Reviews Without a Purchase:
Low Ratings, Loyal Customers, and Deception

Web Appendix

Eric T. Anderson (Northwestern University)

Duncan I. Simester (MIT)

January 2014

List of Tables and Figures

Definition of Variables

Review Characteristics

No Confirmed Transaction	1 if the review does not have a confirmed transaction; and 0 otherwise.
Review Date	The date of the review measures in years since the date of the very first review.
Word Count	The number of words in the review text comments.
Word Length	The average length of the words in the review text comments.
Family	1 if the review contains words describing members of the family; and 0 otherwise.
Repeated !!	1 if the review contains repeated exclamation points; 0 otherwise.

Item Characteristics

Any Fit	1 if the review contained words describing the fit of the product; and 0 otherwise.
Any Feel	1 if the review contained words describing the fit of the product; and 0 otherwise.
Product Age	The number of years between the date of the review and the first transaction for that product.
Prior Units	The number of units of that item sold in the 12 months before the review date.
Average Selling Price	The average selling price of the product in the 12–months before and 12-months after the review date.
% Sold in Retail Stores	The percentage of all unit sales of the item that occurred in the firm's physical retail stores.
% Sold via Catalog	The percentage of all unit sales of the item that occurred in the firm's catalog.

Reviewer Characteristics

Number of Reviews	The number of reviews written by that household.
Items Purchased	The number of items purchased by that household.
Average Item Price	The average price paid for items purchased by that household.
Overall Discount Received	The average % discount received for items purchased by that household.
Discount Frequency	The percentage of items purchased at a discount for that household.
Return Rate	The number of items that the household returned as a percentage of their total purchases.
Years Since First Order	The number of years between the review date and the first purchase by any members of that household.
Number of Children	The number of children in the household.
Married	1 if the head of household is married; and 0 otherwise.
Age	The age of the head of the household.
Estimated Home Value	The estimated value of the household's home.
Est. Household Income	The estimated household income.
Graduate Degree	1 if the head of household has a graduate degree; and zero otherwise.

Summary Statistics

	Mean	Std Dev	Sample Size
Review Characteristics			
No Confirmed Transaction	0.05	0.21	325,869
Review Date	2.01	1.02	325,869
Word Count	52.88	40.53	325,869
Word Length	4.15	0.52	325,405
Family	0.19	0.39	325,869
Repeated !!	0.05	0.21	325,869
Any Fit	0.48	0.50	325,869
Any Feel	0.55	0.50	325,869
Item Characteristics			
Product Age	4.71	5.32	325,754
Prior Units	31,974.41	45,134.87	325,643
Average Selling Price	52.87	55.05	325,643
% Sold in Retail Stores	0.14	0.11	325,754
% Sold via Catalog	0.41	0.14	325,754
Reviewer Characteristics			
Number of Reviews	3.73	8.35	325,869
Items Purchased	169.59	266.97	325,869
Average Item Price	40.05	16.01	325,869
Overall Discount Received	0.08	0.09	325,869
Discount Frequency	0.20	0.19	325,832
Return Rate	0.17	0.18	325,832
Years Since First Order	12.94	6.46	325,869
Number of Children	0.48	0.97	297,099
Married	0.73	0.44	315,488
Age	56.99	14.17	274,767
Estimated Home Value	279,857.70	268,397.20	294,967
Est. Household Income	103,186.80	60,791.58	313,847
Graduate Degree	0.32	0.47	315,488

Pair-wise Correlations: Review Characteristics

	Review Date	Word Count	Word Length	Family	Repeated	Any Fit	Any Feel
No Confirmed Transaction	-0.084**	0.096**	-0.017**	0.011**	0.022**	-0.017**	-0.015**
Review Date		0.035**	-0.016**	-0.004*	-0.011**	0.046**	-0.004*
Word Count			-0.188**	0.148**	0.053**	0.246**	0.138**
Word Length				-0.042**	0.024**	-0.127**	0.068**
Family					0.025**	-0.041**	-0.022**
Repeated !!						-0.013**	-0.006**
Any Fit							0.022**

The table reports Pearson correlation coefficients. The unit of analysis is a review. The sample size is 325,869 except for the correlations with Word Count, where the correlations are 325,405 (a very small number of reviews do not have any text comments). *Significantly different from zero, $p<0.05$, and **significantly different from zero, $p<0.01$.

Pair-wise Correlations: Item Characteristics

	Prior Units	Average Selling Price	% Sold in Retail Stores	% Sold via Catalog
Product Age	-0.275**	0.028*	-0.174**	0.513**
Prior Units		-0.135**	-0.195**	0.268**
Average Selling Price			0.003	0.092**
% Sold in Retail Stores				-0.650**

The table reports Pearson correlation coefficients. The unit of analysis is an item. We restrict attention to items that appear in at least one review. The sample size is 8,270. *Significantly different from zero, $p < 0.05$, and **significantly different from zero, $p < 0.01$.

Pair-wise Correlations: Reviewer Characteristics

	Items Purchased	Average Item Price	Overall Discount Received	Discount Frequency	Return Rate	Years Since First Order	Number of Children	Married	Age	Estimated Home Value	Estimated Household income	Graduate Degree
Number of Reviews	-0.215**	-0.039**	0.091**	0.088**	0.053**	0.060**	-0.007**	-0.001	0.001	-0.003	0.001	0.016**
Items Purchased		-0.099**	0.121**	0.112**	0.155**	0.416**	-0.017**	0.073**	0.059**	0.101**	0.105**	0.112**
Average Item Price			-0.262**	-0.229**	0.096**	-0.091**	-0.058**	-0.033**	-0.022**	0.018**	0.007**	0.002
Overall Discount Received				0.910**	0.068**	0.015**	0.073**	-0.0001	-0.078**	-0.041**	-0.027**	-0.008**
Discount Frequency					0.076**	0.005*	0.081**	0.002	-0.091**	-0.037**	-0.024**	-0.006**
Return Rate						0.065**	0.015	-0.004	-0.045**	0.027**	0.024**	0.015**
Years Since 1st Order							-0.078**	0.101**	0.223**	0.127**	0.142**	0.186**
Number of Children								0.178**	-0.321**	0.036**	0.100**	-0.006**
Married									-0.125**	0.136**	0.199**	0.088**
Age										-0.009**	-0.076**	-0.061**
Estimated Home Value											0.452**	0.188**
Est. Household Income												0.259**

The table reports Pearson correlation coefficients. The unit of analysis is a reviewer. The sample size varies from 178,786 to 213,205.
*Significantly different from zero, $p<0.05$, and **significantly different from zero, $p<0.05$.

Product Ratings: Logistic and OLS Models

		Rating = 1	Product Rating
Review Characteristics	No Confirmed Transaction	0.0431** (0.0035)	-0.2732** (0.0277)
	Review Date	0.0019* (0.0008)	-0.0261** (0.0055)
	Word Count (10s)	0.0005** (0.0002)	-0.0031** (0.0008)
	Word Length	0.0003 (0.0007)	0.0040 (0.0044)
	Family	-0.0058** (0.0015)	0.0724** (0.0078)
	Repeated !!	0.0036 (0.0022)	0.0133 (0.0109)
	Any Fit	0.0025 (0.0014)	-0.0707** (0.0079)
	Any Feel	-0.0106** (0.0014)	0.0631** (0.0079)
Item Characteristics	Product Age (years)	-0.0019** (0.0003)	0.0167** (0.0023)
	Prior Units (10,000s)	-0.0018** (0.0004)	0.0123** (0.0021)
	Average Selling Price ($10)	-0.0002 (0.0002)	0.0001 (0.0011)
	% Sold in Retail Stores	0.0135 (0.0121)	-0.0639 (0.0847)
	% Sold via Catalog	0.0546** (0.0137)	-0.2202* (0.0940)
Reviewer Characteristics	Number of Reviews	0.0003** (0.0001)	-0.0018** (0.0003)

Items Purchased (100s)	-0.0041**	0.0343**
	(0.0003)	(0.0011)
Average Item Price ($10)	-0.0009	0.0080**
	(0.0005)	(0.0018)
Overall Discount Received	0.0243	-0.0306**
	(0.0176)	(0.0654)
Discount Frequency	-0.0112	0.0128
	(0.0077)	(0.0295)
Return Rate	0.0031	-0.0449**
	(0.0032)	(0.0130)
Years Since First Order	0.0075**	-0.0485**
	(0.0002)	(0.0005)
Number of Children	-0.0041**	0.0215**
	(0.0006)	(0.0026)
Married	0.0028*	0.0224**
	(0.0011)	(0.0052)
Age (10s years)	0.0052**	-0.0373**
	(0.0004)	(0.0020)
Est. Home Value ($100,000s)	0.0002	0.0065**
	(0.0002)	(0.0009)
Est. Household Income ($100,000s)	-0.0021*	0.0102*
	(0.0008)	(0.0043)
Graduate Degree	-0.0050**	0.0325**
	(0.0010)	(0.0051)
R^2 or Pseudo R^2	0.0702	0.0773

Model 1 reports the marginal effects from a logistic model where the dependent variables is a binary variable indicating whether the review has a rating = 1. Model 2 reports coefficient from an OLS model where the dependent variable is the product rating. In both models the unit of analysis is a review and the sample size is 273,129 (the sample size reflects missing demographic variables). Both models include a constant, but this is omitted from the table. Standard errors clustered at the item level are reported in parentheses. [†]Significantly different from zero, $p<0.10$, [*]significantly different from zero, $p<0.05$ and [**]significantly different from zero, $p<0.01$.

Product Rating: Fixed Effects Models

In the table below we report the No Confirmed Transaction coefficient from the following OLS fixed effects model where the dependent variable is the product rating on review *i*:

Product Rating = **βX** + β_1*No Confirmed Transaction* + ε.

In this model **X** is the vector of fixed effects and *No Confirmed Transaction* is a binary variable identifying whether there the review did not have a confirmed transaction. The unit of analysis is a review and the sample size is The coefficient of interest is β_1 and we estimate three separate models, with either reviewer fixed effects, item fixed effects or fixed effects for the date of the review. In all three models the unit of analysis is a review and the sample size is 325,869 (although the coefficient of interest is only affected by observations for which there is variation in the *No Confirmed Transaction* within a fixed effect). In all three models the standard errors are clustered at the item level and are reported in parentheses (clustering at the reviewer or review date level results in many clusters and therefore has little impact on the standard errors).

	Reviewer Fixed Effects	Item Fixed Effects	Timing Fixed Effects
No Confirmed Transaction	-0.1523[**]	-0.2720[**]	-0.2157[**]
	(0.0304)	(0.0166)	(0.0249)
R^2	0.2879	0.0983	0.0128

[**]Significantly different from zero, $p<0.01$.

Fit and Feel Analysis at Each Review Rating Level

		Without a Confirmed Transaction	With a Confirmed Transaction	Difference
Any Fit Word	Rating = 1	41.25%	49.89%	-8.64%**
				(1.28%)
	Rating = 2	51.27%	58.46%	-7.19%**
				(1.53%)
	Rating = 3	49.60%	60.90%	-11.29%**
				(1.42%)
	Rating = 4	50.89%	55.30%	-4.41%**
				(1.09%)
	Rating = 5	40.94%	43.55%	-2.61%**
				(0.528%)
Any Feel Word	Rating = 1	30.60%	32.03%	-1.43%
				(1.19%)
	Rating = 2	42.83%	40.75%	-2.08%
				(1.53%)
	Rating = 3	42.63%	44.44%	-1.81%
				(1.44%)
	Rating = 4	53.83%	56.03%	-2.20%*
				(1.09%)
	Rating = 5	56.98%	59.01%	-2.03%**
				(0.52%)
Number of Reviews	Rating = 1	1,680	16,363	
	Rating = 2	1,102	16,752	
	Rating = 3	1,262	20,077	
	Rating = 4	2,179	52,594	
	Rating = 5	9,536	204,324	

The table reports averages for each measure separately according to the rating on the product review. Standard errors are in parentheses. *Significantly different from zero, p<0.05 and **significantly different from zero, p<0.01.

Replicating the Fit and Feel Analysis
Including Controls for Each of the Alternative Explanations

		Without a Confirmed Transaction	With a Confirmed Transaction	Difference
Any Fit Words	Within Item Analysis	44.95%	45.93%	-0.97% (0.66%)
	Within Reviewer Analysis	45.19%	52.26%	-7.07%** (0.78%)
	Matching at the Sub-Category Level	44.63%	47.71%	-3.07%** (0.53%)
	Excluding Store Purchases	45.24%	49.02%	-3.78%** (0.82%)
	Items with Few Store Purchases	38.11%	38.30%	-0.19% (1.65%)
	Underwear Items	56.25%	47.90%	8.35% (4.81%)
Any Feel Words	Within Item Analysis	52.30%	53.69%	-1.39%* (0.64%)
	Within Reviewer Analysis	52.60%	57.20%	-4.59%** (0.77%)
	Matching at the Sub-Category Level	51.72%	55.08%	-3.36%** (0.53%)
	Excluding Store Purchases	50.41%	54.40%	-4.00%** (0.81%)
	Items with Few Store Purchases	43.31%	40.59%	2.72% (1.66%)
	Underwear Items	73.21%	74.48%	-1.27% (4.20%)

The table reports averages for each measure separately for the samples of reviews with and without confirmed transactions. The sample sizes are reported in other tables. Standard errors are in parentheses. [†]Significantly different from zero, p<0.10 and [**]significantly different from zero, p<0.01.

Linguistic Deception Cues: Analysis at Each Review Rating Level

		Without a Confirmed Transaction	With a Confirmed Transaction	Difference
Word Count	Rating = 1	71.40	61.67	9.72**
				(1.12)
	Rating = 2	77.97	61.53	16.44**
				(1.36)
	Rating = 3	73.95	62.00	11.95**
				(1.29)
	Rating = 4	74.32	56.02	18.30**
				(0.96)
	Rating = 5	67.53	48.43	19.10**
				(0.40)
Word Length	Rating = 1	4.12	4.13	-0.01
				(0.01)
	Rating = 2	4.09	4.08	0.004
				(0.01)
	Rating = 3	4.07	4.05	0.02
				(0.02)
	Rating = 4	4.13	4.10	-0.03*
				(0.01)
	Rating = 5	4.11	4.18	-0.07**
				(0.01)
Family	Rating = 1	15.42%	12.63%	2.79%**
				(0.86%)
	Rating = 2	16.79%	12.17%	4.62%**
				(1.03%)
	Rating = 3	16.32%	13.62%	2.70%**
				(1.00%)
	Rating = 4	19.05%	15.99%	3.06%**
				(0.80%)
	Rating = 5	23.10%	21.00%	2.10%**
				(0.43%)

		Without a Confirmed Transaction	With a Confirmed Transaction	Difference
Repeated !!	Rating = 1	8.21%	4.35%	3.86%** (0.54%)
	Rating = 2	4.63%	2.04%	2.59%** (0.46%)
	Rating = 3	3.49%	2.00%	1.48%** (0.42%)
	Rating = 4	4.36%	2.36%	2.00%** (0.34%)
	Rating = 5	7.98%	5.83%	2.15%** (0.25%)
Number of Reviews	Rating = 1	1,680	16,363	
	Rating = 2	1,102	16,752	
	Rating = 3	1,262	20,077	
	Rating = 4	2,179	52,594	
	Rating = 5	9,536	204,324	

The table reports averages for each measure separately according to the rating on the product review. Standard errors are in parentheses. * Significantly different from zero, $p<0.05$ and ** significantly different from zero, $p<0.01$.

**Replicating the Linguistic Deception Cues Analysis
Including Controls for Each of the Alternative Explanations**

		Without a Confirmed Transaction	With a Confirmed Transaction	Difference
Word Count	Within Item Analysis	70.12	58.08	12.04[**] (0.73)
	Within Reviewer Analysis	69.86	67.52	2.34[**] (0.62)
	Matching at the Sub-Category Level	71.82	52.33	19.48[**] (0.43)
	Excluding Store Purchases	70.56	52.04	18.52[**] (0.67)
	Items with Few Store Purchases	65.26	50.51	14.75[**] (1.33)
	Underwear Items	66.97	48.54	18.44[**] (3.63)
Word Length	Within Item Analysis	4.11	4.15	-0.04%[**] (0.01%)
	Within Reviewer Analysis	4.10	4.11	-0.01% (0.01%)
	Matching at the Sub-Category Level	4.10	4.15	-0.05[**] (0.01)
	Excluding Store Purchases	4.11	4.15	-0.05[**] (0.01)
	Items with Few Store Purchases	4.13	4.20	-0.06[**] (0.02)
	Underwear Items	4.13	4.21	-0.08[†] (0.05)

Family	Within Item Analysis	19.72%	20.13%	-0.004% (0.005)
	Within Reviewer Analysis	19.94%	20.03%	-0.09% (0.62%)
	Matching at the Sub-Category Level	20.47%	18.80%	1.67%[**] (0.42%)
	Excluding Store Purchases	20.05%	17.58%	2.46%[**] (0.62%)
	Items with Few Store Purchases	22.82%	20.01%	2.82%[*] (1.36%)
	Underwear Items	12.50%	19.85%	-7.35%[†] (3.82%)
Repeated !!	Within Item Analysis	5.91%	4.27%	1.65%[**] (0.31%)
	Within Reviewer Analysis	6.91%	5.43%	2.65%[**] (0.52%)
	Matching at the Sub-Category Level	7.07%	4.76%	2.32%[**] (0.23%)
	Excluding Store Purchases	5.55%	4.51%	1.05%[**] (0.34%)
	Items with Few Store Purchases	7.32%	4.58%	2.75%[**] (0.72%)
	Underwear Items	7.14%	4.08%	3.06% (1.93%)

The table reports averages for each measure separately for the samples of reviews with and without confirmed transactions. The sample sizes are reported in other tables. Standard errors are in parentheses.
[*]Significantly different from zero, p<0.05 and [**]significantly different from zero, p<0.01.

Replicating the Results for the 10 Largest Product Categories

Product Category	Difference in Mean Rating	Difference in % of Ratings = 1	Chi-Square test	KL Divergence	Number of Reviews Without Prior Trans	Number of Reviews With Prior Trans
1	-0.45** (0.04)	7.62** (0.81)	134.98**	0.0554	814	31,251
2	-0.70** (0.04)	15.13** (0.89)	364.56**	0.1077	992	25,946
3	-0.52** (0.06)	8.04** (1.14)	99.26**	0.0754	421	16,470
4	-0.20** (0.04)	3.23** (0.83)	29.90**	0.0188	573	10,507
5	-0.27** (0.06)	3.38** (1.11)	31.77**	0.0356	317	10,504
6	-0.24** (0.04)	4.73** (0.89)	46.68**	0.0199	902	9,476
7	-0.31** (0.08)	4.48** (1.54)	19.50**	0.0342	217	9,879
8	-0.49** (0.07)	6.03** (1.30)	63.38**	0.0724	306	9,527
9	-0.19** (0.05)	1.51† (0.89)	18.71**	0.0157	482	9,094
10	-0.17** (0.06)	5.56** (1.23)	27.29**	0.0276	387	9,075

The table reports each statistic for the products categories with the 10 largest sales volumes. Standard errors are in parentheses. **Significantly different from zero, $p<0.01$. †Significantly different from zero, $p<0.10$.

Replicating the Results for Items with Different Product Ages

Product Age (Years)	Difference in Mean Rating	Difference in % of Ratings = 1	Chi-Square test	KL Divergence	Number of Reviews Without Prior Trans	Number of Reviews With Prior Trans
Less than 1	-0.30** (0.02)	7.50** (0.36)	544.64**	0.0346	5,232	82,041
1 to 2	-0.25** (0.02)	4.35** (0.46)	174.52**	0.0261	2,713	54,508
2 to 4	-0.21** (0.02)	4.33** (0.41)	176.08**	0.0214	2,989	55,863
4 to 6	-0.19** (0.03)	3.96** (0.55)	76.64**	0.0181	1,578	34,357
6 to 10	-0.27** (0.03)	4.99** (0.61)	93.09**	0.0255	1,229	33,137
Over 10	-0.21** (0.03)	3.58** (0.50)	82.68**	0.0164	1,903	50,204

The table reports each statistic when grouping the reviews according to the age of the item reviewed (at the date of the review). Standard errors are in parentheses. **Significantly different from zero, $p<0.01$.

Replicating the Results for Items with Different Sales Volumes in the Previous 12 Months

Sales Volume (quartile)	Difference in Mean Rating	Difference in % of Ratings = 1	Chi-Square test	KL Divergence	Number of Reviews Without Prior Trans	Number of Reviews With Prior Trans
1 (lowest)	-0.16** (0.02)	4.54** (0.31)	294.48**	0.0178	6,536	74,871
2	-0.26** (0.02)	5.19** (0.39)	281.50**	0.0299	3,732	77,680
3	-0.29** (0.02)	5.62** (0.41)	278.51**	0.0324	2,872	78,540
4 (highest)	-0.31** (0.02)	5.51** (0.41)	271.45**	0.0360	2,397	79,015

The table reports each statistic when grouping the reviews according to the sales volume of the item reviewed (in the year prior to the review date). **Significantly different from zero, $p<0.01$.

Within-Item Analysis (Controlling for Item Differences)

	Without a Confirmed Transaction	With a Confirmed Transaction	Difference
Average rating	4.01	4.23	-0.22[**]
			(0.02)
Rating = 1	10.85%	6.72%	4.13%[**]
			(0.42%)
Rating = 2	7.43%	5.92%	1.50%[**]
			(0.35%)
Rating = 3	9.19%	7.16%	2.03%[**]
			(0.40%)
Rating = 4	14.90%	17.78%	-2.88%[**]
			(0.50%)
Rating = 5	57.63%	62.41%	-4.78%[**]
			(0.67%)
Chi-Square test			163.69[**]
KL Divergence			0.0176

The table reports the average product ratings for reviews with and without a confirmed transaction. Reviews are first averaged at the item level and then averaged across items. The sample includes all of the 3,779 items for which we have review(s) with and without confirmed transactions. Standard errors are in parentheses. [**]Significantly different from zero, $p<0.01$.

i. We identify the 3,779 items that have review(s) with prior transactions and review(s) without prior transactions.

ii. For each product j in this set we calculate (Rating_1$_j^{with}$): among item j's reviews with prior transactions what proportion have a rating equal to 1? We then average Rating_1$_j^{with}$ across the 3,779 items and report this average (10.85%) above.

iii. We also calculate Rating_1$_j^{without}$: among item j's reviews without prior transactions what proportion have a rating equal to 1? We average Rating_1$_j^{ithout}$ across the 3,779 items and also report this average (6.72%) above.

iv. For each item j we also calculate Difference_1$_j$ = Rating_1$_j^{without}$ - Rating_1$_j^{with}$. We average Difference_1$_j$ across the 3,779 items and report this average (4.13%) above.

v. Notice that this difference calculation gives us the within item difference in the proportion of reviews that have a rating equal to 1 when the reviews do not have prior transactions compared to when they do. This explicitly controls for all item differences.

vi. We then repeat this process for all 5 rating levels and also report both a Chi-Square test of the equivalence of the two distributions and the KL Divergence.

Within-Reviewer Analysis (Controlling for Reviewer Differences)

	Without a Confirmed Transaction	With a Confirmed Transaction	Difference
Average rating	4.08	4.27	-0.18** (0.02)
Rating = 1	9.62%	5.61%	4.00%** (0.43%)
Rating = 2	7.62%	6.05%	1.57%** (0.42%)
Rating = 3	8.39%	7.78%	0.61%** (0.45%)
Rating = 4	13.71%	17.15%	-3.43%** (0.57%)
Rating = 5	60.67%	63.41%	-2.74%** (0.73%)
Chi-Square test			216.16**
KL Divergence			0.0163

The table reports the average product ratings for reviews with and without a confirmed transaction. Reviews are first averaged at the reviewer level and then averaged across reviewers. The sample includes all of the 5,234 reviewers for whom we have review(s) with and without confirmed transactions. Standard errors are in parentheses. **Significantly different from zero, p<0.01.

i. We identify the 5,234 reviewers who have written a review(s) with prior transactions and review(s) without prior transactions.

ii. For each reviewer i in this set we calculate ($Rating\_1_i^{with}$): among reviewer i's reviews with prior transactions what proportion have a rating equal to 1? We then average $Rating\_1_i^{with}$ across the 5,234 reviewers and report this average (6.49%) above.

iii. We also calculate $Rating\_1_i^{without}$: among reviewer i's reviews without prior transactions what proportion have a rating equal to 1? We average $Rating\_1_i^{without}$ across the 5,234 reviewers and also report this average (3.14%) above.

iv. For each reviewer i we also calculate $Difference\_1_i = Rating\_1_i^{without} - Rating\_1_i^{with}$. We average $Difference\_1_i$ across the 5,234 reviewers and report this average (3.35%) above.

v. Notice that this difference calculation gives us the within reviewer difference in the proportion of reviews that have a rating equal to 1 when the reviews do not have prior transactions compared to when they do. This explicitly controls for reviewer differences.

vi. We repeat this process for all 5 rating levels and also report both a Chi-Square test of the equivalence of the two distributions and the KL Divergence.

Within-Reviewer Analysis
Reviewers With At Least 3 Reviews Without Confirmed Transactions

	Without a Confirmed Transaction	With a Confirmed Transaction	Difference
Average rating	4.30	4.43	-0.13**
			(4.36)
Rating = 1	6.49%	3.14%	3.35%**
			(0.91%)
Rating = 2	3.84%	4.56%	-0.72%
			(0.83%)
Rating = 3	7.20%	7.51%	-0.31%
			(1.03%)
Rating = 4	17.68%	15.32%	2.36%
			(1.40%)
Rating = 5	64.79%	69.47%	-7.37%**
			(1.77%)
Chi-Square test			9.90*
KL Divergence			0.0147

The table reports the average product ratings for reviews with and without a confirmed transaction. Reviews are first averaged at the reviewer level and then averaged across reviewers. The sample includes all of the 226 reviewers who have at least 3 reviews without a confirmed transaction and at least one review with a confirmed transaction. Standard errors are in parentheses. *Significantly different from zero, $p<0.05$ and **significantly different from zero, $p<0.01$.

Restricting Attention to the Underwear Product Category

	Without a Confirmed Transaction	With a Confirmed Transaction	Difference
Average rating	4.26	4.41	-0.15
			(0.11)
Rating = 1	9.82%	4.95%	4.87%[*]
			(2.12%)
Rating = 2	0.90%	4.27%	-3.38%
			(1.92%)
Rating = 3	8.04%	5.31%	2.72%
			(0.45%)
Rating = 4	16.07%	16.00%	0.07%
			(2.18%)
Rating = 5	65.18%	69.46%	-4.28%
			(4.44%)
Chi-Square test			10.22[*]
Kl Divergence			0.0540

The table reports the average product ratings when we restrict attention to reviews of items in the underwear product category. The sample sizes are 112 (reviews without a confirmed transaction) and 3,088 (reviews with a confirmed transaction). Standard errors are in parentheses. [*]Significantly different from zero, $p<0.05$ and [**]significantly different from zero, $p<0.01$.

Feedback about Customer Service or Shipping Policies
Text Analysis of Customer Service or Shipping Policy Complaints

		Without a Confirmed Transaction	With a Confirmed Transaction	Difference
Customer Service Feedback	All Reviews	1.69% (0.10%)	1.81% (0.02%)	-0.12% (0.11%)
	Rating = 1	3.45% (0.10%)	3.17% (0.02%)	0.28% (0.45%)
	Rating = 2	1.81% (0.40%)	2.05% (0.11%)	-0.23% (0.44%)
	Rating = 3	1.58% (0.35%)	1.71% (0.09%)	-0.12% (0.38%)
	Rating = 4	1.33% (0.25%)	1.60% (0.05%)	-0.27% (0.27%)
	Rating = 5	1.46% (0.12%)	1.74% (0.03%)	-0.28%[*] (0.14%)
Shipping Policy Feedback	All Reviews	0.62% (0.06%)	1.10% (0.02%)	-0.47%[**] (0.08%)
	Rating = 1	1.31% (0.06%)	3.12% (0.02%)	-1.81%[**] (0.43%)
	Rating = 2	0.91% (0.29%)	1.66% (0.10%)	-0.75%[†] (0.39%)
	Rating = 3	0.56% (0.21%)	1.40% (0.08%)	-0.85%[*] (0.34%)
	Rating = 4	0.64% (0.17%)	1.05% (0.04%)	-0.40%[†] (0.22%)
	Rating = 5	0.47% (0.07%)	0.87% (0.02%)	-0.40%[**] (0.10%)
Number of Reviews	All Reviews	15,759	310,110	
	Rating = 1	1,680	16,363	
	Rating = 2	1,102	16,752	
	Rating = 3	1,262	20,077	
	Rating = 4	2,179	52,594	
	Rating = 5	9,536	204,324	

The unit of analysis is a review. Standard errors are in parentheses. [†]Significantly different from zero, $p<0.10$, [*]significantly different from zero, $p<0.05$, and [**]significantly different from zero, $p<0.01$.

Feedback about Customer Service or Shipping Policies
Within Item and Within Reviewer Analysis

	Without a Confirmed Transaction	With a Confirmed Transaction	Difference
Within Item Analysis			
Customer Service Feedback	1.46%	1.96%	-0.49%[**]
	(0.15%)	(0.08%)	(0.17%)
Shipping Policy Feedback	0.57%	1.26%	-0.69%[**]
	(0.10%)	(0.09%)	(0.12%)
Number of Items	3,779	3,779	
Within Reviewer Analysis			
Customer Service Feedback	1.50%	1.66%	-0.17%
	(0.16%)	(0.15%)	(0.21%)
Shipping Policy Feedback	0.57%	0.95%	0.38%[**]
	(0.10%)	(0.11%)	(0.14%)
Number of Reviewers	5,234	5,234	

The unit of analysis is either an item or a reviewer. Standard errors are in parentheses.
[**] Significantly different from zero, $p < 0.01$.

Matching Reviews with Confirmed Transactions at the Sub-Category Level

	Without a Confirmed Transaction	With a Confirmed Transaction	Difference
Average rating	4.04	4.32	-0.28[**] (0.01)
Rating = 1	11.06%	5.38%	5.68%[**] (0.24%)
Rating = 2	6.87%	5.44%	1.44%[**] (0.24%)
Rating = 3	8.23%	6.50%	1.73%[**] (0.26%)
Rating = 4	14.10%	16.88%	-2.79%[**] (0.40%)
Rating = 5	59.74%	65.80%	-6.06%[**] (0.50%)
Chi-Square test			718.48[**]
KL Divergence			0.0271

The table reports the average product ratings for reviews with and without a confirmed transaction when matching reviews with confirmed transactions at the sub-category level. The sample sizes are 9,150 (reviews without a confirmed transaction) and 316,604 (reviews with a confirmed transaction). Standard errors are in parentheses. [**]Significantly different from zero p<0.01.

Distribution of Ratings by Ordering Channel

	Mail or Telephone	Internet	Retail Stores
Average rating	4.31	4.32	4.43
Rating = 1	6.04%	5.17%	4.37%
Rating = 2	5.42%	5.55%	4.27%
Rating = 3	6.36%	6.68%	4.53%
Rating = 4	16.14%	17.19%	17.43%
Rating = 5	66.04%	65.40%	69.39%
Sample Size (number of reviews)	39,551	225,370	5,623

The table reports the average product ratings (for reviews with confirmed transactions) according to the ordering channel in which the confirmed transaction was received. We omit reviews with confirmed transactions in multiple channels.

Excluding Customers Who Purchase in the Firm's Retail Stores

	Without a Confirmed Transaction	With a Confirmed Transaction	Difference
Average rating	4.09	4.33	-0.24**
			(0.02)
Rating = 1	10.67%	5.31%	5.37%**
			(0.37%)
Rating = 2	6.24%	5.39%	0.85%*
			(0.37%)
Rating = 3	8.33%	6.50%	1.82%**
			(0.40%)
Rating = 4	12.96%	16.78%	-3.82%**
			(0.61%)
Rating = 5	61.79%	66.02%	-4.22%**
			(0.77%)
Chi-Square test			282.76**
KL Divergence			0.0260

The table reports the average product ratings for reviews with and without a confirmed transaction. The sample excludes all customers who have purchased in one of the firm's retail stores and customers living within 400 miles of one of the firm's stores. Standard errors are in parentheses. The sample sizes are 3,926 (reviews without a confirmed transaction) and 89,876 (reviews with a confirmed transaction). Standard errors are in parentheses. *Significantly different from zero, $p<0.05$ and **significantly different from zero, $p<0.01$.

Items Where Fewer than 2% of Purchases Occur in the Firm's Retail Stores

	Without a Confirmed Transaction	With a Confirmed Transaction	Difference
Average rating	4.09	4.25	-0.16**
			(0.04)
Rating = 1	11.78%	7.41%	4.37%**
			(0.90%)
Rating = 2	6.69%	5.89%	0.79%
			(0.80%)
Rating = 3	5.63%	6.00%	-0.38%
			(0.80%)
Rating = 4	12.85%	15.62%	-2.78%*
			(1.22%)
Rating = 5	63.069%	65.07%	-2.01%
			(1.62%)
Chi-Square test			30.78**
KL Divergence			0.0128

The table reports the average product ratings for reviews with and without a confirmed transaction. The sample excludes all items where more than 2% of all unit purchases occur in one of the firm's retail stores. Standard errors are in parentheses. The sample sizes are 942 (reviews without a confirmed transaction) and 11,842 (reviews with a confirmed transaction). Standard errors are in parentheses. *Significantly different from zero, $p<0.05$ and **significantly different from zero, $p<0.01$.

Do Items That Have More Sales in Retail Stores Have Lower Product Ratings?
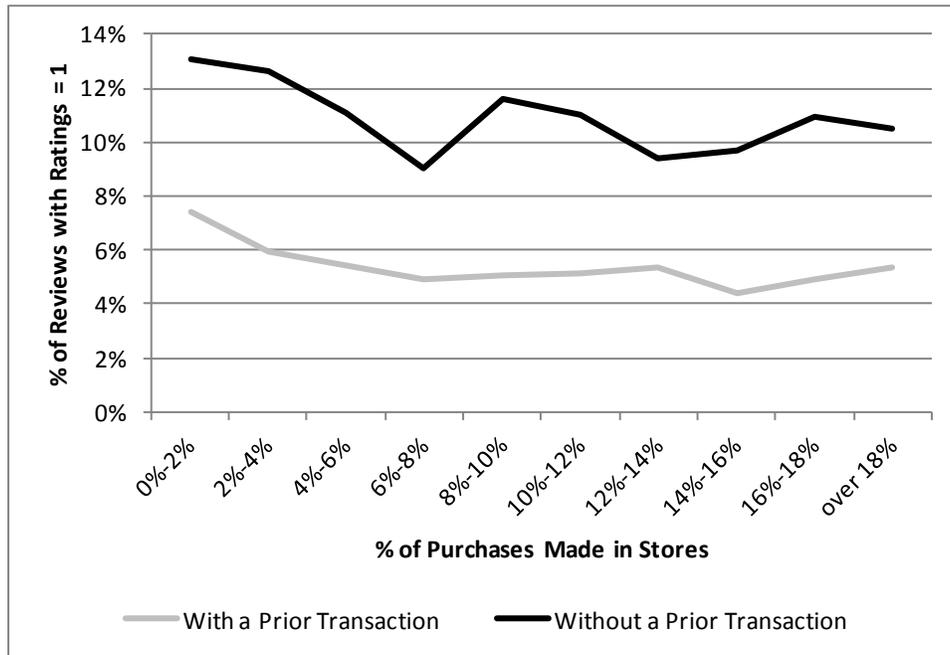
The figure reports the relationship between the proportion of an item's sales that occur in retail stores and the proportion of reviews that have ratings equal to 1. The unit of analysis is a review and the sample includes all 325,869 reviews.

Average Rating by the Year the Review was Written
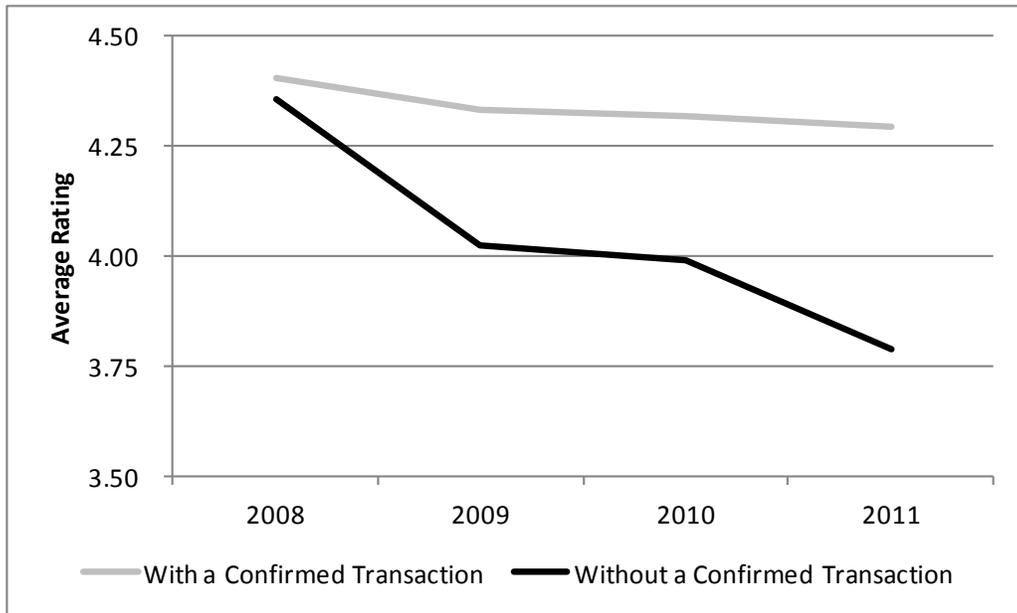
The figure reports the average product ratings for reviews with and without a confirmed transaction, grouped according to the year that the review was written. The unit of analysis is a review and the sample includes all 325,869 reviews.

Predicting Which <u>Reviewers</u> Will Write a Review Without a Confirmed Transaction

	Marginal Effects
Average Review Date	-0.0162**
	(0.0005)
Number of Reviews	0.0117**
	(0.0002)
Items Purchased (100s)	0.0011**
	(0.0002)
Average Item Price ($10)	0.0011**
	(0.0003)
Overall Discount Received	-0.0176
	(0.0149)
Discount Frequency	0.0309**
	(0.0066)
Return Rate	0.0177**
	(0.0026)
Years Since First Order	-0.0001
	(0.0001)
Number of Children	0.0017**
	(0.0005)
Married	-0.0050**
	(0.0012)
Age (10s years)	-0.0085**
	(0.0004)
Est. Home Value ($100,000s)	0.0001
	(0.0002)
Est. Household Income ($100,000s)	0.0002
	(0.0010)
Graduate Degree	0.0008
	(0.0011)
Pseudo R^2	0.0730

The table reports the marginal effects from a logistic model where the dependent variable is a binary variable indicating whether the reviewer has ever written a review without a confirmed transaction. The unit of analysis is a reviewer and the sample size is 178,772 (the sample size reflects missing demographic variables). The model includes a constant, but this is omitted from the table. Standard errors are in parentheses. **Significantly different from zero, $p<0.01$.

Classification Table:
Predicting Which <u>Reviewers</u> Will Write a Review Without a Confirmed Transaction

	Actual: True	Actual: False	Total
Predicted: True	363	220	583
Predicted: False	9,012	169,177	178,179
Total	9,375	169,397	178,772
False Positive	37.74%	220/583	
False Negative	5.06%	9,012/178,719	
Correctly Classified	94.84%	(363+169,177)/178,772	
Benchmark	94.76%	169,397/178,772	

The table reports the classification table from the logistic model reported on the previous page. The benchmark prediction is that none of the reviewers write reviews without a confirmed transaction.

Upset Customers: Text Analysis

	Without a Confirmed Transaction	With a Confirmed Transaction	Difference
By Product Review Rating			
Rating = 1	1.37%	1.65%	-0.28%
			(0.32%)
Rating = 2	1.18%	0.734%	-0.445%
			(0.270%)
Rating = 3	0.872%	0.608%	0.264%
			(0.228%)
Rating = 4	0.872%	0.384%	0.488%**
			(0.139%)
Rating = 5	0.912%	0.479%	0.434%**
			(0.074%)
Number of Reviews			
Rating = 1	1,680	16,363	
Rating = 2	1,102	16,752	
Rating = 3	1,262	20,077	
Rating = 4	2,179	52,594	
Rating = 5	9,536	204,324	

The table reports averages for each measure separately for the samples of reviews with and without confirmed transactions. For the within item analysis, reviews are first averaged at the item level and then averaged across items. For the within reviewer analysis, reviews are first averaged at the reviewer level and then averaged across reviewers. Standard errors are in parentheses. *Significantly different from zero, $p<0.05$ and **significantly different from zero, $p<0.01$.

Upset Customers: Analysis of Subsequent Orders
All Reviews

	(Any) Review Without a Confirmed Transaction	Reviews With a Confirmed Transaction	Difference
All Observations			
Years Until Next Order	0.2603	0.2681	-0.0078[*]
	(0.0042)	(0.0008)	(0.0038)
Purchase Intervals Until Next Order	1.1137	0.9276	0.1861[**]
	(0.0510)	(0.0060)	(0.0290)
No Subsequent Order	12.38%	17.08%	-4.70%[**]
	(0.29%)	(0.07%)	(0.34%)
No Order in Next Purchase Interval	32.12%	34.10%	-1.97%[**]
	(0.45%)	(0.10%)	(0.47%)
No Order in Next Year	13.13%	15.39%	2.26%[**]
	(0.35%)	(0.09%)	(0.38%)
More Orders in Next Year vs. Prior Year	37.21%	29.03%	8.19%[**]
	(0.46%)	(0.11%)	(0.48%)
More Orders in Next Year vs. Prior Average	62.13%	58.89%	3.24%[**]
	(0.50%)	(0.12%)	(0.52%)
Sample Sizes			
Years Until Next Order	11,018	219,019	
Purchase Intervals Until Next Order	11,018	219,019	
No Subsequent Order	12,575	264,140	
No Order in Next Purchase Interval	10,613	212,206	
No Order in Next Year	9,416	176,641	
More Orders in Next Year vs. Prior Year	9,416	176,641	
More Orders in Next Year vs. Prior Average	9,416	176,641	

The unit of analysis is a reviewer x review date. The sample size changes because we restrict attention to observations for which we observe a complete post period. The sample size is also smaller when measuring the time or interval until the next order as we only consider observations where there is a subsequent order. Standard errors are in parentheses.
[*]Significantly different from zero, $p<0.05$ and [**]significantly different from zero, $p<0.01$.

Upset Customers: More Orders in Next Year vs. Prior Year
Adding an Interval Between Prior Period and Review Date

	(Any) Review Without a Confirmed Transaction	Reviews With a Confirmed Transaction	Difference
More Orders in Next Year vs. Prior Year			
No interval	34.25%	25.96%	8.29%**
	(1.44%)	(0.43%)	(1.41%)
2-week interval	36.19%	28.93%	7.26%**
	(1.46%)	(0.45%)	(1.45%)
4-week interval	37.20%	34.94%	2.26%
	(1.47%)	(0.47%)	(1.52%)
6-week interval	37.48%	36.05%	1.43%
	(1.47%)	(0.47%)	(1.53%)
8-week interval	38.03%	36.72%	1.30%
	(1.47%)	(0.47%)	(1.54%)
Sample Sizes	1,086	10,350	

The table reports findings when we add an interval between the end of the pre-period and the review date. We continue to use 12-month pre-periods and post-periods. The unit of analysis is a reviewer x review date. We use observations that include at least one review with a rating equal to 1 and restrict attention to observations for which we observe a complete post period. Standard errors are in parentheses. **Significantly different from zero, p<0.01.

Upset Customers: Multivariate Analysis Including Fixed Reviewer Effects

In the table below we report the *No Confirmed Transaction* coefficient from the following OLS fixed effects model:

$$Y = \boldsymbol{\beta}\mathbf{X} + \beta_1 \textit{No Confirmed Transaction} + \beta_2 \textit{Review Date} + \varepsilon.$$

We use each of the seven post-review outcome measures. The **X** vector is a vector of fixed reviewer effects and *No Confirmed Transaction* is a binary variable identifying whether at least one of the reviews written by that reviewer on that review date did not have a confirmed transaction. The unit of analysis is a reviewer x review date and the sample sizes are indicated in the table below. The sample size changes because we restrict attention to observations for which we observe a complete post period. The sample size is also smaller when measuring the time or interval until the next order as we only consider observations where there is a subsequent order. We estimate two separate models: one using all of the observations, and the other only using observations where at least one of the reviews had a rating of 1. Note that because of the fixed reviewer effects, the coefficient of interest is only affected by observations for which there is variation in the *No Confirmed Transaction* variable within a reviewer). The standard errors are clustered at the reviewer level and are reported in parentheses.

	No Confirmed Transaction Coefficient (β_1)		Sample Size	
	All Observations	Reviews With a Rating = 1	All Observations	Reviews With a Rating = 1
Years Until Next Order	-0.0170** (0.0064)	-0.0082 (0.0561)	230,037	13,879
Purchase Intervals Until Next Order	0.1349 (0.1014)	-0.6492 (1.7031)	230,037	13,879
No Subsequent Order	0.0055 (0.0062)	0.0067 (0.0745)	276,715	16,940
No Order in Next Purchase Interval	0.0138 (0.0140)	0.0043 (0.1869)	222,819	13,817
No Order in Next Year	-0.0033 (0.0059)	-0.0242 (0.0733)	186,057	11,436
More Orders in Next Year vs. Prior Year	0.0342† (0.0189)	0.0130 (0.1996)	186,057	11,436
More Orders in Next Year vs. Prior Average	0.0198 (0.0131)	0.0755 (0.1788)	186,057	11,436

Standard errors are in parentheses, clustered at the reviewer level. †Significantly different from zero, p<0.10, and **significantly different from zero, p<0.01.

Niche Products by Product Rating Level

		Without a Confirmed Transaction	With a Confirmed Transaction	Difference
Prior Units Index	Rating = 1	74.41	100.00	25.59**
				(3.65)
	Rating = 2	77.51	100.00	22.49**
				(4.45)
	Rating = 3	74.84	100.00	25.16**
				(4.19)
	Rating = 4	73.57	100.00	26.43**
				(3.12)
	Rating = 5	68.66	100.00	31.34**
				(1.43)
Niche Items	Rating = 1	22.32%	11.85%	10.47%**
				(0.85%)
	Rating = 2	20.96%	10.59%	10.37%**
				(0.98%)
	Rating = 3	22.03%	10.33%	11.70%**
				(0.91%)
	Rating = 4	25.88%	10.70%	15.19%**
				(0.69%)
	Rating = 5	25.21%	8.47%	16.73%**
				(0.30%)
Very Niche Items	Rating = 1	5.36%	0.89%	4.46%**
				(0.29%)
	Rating = 2	4.54%	0.68%	3.86%**
				(0.29%)
	Rating = 3	5.55%	0.71%	4.80%**
				(0.29%)
	Rating = 4	8.81%	0.72%	8.09%**
				(0.22%)
	Rating = 5	9.94%	0.55%	9.39%**
				(0.10%)

The unit of analysis is a review. The measures are all defined in the main body of the paper. Standard errors are in parentheses. The table uses all of the reviews (sample sizes are reported in other tables).
**Significantly different from zero, $p < 0.01$.

New Products by Product Rating Level

		Without a Confirmed Transaction	With a Confirmed Transaction	Difference
Product Age (years)	Rating = 1	3.17	4.21	-1.04**
				(0.13)
	Rating = 2	3.44	3.98	-0.55**
				(0.15)
	Rating = 3	3.43	4.07	-0.64**
				(0.14)
	Rating = 4	3.73	4.30	-0.57**
				(0.11)
	Rating = 5	4.12	5.03	-0.91**
				(0.06)
New Item	Rating = 1	59.57%	50.95%	8.62%**
				(1.28%)
	Rating = 2	57.83%	51.87%	5.96%**
				(1.56%)
	Rating = 3	54.65%	50.85%	3.80%**
				(1.45%)
	Rating = 4	51.76%	47.71%	4.05%**
				(1.10%)
	Rating = 5	47.67%	41.22%	6.46%**
				(0.52%)
New Category	Rating = 1	1.67%	1.21%	0.46%
				(0.29%)
	Rating = 2	1.73%	1.74%	-0.01%**
				(0.41%)
	Rating = 3	2.15%	1.88%	0.27%
				(0.40%)
	Rating = 4	2.00%	1.53%	0.46%†
				(0.27%)
	Rating = 5	1.31%	0.92%	0.39%**
				(0.10%)

The unit of analysis is a review. The measures are all defined in the main body of the paper. Standard errors are in parentheses. The table uses all of the reviews (sample sizes are reported in other tables). †Significantly different from zero, $p<0.10$ and **significantly different from zero, $p<0.01$.

Which Types of Products Have High or Low Ratings?

	Average Rating	Rating = 1	Rating =2	Rating = 3	Rating =4	Rating = 5	Sample Size
Not Niche Items	4.33	5.36%	5.40%	6.47%	16.56%	66.19%	293,289
Niche Items	4.20	7.04%	6.20%	7.23%	18.69%	60.84%	37,141
Very Niche Items	4.24	6.89%	5.87%	6.98%	16.96%	63.30%	7,813
Sales Volume (Quartile)							
1 (lowest)	4.22	6.61%	6.20%	7.29%	18.41%	61.49%	81,407
2	4.26	6.03%	6.04%	7.17%	17.76%	63.00%	81,412
3	4.34	5.12%	5.32%	6.32%	16.47%	66.76%	81,412
4 (highest)	4.44	4.39%	4.36%	5.42%	14.59%	71.25%	81,412
Item Age							
1 year	4.19	6.86%	6.76%	7.68%	18.21%	60.48%	87,273
1 to 2 years	4.26	5.85%	5.99%	7.33%	18.02%	62.81%	57,221
2 to 4 years	4.36	4.92%	5.19%	6.14%	16.80%	66.95%	58,852
4 to 6 years	4.39	4.71%	4.57%	6.04%	16.11%	68.57%	35,935
6 to 10 years	4.42	4.62%	4.48%	5.46%	15.25%	70.18%	34,366
Over 10 years	4.42	4.85%	4.39%	5.32%	14.63%	70.81%	52,107

The unit of analysis is a review.

Directed to Other Customers vs. the Firm
Analysis at Each Rating Level

		Without a Confirmed Transaction	With a Confirmed Transaction	Difference
Contain Requests Directed to the Firm	Rating = 1	16.01%	5.01%	11.00%** (0.61%)
	Rating = 2	8.62%	3.47%	5.15%** (0.59%)
	Rating = 3	8.56%	2.86%	5.70%** (0.51%)
	Rating = 4	5.42%	1.47%	3.95%** (0.28%)
	Rating = 5	2.44%	1.21%	1.23%** (0.12%)
Contain Advice Directed to Other Customers	Rating = 1	1.37%	1.11%	0.26% (0.27%)
	Rating = 2	2.45%	1.53%	0.92%* (0.39%)
	Rating = 3	2.22%	1.90%	0.32% (0.40%)
	Rating = 4	2.07%	1.35%	0.71%** (0.26%)
	Rating = 5	1.51%	0.92%	0.59%** (0.10%)
Number of Reviews	Rating = 1	1,680	16,363	
	Rating = 2	1,102	16,752	
	Rating = 3	1,262	20,077	
	Rating = 4	2,179	52,594	
	Rating = 5	9,536	204,324	

The table reports averages for each measure separately according to the rating on the product review. Standard errors are in parentheses. *Significantly different from zero, $p<0.05$ and **significantly different from zero, $p<0.01$.

Directed to Other Customers vs. the Firm
Controlling for Item and Reviewer Differences

	Without a Confirmed Transaction	With a Confirmed Transaction	Difference
All Reviews			
Requests Directed to the Firm	5.22%	1.68%	3.54%[**]
	(0.18%)	(0.02%)	(0.11%)
Advice Directed to Other Customers	1.69%	1.10%	0.60%[**]
	(0.10%)	(0.02%)	(0.09%)
Number of Reviews	15,759	310,110	
Within Item Analysis			
Requests Directed to the Firm	4.02%	1.44%	2.58%[**]
			(0.25%)
Advice Directed to Other Customers	1.93%	1.35%	0.58%[**]
			(0.19%)
Number of Items	3,779	3,779	
Within Reviewer Analysis			
Requests Directed to the Firm	5.00%	2.95%	2.04%[**]
			(0.31%)
Advice Directed to Other Customers	1.76%	1.73%	0.03%
			(0.21%)
Number of Reviewers	5,234	5,234	

The table reports averages for each measure separately for the samples of reviews with and without confirmed transactions. For the within item analysis, reviews are first averaged at the item level and then averaged across items. For the within reviewer analysis, reviews are first averaged at the reviewer level and then averaged across reviewers. Standard errors are in parentheses. [**]Significantly different from zero, $p<0.01$.

Change in Revenue and Units: Year-After vs. Year-Before Review Date

Product Rating	Change in Revenue	Change in Units	Sample Size
1	-16.73%**	-14.63%**	647
	(2.09%)	(1.93%)	
2	-15.12%**	-12.39%**	805
	(1.82%)	(1.66%)	
3	-13.06%**	-12.06%**	1,517
	(1.32%)	(1.20%)	
4	-9.72%**	-9.01%**	4,207
	(0.78%)	(0.71%)	
5	-8.94%**	-9.15%**	12,651
	(0.42%)	(0.39%)	

The changes are calculated as a percentage of the midpoint of the pre period and post period outcomes (to ensure that we do not introduce any asymmetry in the magnitude of increases versus decreases). A negative value indicates that revenue (units) was lower in the post period. The unit of analysis is an item x review date. We restrict attention to reviews written at least 1-year after the item was introduced and at least 1-year before the end of the data period. To avoid outliers we also restrict attention to items with at least $1,000 in annual revenue. When there are multiple reviews without confirmed transactions for the same item on the same day we use the average of their product ratings. Observations with a product rating equal to x include all reviewers where the average rating is equal to x plus or minus 0.5. Standard errors are in parentheses. **Significantly different from zero, $p<0.01$.

Change in Revenue and Units: Year-After vs. Year-Before Review Date
Weighted by Number of Reviews

Product Rating	Change in Revenue	Change in Units	Sample Size
1	-16.49%**	-14.32%**	647
	(2.12%)	(1.95%)	
2	-14.58%**	-12.05%**	805
	(1.81%)	(1.65%)	
3	-11.88%**	-11.12%**	1,517
	(1.34%)	(1.23%)	
4	-7.09%**	-6.72%**	4,207
	(0.76%)	(0.68%)	
5	-6.68%**	-7.29%**	12,651
	(0.41%)	(0.38%)	

Observations are weighted by the number of reviews for that item on that date. The changes are calculated as a percentage of the midpoint of the pre period and post period outcomes (to ensure that we do not introduce any asymmetry in the magnitude of increases versus decreases). A negative value indicates that revenue (units) was lower in the post period. The unit of analysis is an item x review date. We restrict attention to reviews written at least 1-year after the item was introduced and at least 1-year before the end of the data period. To avoid outliers we also restrict attention to items with at least $1,000 in annual revenue. When there are multiple reviews without confirmed transactions for the same item on the same day we use the average of their product ratings. Observations with a product rating equal to *x* include all reviewers where the average rating is equal to *x* plus or minus 0.5. Standard errors are in parentheses. [†]Significantly different from zero, $p<0.05$, and [**]significantly different from zero, $p<0.01$.

	Model 1	Model 2	Model 3	Model 4
Post*Rating_1	-9.02%**	-11.07%**	-9.01%**	-11.07%**
	(3.38%)	(3.77%)	(2.80%)	(3.74%)
Post*Rating_2			-6.71%*	-8.72%**
			(2.14%)	(3.18%)
Post*Rating_3			-6.71%*	-5.67%*
			(2.80%)	(2.81%)
Post*Rating_4			-1.22%	-0.57%
			(1.42%)	(2.10%)
Rating_1	4.94%**	6.17%**	5.14%**	6.44%**
	(1.78%)	(2.12%)	(1.79%)	(2.14%)
Rating_2			4.34%**	5.11%**
			(1.44%)	(1.58%)
Rating_3			2.02%†	1.90%
			(1.05%)	(1.37%)
Rating_4			0.14%	-0.12%
			(0.75%)	(1.02%)
Post	-11.43%**	-8.95%**	-11.43%**	-8.95%**
	(1.86%)	(2.34%)	(1.85%)	(2.32%)
Review Date	-2.56%	-0.64%	-11.65%**	-9.89%**
	(1.69%)	(1.68%)	(1.76%)	(1.78%)
Number of Previous Reviews	-11.43%**	-8.95%**	0.05%**	0.04%**
	(1.86%)	(2.34%)	(0.01%)	(0.01%)
Avg. Rating on Previous Reviews	-11.43%**	-8.95%**	7.62%	8.15%
	(1.86%)	(2.34%)	(4.47%)	(4.97%)
No Previous Reviews	-2.56%	-0.64%	32.42%	36.04%
	(1.69%)	(1.68%)	(21.65%)	(22.95%)
Weighting	No weights	Nbr Reviews	No weights	Nbr Reviews
Adjusted R^2	0.8923	0.8974	0.8889	0.8939
Sample size	26,596	26,596	39,644	39,654

The figure reports the coefficients from an OLS model where the dependent variable is: ln(*Revenue*). A constant and fixed item effects were included in all of the models, but are omitted from the table. The unit of analysis is an item x review date in either the pre period or post period. We restrict attention to reviews written at least 12-months after the item was introduced and 12-months before the end of the data period. To avoid outliers we also restrict attention to items with at least \$1,000 in annual revenue. When there are multiple reviews without prior transactions for the same item on the same day we use the average of their product ratings. Observations with a product rating equal to *x* include all reviewers where the average rating is equal to *x* plus or minus 0.5. In Models 1 and 2 we restrict attention to observations with an average rating of 1 or 5. In Models 3 and 4 we include all observations. The observations in Models 2 and 4 are weighted using the number of reviews for that item that day. Standard errors are in parentheses, clustered at the item level. *Significantly different from zero, $p<0.05$ and **significantly different from zero, $p<0.01$.

Text Analysis: Recall and Precision Findings

We randomly selected 500 reviews and asked five coders (PhD students) who had not worked on the project to read 100 reviews each. For each review we asked the coders to indicate:[1]

a. "Does the reviewer comment on the physical fit of the product?"
b. "Does the reviewer comment on the physical feel of the product?"
c. "Does the consumer mention that he/she purchased the item?"
d. "Is the consumer so upset with the firm that he/she is unlikely to purchase again?"
e. "Does the reviewer complain about service?"
f. "Does the reviewer complain about shipping?"

In the table below we compare the coders' responses with the results from the text analysis.

	Text Analysis: No	Text Analysis: Yes	Recall Score
Fit Analysis			
Coder: No	203	33	
Coder: Yes	48	216	82%
Precision Score		87%	
Feel Analysis			
Coder: No	89	28	
Coder: Yes	31	352	92%
Precision Score		93%	
Self-Identified Purchasers			
Coder: No	238	20	
Coder: Yes	41	201	83%
Precision Score		91%	
Upset Customers			
Coder: No	489	1	
Coder: Yes	2	8	80%
Precision Score		89%	
Shipping Feedback			
Coder: No	497	0	
Coder: Yes	0	3	100%
Precision Score		100%	
Service Feedback			
Coder: No	491	0	
Coder: Yes	0	9	100%
Precision Score		100%	

---

[1] Each coder answered questions (a) through (d) for one set of 100 reviews and questions (e) and (f) for another set of 100 reviews.

Reviews Directed to the Firm or Other Customers

We randomly selected 100 reviews:

a. 50 reviews were randomly selected from reviews that text analysis identified were directed at the firm; and

b. 50 reviews were randomly selected from reviews that text analysis identified were directed at other customers.

For each review we asked a coder to indicate:

"Are the reviewers comments directed towards other customers or the firm? (Choose One)"

In the table below we compare the coder's responses with the results from the text analysis.

	Text Analysis: Other Customers	Text Analysis: The Firm	Recall Score
Coder: Other Customers	42	0	100%
Coder: The Firm	2	42	95%
Coder: Neither	6	8	
Precision Score	**84%**	**84%**	