigma_q}\phi(\dfrac{\underline{\delta}+q^e-q}{\sqrt{1-\rho^2}\,\sigma_q})}{\Phi(\dfrac{\underline{\delta}+q^e-q}{\sqrt{1-\rho^2}\,\sigma_q}) + (1-\Phi(\dfrac{\overline{\delta}+q^e-q}{\sqrt{1-\rho^2}\,\sigma_q}))}$$

By the Law of Total Variance and Equation [D5]

$$
\begin{aligned}
\sigma_r^2 &= VarE\left(\left(r_1|\theta_i\right)\right) + E\left(Var\left(r_1|\theta_i\right)\right) \\
&= Var\left(q + \rho\frac{\sigma_q}{\sigma_\theta}(\theta_i - \mu_\theta) + \sqrt{1-\rho^2}\sigma_q\Lambda_0\right) + (1-\rho^2)\sigma_q^2(1+\Lambda_1-\Lambda_0^2) \\
&= \rho^2\sigma_q^2(1-\lambda(\alpha)(\lambda(\alpha)-\alpha)) + (1-\rho^2)\sigma_q^2(1+\Lambda_1-\Lambda_0^2)
\end{aligned}
\qquad \text{[D6]}
$$

Q.E.D.

Proof of Corollary 1:

Given p, q^e, and $\Lambda_0(q^e) = \dfrac{\phi\left(\frac{\bar{\delta}+(q^e-q)}{\sqrt{1-\rho^2}\sigma_q}\right) - \phi\left(\frac{\underline{\delta}+(q^e-q)}{\sqrt{1-\rho^2}\sigma_q}\right)}{\Phi\left(\frac{\underline{\delta}+(q^e-q)}{\sqrt{1-\rho^2}\sigma_q}\right) + \left(1-\Phi\left(\frac{\bar{\delta}+(q^e-q)}{\sqrt{1-\rho^2}\sigma_q}\right)\right)}$

$$
\begin{aligned}
&sign\left(\frac{\partial E(r)}{\partial\bar{\delta}}\right) = sign\frac{\partial\Lambda_0\left(q^e\right)}{\partial\bar{\delta}} \\
&= sign\left(\frac{\partial\phi\left(\frac{\bar{\delta}+(q^e-q)}{\sqrt{1-\rho^2}\sigma_q}\right)}{\partial\bar{\delta}}\left[\Phi\left(\frac{\underline{\delta}+(q^e-q)}{\sqrt{1-\rho^2}\sigma_q}\right) + \left(1-\Phi\left(\frac{\bar{\delta}+(q^e-q)}{\sqrt{1-\rho^2}\sigma_q}\right)\right)\right]\right. \\
&\qquad\left. -\left[\Phi\left(\frac{\bar{\delta}+(q^e-q)}{\sqrt{1-\rho^2}\sigma_q}\right) - \phi\left(\frac{\underline{\delta}+(q^e-q)}{\sqrt{1-\rho^2}\sigma_q}\right)\right]\frac{\partial\left[\Phi\left(\frac{\underline{\delta}+(q^e-q)}{\sqrt{1-\rho^2}\sigma_q}\right)\right] + \left(1-\Phi\left(\frac{\bar{\delta}+(q^e-q)}{\sqrt{1-\rho^2}\sigma_q}\right)\right)}{\partial\bar{\delta}}\right) \\
&= sign\left(\phi'\left(\frac{\bar{\delta}+(q^e-q)}{\sqrt{1-\rho^2}\sigma_q}\right)\left[\Phi\left(\frac{\underline{\delta}+(q^e-q)}{\sqrt{1-\rho^2}\sigma_q}\right) + \left(1-\Phi\left(\frac{\bar{\delta}+(q^e-q)}{\sqrt{1-\rho^2}\sigma_q}\right)\right)\right]\right. \\
&\qquad\left. +\phi\left(\frac{\bar{\delta}+(q^e-q)}{\sqrt{1-\rho^2}\sigma_q}\right)\left[\phi\left(\frac{\bar{\delta}+(q^e-q)}{\sqrt{1-\rho^2}\sigma_q}\right) - \phi\left(\frac{\underline{\delta}+(q^e-q)}{\sqrt{1-\rho^2}\sigma_q}\right)\right]\right) \\
&= sign\left(-\phi\left(\frac{\bar{\delta}+(q^e-q)}{\sqrt{1-\rho^2}\sigma_q}\right)\left(\frac{\bar{\delta}+(q^e-q)}{\sqrt{1-\rho^2}\sigma_q}\right)\left[\Phi\left(\frac{\underline{\delta}+(q^e-q)}{\sqrt{1-\rho^2}\sigma_q}\right) + \left(1-\Phi\left(\frac{\bar{\delta}+(q^e-q)}{\sqrt{1-\rho^2}\sigma_q}\right)\right)\right]\right. \\
&\qquad\left. +\phi\left(\frac{\bar{\delta}+(q^e-q)}{\sqrt{1-\rho^2}\sigma_q}\right)\left[\phi\left(\frac{\bar{\delta}+(q^e-q)}{\sqrt{1-\rho^2}\sigma_q}\right) - \phi\left(\frac{\underline{\delta}+(q^e-q)}{\sqrt{1-\rho^2}\sigma_q}\right)\right]\right) \\
&= sign\left(\phi\left(\frac{\bar{\delta}+(q^e-q)}{\sqrt{1-\rho^2}\sigma_q}\right) - \phi\left(\frac{\underline{\delta}+(q^e-q)}{\sqrt{1-\rho^2}\sigma_q}\right)\right. \\
&\qquad\left. -\left(\frac{\bar{\delta}+(q^e-q)}{\sqrt{1-\rho^2}\sigma_q}\right)\left[\Phi\left(\frac{\underline{\delta}+(q^e-q)}{\sqrt{1-\rho^2}\sigma_q}\right) + \left(1-\Phi\left(\frac{\bar{\delta}+(q^e-q)}{\sqrt{1-\rho^2}\sigma_q}\right)\right)\right]\right)
\end{aligned}
$$

Thus the sign of $\frac{\partial E(r)}{\partial\bar{\delta}}$ depends on the comparison of the $\Lambda_0\left(q^e\right)$ and $\frac{\bar{\delta}+(q^e-q)}{\sqrt{1-\rho^2}\sigma_q}$. When $\frac{\bar{\delta}+(q^e-q)}{\sqrt{1-\rho^2}\sigma_q} < \Lambda_0\left(q^e\right)$, that is, $\bar{\delta}+\left(q^e-q\right) < \Lambda_0\left(q^e\right)\sqrt{1-\rho^2}\sigma_q$, the last expression is positive, $\frac{\partial E(r)}{\partial\bar{\delta}} > 0$. That is, the expected rating increases with $\bar{\delta}$ and vice versa.

Similar to the above results, the sign of $\frac{\partial E(r)}{\partial \underline{\delta}}$ depends on the sign of $\Lambda_0\left(q^e\right)$.

$$sign\left(\frac{\partial E(r)}{\partial \underline{\delta}}\right) = sign\left(\frac{\partial \Lambda_0\left(q^e\right)}{\partial \underline{\delta}}\right)$$

$$= sign\left(-\frac{\partial \phi\left(\frac{\underline{\delta}+\left(q^e-q\right)}{\sqrt{1-\rho^2}\sigma_q}\right)}{\partial \underline{\delta}}\left[\Phi\left(\frac{\underline{\delta}+\left(q^e-q\right)}{\sqrt{1-\rho^2}\sigma_q}\right)+\left(1-\Phi\left(\frac{\bar{\delta}+\left(q^e-q\right)}{\sqrt{1-\rho^2}\sigma_q}\right)\right)\right]-\left[\phi\left(\frac{\bar{\delta}+\left(q^e-q\right)}{\sqrt{1-\rho^2}\sigma_q}\right)\right.\right.$$

$$\left.\left.-\phi\left(\frac{\underline{\delta}+\left(q^e-q\right)}{\sqrt{1-\rho^2}\sigma_q}\right)\right]\frac{\partial\left[\Phi\Phi\left(\frac{\underline{\delta}+\left(q^e-q\right)}{\sqrt{1-\rho^2}\sigma_q}\right)+\left(1-\Phi\left(\frac{\bar{\delta}+\left(q^e-q\right)}{\sqrt{1-\rho^2}\sigma_q}\right)\right)\right]}{\partial \underline{\delta}}\right)$$

$$= sign\left(-\phi'\left(\frac{\underline{\delta}+\left(q^e-q\right)}{\sqrt{1-\rho^2}\sigma_q}\right)\left[\Phi\left(\frac{\underline{\delta}+\left(q^e-q\right)}{\sqrt{1-\rho^2}\sigma_q}\right)+\left(1-\Phi\left(\frac{\bar{\delta}+\left(q^e-q\right)}{\sqrt{1-\rho^2}\sigma_q}\right)\right)\right]\right.$$

$$\left.-\phi\left(\frac{\underline{\delta}+\left(q^e-q\right)}{\sqrt{1-\rho^2}\sigma_q}\right)\left[\Phi\left(\frac{\bar{\delta}+\left(q^e-q\right)}{\sqrt{1-\rho^2}\sigma_q}\right)-\phi\left(\frac{\underline{\delta}+\left(q^e-q\right)}{\sqrt{1-\rho^2}\sigma_q}\right)\right]\right)$$

$$= sign\left(\left(\frac{\underline{\delta}+\left(q^e-q\right)}{\sqrt{1-\rho^2}\sigma_q}\right)\left[\Phi\left(\frac{\underline{\delta}+\left(q^e-q\right)}{\sqrt{1-\rho^2}\sigma_q}\right)+\left(1-\Phi\left(\frac{\bar{\delta}+\left(q^e-q\right)}{\sqrt{1-\rho^2}\sigma_q}\right)\right)\right]-\left[\phi\left(\frac{\bar{\delta}+\left(q^e-q\right)}{\sqrt{1-\rho^2}\sigma_q}\right)\right.\right.$$

$$\left.\left.-\phi\left(\frac{\underline{\delta}+\left(q^e-q\right)}{\sqrt{1-\rho^2}\sigma_q}\right)\right]\right)$$

When $\frac{\bar{\delta}+\left(q^e-q\right)}{\sqrt{1-\rho^2}\sigma_q} < \Lambda_0\left(q^e\right)$, that is, $\bar{\delta}+\left(q^e-q\right) < \Lambda_0\left(q^e\right)\sqrt{1-\rho^2}\sigma_q$, the expected rating decreases with $\underline{\delta}$, $\frac{\partial E(r)}{\partial \underline{\delta}} < 0$, and vice versa.
Q.E.D.

Proof to Corollary 2:

(i) By Proposition 1, the expected consumer rating in a single period is $q+\rho\sigma_q\lambda\left(\alpha\left(p,q^e\right)\right)+\sqrt{1-\rho^2}\sigma_q\Lambda_0\left(q^e\right)-p$ and variance

$$\rho^2\sigma_q^2\left(1-\lambda\left(\alpha\left(p,q^e\right)\right)\left(\lambda\left(\alpha\left(p,q^e\right)\right)-\alpha\left(p,q^e\right)\right)+\left(1-\rho^2\right)\sigma_q^2\left(1+\Lambda_1\left(q^e\right)-\Lambda_0^2\left(q^e\right)\right)\right).$$

Since consumers in different periods are updating their quality beliefs based on Equation [10], $q_{t+1}^2 = \omega q + (1-\omega)E\left(r_t\right)$. If consumers can fully overcome both types of biases, then $q_{t+1}^e = q$. Plug into Proposition 1, then the mean of the rating series keeps the same over time: $q+\rho\sigma_q\lambda\left(\alpha\left(p,q\right)\right)+\sqrt{1-\rho^2}\sigma_q\Lambda_{-}-p$ and variance also does not change over time $\rho^2\sigma_q^2\left(1-\lambda\left(\alpha\left(p,q\right)\right)\left(\lambda\left(\alpha\left(p,q\right)\right)-\alpha\left(p,q\right)\right)\right)+\left(1-\rho^2\right)\sigma_q^2\left(1+\Lambda_1-\Lambda_0^2\right)$. Thus the rating series is stationary.

(ii) If consumers form quality expectation without the biases in the previous periods, $q_{t=1}^e = q$ and $E\left(r_t\right) = q+\rho\sigma_q\lambda\left(\alpha\left(p,q\right)\right)+\sqrt{1-\rho^2}\sigma_q\Lambda_0-p$,
$Var\left(r_t\right) = \rho^2\sigma_q^2\left(1-\lambda\left(\alpha\left(p,q\right)\right)\left(\lambda\left(\alpha\left(,q\right)\right)-\alpha\left(p,q\right)\right)\right)+\left(1-\rho^2\right)\sigma_q^2\left(1+\Lambda_1-\Lambda_0^2\right)$. The correlation between ratings in different time periods $corr(r_t, r_{t+1})$ thus rating series are independent.
Q.E.D.

References

Aigner, D. J. 1971. *Basic Econometrics*, Upper Saddle River, NJ: Prentice Hall

Anderson, E. W. 1998. "Customer Satisfaction and Word of Mouth," *Journal of Service Research* (1:1), pp. 5-17.

Anderson, T. W. 2005. "Origins of the Limited Information Maximum Likelihood and Two-Stage Least Squares Estimators," *Journal of Econometrics* (127), pp. 1-16.

Chakravarti, I. M., Laha, R. G., and Roy, J. 1967. *Handbook of Methods of Applied Statistics*, New York: John Wiley and Sons.

Chen, P-Y., Wu, S-Y., and Yoon, J. 2004. "The Impact of Online Recommendations and Consumer Feedback on Sales," in *Proceedings of the 25th International Conference on Information Systems*, Washington, D.C., pp. 711-724.

Cheng, M-Y., and Hall, P. 1998. "Calibrating the Excess Mass and Dip Tests of Modality," *Journal of the Royal Statistical Society B* (60:3), pp. 579-589

Chevalier, J., and Mayzlin, D. 2006. "The Effect of Word of Mouth on Sales: Online Book Reviews," *Journal of Marketing Research* (43:3), pp. 345-354.

Chintagunta, P., Erdem, T., Rossi, P. E., and Wedel, M. 2006. "Structural Modeling in Marketing: Review and Assessment," *Marketing Science* (25:6), pp. 604-616.

Clemons, E. K., Gao, G., and Hitt, L. M. 2006. "When Online Reviews Meet Hyper Differentiation: A Study of the Craft Beer Industry," *Journal of Management Information Systems* (23:2), pp. 149-171.

Davidson, R., and MacKinnon, J. G. 1993. *Estimation and Inference in Econometrics*, Oxford, UK: Oxford University Press.

Diebold, F., and Kilian, L. 2001. "Measuring Predictability: Theory and Macroeconomic Applications," *Journal of Applied Econometrics* (16:6), pp. 657-669.

Hartigan, J. A., and Hartigan, P. M. 1985. "The DIP Test of Unimodality," *Annals of Statistics* (13:1), pp. 70-84.

Henderson, R., Diggle, P. J., and Dobson, A. 2000. "Joint Modeling of Longitudinal Measurements and Event Time Data," *Biostatistics* (1), pp. 465-480.

Muller, D. W., and Sawitzki, G. 1991. "Excess Mass Estimates and Tests of Multimodality," *Journal of the American Statistical Association* (86), pp. 738-746.

Silverman, B. W. 1981. "Using Kernel Density Estimates to Investigate Multimodality," *Journal of the Royal Statistical Society* (43:1), pp. 97-99.

Stephens, M. A. 1974. "EDF Statistics for Goodness of Fit and Some Comparisons," *Journal of the American Statistical Association* (69), pp. 730-737.

Thode Jr., H. C. 2002. *Testing for Normality*, New York: Marcel Dekker.