

All Reviews are Not Created Equal:
The Disaggregate Impact of Reviews and Reviewers at Amazon.com

Pei-yu Chen,[‡] Samita Dhanasobhon,[†] and Michael D. Smith[†]

{pychen, sdhanaso, mds}@andrew.cmu.edu

This Version: July 2006

Available from <http://ssrn.com/abstract=918083>

[†] H. John Heinz III School of Public Policy and Management, Carnegie Mellon University, Pittsburgh, PA, 15213. (Dhanasobhon: 412-268-1663; Smith: 412-268-5978)

[‡] Tepper School of Business, Carnegie Mellon University, Pittsburgh, PA, 15213. (Chen: 412-268-1291)

All Reviews are Not Created Equal:

The Disaggregate Impact of Reviews and Reviewers at Amazon.com

ABSTRACT

Online product review networks play an important role in Internet commerce by transmitting information that customers can use to evaluate products in digital marketplace. These networks frequently include an explicit social component allowing consumers to view both how community members have rated individual product reviews and the social status of individual reviewers. We extend this prior work by analyzing how these social factors impact consumer responses to disaggregate review information. To do this, we use a new dataset that allows us to control for the degree to which other community members found the review helpful, and the reputation of the reviewer in the community.

We find that reviews that the community finds helpful have a stronger impact on consumer purchase decisions than other reviews. Moreover, these reviews have a stronger impact on less popular books than more popular books, where consumers may have more information about the product. We also find evidence consistent with the hypothesis that featured reviews have a stronger impact on consumer purchase decisions than other reviews do. Overall, our results suggest that the micro-level dynamics of community interactions are valuable in signaling quality — over-and-above the aggregate-level summary quality scores.

Keywords: *Electronic Commerce, Recommendation System, Digital Word-of-Mouth*

1. Introduction

Digital networks for product information have redefined traditional “word-of-mouth” social networks by allowing consumers to easily share their opinions and experiences with other members of large-scale online communities (Dellarocas 2003). Many online retailers, such as Amazon.com and BarnesandNoble.com, are augmenting their product markets by building online communities to provide product reviews to other consumers. Likewise, many auction sites, such as Ebay.com, allow consumers to rate the product sellers. Such information sharing has the potential to reduce the uncertainty consumers face regarding the quality of a product or a seller.

Several papers in the literature have shown that large-scale information sharing in digital networks may help communicate *product/seller quality* and build trust between *buyers* and *sellers* in online markets (Ba and Pavlou 2002; Resnick and Zeckhauser 2002; Dellarocas 2003; Chen and Wu 2005; and Chevalier and Mayzlin 2006). For example, Resnick (2002) shows that seller reviews in eBay influence the probability of a sale, while Chevalier and Mayzlin (2006) find that product reviews at Amazon.com impact book sales. Dellarocas (2005) shows that strategic manipulation of consumer reviews can either increase or decrease the information value of a review to consumers.

This raises the question of *why* consumers would trust the information provided by strangers they may have never met and how trust is formed among consumers themselves? Credibility is a critical issue in effective information sharing, which involves information reliability and consumer trust. There is an extensive literature in the field of social psychology that shows the importance of credibility in influencing the impact of a persuasive message, where credibility

can be based either on the reputation of the author or the content of the message (see Cialdini 2000). Indeed, in offline social networks, consumers usually attach different weights to different information sources according to their social-ties and knowledge about the source and the information.

Thus, to unleash the full potential and benefits of information sharing in online community, it is essential to have an effective mechanism that can help consumers gauge information reliability and enhance consumer trust. To this end, many retailers have invested in rating systems that allow consumers to provide and read reviews not only on the product *per se*, but also on the credibility of the review message and the reviewer. For example, Amazon.com not only lets its customer post reviews of products, it also allows customers to vote on whether posted reviews were helpful to them in making a purchase decision. The proportion of helpful votes a review receives can serve an indicator for the quality of the reviews to other consumers. Furthermore, Amazon.com identifies individual reviewers based on a ranking system where reviewers who post more reviews and have a higher number of helpful votes are singled out to other community members.

Most of the existing empirical literature on online word-of-mouth focuses on aggregate numerical review scores. However, while providing an important contribution, this only tells part of the story. As noted by Resnick et al. (2000): “these simple numerical ratings fail to convey important subtleties of online interactions. For example, ... what were the reputations of the people providing the feedback?” It remains an open empirical question to what extent community evaluations of individual reviews and individual reviewers influence consumer purchase decisions online. Our research aims to bridge this gap by considering the role of

disaggregate information credibility in consumer decision-making along two dimensions: the quality of the content and the source of the information.

To do this we use consumer reviews for books sold at Amazon.com to analyze how the content of reviews and the reputation of the reviewer impact consumers' responses. We selected Amazon.com, because it is one of the largest online retailers and has one of the most active reviewing communities online (Chevalier and Mayzlin 2006). Within Amazon's online review communities, we consider three primary measures of information credibility: First, the quality of the content, which can be indexed by how helpful the community found the review (Amazon lists "x of y people found this review helpful," see Figure 1). Second, the quality of the information source, i.e., the reputation of the reviewer. Amazon identifies its "top" reviewers (see Figure 1), and there is anecdotal evidence that these popular reviewers can have a large impact on book sales (e.g., Paumgarten 2003). Third, we analyze the impact of whether reviews are highlighted when they are displayed on the product page. Amazon moves the top two most helpful reviews to the "spotlight" position at top of all reviews (see Figure 2). These "spotlight" reviews are set apart from the other reviews and are shown before other reviews, so they may have a relatively stronger effect on book sales than other reviews do.

To analyze these questions, we collect data daily on product sales levels and customer reviews from Amazon.com's web pages. Our data include 50,626 observations of 535 newly released book titles, collected over 195-day period from November 11, 2005 to May 25, 2006. In addition to confirming Chevalier and Mayzlin's (2006) prior finding that higher average star ratings of books are associated with higher sales, we find that reviews with high proportion of helpful votes (i.e., quality reviews), and featured reviews are associated with increased sales even after

controlling for average star ratings. Further we find that this review information has a stronger impact on less popular books than for popular books. Finally, we find no evidence that the social status of reviewers has an impact on consumer purchase behavior. Together these results suggest that consumers associate different weights to different messages they receive in making purchase decisions, and reviews are more important for consumers when less outside information is available on the product (as in the case of less popular books).

This research makes two unique contributions. First, this study uses real data in an online setting to unpack the mechanisms that drive people to trust and respond to online product reviews. To the best of our knowledge, this has not been studied empirically in the literature. Second, this research allows us to understand what form of micro-level dynamics in community interactions may be valuable in signaling quality, over and above aggregate-level summary quality scores. This has important implications. When only the aggregate measures are available (i.e. where an uninformative review from a new community member carries as much weight as an informative review from an established and respected community member) it may be easier for self-interested parties to manipulate review results. However, if the micro-level dynamics of reputation communities are important factors in determining product sales, it would be harder for self-interested parties to manipulate reviews, making the reputation system more reliable. Thus, understanding the micro-level dynamics of virtual communities has important implications for designing a more reliable reputation system that is less subject to manipulation.

The remainder of this paper proceeds as follows. In Section 2, we review the prior literature as it relates to our research setting. In Section 3, we use this literature to develop our research

framework and hypotheses. We present our data and model specifications in Sections 4 and 5 respectively. We present our empirical results in Section 6 and conclude in Section 7.

2. Literature Review and Research Framework

In Figure 3, we propose a research framework to study consumer demand in the presence of an online product review system displaying information about both review and reviewer quality. Specifically, we draw on the theoretical literature on information systems, economics, marketing and social psychology to consider consumers' purchase decisions as influenced by the perceived quality of the product (relationship A), the perceived quality of the seller (B), and the quality of the review information (C) and the information source (D).

In digitally mediated markets, absent reputation systems, consumers can face higher uncertainty about product quality, since there are fewer quality and trust cues available than what is possible in brick-and-mortar markets. To compensate for the lack of quality and trust cues in online markets, many retailers provide rating systems for consumers to rate products and/or allow consumers to write reviews about the quality of products. Such information sharing among consumers provides the potential for consumers to reduce consumer uncertainty about product quality (Dellarocas, 2003). The search literature has shown that better information on product quality has an impact on consumers purchase decision (Stiglitz, 1989), and this is reflected in relationship A in Figure 3.

The prior literature has also shown that when product quality cannot easily be verified, the reputation of the supplier of the product may be used by consumers as an indication for the quality of the product (Resnick et al. 2000). As a result, the perceived quality of the product

supplier/seller will have an impact on consumers' purchase decisions and the resulting product sales (relationship B in Figure 3). This is especially true for auction sites, where information sharing on particular products is less relevant because each "product" is a combination of product and seller, and thus is essentially different.

On the other hand, the social psychology literature has shown that consumers' purchase decisions are influenced by their attitude toward a product, which may be influenced by the external messages they receive, and that the magnitude of this influence depends on six basic principles of influence on persuasion and attitude change: reciprocation, commitment and consistency, social validation, liking, authority, and scarcity.

Pervasiveness of free samples is a good example of reciprocation because people tend to repay what they have received. People honor commitment and consistency, so they are more willing to act upon prior commitment. For example, people will typically purchase a car even if the price was raised after they decided to buy it. Social validation derives from what many individuals do. People are more likely to agree with an idea held by many individuals, because we perceive the idea to be more correct. Liking is the concept of personal preference. People prefer to say yes to individuals they know and like. People also tend to obey individuals who possess legitimized or non-legitimized authority figures, such as titles, clothing, and automobiles. Scarcity, such as putting "limited edition" or "limited time only" on product offers, can encourage sales.

Overall, this literature suggests that credibility influences the impact of a persuasive message, where credibility can be based either on the reputation of the information source or the content of the message in itself. These factors are shown as relationships C and D in Figure 3. Note that there is a subtle difference between the reputation of a review supplier and the reputation of a

product supplier, because the product supplier is direct beneficiary of high product sales, while the review supplier is not necessarily a direct beneficiary of high sales although they may both result in higher sales.

While there is strong theoretical support for the research framework presented in Figure 3, the extant empirical research has only addressed a subset of these relationships. Most of empirical literature in IS, economics and marketing focuses on the relationships between reviews on products and sales (relationship A), and reviews on suppliers and sales (relationship B). Among this stream of the literature, Resnick and many others have shown that reviews of product suppliers represent a good proxy for the reputation of product supplier, and have an impact on product sales characteristics such as the price premium and the probability of a sale. Findings on the relationship of product ratings and sales are mixed, with Chevalier and Mayzlin (2006) showing that an improvement in the review score of a book leads to an increase in relative sales at Amazon.com, while Chen and Wu (2005) and Duan et al. (2005) show that high product ratings do not necessarily lead to increased sales.

With regard to the empirical research in the social psychology literature, in a review of the literature, Guadagno and Cialdini (2003) note that only two of six principles — authority and consistency — have been examined empirically in an online context. They further observe that status is a meaningful social category in online markets and can translate into higher compliance, particularly when the influence agent is a high-status in group member. Guadagno and Cialdini (2003) also note that there is a reason to believe that influence attempts online may not function in the same way as influence attempts in other contexts.

Thus, we believe research is needed to address the impact of social cues in review communities on sales at a disaggregate level. This paper aims to bridge this gap in the literature by developing a theoretical model and empirical results that not only study the impact of product/seller reviews on sales but also unpack the micro-level impact of consumer interactions on sales and how credibility might be built in the online context. Specifically, we use average book ratings to control for product/supplier quality (relationships A & B in Figure 3), and we use reviews that receive high helpful votes by other consumers and spotlight reviews to indicate the quality of reviews that may be used by consumers to form their purchase decision (relationship C in Figure 3) and we measure the quality of reviewer (i.e., information source) by their standing in the Amazon.com community as ranked by Amazon.com (relationship D in Figure 3).

In Section 3, we apply this framework to develop the theoretical hypotheses we test in this study.

3. Hypotheses

As noted above, the goal of this paper is to examine how the quality of online product reviews and the reputation of online reviews impact how customers make purchase decisions in the presence of a word-of-mouth review system. To this end, we develop the following hypotheses following our research framework discussed above (Figure 3).

H1: Higher product ratings are positively associated with higher sales.

Higher ratings of a book imply that other people like the book and may be considered as a social validation. According to the social psychology literature (Cialdini 2000), social validation has an influence on consumers' attitude. Accordingly, consumers may be more willing to purchase a book that has acquired a social validation than one without, *ceteris paribus*. As noted above, this

result has been studied in the prior literature with Chevalier and Mayzlin (2006) finding that higher product ratings lead to higher sales in the context of Amazon.com.

H2: Reviews with a high proportion of helpful votes have a relatively higher impact on sales than reviews with a low proportion of helpful votes do.

Cialdini (2000) identified social validation as one of the six basic principles of influence of a persuasive message. Reviews that have high proportion of helpful votes vouch for the quality of the review by indicating that the community validates those reviews. Hence, reviews with a high proportion of helpful votes may signal higher reliability of the review's content and thus have higher influence on consumer decision-making and product sales.

H3: Reviews by more prominent members of the community have higher impact on sales than reviews by other consumers.

Cialdini (2000) identified authority or reputation as another basic influential principle. In our setting, more prominent reviewers may have higher influence over consumer decisions. We will use the rank of the reviewers as a proxy for the reputation of the reviewers. For "top reviewers," (so identified by Amazon based on the number of prior reviews they have posted and the number of helpful votes those reviews have received) this rank is identified next to the reviewer's name in the review display, and thus is readily available to customers viewing the reviews.

H4: Consumer ratings with high proportion of helpful votes have a larger impact on less popular books than on more popular ones.

Reviews for more popular books (for example New York Times Bestsellers), are readily available to consumers through other channels (e.g., book clubs, newspapers), and thus consumers may come to Amazon with a strong prior belief about the quality of these books. For less popular books, particularly newly released titles, consumers will have fewer quality cues to rely on and therefore may place a higher weight on the reviews available at the Internet retailer. Because of this, we hypothesize that reviews will have a larger impact on less popular books than more popular books.

H5: “Featured” reviews (those listed first) have a larger positive marginal impact on sales than other reviews do.

Prior research in online markets has shown that consumers perceive relatively high costs associated with processing information online (Brynjolfsson, Dick, and Smith 2005) and that the order information is displayed to the consumer has a disproportionately strong impact on their behavior (Smith and Brynjolfsson 2001). Both of these results are reasonable in our setting given that consumers who have limited time may spend relatively more time reading reviews displayed first in the list of reviews. Because of this, we hypothesize that, *ceteris paribus*, reviews that are highlighted at the top of review listings will have a larger impact on sales than other reviews do.

4. Data

Our data are collected from publicly available information on Amazon.com. Data were collected using Perl scripts to parse data from the relevant HTML pages and, where possible, from the XML data feed Amazon.com provides to its developers. Our data consist of 535 (20 titles x 28 weeks) new books released over a 195 day period from November 11, 2005 to May 25, 2006. We focus on newly released titles, because consumer opinions are less well formed for these

products, making the product reviews more important for consumer purchase decisions. There are also significant changes in the number of reviews for these titles (initially zero, increasing over time), providing an additional source of variation in our data.

To create our sample of books, we first collect the list of all upcoming book releases as listed by Buy.com. We randomly selected 20 unique titles from the list of titles in each week. For each title, we begin collecting data on the first day the book is released.¹ For this sample, we extracted generic information of each book, such as its International Standard Book Number (ISBN), title, author, release date, and genre from Amazon.com. This data should be constant across sites for a particular book, and we collected it from Amazon.com out of convenience. In addition, for each book in our sample we collected daily information from Amazon.com on the price, sales rank, and the time until the book would ship.

Following the literature (Chevalier and Goolsbee 2003; Brynjolfsson, Hu, and Smith 2003), we use the sales rank listed at Amazon.com as a proxy for product sales.² This prior work has shown that the relationship between sales rank and sales follows a Pareto distribution:

$$Quantity = \beta_1 Rank^{\beta_2} \tag{1}$$

This relationship can be parameterized either by direct observation of sales levels and resulting sales ranks for a number of titles, data that typically is available from Amazon's suppliers (see Brynjolfsson, Hu, and Smith 2003), or by means of an experiment (see Chevalier and Goolsbee 2003). Lacking direct supplier data, we used the experiment proposed by Chevalier and

¹ Amazon.com does not allow consumers to post reviews before the book is released, so beginning to collect data prior to release would not provide any additional information for the purposes of our study.

² This technique has also been applied in a variety of other studies, including Chevalier and Mayzlin (2004); Ghose, Smith, and Telang (2006); and Ghose and Sundararajan (2005).

Goolsbee to parameterize this relationship, yields a slope parameter of $\beta_2 = -0.954$.³ This estimate is in the range of coefficient values reported by other studies in the literature (albeit at the high end of this range).⁴

Finally, we collected daily information regarding the reviews posted at Amazon for each book in our sample. For each book review, we collected the review’s posting date, the full text of the review, the 1 to 5 star rating given in each review, the identity of the reviewers, whether the reviewer was identified as a “top” reviewer (see Figure 1),⁵ the number and proportion of helpful votes (see Figure 1), and whether the review was highlighted as a “spotlight” review (see Figure 2).⁶ Table 1 provides summary statistics for our data.

Table 1: Summary Statistics

<i>Variable</i>	<i>Obs</i>	<i>Mean</i>	<i>Std. Dev.</i>	<i>Min</i>	<i>Max</i>
No. of days since release	50,626	62.54	44.66	0	188
Sales rank	41,979	867,894	986,367	4	4,208,216
Price	48,875	33.62	44.01	3.95	687
No. of reviews	33,816	3.33	9.73	0	100
Amazon star rating	16,065	4.38	0.69	1	5
Average star rating of more than 80% helpful vote	12,796	4.52	0.70	1	5
Average star rating of top 1000 reviewers	7,768	4.54	0.56	1	5
Average Stars Rating of Spotlight Review	5,872	4.41	0.59	1	5

³ We conducted our experiment on February 14, 2006, by ordering 7 copies of two book titles from different buyer accounts. We picked two book titles that had steady movement in ranks for six months, and tracked the movement of rank for 24 hours after we bought those items. The rank of one book title jumped from 662,973 to 5,521 and the rank of another title jumped from 868,303 to 5,529.

⁴ These coefficients include -0.834 (Weingarten 2001), -0.855 (Chevalier and Goolsbee 2003), -0.871 (Brynjolfsson, Hu, and Smith 2003), -0.877 (Ghose, Smith, and Telang 2006), and -0.952 (Poynter 2000). Using a lower coefficient value (for example -0.834) would only affect our price elasticity result, not our main findings.

⁵ Top reviewers are selected by Amazon based on the number of reviews they post across all products on Amazon’s site and the number of helpful votes they receive from other community members for their reviews.

⁶ As noted above, Amazon.com highlights two reviews for the “spotlight” position at the top of the review listings based on the number of helpful votes assigned to that review.

5. Model Specification

We adopt the same *difference-in-difference* strategy used in Chevalier and Mayzlin (2006), while incorporating measures for the quality of the content of the reviews and the standing of the reviewers in the community. Chevalier and Mayzlin define the book's sales rank as a function of a book fixed effect (v_i) and other factors that may impact the sales of a book and use a constant elasticity demand specification.

Specifically, to study the impact of reviews and the quality of reviews on sales, we consider the following model:

$$\text{Error! Objects cannot be created from editing field codes.} \quad (2)$$

where $rank_i^t$ is the sales rank of book i at Amazon.com at time t ; P_i^t is the price of book i at time t and α is the own-site price effect; S is the vector capturing the shipping times promised for book i and Π captures the effect of the shipping time; v_i is a book fixed effect, summarizing the impact of other (unobserved) variables, such as the inherent popularity of the book subject, the author, and unobserved marketing variables etc., that contribute to book sales; R is a vector summarizing review activities and Γ measures its impact; and ε_i^t is random effects summarizing all other unknown variables.

Taking the difference of equation (2) at time t and time 0, allows us to eliminate the unobserved book fixed effect (v_i) and provides us with the specific model we estimate:

$$\Delta r_i^t \equiv \ln(rank_i^t) - \ln(rank_i^0) = \alpha (\ln(P_i^t) - \ln(P_i^0)) + \Pi (S^t - S^0) + \Gamma (R^t - R^0) + \varepsilon_i^t \quad (3)$$

Within this model, we use different measures of R to fit the model. In our base model we use the number of reviews and the average star rating across all reviews. We extend this to measure information quality by adding variables for the average star rating of reviews with a high proportion of helpful votes, the average star rating of top reviewers, and the average star rating of spotlight reviews. We define the average star rating of the high proportion of helpful votes as the average star rating of the reviews that have more than 80% of helpful votes. The average star rating of top reviewers is the average star rating of reviewers who are ranked by Amazon in the top 1000 reviewers on the site. We use this cutoff because reviewers in the top 1000 are identified with a logo next to their name in the review listing. The average star rating of spotlight reviews is the average star rating of reviews that are in the spotlight review section.

6. Results

We now fit these empirical models to our data. Our results are shown in Tables 2 and 3. In these results, our control variables are consistent across specifications and have the expected signs. The coefficient on price is positive and significant across all specifications meaning that, as expected, when price rises, sales rank rises and sales fall. Multiplying these coefficients by the sales-rank coefficient β_2 estimated above yields an own price coefficient in the range of -0.5249 to -0.8132, which is in the middle of the range of own price elasticities for Amazon found in prior studies (Chevalier and Goolsbee 2003; Ghose, Smith, and Telang 2006). The coefficient on the number of days since release is positive, suggesting that these books follow a normal sales lifecycle of declining sales over time. The coefficient on “ships within 24 hours” suggests that faster shipping is associated with higher sales compared to longer shipping times. Finally, the coefficient on the log of the number of reviews is positive which, because this is a first-

difference model, suggests that the marginal impact of an additional review declines with the number of reviews, a result that is also consistent with expectations.

Table 2: Statistical Analysis

<i>Term</i>	<i>Model 1</i>	<i>Model 2</i>	<i>Model 3</i>	<i>Model 4</i>
Ln (Price)	0.5502** (0.0826)	0.6069** (0.0890)	0.7208** (0.1101)	0.7437** (0.1134)
Ln (Days since release)	0.0696** (0.0102)	0.0504** (0.0113)	0.1233** (0.0155)	0.0952** (0.0169)
Ship within 24 hours	-0.5150** (0.0517)	-0.5315** (0.0568)	-0.9677** (0.1081)	-0.8551** (0.1225)
Ln (No. of reviews)	0.6404** (0.0280)	0.6833** (0.0344)	0.7990** (0.0545)	0.7955** (0.0637)
Overall average star rating	-0.3063** (0.0419)	-0.3751** (0.0575)	-0.1649* (0.0748)	-0.4382** (0.1101)
Average star rating of high helpful vote		-0.5000** (0.0522)		-0.4236** (0.0681)
Average star rating of top reviewers			0.0408 (0.0689)	0.0801 (0.0732)
N	13,783	10,798	6,799	5,830
R-Squared	0.0883	0.0982	0.1022	0.1027

Notes. Dependent variable is $\ln(\text{rank}) - \ln(\text{rank at the first day the book was released})$. All models run with book-type fixed effects. Standard errors are in the parentheses. * $p < 0.05$; ** $p < 0.01$.

With respect to our review variables of interest, we find that higher overall star ratings have a positive impact of sales (negative impact on sales rank). This is consistent with Chevalier and Mayzlin's (2006) prior findings in the context of Amazon's marketplace and is consistent with Hypothesis 1.

In Model 2 of Table 2, we add a variable for the average star rating among reviews with more than 80% helpful votes. The coefficient on this variable is negative and significant in all model specifications including this variable. This finding suggests that reviews that are identified as very helpful make additional contributions to sales beyond average star ratings (which take into

account every review posted on the site and weighs each review equally). Specifically, Model 2 of Table 2 shows that, while average star ratings are associated with higher sales (-0.3751), high ratings of quality reviews are associated with an additional marginal increase in sales (-0.5). Overall, this finding shows that, consistent with hypothesis 2, quality reviews (i.e., reviews with high helpful votes) are associated with higher sales than other reviews, and consumers attach more weight to these quality reviews in making purchase decisions.

Table 3: Additional Statistical Analysis

<i>Term</i>	<i>Model 5</i>	<i>Model 6</i>	<i>Model 7</i>
Ln (Price)	0.8524** (0.1698)	0.5600** (0.1035)	1.2232 (0.6472)
Ln (Days since release)	0.4767** (0.0243)	-0.0802** (0.0124)	-0.0116 (0.0191)
Ship within 24 hours	-0.5338* (0.2690)	-0.3555** (0.0563)	-0.4336** (0.1399)
Ln (No. of reviews)	0.3580** (0.0622)	0.5956** (0.0419)	1.1048** (0.0719)
Overall average star rating	-0.5241** (0.1072)	-0.2114** (0.0665)	-0.1672 (0.1482)
Average star rating of high helpful vote	-0.3014** (0.0742)	-0.7062** (0.0717)	-0.2979** (0.0711)
Average star rating of spotlight reviews			-0.7216** (0.0705)
N	2,957	7,841	4,750
R-Squared	0.2971	0.0567	0.1069

Notes. Dependent variable is $\ln(\text{rank}) - \ln(\text{rank at the first day the book was released})$. All models run with book-type fixed effects. Standard errors are in the parentheses. * $p < 0.05$; ** $p < 0.01$.

In Model 3 of Table 2 we include a variable for the average star rating among reviews from the top 1000 reviewers at Amazon.com. Reviews from these individual are specifically flagged on in the review listings (see Figure 1), and thus might have a larger impact on consumer behavior. However, in our results the coefficient on this variable is small and statistically insignificant.

Thus, we fail to accept Hypothesis 3 that reviews from more prominent members of the community will be more influential than other reviews. This result may imply that customers do not trust top rank reviewers as much as we expected, or it may imply that to become a top reviewer you have to review so many products that you can't show the specific product expertise that is expected by the community. We discuss this finding in more detail in the discussion section.

To test Hypothesis 4, that customer reviews with a high proportion of helpful votes have a larger impact on less popular books than on more popular ones, we divide our sample into books that have a sales rank of less than 100,000 (Model 5) and books that have a sales rank of greater than 100,000 (Model 6). We choose 100,000 because this is number of unique titles normally carried by Barnes and Noble superstores (Brynjolfsson, Hu, and Smith 2003).⁷ Our results are still consistent with Hypothesis 2, that reviews with high proportion of helpful votes have an additional marginal impact on sales. Our results also support our hypothesis that those reviews have a stronger impact on less popular books more than on more popular books.

Finally, we test Hypothesis 5 that spotlight reviews have additional positive marginal impact on sales. The result is shown in Model 7 of Table 3. This model shows that the reviews in the spotlight position have additional impact on product sales. We also note that in this model the coefficient on overall star rating is insignificant, suggesting the most of the consumer response is being explained by the combination of spotlight reviews and other reviews with high helpful votes. Also, the degree of impact of the average rating of spotlight reviews is stronger than the

⁷ We note that our qualitative result is not sensitive to the choice of rank cutoff.

impact of the average rating of reviews that have high helpful votes. This result also implies that customers rely more on spotlight reviews than other overall reviews.⁸

7. Discussion and Conclusion

Online feedback mechanisms and virtual communities have become increasingly important to consumers' decision making in online markets. However, most of the extant literature on online feedback mechanisms focuses on product or seller reviews using aggregate measures of quality and reputation. Less is known in the literature about how the quality of individual reviews and the reputation of individual reviewers influence the community's perception of the validity of the opinions expressed in the review. While the social psychology literature has shown that credibility influences the impact of a persuasive message, where credibility can be based either on the reputation of the author or the content of the message, this theory has not been examined empirically in the context of online markets (Guadagno et al. 2003). In extending these two streams of the literature, the goal of this paper is to examine the micro-level impact of reviews — specifically the quality of online product reviews and the reputation of the reviewers — on sales.

Our results show that while higher book ratings are associated with higher book sales, higher quality reviews (i.e., reviews with the high proportion of helpful votes) strengthen this impact by creating additional sales. This result suggests that consumers may consider quality reviews more important in making purchase decisions. Furthermore, these reviews affect non-popular books more than popular books. We also find suggestive evidence that “spotlight” reviews have a stronger effect on book sales than overall reviews do. However, contrary to our expectations, we find no evidence that the reputation of reviewers (i.e. top reviewers) is an important factor in

⁸ We note, however, that this result should be interpreted with caution because of the relatively high correlation between spotlight reviews and reviews with a high proportion of helpful votes.

consumers' purchase decisions. We speculate that this may be due to the fact that top 1,000 reviewers must review so many products that they aren't likely to have the requisite expertise in any product to make a significant impact on customer purchase decisions.

This research makes two unique contributions to the literature. First, this study uses a new dataset from a working online market to unpack the mechanisms that drive people to trust and respond to product reviews. To the best of our knowledge, this has not been studied empirically in the literature. Second, this research allows us to understand what form of micro-level dynamics of community interactions may be valuable in signaling quality — in addition to the aggregate-level summary quality scores. Our results suggest that social validation is very important in the online community.

For online retailers, our research suggests that community peer-rating systems provide an important signal of trust and can facilitate commerce. As noted above, the fact that the content of reviews matters to consumer purchase decisions should strengthen the reliability of online review systems by making it harder for self interested parties to manipulate the ratings. Ratings that provide a simple 5-star (or 1-star) review, while having equal weight in the overall average star rating listed on Amazon's site, do not have as much influence on consumer response as more detailed reviews that have been rated as "helpful" by other members of the community.

Our research also shows that reviews have the more impact on less popular books than other titles. An implication of this finding for the publishing industry is that online review systems may play an important role in the development of "long tail" markets. Recent papers in the academic literature (Brynjolfsson, Hu, and Smith 2003) and popular press (e.g., Anderson 2004; Anderson 2006) have discussed the impact of the increased product variety available in online

markets on consumer surplus and industry structure. The Internet allows retailers to “stock” far more products than what would be possible in a typical brick-and-mortar environment. In the case of bookstores, Internet retailers can stock nearly all of the approximately 3 million books in print and numerous out-of-print titles while a typical brick-and-mortar bookstore can only stock 40,000 to 100,000 titles (Brynjolfsson, Hu, and Smith 2003). However, in the absence of reliable product information, it may be difficult to credibly signal the quality of these products to consumers. Online product review systems may be able to serve this function and play an important role in extending long tail markets — which can have important spillover effects for authors and publishers (Brynjolfsson, Hu, and Smith 2006).

The future research could extend our results by using different micro-level data, such as direct text analysis of the reviews, causality between prominent reviewers and number of helpful votes, or the effect of reviews across book genres. Future research could also extend our results by analyzing the impact of review communities in other contexts, such as online job markets, restaurants, or professional services.

References

- Anderson, C. 2004. The Long Tail. Wired Magazine. October.
- Anderson, C. 2006. The Long Tail: Why the Future of Business Is Selling Less of More. Hyperion Press, New York, NY.
- Ba S., P. Pavlou. 2002. Evidence of the effect of trust building technology in electronic markets: Price premiums and buyer behavior. *MIS Quarterly* **26**(3) 243-268.
- Brynjolfsson, Erik, Astrid A. Dick, Michael D. Smith. 2004. Search and Product Differentiation at an Internet Shopbot. Working Paper, Carnegie Mellon University, Pittsburgh, PA.
- Brynjolfsson, Erik, Yu Hu, Michael Smith. 2003. Consumer Surplus in the Digital Economy: Estimating the Value of Increased Product Variety. *Management Science*, **49**(11) 1580-1596.
- Brynjolfsson, E., Y. Hu, M. Smith. 2006. From Niches to Riches: Anatomy of the Long Tail. *Sloan Management Review* 47(4) 67-71.
- Chen, P-Y., S-Y. Wu. 2005. The Impact of Online Recommendations and Consumer Feedback on Sales. Working Paper, Carnegie Mellon University, Pittsburgh, PA.
- Chevalier, J., A. Goolsbee. 2003. Measuring prices and price competition online: Amazon.com and BarnesandNoble.com. *Quantitative Marketing and Economics* **1**(2) 203-222.
- Chevalier, J., D. Mayzlin. 2006. The Effect of Word of Mouth on Sales: Online Book Reviews. *Journal of Marketing Research*. Forthcoming (August).
- Cialdini, R.B. 2000. *Influence: Science and Practice* (4th ed). New York: HarperCollins.
- Dellarocas, C. 2003. The Digitization of Word of Mouth: Promise and Challenges of Online Feedback Mechanisms. *Management Science* **49**(10) 1407-1424.

Dellarocas, C. 2005. Strategic Manipulation of Internet Opinion Forums: Implications for Consumers and Firms. Working Paper. University of Maryland, College Park, MD.

Duan, W., B. Gu and A. B. Whinston, 2005. Do Online Reviews Matter? - An Empirical Investigation of Panel Data. Working paper. University of Texas at Austin, Austin, TX.

Ghose, A., A. Sundararajan. 2005. Software Versioning and Quality Degradation? An Exploratory Study of the Evidence. Working Paper. New York University, New York.

Guadagno, R. E., R.B. Cialdini. (2003). Online persuasion and compliance: Social influence on the Internet and beyond. In Y. Amichai-Hamburger (Ed.), *The social net: The social psychology of the Internet*. Oxford, UK: Oxford University Press.

Paumgarten, N. 2003. No. 1 Fan Dep't Acknowledged. *The New Yorker*, May 5.

Poynter, Daniel. 2000. "Publishing Poynters," April-June 2000, (Available at <http://parapub.com/getpage.cfm?file=newsletter/News0400.html>), accessed February 26, 2003.

Resnick, P., R. Zeckhauser. 2002. Trust among strangers in Internet transactions: Empirical analysis of eBay's reputation system. *The Economics of the Internet and E-Commerce. Advances in Applied Microeconomics*, Vol. 11. JAI Press, Greenwich, CT.

Resnick, P., R. Zeckhauser, E. Friedman, K. Kuwabara. 2000. Reputation Systems. *Communications of the ACM* **43**(12) 45-48.

Smith, Michael, Erik Brynjolfsson. 2001. Customer Decision Making at an Internet Shopbot: Brand Still Matters. *The Journal of Industrial Economics* **49**(4) 541-558.

Weingarten, Gene. 2001. Below the beltway, *Washington Post*, June 17, 2001.

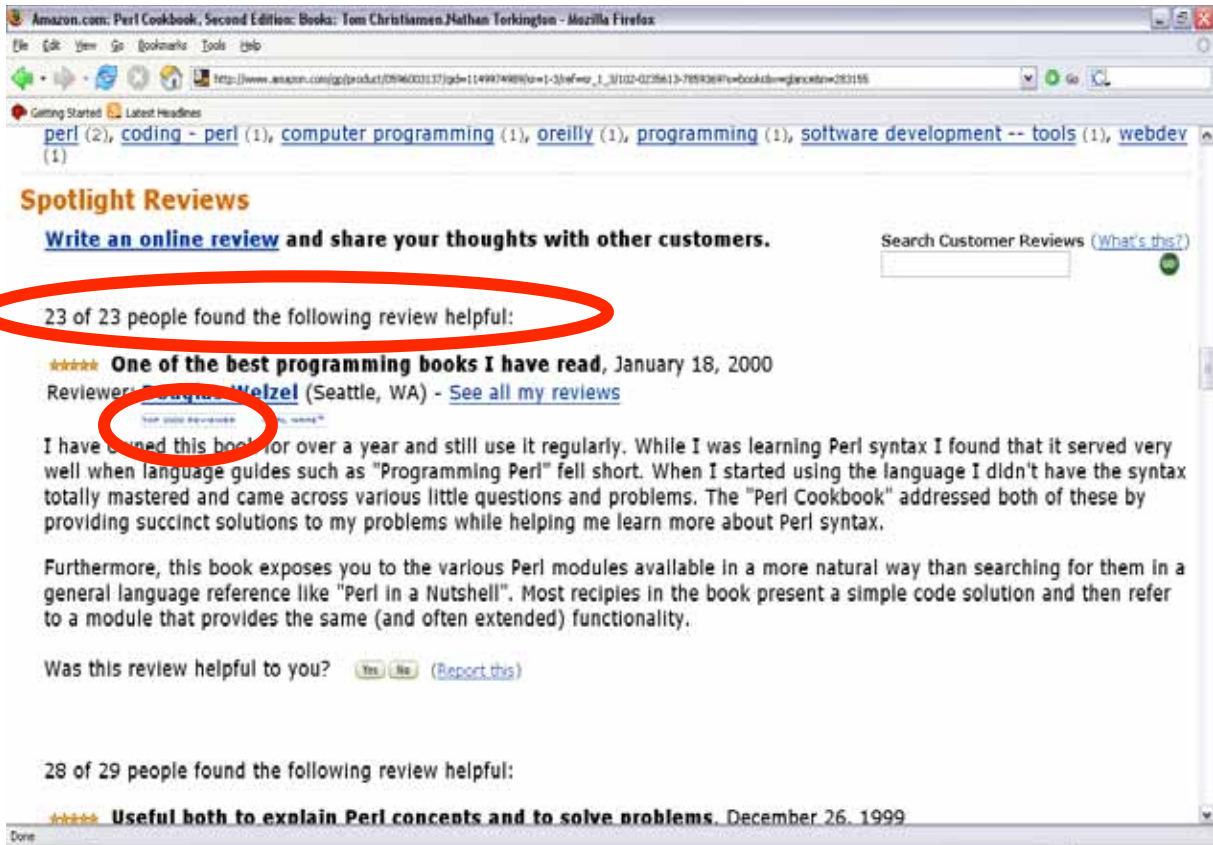


Figure 1: Sample Number of Helpful Votes and Top Reviewer at Amazon.com

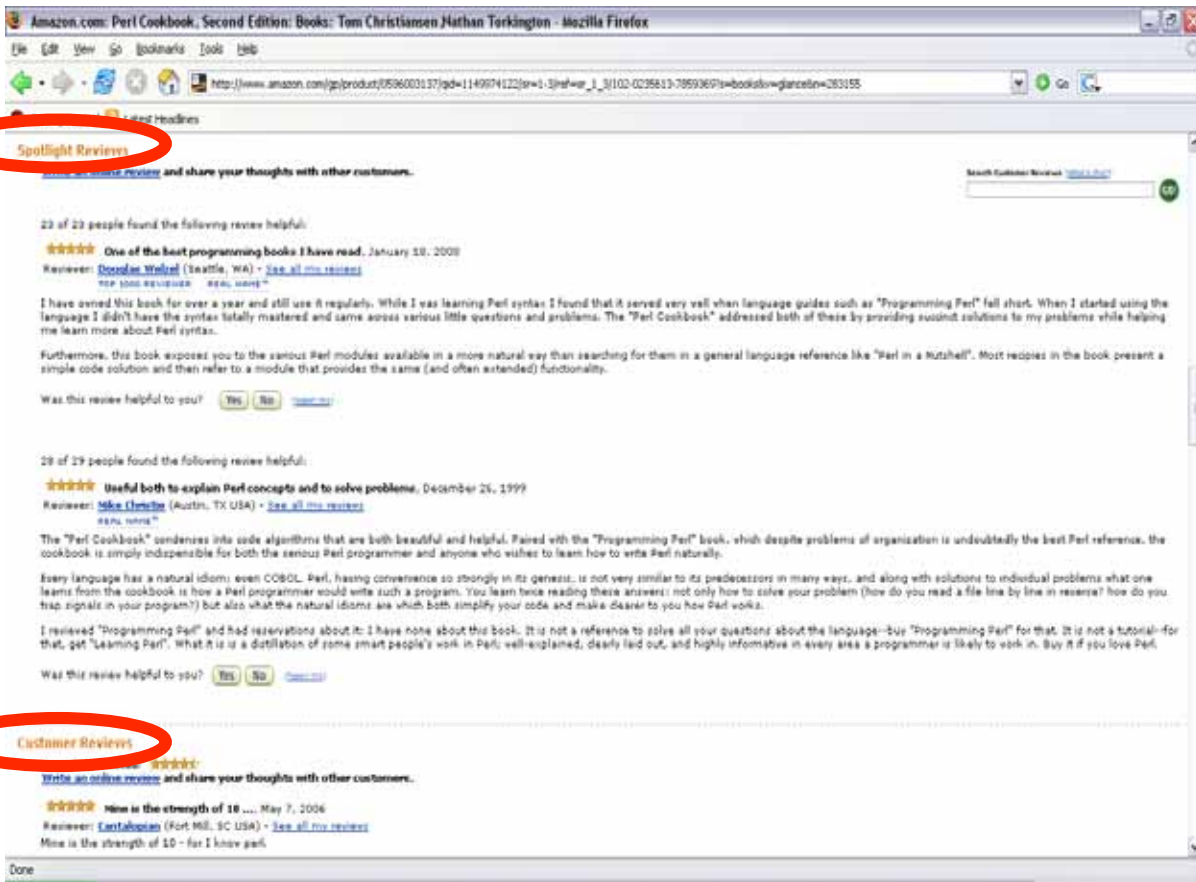


Figure 2: Sample Spotlight Reviews and Customer Reviews at Amazon.com

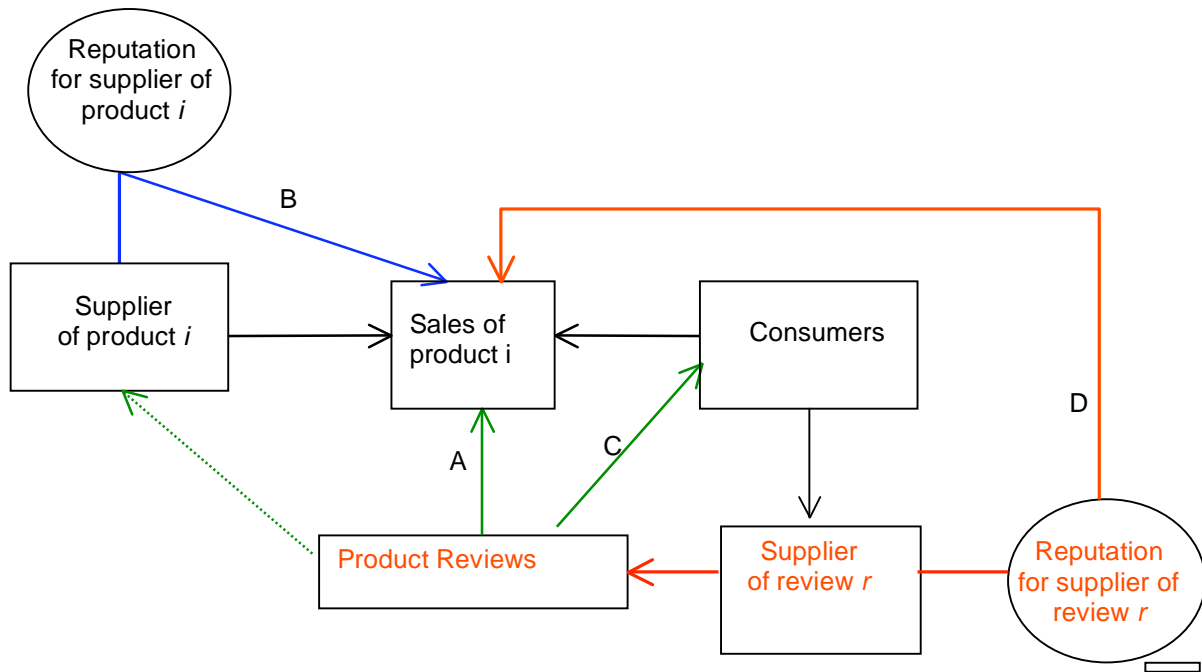


Figure 3: Research Framework