THE EFFECT OF WORD OF MOUTH ON SALES: ONLINE BOOK REVIEWS *

Judith Chevalier
judith.chevalier@yale.edu

Dina Mayzlin
dina.mayzlin@yale.edu

Yale School of Management
135 Prospect Street
P.O. Box 208200
New Haven, CT 06520-8200

April 28, 2004

* This is a preliminary draft. Please do not quote without the authors' permission. We would like to thank participants at Cornell, the Marketing Science Conference at the U. of Maryland, NYU, the University of Indiana, Yale Applied Micro lunch, MIT Marketing Seminar, Sharon Oster, Jackie Luan and David Godes for helpful comments. Both authors contributed equally and their names are listed in alphabetical order.

THE EFFECT OF WORD OF MOUTH ON SALES: ONLINE BOOK REVIEWS

Literature in marketing and economics identifies "word of mouth" or, between-customer communication, as a probable driver of consumer decision-making. However, establishing a causal link between "word of mouth" and consumer choice is difficult. Despite this, many Internet retailers have created business models that encourage between-customer communication. In this paper, we examine the effect of voluntarily-supplied customer reviews on subsequent sales of books at Amazon.com and Barnesandnoble.com. We find that customer reviews are overwhelmingly positive on both sites, but that there are longer reviews and more reviews at Amazon.com. We find that an improvement in reviews for a book at one site leads to a relative increase in the sales of that book at that site. Consistent with the hypothesis that buyers suspect that many reviewers are authors or other biased parties, we find that the marginal (negative) impact of 1-star reviews is greater than the (positive) impact of 5-star reviews. The results suggest that new forms of customer communication on the Internet have an important impact on customer behavior.

Keywords: advertising, word-of-mouth, source credibility, Internet marketing

1. Introduction

Online user reviews have become an important source of information to consumers, substituting and complementing other forms of word of mouth communication about the quality of various products. Indeed, one of the main ways in which a retail website can be said to differ consequentially from a mail-order catalog is that the website can provide many pages of descriptive information at essentially zero marginal cost. Consequently, many managers believe that a Web site needs to provide community content in order to build brand loyalty (See, for example, McWilliams (2000) or Fingar, Kumar, and Sharma (2000)). Despite this widespread belief, to our knowledge, there is no literature documenting that community content plays any role in consumer decision-making. Such a finding, it seems, is a necessary prerequisite for content provision to be a profitable strategy. A finding that recommender systems affect consumer purchasing decisions is also a potentially important input into any analysis of the consumer surplus created by ecommerce as a retail medium.

There are many reasonable arguments as to why making investments in providing such content could potentially be a poor strategy for ecommerce firms. First, it is not clear why users would bother to take the time to provide reviews for which they are not in any way compensated. Second, even if user reviews are provided, rival retailers can free ride on them; there is nothing to stop a consumer from utilizing the information provided by one website to inform purchases made elsewhere. Third, by providing user reviews, a site cedes control over the information displayed; unfavorable reviews created by either legitimate users or by biased interested parties may depress sales. Note that this may be less of a threat to a retailer that sells many different brands as opposed to a manufacturer. Similarly, since authors and publishers can freely proliferate favorable reviews for their own books, positive reviews may not be credible and may not function to stimulate sales.[1] Last, online user reviews may not be useful, and may not stimulate sales due to the sample selection bias that is inherent in an amateur review process. That is, a consumer only chooses to read a book or watch a movie if she perceives that there is a high probability that she will enjoy the experience. In the presence of consumer heterogeneity, this implies that the pool of reviewers will have a positive bias in their evaluation compared to the general population. Thus, positive reviews may simply be discounted by potential buyers.[2]

In this study, we characterize patterns of reviewer behavior, and examine the effect of consumer reviews on firms' sales patterns. In particular, we use publicly available data from the two leading online booksellers, Amazon.com (Amazon) and BN.comandNoble.com (BN.com), to construct measures of each firm's sales of individual books. Both BN.com and Amazon allow for customers to post reviews on the site. However, Amazon's investments in "collaborative" consumer content have been more extensive and much-imitated by other Internet retailers. By focusing on the differences between the two sites' sales of the same books, we examine the relationship between the customer reviews at each site and firm sales, controlling for other drivers of book sales.

---

[1] See Mayzlin (2003) for a theoretical treatment of recommendation systems where firms can anonymously post reviews.

[2] In a very different context, Resnick and Zeckhauser (2002) find that 99% of the feedback ratings on ebay.com are positive.

There has been a long-held belief in marketing[3] that word of mouth drives sales. Indeed, this work contributes to the broader literature on the link between customer word of mouth and sales, which has been demonstrated in several studies. For example, researchers have used word of mouth to explain the adoption of high-yield varieties of seeds among farmers (Foster and Rosenzweig (1995)), the adoption of tetracycline among physicians (Coleman (1966)), as well as evolution of the ratings of new TV shows (Godes and Mayzlin (2003)).

However, these studies have an important limitation in that they do not determine the direction of causality between word of mouth and product sales. For example, Foster and Rosenzweig (1995) infer that word of mouth was influential from the pattern of adoption: their method does not allow for establishing the direction of causality. Godes and Mayzlin (2003) demonstrate that TV shows with more dispersed conversations experience higher future ratings. However, the authors cannot rule out that the quality of the shows is driving both the conversations and the ratings. In general, the studies that have tracked sales and word of mouth over time suffer from an inability to rule out the alternative hypothesis that word of mouth may simply be correlated with total sales.

Theoretically, causality may work in either direction. For example, in herding models such as Banerjee (1992) and Bikhchandani et al. (1991), relatively small differences in signals received by the customers who initially sample the product may have lasting long-range consequences on the success or failure of a product – the early trials drive total sales. In their model, then, word of mouth is an important driver of sales. Alternatively, it may be a (noisy) signal of over-all performance since a product's quality is revealed with time or,

[3] Katz and Lazarsfeld (1955) cited word of mouth as an important source of information for certain purchase decisions.

perhaps, it may be correlated with an omitted variable, such as a successful advertising campaign. In this sense, word of mouth is an early measure of a product's success but not necessarily its driver.[4] For example, Van den Bulte and Lilien (2001) re-analyze Coleman's data to demonstrate that word of mouth influence on tetracycline adoption was over-estimated in the original study due to a lack of control for the marketing efforts of the drug companies.

In this study, we are better positioned to establish the causality between word of mouth and sales by comparing the sales and changes in sales of a given book across the two booksellers. Consider a highly publicized and anticipated book release, such *as Harry Potter and the Order of the Phoenix*, the fifth book in the popular Harry Potter series. We would not be surprised to find both high sales of this book and numerous positive reviews posted online. Clearly, we would not want to interpret the reviews as "causing" the sales of the book. However, the "traditional" methodologies employed in the prior literature would be to examine the sales of this book and the online reviews either through time, or in comparison to other books. These methodologies would suggest a positive relationship between customer reviews and book sales, but this relationship would not necessarily be causal in this case.

In our methodology, we examine the relationship between market shares and customer reviews for a *given* book *across* the two sites. By focusing on the *differences* between the market share of the book at the two sites, we are able to control for shocks to word of mouth and to sales that are common to both booksellers and, instead, focus on the idiosyncratic shocks alone. Consider a book that is generally well-reviewed and well liked. If a cranky consumer

[4] For example, Eliashberg and Shugan (1997) show that critical acclaim seems to serve as an early indicator of a movie's over-all box office success.

posts a negative review of that book on Amazon, but doesn't post that review on BN.com, will the market share of the book at Amazon fall relative to the market share of the book at BN.com?[5] This "ideal experiment" is the basis of our empirical strategy. Of course, data limitations force our analysis to differ somewhat from the ideal experiment, as we discuss later. However, we observe the same books, their customer reviews, and a proxy for each book's market share at each site. Our large database of books also allows us to control for other important factors that might affect the relative market share of a particular book across sites, such as differences across the sites in the price of the book or differences in the speed with which the book has been promised to be shipped. Furthermore, in order to partially rule out the hypothesis that the differences in word of mouth across sites are driven by unobservable underlying differences in the two populations, we show that the two sites are very similar in terms of customer preferences and reviewer behavior across broad categories of books in our sample. Finally, as an additional robustness check on the direction of causality, we obtain a second time point of data and examine the difference in the change in reviews and change in sales across sites.

Our user review data contains a star rating provided by the reviewer as well as a text description. In this paper, we focus our analysis on the star ratings. In fact, operationally, the star ratings provide an excellent opportunity to measure the valence of comments without analyzing the comments themselves, a very difficult task as demonstrated in Godes and Mayzlin (2003). We examine the incremental sales effects of having reviews for a

---

[5] Here, we define market share of a book as the sales of that book relative to the sales of all other books on that website. Of course, Amazon's total sales are greater than total sales at BN.com. However, we can say that a book is "more popular" at Amazon if it is the 5th most popular book at Amazon and the 10th most popular book at BN.com.

[7] The data on Amazon was collected on 5/6/2003 and the data on BN.com and Noble was collected on 5/7/2003.

particular book versus not having reviews and also the differential sales effects of positive and negative reviews.

Briefly, our results highlight some interesting characteristics of reviewer behavior. Reviews tend to be very positive on average, especially at BN.com. Under various specifications, we show that, if a particular book has more reviews and higher-starred reviews at one site, that book will tend to have a higher market share at that site. Our results on review lengths suggest that consumers actually read and respond to written reviews, not merely the average star summary statistic. Finally, we show that the results are robust to time series specifications designed to explore the reverse causality hypothesis. Our paper also contributes to the growing literature examining the economics of bookselling online. Recent contributions to this literature include Brynjolfsson and Smith (2000), Clay et. al. (2001), and Chevalier and Goolsbee (2003a, 2003b).

The rest of the paper is organized as following. In Section 2, we describe the data. In Section 3, we describe the methodology and present results on the distribution of reviews and sales across sites, providing further insight into the reviewing process. In Section 4, we discuss in more detail the model specification. In Section 5, we present our empirical analysis of the effect of word of mouth reviews on product sales. In Section 6, we present some of the limitations of our data set and obtain additional data in order to rule out arguments of reverse causality. In Section 7, we conclude.

2. Data

Our data consists of individual book characteristics and user review data that were collected from the public Web sites of Amazon and BN.com. The goal was to generate a representative sample of sites' sales. Since we do not have access to this proprietary data, we approximate a random sample of sales in the following way. First, we collect a random sample of books released. In order to maximize the probability that a book would be available on both Amazon and BN.com and Noble, we focus on a set of relatively recent books: titles that were released in the last five years. One shortcoming of a random sample of published books is that it overweighs books that have very few sales. One possible bias inherent in over-sampling these small titles is that word of mouth may be especially influential on the sales of these books, since there is little a prior awareness of these titles. Thus, in addition, we also extract a sample that consists of books that appeared at least once on a bestseller list. Hence, the sample was generated from two sources:

1) A random sample of books selected from a catalog "Global Books in Print" that were published in 1998-2002. (See Appendix for description of algorithm to generate the sample).
2) Publisher's Weekly bestseller lists: titles that appeared in the lists from 1/14/1991 to 11/11/2002.

Since a given book can be released in many different formats (such as hardback, paperback, etc), we use data from Bowker's Global Books in Print.com to obtain a listing of all possible English-language format releases of a given book. We discarded digital and audio format releases. Fortunately, each title-format combination has a unique International Standard Book Number (ISBN), assigned under the auspices of the International ISBN Agency in

Berlin. Though it may not be apparent to the casual user, both Amazon and BN.com and Noble.com use the ISBN numbers to organize the cataloguing of books on their web sites.

Over a two-day period in May of 2003, we searched the two Web sites to extract a body of data for each of the ISBN numbers in our sample.[7] Our extraction included: the title, author, publisher, release date, and format type of the book. We also gathered information on the price charged for the book at each website, the promised time until the book would ship, and data for the most recent 500 reviews of the book posted on the website (we extracted the number of stars assigned, the date the review was posted, and the full text of the review). Most books have far fewer than 500 reviews, but for those with more than 500 reviews, we also extracted the total number of reviews posted, as well as the average number of stars assigned overall.

Last, both BN.com and Amazon provide a "sales rank" for each book on the site. These sales ranks reflect the total sales of that book at that site relative to the sales of other books at that site. Note that books with *higher* sales are associated with *lower* ranks.[8] Chevalier and Goolsbee (2003a) report that Amazon claims that for books in the top 10,000 ranks, the rankings are based on the last 24 hours and updated hourly. For books ranked 10,001-100,000, the ranks are updated once per day. For books ranked greater than 100,000, the sales ranks are updated once per month (Amazon, 2000). Based on this system then, books that have not been purchased in the past month would not be ranked. Many hundreds of thousands of books, however, have a rank but almost certainly have less than one sale per

[8] In this study, we refer to higher-ranked books as books with higher numerical values of ranks or lower sales. This differs from every-day usage, where, for example, the "highest-ranked" sports team is ranked number 1.

month. Italie (2001) claims that for these rarely purchased books, Amazon bases the rank on the total sales since Amazon's inception. BN.com claims to update all the rankings daily (BN.com, 2000).[9] Thus, with the exception of the books that have very high ranks (low sales) on Amazon, the rankings represent a *current* snapshot of sales.

Amazon and BN.com provide identical reviews for all of the different formats of a given title. Since we do not want the dataset to include duplicate information, we examine sales and reviews only for the most popular ISBN (format) within a title. We then exclude from our analysis those books for which the most popular ISBN (format) within the title is different at Amazon and BN.com. That is, if the hardcover is the better seller at Amazon and the paperback is the better seller at BN.com, we exclude the book from our sample. This creates a sample of 2505 ISBN codes. Since we are not aggregating across books, we can use the sales ranks as is in our analysis, and discuss the impact of reviews on sales ranks directly. However, an extension of the methodology described in Chevalier and Goolsbee (2003a) and Schnapp and Allwine (2001) will allow us to also calibrate the sales rank relationships into total sales relationships. Any such calibration will be "back of the envelope", but will give us an opportunity to understand very approximate magnitudes.

We only include in our sample those books that are listed as "available" at both sites. Finally, we are forced to address the problem that BN.com only provides sales ranks for approximately 650,000 books and address the issue of "stale ranks" on Amazon. There are books at BN.com that are available for purchase but for which the rank is "too high" (sales

[9] Since BN.com provides rankings on tens of thousands of books that average far less than one sale per day, this statement cannot be completely accurate. They would not provide us any more detail on their ranking system (despite repeated requests).

are too low) to be disclosed. Amazon does not censor their sales ranks and they appear to range upwards of one million. If we were to use as our sample all books with prices and ranks at both sites, our sample would contain a large number of books that are relatively popular at BN.com, and relatively unpopular at Amazon. However, books that are relatively popular at Amazon and relatively unpopular at BN.com would not appear in the sample, as they have been censored out by BN.com's rank reporting strategy. To address this asymmetry, we remove those books with ranks above 650,000 at Amazon. More importantly, removing these books serves to remove books for which the ranks are updated very infrequently. As we argued earlier, for books with very high ranks, the ranks no longer represent a snapshot of current sales. Due to this, the sales could have preceded the posting of reviews on the site, in which case we would want to avoid concluding that customer reviews had any causal relationship to sales. The final sample contains 2394 observations, 1093 of which have reviews posted at both sites.

Table 2 presents the summary statistics for our data. The average sales ranks and the average prices in the sample are very similar across the two sites. Most of the books have a promised delivery of 24 hours (96% at Amazon and 88% at BN.com). However, Amazon and BN.com use other shipping categories such as "usually ships in 2-3 days" or "Special order: usually ships in 1-2 weeks." The two notable differences across the two sites are: 1) BN.com prices are significantly higher (as can be shown in a paired t-test),[10] 2) Amazon has more reviews than BN.com.

[10] While BN.com is currently more expensive than Amazon.com, this has not always been true historically. See Chevalier and Goolsbee (2003a), for example.

3. The Reviewing Process and the Distribution of Reviews

In this section, we provide information about the characteristics of reviews and compare the differences in preferences for different categories of books across the two sites.

Table 3 presents the cumulative distribution function on the number of reviews across the two sites. As is expected from the summary information, BN.com has a much higher fraction of books with zero reviews compared to Amazon (54.22% versus 12.61%). The median of the distribution on the number of reviews on Amazon is 11. However, both sites contain a few books with an enormous number of reviews. The two most reviewed books at each site in our sample were part of the "Harry Potter" series (by J.K. Rowling): *Harry Potter and the Goblet of Fire* with 4457 reviews on Amazon, and *Harry Potter and the Prisoner of Azkaban* with 956 reviews on BN.com.

On both Amazon.com and BN.com, there is a burst of reviewing activity when a book is new. For books in our sample for which at least 365 days had passed at the time we collected our data, about 17% percent of the Amazon reviews that were posted during the first year of reviewing for a book were posted during the first 30 days of reviewing. For BN.com, approximately 12% of the first year's reviews were posted in the first 30 days.

It is interesting to compare the frequency with which reviews are posted on Amazon and BN.com to review frequencies in other contexts. For example, Resnick and Zeckhauser (2003) find that over half of buyers on Ebay.com provide some feedback on a completed transaction. In contrast to Ebay, where a transactor can post feedback once and only once per completed transaction, the number of reviews at Amazon and BN.com may, in principle, be unrelated to those sites' past sales of the books. Customers who purchased a book elsewhere could post a review, and it is fairly simple for customers to post multiple reviews of a book. Nonetheless, these data suggest that review posting is, relative to feedback provision at Ebay, quite rare. For example, Amazon reported that its cumulative pre-orders of *Harry Potter and the Goblet of Fire* totaled 350,000 one minute *before the book's release* in July of 2000. It continued to be on the USA Today national top-ten bestseller list for the next two and half years, so it is likely that Amazon sold many more copies of the book after its release. The 4457 reviews of the book posted on the site, then, are very small in comparison to the site's overall sales of the book.

Next, we present the results on the distribution of star ratings in our sample, conditioning on a book having non-zero reviews on both sites. As Table 4 demonstrates, the average star ratings on both sites are quite high. It is interesting to compare the distribution of reviews in this paper to the distributions found in other contexts. For example, Resnick and Zeckhauser (2003) find that 48.3% of buyers on Ebay.com provide no feedback on transactions, 51.2% provide positive feedback, and only 0.5% provide negative or neutral feedback. In this sense, the reviews in this paper have a lot more variance in ratings than the feedback on Ebay.com. There are a number of reasons that can be used to explain this difference, including the fact that on Ebay both sellers and buyers rate each other, which can result in an incentive to post positive reviews by the buyer that are in turn reciprocated by the seller. Godes and Mayzlin (2003) find that in their sample of online conversations about TV shows, within the sub-sample where conversations could be described as either negative or positive, about 70% of posts were in fact positive. Thus, in all three settings, despite a predominance of positive reviews, there is some variance on the valence of reviews.

In addition, the reviews on BN.com are significantly more positive than the reviews on Amazon. An implication of this may be that consumers may be more skeptical when reading a 5-star review on BN.com, compared to a 5-star review on Amazon, which would imply that in our estimation we should account differentially for the effect of star ratings on the two sites. However, despite this general upwards bias, a significant number of reviews have 1 – 3 stars.

Beyond the ratings given by the reviewers, there might be additional information contained in the message text. Unfortunately, reading the reviews is an extremely costly task, and the measures obtained are very noisy as is shown by Godes and Mayzlin (2003). However, one relatively cost-effective measure of the review text is the length (total number of typed characters) contained in the review. A priori, it is not completely clear how to interpret this measure. One possibility is that a longer review represents more effort on the part of the reviewer. Another possibility is that a longer explanation is required to support a “mixed” review. We find partial support for the latter interpretation: Table 5 shows that, at both sites, 1-star and 5-star reviews are much shorter than 2-star, 3-star, and 4-star reviews. Another pattern that emerges is that Amazon reviewers post longer reviews at all star levels than do their peers at Bn.com. This can be due to several reasons, such as Amazon’s efforts to elicit more nuanced reviews from its consumers.

4. Model Specification

Chevalier and Goolsbee (2003a) and Schnapp and Allwine (2001) provide evidence that the distribution of book sales ranks and book sales follow a Pareto distribution. In the Pareto distribution, the probability that an observation, s, exceeds some level, S, is an exponential function

$$\Pr(s > S) = (k / S)^{\theta}$$

where k and θ are the parameters of the distribution. The most important parameter is θ, the shape parameter that indicates the relative frequency of large observations. If θ is 2, for example, the probability of an event decreases in the square of the size. With a value of 1, it decreases linearly.

If there are a sufficient number of books to eliminate discreteness problems, the probability that a book's sales exceed some level S can be approximated as (Rank-1)/(Total Number of Books). Taking logs, we can translate sales quantities into ranks according to

$$\ln(Rank - 1) = c - \theta \ln(Sales)\ . \qquad (1)$$

We consider sales-rank relationships at two websites, Amazon.com and BN.com. To do this, following Chevalier and Goolsbee (2003a), we allow c to differ across websites—c^A and c^B, but assume that the shape of the Pareto distribution, parameterized by θ, to be the same for all websites. Chevalier and Goolsbee justify this assumption by showing that the Pareto distribution is a good fit for book sales generally, and for Amazon’s sales specifically. Chevalier and Goolsbee provide estimates of the Pareto parameter in the range of 0.9 to 1.5. In the calibrations that follow, we will illustrate results using a Pareto parameter of 1.3.

Consider a book i that is sold on Amazon and BN.com. We hypothesize that the quantity of books sold at a site is a function of a book fixed effect (which may include either off-line promotion, the quality of the book, or the popularity of the author) as well as other factors. That is,

$$\ln(\text{quantity}^A_i) = \nu_i + \alpha^A \ln(P^A_i) + \gamma^A \ln(P^B_i) + X\Gamma^A + S\Pi^A + \varepsilon^A_i \quad (2)$$

$$\ln(\text{quantity}^B_i) = \nu_i + \alpha^B \ln(P^B_i) + \gamma^B \ln(P^A_i) + X\Gamma^B + S\Pi^B + \varepsilon^B_i \quad (3)$$

where the superscripts A and B refer to Amazon and BN.com respectively and the subscript i indexes the book title. We adopt the familiar constant elasticity (log-log) demand specification. P denotes price. X denotes the vector of review variables from both sites; we allow Amazon reviews to affect BN.com's customers and BN.com's reviews to affect Amazon's customers. S is a vector of dummy variables summarizing the shipping times promised by each website for each book. For each book, S has a 1 for the promised ship time category at Amazon and a 1 for the promised ship time category for that book at BN.com.

Note that we would expect the unobservable book fixed effect, $\nu_{i,}$ to be correlated with the review variables. Thus, omitting this effect would bias the coefficients on the review variables. In order to eliminate ν_i, we can estimate (2) – (3),

$$\ln(\text{quantity}^A_i) - \ln(\text{quantity}^B_i) = \beta^A \ln(P^A_i) + \beta^B \ln(P^B_i) + X\Gamma + S\Pi + \varepsilon_i \quad (4)$$

Because S is exhaustive of all of the shipping time categories, we do not include a constant term in the regression. In the interest of space, we don't present the parameters of Π in the tables in Section 5.

The only complication in estimating such a specification is that we do not, in fact, have direct data on sales, but rather, only data on sales rank. However, to translate this model into one that we can use rank data for requires us only to substitute the log sales rank for the log quantity in equation (4) according to the Pareto relationship in (1):[11]

$$\ln(rank^A_i) - \ln(rank^B_i) = -\theta\beta^A \ln(P^A_i) - \theta\beta^B \ln(P^B_i) - \theta X\Gamma - \theta S\Pi - \theta\varepsilon_i \qquad (5)$$

In other words, estimating the equations using log ranks, r, rather than actual quantities, yields coefficients equal to minus one times the correct coefficients in the sales regression, but scaled up by the Pareto shape parameter, θ.

To ensure that (5) is identified, we need to make the following assumptions:

A1) The universe of books offered at the two sites is the same.

A2) The two sites draw from the same population of consumers.

A3) The reviews generated on a site are more likely to impact that site's customers.

A4) There is variance in consumer reviews across the two sites.

Assumption (A1) ensures that comparison in ranks across the two sites is meaningful. To examine the reasonableness of A1, we examined the availability of books at BN.com that were unavailable at Amazon and vice versa. We found that, in a sample of 20,228 ISBNs

[11] Note that we use ln(rank) rather than ln(rank-1) in the empirical specifications. We do this for simplicity and to match the specification estimates in Schnapp and Allwine (2001). Given that the mean Amazon ranking in our data is around 129,000, calculating ln(rank) rather than ln(rank-1) is immaterial.

(the starting point for the sample in this paper), 9577 were unavailable at Amazon.com and 9301 were unavailable at BN.com. Of these, 7912 were unavailable at both sites. Books that were unavailable at BN.com but available at Amazon.com had an average Amazon.com ranking of 817,551, indicating very low sales, low enough to be censored out of our specifications below.

Assumption (A4) appears to hold in our sample. In our sample of 2394 books, the correlation between a book's average star on the two sites is relatively low: 0.329 (see Table 7). We have no mechanism for proving that (A3) holds. Indeed, there is some evidence that reviews do exert influence beyond the site on which the reviews are posted. For example, the success of a recently released best-seller "DaVinci Code" was attributed partly to an endorsement by a prolific Amazon reviewer: Francis McInerney.[12] However, if reviews had an equal impact on both sites' customers, the effect of reviews would cancel out in Equation (5): we would observe no effect of reviews on rank difference. In this sense, to the extent that (A3) fails to hold, our measures of the relationship between reviews and sales at a particular site are an underestimate of the true influence on reviews on sales.

Finally, demonstrating that (A2) holds is crucial for identifying Equation (3). That is, suppose that BN.com customers prefer fiction to non-fiction, while Amazon customers have opposite preferences. We would observe higher ratings and higher market shares for fiction books on Amazon, and higher ratings and higher market shares for non-fiction books

---

[12] Paumgarten, N. "No. 1 Fan Dep't Acknowledged," www.newyorker.com, Issue of 2003-05-05, posted 2003-04-28.

[14] In fact, each book may contain up to 6 subjects. We used first subject only in this study.

and BN.com. However, the inferred link between ratings and sales would be essentially due to differences in preferences across the users of the two sites.

In terms of the equations above, suppose that (1) and (2) also contain a book-site fixed effect that is correlated with the review variables:

$$\ln(rank^A_i) = \mu_i^A + \nu_i + \alpha^A \ln(P^A_i) + \gamma^A \ln(P^B_i) + X\Gamma^A + S\,\Pi^A + \varepsilon^A_i \qquad (6)$$

$$\ln(rank^B_i) = \mu_i^B + \nu_i + \alpha^B \ln(P^B_i) + \gamma^B \ln(P^A_i) + X\Gamma^B + S\,\Pi^B + \varepsilon^B_i \qquad (7)$$

In this case, these effects would not cancel out in (as in 5) and their omission would bias the coefficients on review variables:

$$\ln(rank^A_i) - \ln(rank^B_i) = \mu_i^A - \mu_i^B + \beta^A \ln(P^A_i) + \beta^B \ln(P^B_i) + X\Gamma + S\Pi + \varepsilon_i \qquad (8)$$

To eliminate $\mu_i^A - \mu_i^B$, we need to obtain another time point of data and to difference (8) across time. This is our strategy in Section 6, where we relax (A2) and estimate:

$$\Delta[\ln(rank^A_i) - \ln(rank^B_i)] = \beta^A\, \Delta\ln(P^A_i) + \beta^B\, \Delta \ln(P^B_i) + \Delta X\Gamma + \Delta S\Pi + \varepsilon_i \qquad (9)$$

However, for most of the paper, we assume that (A2) holds. There is considerable evidence to suggest that, in fact, this is a reasonable assumption. Below we demonstrate that the two sites appear to have similar reviewer and purchasing preferences. As a first cut, the correlation between log ranks of individual book titles across the two sites is very high, 0.825 for the 2394 books in the sample. In addition, to address the issue of different subject preferences, we collect data on book subjects. The book's subject is in most cases classified using the system provided by Book Industry Standards and Communications (BISAC) and is available on Bowker's Global Books in Print.com.[14] In cases where the BISAC subject was not available, we used the subject classification on Amazon. Further, using the original sample of 6429 titles, we aggregated the subjects into broader categories. The complete classification is available in the Appendix.

In Columns 2 and 3 of Table 6, we present the results of normalized mean ratings for each category. The normalized mean rating for category j at site k is defined as the (mean star rating for books in category j at site k – overall mean star rating across books at site k)/overall standard deviation of star ratings for books at site k. In constructing this measure, we use only the sample of books that have non-zero number of reviews at both sites.

Reviewing patterns for different categories of books are remarkably similar across sites. In particular, we find that the signs for normalized mean ratings are identical for all categories, and the magnitudes are similar across the two sites for most categories. On both sites, for example, juvenile fiction is the highest rated category. That is, reviews posted for books in the juvenile fiction category are typically very positive on both sites. On both sites, the least liked books are in the "serious non-fiction" category.

Finally, we also show that the sales within a category are similar across the two sites. That is, for each individual book in the sample we can assign a number 1-4 based on the quartile in which it falls in the distribution of Amazon log rank (we can perform the same procedure based on BN.com log rank). Column 4 of Table 6 presents the percentage of books (by category) for which the Amazon log rank quartile and BN.com quartile match. The percentage of match exceeds 70% for all categories. This, along with similarity of reviewer preferences, demonstrates a lack of obvious differences in purchasing preferences across the users of the two sites. Thus, given the evidence of similarity of consumers across sites, we proceed in Section 5, assuming that (A2) holds.

5. The effect of reviews on sales

In this section, we examine the relationship between a book's customer reviews and its sales rank across sites:

$$\ln(\text{rank}^A_i) - \ln(\text{rank}^B_i) = \beta^A \ln(P^A_i) + \beta^B \ln(P^B_i) + X\Gamma + S\Pi + \varepsilon_i \qquad (5)$$

Table 7 presents the correlation matrix for the full sample of 2394 observations. Table 9 presents the estimation results for this sample. Column one of Table 9 presents the results for a regression in which no review variables are included, only prices at both sites and the shipping dummies. The price coefficients reflect a combination of own- and cross-price elasticities at both sites. The price coefficient for Amazon is positive and statistically significant, suggesting that, when prices rise, sales ranks at Amazon become larger, that is, sales fall. The price coefficient is negative for BN.com. This is as expected; recall that the left hand side variable is ln(rank) at Amazon *minus* ln(rank) at BN.com. Again, when prices rise at BN.com, sales ranks become larger, that is, sales fall at BN.com relative to Amazon. The absolute value of the price coefficient is larger at BN.com, suggesting that sales ranks respond more to prices at BN.com than at Amazon. This is consistent with the findings in Chevalier and Goolsbee (2003a) that demand is more elastic at BN.com than Amazon.

For approximate magnitude measures, we refer to Schnapp and Allwine (2001). They have proprietary data from a single publisher from May of 2001 relating that publisher's sales at

Amazon to that publisher's sales ranks. They fit the sales-ranks relationship for a subsample of the publisher's titles as:

$\ln(\text{sales}^{AMZN}) = 9.61 - 0.78 \ln(\text{rank}^{AMZN})$

Notice that this translates into a Pareto parameter of 1.3. While they do not provide R-squareds or other measures of fit, the scatterplots they supply suggest that the fit is very good and suggests no obvious objection to the underlying distributional assumption. Since this dates from 2001, we scale up their sales estimates by 24%, the growth in Amazon's North American sales in the two years intervening between the time of our sample and the time of their sample. BN.com does not report data to publishers in a way that allows them to make such a comparison. We assume that the basic shape of the rank to sales relationship is the same at BN.com as it is at Amazon, but that it is scaled down to reflect the fact that BN.com's total sales equal about 15% of Amazon's North American sales.

Using the relationship between ranks and sales, we construct an example will give a general sense of the magnitudes of the price elasticities. Consider a book whose other characteristics led to a sales rank of 500 at both Amazon and BN.com with a price of $10 at both sites. This sales rank corresponds to sales of roughly 145 copies per week at Amazon.com and 21 per week at BN.com. Increasing the price at Amazon to $12 would be predicted to change the difference in the log ranks at both sites to 0.28 , as for example would occur if the rank at Amazon moved to 580 and the rank at BN.com moved to 437. What does that mean for sales? The calibrations described above that extend the results from Schnapp and Allwine (2001), suggest that, in the example above, Amazon's sales of the book would fall from approximately 145 per week to 129 per week, while BN.com's sales would rise from approximately 21 units per week to 24.

Column 2 includes measures of the total number of reviews for each book. The variables include the natural log of the total number of reviews at Amazon and the natural log of the total number of reviews at BN. These are set to zero when the number of reviews equals zero. We also include dummies, one that takes the value one when a title at Amazon has no reviews (and zero otherwise) and one that takes the value one when BN.com has no reviews (and zero otherwise). These results suggest that ranks are lower (sales higher) at Amazon when Amazon has more reviews, and that ranks are lower (sales higher) at BN.com when BN.com has more reviews. This is consistent with evidence from a different data sample in Chevalier and Goolsbee (2003b). The magnitudes are non-trivial. Consider a book with no reviews at either site whose price and other characteristics would suggest a sales rank of 500 at both sites. The posting of an additional 3 reviews at Amazon, if it didn't alter the sales rank at BN.com, would be expected to lower the sales rank to number 327, implying incremental sales of approximately 57 books per week.

It is important to consider how to judge these results in the presence of possible endogeneity bias. To consider this, it is important to recall that the sales ranks represent "current" sales, while the reviews are largely older. Fewer than 2% of the reviews in our sample were written during the previous two months. Thus, the reviews predate the measured sales temporally. In unreported results, we recalculate all results in the paper using only data from reviews that are more than 2 months old. This leads to almost numerically identical results. However, the results could be an artifact of an endogeneity bias if there are differences in preferences across the BN.com user population and the Amazon user population. Our analysis of reviewing behavior does not expose such differences in preferences, but nonetheless, concerns may remain. We argue that our "average star" results below are less vulnerable to endogeneity concerns than the "number of reviews" specifications, and, further, that the time-series results in Section 6 will help allay endogeneity concerns.

The specification in Column 2, however, might be somewhat misleading in that obviously, not all reviews are created equal. As the summary data in Section 3 showed, reviews are, on average quite enthusiastic, with at least half of the reviews being 5 stars on both sites. Thus, by including the number of reviews in Column 2, but omitting their content, we are implicitly measuring the effect, on average, of new favorable reviews being posted. Column 3 of Table 9 improves upon this specification, by including the average star value of the book's customer reviews at each site in the regression. Note that the sign of the coefficient on the Amazon no reviews dummy changes between Column 2 and Column 3. This is due to the inclusion of average star value in Column 3. The coefficient on the no reviews dummy must be interpreted keeping in mind that a book that gets its first review also experiences a change in its average star rating (from zero to a positive number). Suppose that a book has no reviews on either site. If it gets one Amazon review with 1, 2 or 3 stars, its rank on Amazon will rise (sales fall), assuming that its rank on BN.com stays constant. If, on the other hand, it gets a positive review: 4 or 5 stars, its rank on Amazon will fall (sales rise). As expected, for both sites, the coefficients for the average star value suggest that sales improve when books are rated more highly, but the effect is statistically insignificant for BN.com. To illustrate the magnitude of the effects, consider a book with four 5-star reviews at both Amazon and BN.com and a rank of 500 at both sites. Now imagine that one of the 5 star reviews at Amazon were changed to a 1-star review. The coefficients imply that, if BN.com's ranking of the book were unchanged by this review change, the rank at Amazon would be expected to fall to 603, an estimated change in sales of about 20 books per week.

Column 4 focuses on a different way of measuring review valence. The fraction of reviews that are 1 star reviews and the fraction of reviews that are 5 star reviews are included for each site. As expected, the coefficients suggest that 5 star reviews improve sales and 1 star reviews hurt sales in a statistically significant way at Amazon. The coefficient for 1 star reviews for BN.com is of the expected sign and statistically significant at the 6 percent level. However, the coefficient for 5 star reviews is almost zero but of the "wrong" sign. Nonetheless, it is interesting to note that the 1 star reviews have large coefficients in absolute value, relative to the 5 star reviews, indicating that the relatively rare 1 star reviews carry a lot of weight with consumers. This result also makes sense when one considers the credibility of 1-star and 5-star reviews. After all, the author or other interested party may "hype" his or her own book by publishing glowing reviews on these websites.[15] While the author can post a large number of meaningless 5 star reviews cheaply, he or she cannot prevent others from posting 1-star reviews.[16]

The robustness of the estimates in Table 9 is further examined in Table 10 (the pairwise correlations are presented in Table 8 for this sample). In particular, in Table 10 we examine only the subsample of 1093 books that have at least one review on each site. We drop the "no review" variables, but measure the impact of number of reviews and star rankings for

[15] For one well-publicized example in economics, see Morin (2003).

[16] One could argue that posting 1 star reviews of competing books could be a reasonable strategy for an author. We acknowledge that this may be true, although it is not at all clear that two books on the same subject, for example, are substitutes rather than complements.

this subsample. The results are similar to those presented above. However, the coefficient magnitudes and significance levels for the variables measuring star rankings are somewhat larger, emphasizing the importance of having higher star rankings for this subsample.

Finally, we examine the relationship between review lengths and sales. To do this, we repeat the specification in Table 11, including the natural log of the average length of all of the reviews for each book at each site. The coefficient is positive and statistically significant at Amazon, negative and insignificant at Bn.com. This suggests, controlling for the star rating of the book, longer reviews depress the site's relative share. We check the robustness of this result by replacing the average star measures with the fraction of 1 star, 2 star, 3 star, 4 star and 5 star reviews. We find these results quite robust.

There are (at least) two possible interpretations of this result. The first, which we view as the less likely, is that encouraging longer, more useful, more nuanced reviews is in fact harmful to sales. More likely, however, is that, within each site, the length of the review is correlated with the enthusiasm of the review in ways that are not captured by the star measures. For example, even within the realm of the statistically dominant 5 star reviews, there could be differing degrees of enthusiasm. That is, some "read like" 4.5 star reviews, while some read more like 5-star reviews. The ones that read like 4.5 star reviews might on average be longer since they are more likely to be mixed – to mention the negative as well as positive aspects of the book. We find some evidence for this in our data. Consider the subsample of 1093 books with at least one review at both sites. Within that group, consider the subsample of 5 star reviews. The average length of these 5 star reviews at Amazon is 796 characters for books whose average Amazon star rating is 4 or greater, and is 849 characters for books whose average Amazon rating is less than 4. Similarly, the average review length at Bn.com is 491 for 5 star reviews for a book for which the average rating is 4 or greater, and 672 for 5 star reviews for a book for which the average rating is less than 4. Assuming that the books with the lower average ratings have the "less enthusiastic" 5 star reviews, this at least suggests that even within the 5-star category, review length is correlated with the reviewer's level of enthusiasm for the book. Regardless of the interpretation of the length results, the results do seem to suggest that customers read and respond to the review content at each site. However, longer reviews do not necessarily stimulate sales.

6. The effect of changes in reviews on change in sales

As discussed earlier, omitted book-site fixed effects could bias the results above. Such a situation could arise if, for example, the customer populations at the sites differed in their reviewing and purchasing preferences, or, if editorial reviews posted at one site but not the other led consumers to both post more positive reviews and to buy more books. In this section address these concerns by obtaining additional data and utilizing a "differences in differences" specification. The "differences in differences" specification below should further lessen such concerns, as we argue in Section 4.

We collect review data for May 8 – July 8, 2003, and the ranks, prices, and shipping data as posted on Aug 8, 2003[17] for the sample of 2394 books analyzed in the previous Section. The specification we estimate is:

$$\Delta[\ln(\text{rank}^A_i) - \ln(\text{rank}^B_i)] = \beta^A \Delta\ln(P^A_i) + \beta^B \Delta \ln(P^B_i) + \Delta X\Gamma + \Delta S\Pi + \varepsilon_i \qquad (7)$$

[17] That is, $\Delta\ln(P^B$ for book i)= $\ln(P^B$ posted in August for book i)- $\ln(P^B$ posted in May for book i), while $\Delta\ln$(Number of reviews^B on Amazon for book i)= ln(Number of reviews^B in July for book i)- ln(Number of reviews^B in May for book i).
The one month gap between the last review data collected and the rank data collected was to eliminate the possibility that the sales that possibly generated the reviews were included in the dependent measure.

The differenced data controls for editorial content posted on one site but not the other, as long as the editorial content is the same at both time points. Indeed, we do not expect that there would be a lot of changes in editorial content over time since most professional reviews appear either right before or soon after a book is published. Thus, differencing the data across time should largely eliminate the impact of editorial reviews.

Finally, as we argued in Section 4, this specification eliminates a book-site "fixed effect." This specification would not eliminate time-variant population differences, but large effects of this sort seem unlikely since the difference between the two data points is just three months.

Out of the sample of 2394 books, only 2091 books were available at both sites in the second period and contained rank information at both sites. In the two months under consideration, new reviews were posted for 766 titles at Amazon and for 324 titles at BN.com. There are only 276 titles that contained new reviews at both sites. In this set of 276, an average of 8.88 reviews was posted on Amazon (with an average star of 3.87) and an average of 3.94 reviews was posted on BN.com (with an average star of 4.40). Thus, one limitation of our analysis (which biases against significance) is that we have relatively little new reviewing activity.

The results of the estimation are presented in Table 12. Columns 1 and 2 present estimation results that include differences in average stars and number of reviews for the whole sample of 2091 books and the sample of 276 books that had reviews at both sites. Columns 3 and 4 present results that include differences in fraction of 1-stars and 5-stars for these two samples respectively. We can compare Columns 1 and 3 of Table 12 to Table 9 (Columns 3 and 4 respectively), and Columns 2 and 4 of Table 12 to Table 10 (Columns 3 and 4 respectively). As we can see, the magnitudes on price elasticities in Table 12 are lower than in the previous tables. This may be due to relatively little variance in prices over time. In contrast, most of the coefficients on review variables are actually higher in magnitude, even though some are no longer significant.

Qualitatively, most of the results of the previous Section are replicated. Thus, an increase in average star on Amazon over time results in higher relative share of the book on Amazon over time (one month after the reviews under consideration have been posted). Similarly, the opposite holds for change in average star on BN.com. The results for fraction of 5-stars and 1-stars are also consistent with this intuition. We again find evidence that 1-star reviews have a bigger impact than 5-star reviews on the same site. As expected, an increase in the difference in the number of reviews on Amazon over time is associated with a greater relative share of the book an Amazon over time. The only exception we find is for the difference in number of reviews on BN.com over time. The coefficient is, surprisingly, of the wrong sign (albeit, it is only significant in Column 4). However, it is important to notice that the difference in the change in the number of reviews at Amazon and the change in the number of reviews at BN.com continues to be negative. Thus an increase in the number of reviews at Amazon relative to BN.com continues to improve sales at Amazon relative to BN.com. Separately identifying the coefficients for the change in the number of reviews at Amazon and the change in the number of reviews at BN.com may simply be impossible given the limited number of new reviews in the sample.

7. Conclusion

We analyze reviewing practices at Amazon and BN.com. We find that customer reviews tend to be very positive at both sites, that they are more detailed at Amazon, and that the relative popularity of different types of books is very similar across sites. Our regression estimates suggest that the relative market share of a book across the two sites is related to differences across the sites in the number of reviews for the book and in differences across the sites in the average star ranking of the reviews.

This evidence suggests that customer word-of-mouth has a causal impact on consumer purchasing behavior at two Internet retail sites. We believe that this has not been shown before. That customer content impacts sales is certainly a prerequisite for differences in customer content quality to have any impact on differences in revenues or profitability across retailers. Our evidence however, stops short of showing that the retailer profits from providing such content. For example, there is nothing in our evidence that shows that customer reviews do not merely move sales around across books within a site. Since Amazon has many more reviewers than rivals, its reviews are on average quite lengthy, and its reviews are on average quite positive, it seems plausible to at least speculate that the total number of books sold at Amazon is higher than it would have been absent the provision of customer review features. Further, and more interestingly, our results show that customers certainly behave *as if* the fit between customer and book is improved by using reviews to screen purchases. One interesting extension to this research would be to examine whether improving a customer's satisfaction with his or her purchases affects subsequent customer loyalty.

7. Tables

Table 1: Initial Sample of Books

Source	**# Unique titles**
Books in Print	3,617
Publisher's Weekly	2,812
Total	6,429

Table 2: Summary data.
The sample is all books in our database for which the most popular format of the book at Amazon is the same as the most popular format of the book at BN.com.

Summary information

Variable	Mean	Std. Dev.	Min	Max
Amazon sales rank	129467.50	169227.30	7	645406
BN.com sales rank	120872.70	156829.50	6	647611
Amazon price	13.96	14.39	3.25	250
BN.com price	15.50	14.73	3.25	250
Amazon no of reviews	61.27	180.27	0	4457
BN.com no of reviews	12.87	44.60	0	956
Shipping Dummies				
Amazon, up to 24 hours	0.959			
Amazon, 2-7 days	0.024			
Amazon, more than a week	0.005			
Amazon, special order	0.012			
BN.com, 24 hours	0.882			
BN.com, more than 24 hours	0.118			
Number of observations	2394			

Table 3: CDF on the number of reviews.
The sample is all books in our database for which the most popular format of the book at Amazon is the same as the most popular format of the book at BN.com.

Amazon

x	Prob(no of reviews < = x)
0	12.61
1	22.18
3	34.21
5	40.02
11	50.79
20	60.69
37	70.47
64	80.16
146	90.02
280	95.03
4457	100

BN.com

x	Prob(no of reviews < = x)
0	54.22
3	64.04
5	70.97
12	81.12
29	90.23
61	95.03
956	100

Table 4: The distribution of stars.
The sample is as in Tables 2-3 with the additional restriction that non-zero reviews have been posted at both sites.

Amazon

Star Rating	Percentage
1 star	8.97
2 stars	7.53
3 stars	10.56
4 stars	19.89
5 stars	53.05

Average Rating	4.01 stars

BN.com and Noble.com

Star Rating	Percentage
1 star	3.44
2 stars	4.07
3 stars	6.00
4 stars	19.27
5 stars	67.22

Average Rating	4.45 stars

Table 5: Average review length by site and number of stars

	Amazon	Bn.com
1 star reviews	765	558
2 star reviews	916	599
3 star reviews	997	566
4 star reviews	949	577
5 star reviews	812	508
Overall	854	529

Table 6: The similarity in ranks and reviews across sites (by subject category).
The sample is as in Table 4.

Category	# books	Amzn avg star	BN avg star	Logrank Quartile Match
Adult Fiction	669	-0.261	-0.169	75.04%
Adult Non-Fiction	101	0.162	0.103	77.23%
Do-it-yourself	18	0.290	0.328	88.89%
Entertainment	8	0.301	0.195	75.00%
Juvenile	144	0.782	0.557	71.53%
Language & Arts	20	0.406	0.260	70.00%
Serious Non-fiction	20	-0.617	-0.697	75.00%
Self-Improvement	60	0.371	0.130	76.67%
Social Science	48	0.374	0.304	83.33%
Travel	5	0.370	0.260	80.00%

Table 7: Pairwise correlation matrix for the same sample as in Tables 2-3.

	ln(Amzn rank) - Ln(BN rank)	Amzn ln(price)	BN ln(price)	Amzn ln(no. of reviews)	BN ln(no. of reviews)	Amzn avg star	BN avg star
ln(Amzn rank) - ln(BN rank)	1.000						
Amzn ln(price)	-0.158	1.000					
BN ln(price)	-0.215	0.965	1.000				
Amzn ln(no. reviews)	-0.070	-0.187	-0.200	1.000			
BN ln(no. reviews)	0.033	-0.146	-0.159	0.797	1.000		
Amzn avg star rating	-0.104	-0.160	-0.159	0.335	0.170	1.000	
BN avg star rating	0.041	-0.207	-0.221	0.628	0.633	0.329	1.000

Table 8: Pairwise correlation matrix for the same sample as in Table 4.

	ln(Amzn rank) - ln(BN rank)	Amzn ln(price)	BN ln(price)	Amzn ln(no. of reviews)	BN ln(no. of reviews)	Amzn avg star	BN avg star
Ln(Amzn rank) - ln(BN rank)	1.000						
Amzn ln(price)	-0.241	1.000					
BN ln(price)	-0.305	0.958	1.000				
Amzn ln(no. reviews)	-0.159	0.086	0.079	1.000			
BN ln(no. reviews)	-0.023	0.011	0.005	0.785	1.000		
Amzn avg star rating	-0.102	-0.059	-0.061	-0.183	-0.087	1.000	
BN avg star rating	0.006	-0.078	-0.091	-0.136	-0.023	0.611	1.000

Table 9: The effect of reviews on sales.
This table shows regressions in which each data point is a book sold at both Amazon and BN.com. The sample is as in Table 2-3.
Dependent variable is the difference between the sales rank of the book at Amazon and the sales rank of the book at BN.comand .
Dependent variable: ln(Amazon sales rank) - ln(BN.com sales rank)

	(1)	(2)	(3)	(4)
Amazon ln(price)	1.574***	1.568***	1.564***	1.549***
	(0.160)	(0.156)	(0.155)	(0.156)
BN ln(price)	-1.821***	-1.863***	-1.859***	-1.845***
	(0.148)	(0.145)	(0.144)	(0.145)
Amazon ln(no. of reviews)		-0.191***	-0.218***	-0.208***
		(0.023)	(0.024)	(0.024)
BN ln(no. of reviews)		0.118***	0.133***	0.132***
		(0.033)	(0.033)	(0.033)
Amazon no reviews dummy		0.234***	-0.586***	0.072***
		(0.084)	(0.187)	(0.109)
BN no reviews dummy		-0.253***	-0.147	-0.337**
		(0.082)	(0.100)	(0.131)
Amazon average star rating			-0.187***	
			(0.038)	
BN average star rating			0.025	
			(0.017)	
Amazon fraction reviews 5 star				-0.260***
				(0.100)
BN fraction reviews 5 star				-0.127
				(0.149)
Amazon fraction reviews 1 star				0.508**
				(0.256)
BN fraction reviews 1 star				-0.884*
				(0.467)
No. observations	2394	2394	2394	2394
includes shipping dummies?	y	y	y	y
R-squared	0.088	0.131	0.140	0.139
*** $p < 0.01$				
** $p < 0.05$				
* $p < 0.10$				

Table 10: The effect of reviews on sales.
This table shows regressions in which each data point is a book sold at both Amazon and BN.comand. The sample is as in Table 4.
Dependent variable is the difference between the sales rank of the book at Amazon and the sales rank of the book at BN.comandNoble.com.

	(1)	(2)	(3)	(4)
Amazon ln(price)	2.136***	2.218***	2.208***	2.183***
	(0.339)	(0.332)	(0.327)	(0.328)
BN ln(price)	-2.644***	-2.640***	-2.644***	-2.617***
	(0.291)	(0.285)	(0.281)	(0.282)
Amazon ln(no. of reviews)		-0.336***	-0.381***	-0.377***
		(0.049)	(0.049)	(0.050)
BN ln(no. of reviews)		0.221***	0.241***	0.243***
		(0.052)	(0.051)	(0.052)
Amazon average star rating			-0.444***	
			(0.079)	
BN average star rating			0.133	
			(0.087)	
Amazon fraction reviews 5 star				-0.723***
				(0.236)
BN fraction reviews 5 star				0.083
				(0.188)
Amazon fraction reviews 1 star				1.194**
				(0.505)
BN fraction reviews 1 star				-0.986*
				(0.566)
No. observations	1093	1093	1093	1093
includes shipping dummies?	y	y	Y	y
R-squared	0.147	0.184	0.211	0.209

Table 11: The Effect of review length on book market shares.
The sample is as in Table 4.
Dependent variable ln(rank) at Amazon minus ln(rank) at Bn.com.

	(1)	(2)
Amazon ln(price)	2.162***	2.128***
	(0.325)	(0.326)
BN ln(price)	-2.700***	-2.673***
	(0.280)	(0.281)
Amazon ln(no. of reviews)	-0.419***	-0.415***
	(0.0502)	(0.0503)
BN ln(no. of reviews)	0.269***	0.269***
	(0.0518)	(0.052)
Amazon average star rating	-0.422***	
	(0.0791)	
BN average star rating	0.158*	
	(0.0876)	
Amazon fraction reviews 5 star		-0.464*
		(0.242)
BN fraction reviews 5 star		0.110
		(0.188)
Amazon fraction reviews 1 star		1.594***
		(0.512)
BN fraction reviews 1 star		-1.067*
		(0.562)
Amazon ln(average rev length)	0.555***	0.580***
	(0.146)	(0.151)
Bn ln(average rev length)	-0.0351	-0.038
	(0.0917)	(0.0920)
No. observations	1093	1093
Includes shipping dummies?	Y	Y
R-squared	0.222	0.221

Table 12: The effect of change in reviews on change in sales.
The sample in (1) and (3) is the same as in Table 4, with the additional restriction that the books had to be available in both times. The sample in (2) and (4) consists of books that, in addition, had non-zero posted reviews at both sites in May 8 – July 8, 2003.
Dependent variable is Δ (ln(rank) at Amazon minus ln(rank) at Bn.com).

	(1)	(2)	(3)	(4)
Amazon Δ ln(price)	0.09465	1.576	0.106	1.416
	(0.234)	(0.870)	(0.234)	(0.890)
BN Δ ln(price)	-1.415***	-1.544***	-1.417***	- 1.444***
	(0.206)	(0.519)	(0.206)	(0.520)
Amazon Δ ln(no. of reviews)	-0.765**	-1.442	-0.668**	-1.022
	(0.343)	(0.992)	(0.328)	(0.951)
BN Δ ln(no. of reviews)	-0.309	-0.529	-0.568	-1.142*
	(0.335)	(0.707)	(0.362)	(0.599)
Amazon Δ avg star rating	-0.453*	-2.373*		
	(0.271)	(1.300)		
BN Δ avg star rating	0.0159	0.593*		
	(0.0158)	(0.333)		
Amazon Δ fraction reviews 5 star			-0.177	-3.799
			(0.539)	(3.203)
BN Δ fraction reviews 5 star			1.171**	1.092
			(0.591)	(1.192)
Amazon Δ fraction reviews 1 star			2.495*	4.167
			(1.291)	(8.801)
BN Δ fraction reviews 1 star			-1.112	-3.622
			(1.742)	(2.876)
No. observations	2091	276	2091	276
includes Δ shipping dummies?	Y	Y	Y	Y
includes Δ presence reviews?	Y	Y	Y	Y
R-squared	0.0379	0.0967	0.0404	0.0951

8. Appendix

<u>Criteria to select a random sample of books in print</u>

1) Only those entries with "A" for the first letter of the author's last name.
2) Only books published in 1998, 1999, 2000, 2001, 2002.
3) Search for each of the following "keyword in title" choices: "the," "of," and, "a."

For keyword in title = "the"; "of"; "and"; "a";
For publication year=1998-2002;

a. search hardcover fiction
b. search softcover fiction
c. search hardcover nonfiction
d. search softcover nonfiction

Categorization

BISAC	CATEGORY
FICTION	ADULT FICTION
POETRY	ADULT FICTION
TRUE CRIME	ADULT NON-FICTION
CURRENT EVENTS	ADULT NON-FICTION
BIOGRAPHY & AUTOBIOGRAPHY	ADULT NON-FICTION
RELIGION	ADULT NON-FICTION
HOUSE & HOME	DO IT YOURSELF
CRAFTS & HOBBIES	DO IT YOURSELF
GARDENING	DO IT YOURSELF
FOREIGN LANGUAGE STUDY	DO IT YOURSELF
COOKING	DO IT YOURSELF
GAMES	ENTERTAINMENT
HUMOR	ENTERTAINMENT
PETS	ENTERTAINMENT
JUVENILE NONFICTION	JUVENILE
JUVENILE FICTION	JUVENILE
PHOTOGRAPHY	LANGUAGE & ARTS
ANTIQUES & COLLECTIBLES	LANGUAGE & ARTS
ARCHITECTURE	LANGUAGE & ARTS
MUSIC	LANGUAGE & ARTS
PERFORMING ARTS	LANGUAGE & ARTS
ART	LANGUAGE & ARTS
LANGUAGE ARTS & DISCIPLINES	LANGUAGE & ARTS
LITERARY COLLECTIONS	LANGUAGE & ARTS
LITERARY CRITICISM	LANGUAGE & ARTS

DRAMA	LANGUAGE & ARTS
BUSINESS & ECONOMICS	NON-FICTION SERIOUS
LAW	NON-FICTION SERIOUS
REFERENCE	NON-FICTION SERIOUS
MATHEMATICS	NON-FICTION SERIOUS
SCIENCE	NON-FICTION SERIOUS
TECHNOLOGY	NON-FICTION SERIOUS
COMPUTERS	NON-FICTION SERIOUS
STUDY AIDS	SELF-IMPROVEMENT
EDUCATION	SELF-IMPROVEMENT
SELF-HELP	SELF-IMPROVEMENT
FAMILY & RELATIONSHIPS	SELF-IMPROVEMENT
MEDICAL	SELF-IMPROVEMENT
BODY, MIND & SPIRIT	SELF-IMPROVEMENT
SPORTS & RECREATION	SELF-IMPROVEMENT
HEALTH & FITNESS	SELF-IMPROVEMENT
HISTORY	SOCIAL SCIENCE
PHILOSOPHY	SOCIAL SCIENCE
PSYCHOLOGY	SOCIAL SCIENCE
POLITICAL SCIENCE	SOCIAL SCIENCE
SOCIAL SCIENCE	SOCIAL SCIENCE
TRANSPORTATION	TRAVEL
TRAVEL	TRAVEL
NATURE	TRAVEL

References

Avery, C.; Resnick, P.; Zeckhauser, R. (1999) "The Market for Evaluations," *American Economic Review*. 89 (June), 564-84.

Brynjolffson, E. and Smith, M. (2000) "Frictionless Commerce? A Comparison of Internet and Conventional Retailers", Management Science 46, 563-585.

Chevalier, J. and Goolsbee, A. (2003a) "Measuring prices and price competition online: Amazon.com and BN.comandNoble.com", *Quantitative Marketing and Economics* I(2), forthcoming.

Chevalier, J. and Goolsbee, A. (2003) "Valuing Internet Retailers: Amazon.com and BN.comandNoble.com", Yale School of Management working paper.

Clay, K., Krishnan, R. and Wolff, E. (2001) "Prices and Price Dispersion on the Web: Evidence from the Online Book Industry", *The Journal of Industrial Economics* 49, 521-540.

Coleman, J. S., E. Katz, and H. Menzel (1966). "Medical Innovation: A Diffusion Study," Indianapolis, Indiana: Bobbs-Merrill.

Eliashberg, J. and Shugan, S. "Film critics: Influencers or predictors?" *Journal of Marketing*, Vol. 61, No. 2 (April 1997), 68-78.

Ellison, G.; Fudenberg, D. (1995) "Word-of-Mouth Communication and Social Learning," *The Quarterly Journal of Economics*. 110 (1), 93-125.

Fingar, P.; Kumar, H.; Sharma, T. (2000) *Enterprise E-Commerce*, Meghan-Kiffer Press.

Foster, A. and M. Rosenzweig (1995, December). "Learning by doing and learning from others: Human capital and technical change in agriculture," *Journal of Political Economy*. 103(6), 1176–1210.

Godes, D. and Mayzlin, D. (2003) "Using Online Conversations to Study Word of mouth Communication," Yale SOM working paper.

Katz, E. and P. F. Lazarsfeld (1955). Personal Influence. Glencoe, IL: Free Press.

Mayzlin, D. (2003) "Promotional Chat on the Internet," Yale SOM working paper.

McWilliam, G. (2000) "Building Strong Brands through Online Communities," *MIT Sloan Management Review*, 41 (3), 43-54.

Morin, R. (2003) "Scholar Invents Fans to Answer his Critics," *Washington Post*, 2/1/2003, page C01.

Resnick, P. and Zeckhauser, R. (2002). "Trust Among Strangers in Internet Transactions: Empirical Analysis of eBay's Reputation System." *The Economics of the Internet and E-Commerce.* Michael R. Baye, editor. Volume 11 of Advances in Applied Microeconomics. Amsterdam, Elsevier Science.

Schnapp, M. and Allwine, T. (2001) "Mining of book data from Amazon.com", Presentation at the UCB/SIMS web mining conference, http://www.sims.berkeley.edu/resources/affiliates/workshops/webmining/slides/ORA.ppt

Van den Bulte, C. and G. Lilien (2001). Medical Innovation revisited: Social contagion versus marketing effort. *American Journal of Sociology* 106 (5), 1409–35.