

## **Internet Exchanges for Used Books:**

An Empirical Analysis of Product Cannibalization and Welfare Impact

Anindya Ghose,<sup>‡</sup> Michael D. Smith,<sup>†</sup> and Rahul Telang<sup>†</sup>

aghose@stern.nyu.edu  
{mds, rtelang}@andrew.cmu.edu

This Version: September 2005

Available from <http://ssrn.com/abstract=584401>

Acknowledgements: The authors thank Al Greco, Paul Kattuman, Otto Koppius, Hank Lucas, Ivan Png, Susan Siegel, Hal Varian, Bruce Weber, Andy Whinston, and participants at the 2005 International Industrial Organization Conference, the 2004 International Conference on Information Systems, the 2004 INFORMS Annual Conference, the 2004 MISRC/CRITO Symposium on the Digital Divide, the 2003 International Conference on Information Systems, Indiana University's Kelley School of Business, the University of California at Davis, the New York University's Stern School of Business, and the Tepper School of Business and H. John Heinz School of Public Policy and Management at Carnegie Mellon University for valuable comments on this research. The authors thank representatives of the Book Industry Study Group, MIT Press, Prentice Hall Publishing, and various other publishers for providing valuable information on the book industry. We also thank David Dewey, Samita Dhanasobhon, Steve Gee, Stoyan Arabadjyski, Sumiko Hossain, Jae Hong Park, Nat Luengnaruemitchai, and Sriram Gollapalli for outstanding research assistance. Michael Smith acknowledges the generous financial support of the National Science Foundation through CAREER Award IIS-0448516.

---

<sup>†</sup> H. John Heinz III School of Public Policy and Management, Carnegie Mellon University, Pittsburgh, PA, 15213. (Smith: 412-268-5978; Telang: 412-268-1155)

<sup>‡</sup> Stern School, New York University, New York, NY, 10012. (212-998-0807)

# **Internet Exchanges for Used Books:**

## An Empirical Analysis of Product Cannibalization and Welfare Impact

### **ABSTRACT**

Information systems and the Internet have facilitated the creation of used product markets that feature a dramatically wider selection, lower search costs, and lower prices than their brick-and-mortar counterparts do. The increased viability of these used product markets has caused concern among content creators and distributors, notably the Book Publishers Association and Author's Guild, who believe that used product markets will significantly cannibalize new product sales.

However, this proposition, while theoretically possible, is based on speculation as opposed to empirical evidence. In this paper, we empirically analyze the degree to which used products cannibalize new product sales for books — one of the most prominent used product categories sold online. To do this, we use a unique dataset collected from Amazon.com's new and used book marketplaces to measure the degree to which used products cannibalize new product sales. We then use these estimates to measure the resulting first-order changes in publisher welfare and consumer surplus.

Our analysis suggests that used books are poor substitutes for new books for most of Amazon's customers. The cross-price elasticity of new book demand with respect to used book prices is only 0.088. Because of this only 16% of used book sales at Amazon cannibalize new book purchases. The remaining 84% of used book sales apparently would not have occurred at Amazon's new book prices. Further, our estimates suggest that this increase in book readership from Amazon's used book marketplace increases consumer surplus by approximately \$67.21 million annually. This increase in consumer surplus, together with an estimated \$45.05 million loss in publisher welfare and a \$65.76 million increase in Amazon's profits, leads to an increase in total welfare to society by approximately \$87.92 million annually from the introduction of used book markets at Amazon.com.

**Keywords:** *Publisher Welfare, Retailer welfare, Consumer Surplus, Price competition, Used Books Sales, Electronic Markets*

## 1. Introduction

“...as a leader in the bookselling industry, Amazon’s [used book] sales practices can have a significantly deleterious effect on new book sales. If your aggressive promotion of used book sales becomes popular among Amazon’s customers, this service will cut significantly into sales of new titles, directly harming authors and publishers.”

*Author’s Guild and Association of American Publishers, Open Letter to Jeff Bezos (CEO Amazon.com), dated April 9, 2002*

“We’ve found that our used books business does not take business away from the sale of new books. In fact, the opposite has happened. Offering customers a lower-priced option causes them to visit our site more frequently, which in turn leads to higher sales of new books while encouraging customers to try authors and genres they may not have otherwise tried. In addition, when a customer sells used books, it gives them a budget to buy more new books.”

*Jeff Bezos, open letter to Amazon.com’s used booksellers, dated April 14, 2002.*

Globally networked Information Systems can reduce the search and transaction costs for buyers and sellers to locate and trade products (Bakos 1997), and can thereby facilitate the creation of technology-mediated electronic exchanges (Malone, Yates, and Benjamin 1987). These exchanges allow sellers to easily reach a worldwide market, and allow buyers to easily locate items that frequently would be unavailable in traditional physical stores.

Several papers in the literature have empirically analyzed the impact of lower search costs on *retailer-level* competition (e.g., Brynjolfsson and Smith 2000, Chevalier and Goolsbee 2003). In this paper, we empirically analyze the impact of lower consumer search costs on *product-level* competition — specifically competition between new and used books. The Internet facilitates the creation of side-by-side new and used product markets, allowing consumers shopping for new products to easily locate competing used product offers.

There is, of course, nothing new about the sale of used products. In the United States, the First Sale Doctrine (17 U.S.C. §109) allows for the resale of copyrighted works such as books, and used bookstores are common in physical settings. Rather, electronic exchanges alter the scale, scope, and efficiency of what is possible with regard to the sale of used products. Thus, while brick-and-mortar bookstores have high search costs, limited inventory capacity, limited geographical coverage, and relatively high prices, IT-enabled markets for used books offer low search costs, nearly unlimited (virtual) inventory capacity, global coverage, and — through competition among sellers — relatively low prices. These market characteristics are clearly attractive for consumers. Internet sales of used books made up an estimated 67% of all used book sales in 2004 (Wyatt 2005), versus a penetration of *only* 12.7% for Internet sales of new books.

These changes in used book sales have also raised concerns among publishers and authors. Groups such as the Author's Guild and the Association of American Publishers reason that since authors and publishers are only paid for the initial sale of a product, the lower search costs, increased selection, and lower prices of online used book markets will cannibalize new book sales and cut significantly into both publisher revenues and author royalty payments.

However, to date, the publishing industry has been unable to precisely measure the degree to which used books cannibalize new book sales because key sales data have been unobservable in online markets. For example, Tedeschi (2004) quotes Paul Aiken, the Executive Director of the Author's Guild as saying “[t]here has always been used-book sales, but it's always been a background noise sort of thing. Now it's right there next to the new book on Amazon...We think it's not good for the industry and it has an effect, *but we can't measure it*” (emphasis ours). Likewise, Kelly Gallagher, chairperson of the Book Industry Study Group's Research Committee ob-

serves “everybody has anecdotal evidence to show used books’ cannibalization of new books, *but we don’t have any accurate numbers*” (emphasis ours, Publishing Trends 2004).

Therefore, the motivation of this study is to provide direct empirical estimates of the degree to which used books cannibalize new book sales, and analysis of the resulting impact on publisher welfare, consumer surplus, and social welfare. Our paper contributes to the literature in that, while a variety of analytic models have analyzed competition to new product sales from used product markets, ours is the only paper we are aware of to empirically analyze the elasticity of new product demand with respect to used product prices and the resulting changes in new and used product sales and overall surplus. This analysis also contributes to the literature by providing publishers and industry analysts with a methodology to conduct similar analyses based on data that can be readily collected from Internet markets, which as noted above, is a capability publishers and content creators have been lacking.

We use Amazon.com’s used book market as our setting to answer this question because it is one of the most prominent used and new book marketplaces online (Brynjolfsson, Hu, and Smith 2003, Brynjolfsson and Smith 2000) and it lists new and used products side-by-side on their product pages. We collect a unique dataset from Amazon.com’s new and used marketplaces documenting prices and quantities sold for new and used books. Our data cover two samples: one collected from September 2002 to March 2003 and one collected from April to July 2004. Together these samples include 41,994 observations for 393 individual book titles.

Our data suggest that used books are not a strong substitute to new books for most of Amazon’s customers. The cross-price elasticity of demand for new books with respects to used book prices is only 0.088. Because of this only 16% of used book sales at Amazon cannibalize new book

purchases. The remaining 84% of used book sales apparently would not have occurred at Amazon's new book prices.

This increased access to low priced (used) book titles increases consumer surplus by \$67.21 million annually. While much of the discussion of the value of the Internet to consumers has revolved around lower prices, consumers' benefit from access to used book markets is nearly as large as consumers' benefit from access to lower prices on new books (Brynjolfsson, Hu, and Smith 2003).<sup>1</sup>

With respect to retailer losses, our results show that without raising prices, book publishers lose approximately \$45.05 million annually in gross profit from the presence of Amazon's used book markets. This represents approximately 0.3% of total publisher gross profits in 2003.<sup>2</sup> It is also important to note that this figure does not take into account any secondary revenue that might accrue to authors from increased readership. Moreover, this number does not take into account potential increased sales of new books due to an increase in consumers' valuations caused by the presence of a viable used book market (Ghose, Telang and Krishnan 2005).

We also note that the overall impact of used book markets on social welfare is overwhelmingly positive. The consumer surplus gains and producer surplus losses, together with an estimated \$65.76 million increase in Amazon.com's gross profit from the used book exchanges, suggests a net welfare gain to society of approximately \$87.92 million from these used book exchanges.

The remainder of this paper proceeds as follows. In Section 2, we discuss the literature relevant to our present work. In Section 3, we develop an analytic model of used product markets to high-

---

<sup>1</sup> Brynjolfsson, Hu, and Smith (2003) find that the consumer surplus gain from access to low new book prices on the Internet (versus brick-and-mortar stores) is between \$100.5 million and \$103.3 million annually.

<sup>2</sup> The Book Industry Study Group (2004) places publisher revenue at \$26.0 billion dollars and Brynjolfsson, Hu, and Smith (2003) place publisher gross margins at 60%.

light the importance of empirical measurements in determining the degree to which used products cannibalize new product sales. In Section 4, we compare the characteristics of the brick-and-mortar used book market to the Internet used book market to show that the Internet may have a significant effect on used book sales. We describe our data in Section 5 and present our empirical model and results in Section 6. We discuss the implications of our results in Section 7.

## **2. Literature**

Our research is related to different streams of extant work. The first stream of relevant literature relates to implications of concurrent availability of new and used goods. The difficulty of maintaining monopoly power on durable goods is due in part to the problem of time-inconsistency, first pointed out by Coase (1972). He conjectured that if a firm were to exploit its residual demand in future periods, then rational consumers would anticipate this behavior and price would rapidly fall to the competitive level. The interrelationship between the markets for new and used goods was pointed out by Benjamin and Kormendi (1974) and Liebowitz (1982). They argued that a monopolist can maintain market power by restricting the used market. Using the textbook market as an example, Miller (1974) suggests that the opening of secondary markets will force publishers to increase new good prices in order to extract the maximum possible profit from the onetime sale of a new good. Further research in this area has shown that a monopolist can avoid the commitment problem by leasing as opposed to selling (Bulow 1982), depreciation reduces the monopolist's incentive to cut price (Bond and Samuelson 1984). The main argument of these papers is that second-hand markets need not hurt the manufacturer because they will anticipate the resale value of their product and thereby, increase the new good price accordingly.

A more recent stream of the literature uses analytic models to show that secondary markets also create a substitution effect since new goods face competition from used goods (Anderson and

Ginsburgh 1994, Hendel and Lizzeri 1999). When considering the impact of used good markets in competitive new good markets, upstream suppliers such as book publishers can benefit from secondary markets under some conditions as shown by Ghose, Telang and Krishnan (2005).

Recent empirical work in the context of used goods has used aggregate data to show that textbook sell-through —new sales as a proportion of total sales — declines radically from 90% in the first year of a textbook’s publication to 45% in the second year and 10% in the third year (Greco 2005, p. 185-186 and Greco et al. 2005). A more recent study also uses textbook data to show that students are forward looking when making their purchases — and that their value of a textbook declines when the release of a new edition will foreclose on the resale market for a new textbook purchase (Chevalier and Goolsbee 2005).

A third stream of literature relevant to our study is research developing techniques to estimate welfare effects from the introduction of new goods. Classic economic theory shows that if the price of an existing good changes from  $p_0$  to  $p_1$ , the resulting change in welfare is given by how much the consumer would pay, or would need to be paid, to be just as well off after the price change as they were before the price change. This measure corresponds to Hicks’ (1942) compensating variation measure. To measure the welfare change from the introduction of a new good, Hausman (1997a) modifies Hick’s measure by using the product’s “virtual price” — the price that would set demand to zero — as  $p_0$  and the introductory price as  $p_1$ . This technique has been applied to measure welfare gains for new goods ranging from Honeynut Cheerios (Hausman 1997b) to increased product variety on the Internet (Brynjolfsson, Hu and Smith 2003). Related techniques have also been used to analyze the welfare impact from online auction sites (Bapna, Gank and Shumueli 2005).

Finally, our research draws on the literature relating to competition on the Internet (e.g., Brynjolfsson and Smith 2000; Clay, Krishnan and Wolff 2001; Baye, Morgan, and Scholten 2004), and specifically the direct measurement of consumer price sensitivity. Papers in this literature have shown that BarnesandNoble.com seems to face much stronger competition from Amazon.com than Amazon does from Barnes & Noble (Chevalier and Goolsbee 2003). Various papers in this literature have also analyzed the own price elasticity for offers listed at shopbots, finding elasticity measures ranging from -6 to -10 for shopbots listing books (Brynjolfsson, Dick, and Smith 2004) and PDAs (Baye et al. 2004) sold by differentiated sellers to -50 for a shopbot listing computer motherboards and memory modules sold by undifferentiated sellers. Elasticity measures at Internet shopbots are relevant for our study because the display of information at these services is comparable to the information display in Amazon's used book market place.

### **3. Theoretic Analysis**

In this section, we develop a model to analyze the impact of secondary markets on welfare to publishers, retailers, and consumers. Our model spans two periods and consists of a single publisher,  $S$  selling a new good through a retailer at a wholesale price of  $w$ , to a unit mass of consumers. The book is new when it is marketed in period 1 and can be sold as used in period 2 when it undergoes some degradation in quality. The retailer opens a secondary (used book) market where consumers can buy and sell used goods. Whenever a consumer sells the used book, the retailer gets a commission  $k$  per used good sold while the rest  $(1-k)$  is the gain to the consumer. For example, Amazon charges a 15% commission (as a fraction of the used book sale price) to each used book seller while the remaining 85% of the used book sale price goes to the consumer.

Further, let  $\theta$  be a consumer's valuation for a good, where  $\theta \sim U[0,1]$ . The type parameter  $\theta$  indicates a consumer's marginal valuation for quality. For any given quality, a consumer with a

higher  $\theta$  is willing to pay more for the product than one with a lower  $\theta$ . Without loss of generality, let 1 denote the quality of the new good and  $q$  denote the quality of the used good, where  $0 < q < 1$ ; thus  $q$  can be interpreted as the degree of quality degradation of new book. If a consumer purchases a book of quality  $q$  at price  $p$ , her utility is  $U(\theta) = \theta q - p$ .

The game is modeled as a multi-stage process across a single period. First the retailer chooses an optimal new good price given the per unit wholesale price  $w$  of the supplier. Then market forces determine the optimal price of used goods from clearance conditions. Finally, consumers demand is realized. We consider a Subgame Perfect Equilibrium of this game using backward induction.

### 3.1. No Secondary Electronic Market

We begin by modeling the case where only new goods are sold in the marketplace. We denote the price of the new good in the absence of a used good market as  $P_N$ . Thus, consumers buy a new book as long as they get positive surplus; i.e., as long as  $\theta - P_N > 0$ . Hence, the demand for a new book is  $D_N(P_N) = 1 - \theta = 1 - P_N$ . Under these conditions the profit of the publisher is,

$$\pi_S = D_N * w = (1 - P_N)w \quad (1)$$

and profit to the retailer is

$$\pi_R = D_N * (P_N - w) = (1 - P_N)(P_N - w) \quad (2)$$

In equilibrium, the retailer maximizes this profit by setting  $P_N^* = \frac{1+w}{2}$ . At this price, the publisher makes a profit of  $\pi_s = \frac{w(1-w)}{2}$ , while the total profit of the retailer is

$$\pi_R = D_N * (P_N - w) = (1 - P_N)(P_N - w) = \frac{(1-w)^2}{4} \quad (3)$$

Note from (1) that in equilibrium the publisher will set  $w^* = 1/2$ .

### 3.2. Retailer establishes a Secondary Electronic Market

The presence of a secondary market allows new book buyers to sell them later. Rational consumers take this into account in their utility function in the buying process. As a result, the retailer (and/or publisher) should be able to sell the new book at a higher price. This is the *price-increase effect* as outlined in the prior literature (e.g., Miller 1974; Swan 1980).

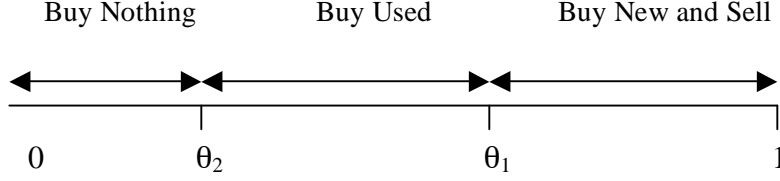
However, Waldman (1997) and Hendel & Lizzeri (1999) argue that the used book market also creates a *substitution effect*. The substitution effect arises from the fact that new books face competition from used books. Accordingly, some consumers who previously would have bought a new book will shift to the used-book market, cannibalizing new book sales. The price-increase effect will increase publisher welfare, while the cannibalization effect will decrease publisher welfare. The actual impact of used book markets on publisher welfare depends on the actual elasticity and propensity to resell books observed in the market.

The figure below describes the segmentation of the market based on consumer types. Let  $P_N^U$  and  $P_U$  denote the new and used book prices, respectively, in the presence of used book markets. An artifact of secondary markets is that not all consumers resell their used books. There could be multiple reasons for some consumers to hold on to their used book. For example, not all consumers may be aware of the existence of the used book market, or the transactions and search costs faced by some consumers could be sufficiently large that they don't have an incentive to participate in the used book markets. In short, the net utility to some consumers from keeping the book might exceed that from selling it. To account for this fact, we assume that only a proportion  $\alpha$  of new good buyers eventually sell their used goods.<sup>3</sup> Hence, the expected revenue from a used

---

<sup>3</sup> We do not model flows within the used book markets, i.e. used book consumers who resell their used books.

book sale (which is equivalent to  $(1-k)P_U$ ) realized by consumers is  $\alpha(1-k)P_U$ , where  $\alpha \in (0,1)$



Hence, the corresponding expected utilities derived from various strategies are as follows:

- (i) Buy new good and sell it:  $\theta - P_N^U + \alpha(1-k) P_U$
- (ii) Buy used good:  $\theta - P_U$
- (iii) Buy nothing: 0

Thus, higher valuation consumers buy new goods and lower valuation consumers wait and buy used goods later. It is important to recognize that, in our model, the number of consumers in these groups emerges endogenously. This ensures that clearance conditions will equalize demand and supply of used goods at all times. The new profit of the publisher is given by

$$\pi_S^U = D_N^U w^U = (1-\theta_1)w^U \tag{4}$$

where  $D_N^U$  denotes the demand for new books in the presence of used books. Similarly, profits for retailer can be written as

$$\pi_R^U = D_N^U (P_N^U - w^U) + D_U P_U k = (1-\theta_1)(P_N^U - w^U) + (\theta_1 - \theta_2)P_U k . \tag{5}$$

Here the  $D_U$  is demand for the used books. Note that retailers, like Amazon, also enjoy the benefits of increased used book sales on its marketplace through the commission  $k$  that it charges to used-book sellers. This used book demand effect is not available to publishers. Thus, in general, retailers are more likely to benefit from used book markets than publishers are.

By comparing (1) and (4), we get the publisher's loss/gains from the establishment of used book market as:

$$\pi_S^U - \pi_S = D_N^U w^U - D_N w = \underbrace{(D_N^U - D_N) w}_{\text{Substitution effect}} + \underbrace{(w^U - w) D_N^U}_{\text{Price increase effect}} \quad (6)$$

Similarly, for the retailer, the loss/gain from the establishment of the used book market (after substituting  $R^U$  for  $(P_N^U - w^U)$  and  $R$  for  $(P_N - w)$  and comparing (2) and (5)) is given by

$$\pi_R^U - \pi_R = D_N^U R^U - D_N R + D_U k P_U = \underbrace{(D_N^U - D_N) R}_{\text{Substitution effect}} + \underbrace{(R^U - R) D_N^U}_{\text{Price increase effect}} + \underbrace{D_U k P_U}_{\text{Used-book demand effect}} \quad (7)$$

Thus, consistent with the prior literature, used book markets have two countervailing effects on publisher welfare. On one hand, used books cannibalize the sales of new books, reducing publisher welfare. On the other hand, the presence of the used book market may lead to increased consumer willingness to pay for new books, and as a result higher prices. Which of these two effects dominates depends on the actual behavior of customers in the market — notably the sensitivity of consumers to used book prices, which can be measured by the own- and cross-price elasticities of demand observed in the market.<sup>4</sup> However, while these effects have been well studied in the analytic literature, we are aware of no papers in the empirical literature that have directly measured them. Moreover, as noted above, direct estimates of the cannibalization of new books by used markets is a very important — and currently unavailable — measure for publish-

---

<sup>4</sup> There may be other potential effects which can impinge upon the outcome as well (for example, the secondary market may provide more incentives for users to buy other new books). However, identifying disaggregate individual effects (like customer heterogeneity) is not possible in our setting, due to lack of appropriate data and this is a limitation of our study.

ers and authors. Thus, a key contribution of our research is to develop and implement a methodology for creating these measures.<sup>5</sup>

To do this we note that the relevant prices and margins in (6) and (7) can typically be observed through secondary sources. Likewise, recent research results (Brynjolfsson, Hu, and Smith 2003) can shed light on Amazon's used book sales ( $D_U$ ). However, the substitution effect cannot be measured directly. To measure this effect we note that under the standard definition of elasticity

( $\eta = \frac{\Delta Q}{\Delta P} \frac{P}{Q}$ ), the substitution effect can be measured as:

$$D_N^U - D_N = \Delta Q = D_N \times \eta \times \left( \frac{P_N^U - P_U}{P_N^U} \right) \quad (8)$$

This shows that measuring the cross-price elasticity of new book demand with respect to used book prices ( $\eta$ ) is critical to being able to measure the substitution effect, and therefore the change in publisher and retailer surplus. One contribution of our work is implementing a methodology to measure this elasticity in the context of Internet sales.

Thus, the key insight from this model is that the extent of gains (and losses) critically depends on the effects outlined in (6) and (7). However, the substitution effect, price increase effect, and used book demand effects are inherently empirical in nature. Therefore, the theoretical model provides us with key measures that need to be quantified in order to assess the benefits of used book markets to different parties involved in these transactions. Hence, in our empirical analysis we estimate the extent of substitution effect (or cannibalization effect) and extent of used book

---

<sup>5</sup> In the Appendix, we calculate optimal values of  $P_U$ ,  $P_N$  and the profits of the retailer and publisher. This shows that for a wide range of parameter values retailers tend to benefit from the used book markets while the publishers tend to lose. Our model does not consider competition between retailers and assumes specific functional forms for different parameters in the utility function and the results should be interpreted in the context of these assumptions.

demand effect. With this information, we then quantify the gains (and losses) to publishers and retailers based on real-world data from an Internet used book market.

#### **4. Analysis of Brick-and-Mortar versus Internet Used Book Markets**

As noted above, there is nothing new about the sale of used books. Rather, what changes on the Internet is a radical increase in the variety of used books offered for sale and a radical decrease in the prices of these used books and the associated search costs for consumers searching for both new and used titles to locate these books. To illustrate the magnitude of these changes, in June 2004 we generated a list of 30 randomly selected books from the October 2002 Bowker Books in Print listings, and 30 randomly selected books from the 2002 New York Times bestseller lists. These lists are useful because they are old enough to include books that would generally be available in physical bookstores, and they contain a mix of popular and less popular titles.

We searched for these books at Amazon.com and found that at least one used copy was available for each of the 60 books we sampled — even though 13 of the books were out of print and did not have new copies available from Amazon itself. Moreover, there were an average of 22.6 competing used book offers for each random book, and an average of 241.3 competing offers for each former New York Times bestseller. These multiple offers create competition among sellers to lower their prices. As a result, the random books have an average 40.1% discount off the new book list price and the former bestsellers have an average 84.5% discount off list price. As a point of comparison, Amazon's new books had an average 9.1% and 30% discount off list price for these random and former bestseller titles.

To understand how the used title selection at Amazon would compare to the selection of a typical brick-and-mortar used bookstore, we searched the catalogs of four brick-and-mortar used book-

stores in the Pittsburgh area advertising themselves as having a “general selection.” Three of the four stores did not carry any of the 60 books on our list. The fourth, Eljay’s Used Books, carried none of the random books and only six of the former bestsellers. Moreover, Amazon’s used price was an average of 75% (\$8.16) lower than the price at Eljay’s for these six books.

As another point of comparison, we used ABEbooks.com to search for the used bookstore in the United States with the widest selection of titles in our sample. ABEbooks catalogs the inventory of 7,680 used booksellers in the United States.<sup>6</sup> According to ABEbooks’ listings, Powell’s Books of Portland, Oregon had the best selection of any of these 7,680 booksellers for the books in our sample — but still only carried 11 of the 30 random titles and 29 of the 30 former bestsellers. Moreover, the used price at Amazon averaged 38% (\$4.93) lower than the Powell’s price on the random books and 67% (\$7.03) lower than the Powell’s price on the former bestsellers.

Thus, a used book shopper at Amazon.com would have lower search costs to locate a book, significantly larger selection (both in terms of availability and the number of competing offers), and dramatically lower prices than they are likely to find at their local used bookstore (even if they are fortunate enough to live close to Powell’s Books in Portland). Moreover, because new and used books are listed side-by-side in many Internet markets, new book shoppers can more easily become aware of used book offerings than they could in a typical brick-and-mortar bookstore, and might be tempted by the wide selection and low prices to buy a used book instead of a new book. Together, these factors may be what is driving the penetration of used book sales through the Internet channel, which are growing at a rate of 30% per year (Wyatt 2005) and accounted for 54.4% of all used book sales in 2003 (Siegel and Siegel 2004) and 67% of all used book sales

---

<sup>6</sup> As a point of comparison, Siegel and Siegel (2004) estimate that there are between 8,000 to 10,000 used booksellers in the United States.

in 2004 (Wyatt 2005). As a point of comparison, new book sales on the Internet accounted for *only* 12.7% of total book sales in 2003 (Rappaport 2004).

## **5. Data**

Our data are compiled from publicly available information on new and used book prices and sales ranks at Amazon.com (Figures 2 and 3). The data are gathered using automated Perl scripts to access and parse HTML pages downloaded from the retailer.<sup>7</sup> The data were collected in two separate samples. The first was collected over a 180-day period from September 2002 to March 2003 and includes 273 individual book titles. This panel of books includes an equal number of books from each of five major categories — New York Times best sellers, former New York Times best sellers, Amazon Computer best sellers, best selling Text Books, and “New and Upcoming Books.” New York Times bestsellers were selected at random from the current New York Times list at the beginning of the sample, and were replaced as they were removed from the list. Former bestsellers were selected at random from the full list of books appearing in New York Times bestseller listings in the year 1999. Computer bestsellers and new and upcoming books were selected at random from the respective list at Amazon.com. Finally, bestselling textbooks were selected from the facultyonline.com bestseller list.

### **[Insert Figures 2 and 3]**

In early 2004 Amazon.com added a new variable to their XML data feed to developers, allowing us to obtain accurate measures of their used book sales (which we describe in more detail below). At this point, we created a similar sample of books, drawing 40 books from each of four categories: current bestsellers, former bestsellers, new and upcoming, and random. New and up-

---

<sup>7</sup> A number of papers in recent work have used PERL or Java scripts to scrape data from Internet websites (for example, Bapna, Goes and Gupta 2003).

coming books were selected in the same way as the first sample. Current and former bestsellers were drawn from the current list of Amazon bestselling books and Amazon's top selling books in 2002. Finally, random books were selected at random from all Amazon.com titles listed in the "browse" section (which we believe includes all titles offered for sale by Amazon). In this sample, we dropped 15 books (10 random titles and 5 former bestsellers) that were out of print and therefore did not have new Amazon prices. These data were collected over an 85-day period from April to July 2004. Our total data sample includes 41,994 observations of 393 titles.

For each sample, we collect data on the Amazon.com sales rank and new book price, and the book prices charged by Amazon.com marketplace sellers. Our marketplace data include the price, condition, and seller rating for each used book listed for sale. Book condition is self-reported by the seller and can be either "like new," "very good," "good," or "acceptable." We also collect the seller rating for each used bookseller, which is a 1-5 star measure of the reported experiences of prior buyers with each seller.

In addition to these variables, in the second sample we were able to use Amazon's XML data feed to collect the number of used books sold. We do this by using the unique product identifier for each product listed in the used book market. This unique identifier was added to the XML feed sometime between March 2003 and April 2004, making it only available in the second half of our sample. Using this product identifier, we infer that a sale has occurred when an identifier that appeared in the previous collection period does not appear in the current collection period's XML listings. We collected this data once every 2 hours for books with a sales rank lower than 10,000, and every 6 hours for all other titles. In our empirical estimates below, in sessions where multiple sales are observed in between two collection periods (11% of the sessions in our sample), we assume that the books were sold in order of price. This is consistent with the strong

preference we observe in our data for low priced books (20% of the sales in our data occur for the lowest priced book in the session). To the extent that this assumption is incorrect, it will inflate our estimates of the own price elasticity of used book demand and mean that our consumer surplus estimates represent a lower bound on the true consumer surplus gain.

**Table 1: Summary Statistics**

| <i>Variable</i>                  | <i>Obs.</i> | <i>Mean</i> | <i>St. Dev.</i> | <i>Min</i> | <i>Max</i> |
|----------------------------------|-------------|-------------|-----------------|------------|------------|
| Amazon Rank                      | 41,994      | 74,169      | 238,283         | 1          | 2,556,356  |
| List Price                       | 41,994      | 30.60       | 30.05           | 5.99       | 299.99     |
| Amazon Price                     | 41,994      | 24.04       | 26.47           | 1.95       | 209.99     |
| Best “Like New” Used Price       | 41,492      | 15.16       | 21.32           | 0.01       | 194.25     |
| Best “Very Good” Used Price      | 33,517      | 11.26       | 18.41           | 0.01       | 207.60     |
| Best “Good” Used Price           | 31,939      | 11.24       | 16.42           | 0.01       | 200.00     |
| Best “Acceptable” Used Price     | 20,368      | 7.86        | 15.70           | 0.01       | 222.35     |
| Best Used Price (All Conditions) | 41,994      | 13.14       | 19.16           | 0.01       | 151.95     |
| Seller Rating                    | 41,994      | 3.97        | 1.65            | 1          | 5          |
| Count of Used Books              | 41,994      | 81.15       | 131.78          | 1          | 753        |
| Days Since Release               | 41,994      | 717.87      | 1,336.94        | 1          | 21,235     |

Finally, for the second sample, we collect the number of lifetime ratings for each seller at the start of our data collection.<sup>8</sup> From the number of lifetime ratings, we generate two additional variables: a dummy variable identifying when a particular seller has zero lifetime ratings, and a dummy variable for the top ten sellers in our sample on the basis of the most lifetime ratings.<sup>9</sup> Table 1 lists summary statistics for our data. All prices represent the lowest price for each ISBN.

## 6. Results

### 6.1. Sales Rank to Quantity Calibration

Until recently it was difficult to calculate the price elasticity for products sold on the Internet because, while the price of individual items could be readily observed, the quantity sold remained

<sup>8</sup> This variable was constructed on April 13, 2005, one year after the start of the collection of the second sample, by subtracting the number of seller ratings over the most recent 365 days from their lifetime number of seller ratings.

<sup>9</sup> These were exclusively large, professional, used booksellers. Powells.com and ABEBooks.com were the top two booksellers in our list in terms of the most lifetime ratings.

unobserved. Two recent papers address this problem, providing a way to map the observable Amazon.com sales rank to the corresponding number of books sold. In both cases, the authors find on a stable relationship between the ordinal sales rank of a book and the cardinal number of sales, using roughly the following Pareto relationship:

$$Quantity = \delta \cdot Rank^{\beta} \tag{9}$$

Chevalier and Goolsbee (2003) calibrate this relationship using a creative and easily executed experiment where the authors obtain a book with a known number of weekly sales, purchase several copies of the book in rapid succession from Amazon.com, and track the Amazon sales rank before and shortly after their purchase. Using these two points, they estimate  $\beta=-0.855$ .<sup>10</sup> They also estimated this parameter from similar sales-rank experiments conducted by Weingarten (2001) and Poynter (2000) as -0.952 and -0.834 respectively.

Brynjolfsson, Hu and Smith (2003) calibrate this relationship using data from a publisher mapping the Amazon sales rank to the number of copies the publisher sold to Amazon in the summer of 2001. The data include weekly sales and rank observations for 321 books with sales ranks ranging from 238 to 961,367. Using these data they estimate  $\beta = -0.871$ .

For the purposes of this paper, we will use Brynjolfsson, Hu, and Smith's estimate because is based on 861 data points as opposed to 2 data points in the experiments; however, our results are not particularly sensitive to this choice versus one of the other estimated parameter values.

Note that the  $\beta$  parameter will be stable over time as long as customers' relative tastes for popular and obscure books do not change. Increases in sales over time (holding tastes constant) will only shift the  $\delta$  parameter, which is a scaling parameter that does not impact our elasticity calcu-

---

<sup>10</sup> Note that the  $\theta$  reported by Chevalier and Goolsbee corresponds to  $-1/\beta$ .

lations. Also note that any shifts in customer preferences for new books that resulted from the introduction of used book marketplaces would be incorporated into our parameter estimate as the sales-rank data used by both Chevalier and Goolsbee (2003) and Brynjolfsson, Hu, and Smith (2003) were gathered well after the December 2000 introduction of Amazon's used marketplace.

We also empirically confirmed that during the time of our study Amazon's sales rank was calculated based only on new book sales as opposed to new and marketplace sales.<sup>11</sup> To do this, we located a book with a high sales rank and observed the rank of this title over the course of several weeks. During our observation period, the sales rank of the book varied between 596,625 and 606,439, and based on the movement of the rank, the book appeared to have 1 sale every 2 weeks. Having established the initial rank of the book, we then listed five used books for this title in Amazon's used book marketplace and purchased them on Monday, October 23, 2002 using 5 different Amazon.com buyer accounts. The sales rank before we made the purchase was 599,352 and it remained stable until October 30 when it increased to 601,457. On October 30, we purchased 5 copies of the new book using the same Amazon accounts and the next morning the sales rank was 4,647. We infer from this marketplace sales were not included in Amazon's sales rank figures during our study period. (Note that assuming this book had 1 sale every two weeks at a rank of 599,352, our estimated  $\beta$  parameter for this experiment would be  $-.877$ .)

## **6.2. Estimation**

We run two regressions to analyze the structure of Amazon's new and used book marketplaces. First, we analyze the impact of new and used prices on new book sales by estimating models of the form:

---

<sup>11</sup> On or about October 14, 2004 Amazon changed the way it calculates sales rank, and now appears to include both new and marketplace (e.g., used) sales in the sales rank calculation.

$$\text{Log}(\text{Rank}_{bt}) = c + \Gamma \cdot \text{Log}(\text{Amazon Price}_{bt}) + \Psi \cdot \text{Log}(\text{Used Price}_{bt}) + \Omega'X + \varepsilon_{bt} \quad (10)$$

where,  $b$  and  $t$  index book and date. The dependant variable is the log of rank. The independent variables are Amazon price (*AmazonPrice*), the lowest priced used book in the market (*Used-Price*), and a vector of other control variables ( $X$ ). Our control variables include the log of the time since the book was released, the condition of the lowest priced used book, the seller rating for the lowest priced used book, and the log of the number of used books offered for sale for a particular book.  $\Gamma$ ,  $\Psi$ , and  $\Omega'$  are the parameter vectors to be estimated.

Note that because of the structure of this industry, quantity and price are not jointly determined, and thus we do not face the endogeneity concerns that would normally arise in demand regressions. With regard to Amazon's own price, because books are produced in large printings prior to going to market, the quantity of new books Amazon can sell is predetermined (and usually virtually infinite) at the time Amazon sets their price. Likewise, used price is not a function of current period sales at Amazon, as used copies typically would take some time before they enter the used book market.<sup>12</sup> This follows the standard approach taken in the literature for demand estimation of Internet book sales (see for example Chevalier and Goolsbee (2003)).

For the second regression, we use the fact that we observe each of the marketplace offers shown to Amazon's used book customers along with the offer each customer chose to purchase. Because of this, we can use the multinomial logit model (Ben-Akiva and Lerman 1985, Guadagni and Little 1983) to determine the sensitivity of customers to the parameters of the offered products. Specifically, under the multinomial logit model we assume that used book customers maximize an indirect utility function of the form:

---

<sup>12</sup> Note that we also include count of used book to control for the supply of used books across time and books.

$$u_{ij} = z_j \theta + \varepsilon_{ij} \tag{11}$$

where  $u_{ij}$  represents the utility of user  $i$  for offer  $j$ , which is a linear combination of the observed product characteristics ( $z$ ) and their associated parameters ( $\theta$ ) and a mean zero random disturbance ( $\varepsilon_{ij}$ ). Under the assumption that  $\varepsilon_{ij}$  follows a type-I extreme value distribution, if consumers select the offer that maximizes their utility, then the conditional probability that offer  $j$  will be selected is given by a standard multinomial logit equation:

$$P_j = \frac{\exp(z_j \theta)}{\sum_{r=1}^J \exp(z_r \theta)} \tag{12}$$

Table 2 presents regression results of the impact of new and used price on sales rank (equation 10). To estimate this model, we use OLS with book-level fixed effects.<sup>13</sup> Models 1-4 in Table 2 progressively add control variables to check the stability of our parameters of interest. Note that the parameters of interest (Amazon and used prices) have the expected signs (recall that an increase in sales rank implies a decrease in sales), are precisely estimated, and the parameter estimates and associated standard errors are stable across specifications, suggesting that the estimates are robust and that multi-collinearity is not a significant problem in the model.<sup>14</sup>

---

<sup>13</sup> The use of book-level fixed effects is equivalent to a first differences approach.

<sup>14</sup> We selected the minimum used price across all conditions because this is the only used price shown on Amazon’s new book pages (see Figure 2). We are unable to include separate price coefficients for each book condition because of collinearity among these variables. Including a price coefficient for the lowest priced “like new” condition book— arguably the closest substitute to a new book — would result in a coefficient estimate of -0.069, lower than the -0.102 we obtain for the minimum used price across all conditions. Using this coefficient instead of minimum price, would reduce the loss faced by publishers in our welfare calculations.

**Table 2: Results for New Book Market**

| <i>Indep. Vars.</i>         | <i>1</i>            | <i>2</i>            | <i>3</i>            | <i>4</i>            |
|-----------------------------|---------------------|---------------------|---------------------|---------------------|
| Constant                    | -2.059**<br>(0.161) | -2.067**<br>(0.161) | -2.078**<br>(0.161) | -2.161**<br>(0.162) |
| Ln(Amazon Price)            | 1.347**<br>(0.048)  | 1.347**<br>(0.048)  | 1.347**<br>(0.048)  | 1.345**<br>(0.048)  |
| Ln(Min. Used Price)         | -0.105**<br>(0.010) | -0.105**<br>(0.010) | -0.105**<br>(0.010) | -0.102**<br>(0.010) |
| Ln(Days Since Release)      | 1.142**<br>(0.015)  | 1.140**<br>(0.015)  | 1.140**<br>(0.015)  | 1.120**<br>(0.015)  |
| Condition Rating            |                     | 0.009*<br>(0.004)   | 0.008<br>(0.004)    | 0.008<br>(0.004)    |
| Seller Rating               |                     |                     | 0.003<br>(0.003)    | 0.003<br>(0.003)    |
| Ln(Number of Used for Sale) |                     |                     |                     | 0.057**<br>(0.011)  |
| No. of Observations         | 41,994              | 41,994              | 41,994              | 41,994              |
| Pseudo R <sup>2</sup>       | 0.229               | 0.229               | 0.229               | 0.228               |

The dependent variable is ln(sales rank). Standard errors are listed in parenthesis; \*\* and \* denote significance at 0.01 and 0.05, respectively. All models use book-level fixed effects.

The other control variables suggest that, as expected, sales of new books decrease over time and older books and books with more used copies for sale have higher rank. It might be initially surprisingly that both seller rating and condition are insignificant in our regressions. However, it is important to realize that this does not necessarily mean that condition and/or rating are unimportant to used book purchasers, but rather that the condition and rating of the lowest priced book aren't what is driving the cannibalization of new book sales.

However, while condition and seller rating do not appear to be driving the cannibalization of new book sales, they do have a strong impact on the choices of used book customers. This can be seen in Table 3, which provides estimates of the taste parameters ( $\theta$  in equation 11) for Amazon.com's used book customers. These parameters are precisely estimated and relatively stable across specifications. The signs of the parameters suggest that, as expected, higher priced used books are less likely to be purchased, and books in "very good" condition are preferred to

“good” condition books, which are in turn preferred to “acceptable” books. It is surprising that each of these conditions is preferred to “like new” condition books, the referenced category. This could be driven by the fact that used book customers are very sensitive to price and “like new” condition books carry much higher prices than the other used book categories. Finally, we see from Table 3 that seller characteristics matter for customer choice. Sellers with higher ratings and more lifetime ratings are preferred to other sellers, sellers with at least one rating are strongly preferred to sellers with no ratings, and the top 10 sellers in our sample (who are exclusively large professional book sellers) are preferred to other sellers.

**Table 3: Multinomial Choice Results for Used Book Market**

| <i>Indep. Vars.</i>              | <i>1</i>            | <i>2</i>            | <i>3</i>            | <i>4</i>            | <i>6</i>            |
|----------------------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| Used Price                       | -0.055**<br>(0.001) | -0.054**<br>(0.001) | -0.054**<br>(0.001) | -0.055**<br>(0.001) | -0.055**<br>(0.001) |
| “Very Good”<br>Condition (0/1)   |                     | 0.210**<br>(0.011)  | 0.191**<br>(0.011)  | 0.184**<br>(0.011)  | 0.184**<br>(0.011)  |
| “Good”<br>Condition (0/1)        |                     | 0.186**<br>(0.015)  | 0.163**<br>(0.015)  | 0.147**<br>(0.015)  | 0.148**<br>(0.015)  |
| “Acceptable”<br>Condition (0/1)  |                     | 0.095**<br>(0.027)  | 0.068*<br>(0.027)   | 0.053<br>(0.027)    | 0.051<br>(0.027)    |
| Rating (1-5<br>Stars)            |                     |                     | 0.032**<br>(0.003)  | 0.013**<br>(0.003)  | 0.011**<br>(0.004)  |
| Ln(Lifetime<br>Ratings +1)       |                     |                     |                     | 0.014**<br>(0.002)  | 0.008**<br>(0.002)  |
| Top 10 Selling<br>Merchant (0/1) |                     |                     |                     |                     | 0.042*<br>(0.017)   |
| No Seller Rat-<br>ings (0/1)     |                     |                     |                     |                     | -0.058*<br>(0.23)   |

The dependent variable is whether a used book is sold. Standard errors are listed in parenthesis; \*\* and \* denote significance at 0.01 and 0.05, respectively. We observe 9.8 million offers across 56,091 choice sets.

We can use the results of these two regressions to calculate the relevant own- and cross-price elasticities in the new and used markets. With respect to the new market, one can easily show from (9) and (10) that own- and cross-price elasticity are given by  $\beta\Gamma$  and  $\beta\Psi$  respectively. Thus,

using  $\beta = -0.871$ , we see that Amazon’s own price elasticity is approximately -1.17, while the cross price elasticity of new book sales to used books prices is approximately 0.088.<sup>15</sup> Both results have the expected signs. Amazon’s own price elasticity is close to -1, which is consistent with what one might expect from a firm with significant market power. The cross-price elasticity estimates are quite low, suggesting that used books are not a strong substitute for new books for most of Amazon’s customers.

To calculate the own price elasticity of used book offers, we note that under the multinomial logit model, the own price elasticity of demand for an individual offer is given by (Ben-Akiva and Lerman 1985, p. 111):

$$\eta_{jk} = \alpha_k p_k (1 - P_j) \tag{13}$$

where  $\alpha_k$  is the estimated parameter of used price,  $p_k$  is used price itself, and  $P_j$  is the conditional choice probability defined in equation (11). Using (13), we can calculate the average own price elasticity imputed for offers in each session and take the average of this across sessions to obtain an own price elasticity of -4.87.<sup>16</sup>

This elasticity, as expected, is high — suggesting that the used book market place is competitive and small price changes have large impact on the probability of book being sold. Similar, though slightly larger, estimates for demand elasticities are also found by other researchers in the context of shopbots — settings where the layout of offers is very similar to Amazon’s used book

---

<sup>15</sup> Cross price elasticities for the old and new samples are .089 and .079 respectively. Own price elasticities for the old and new samples are -1.16 and -1.40 respectively.

<sup>16</sup> Calculating elasticity at the session level as opposed to the offer level is common in the literature and we believe is appropriate in our case because it imputes less weight to sessions with a large number of offers. The offer level elasticity in our data is -4.14. Using the offer level elasticity instead of the session level elasticity would increase both the resulting consumer surplus and total welfare in our results.

marketplace.<sup>17</sup> The fact that our elasticity is slightly lower is likely due to the degree of differentiation in offerings across the studies: In our study both products (quality) and sellers (ratings) are differentiated, while prior studies of shopbot elasticity were conducted in the context of either undifferentiated products (Brynjolfsson, Dick, and Smith 2004) or both undifferentiated sellers and undifferentiated products (Ellison and Ellison 2004).

### **6.3. Welfare Estimations**

#### **6.3.1. *Publisher Welfare***

As noted in Section 3, publishers are worse off in the presence of a used book marketplace if there is no increase in the price of new books after the introduction of the used book markets. Conversations with representatives of three major publishers revealed no changes in the wholesale prices of books following the introduction of Internet used book marketplaces. Al Greco, the author of the Book Industry Study Group's annual Book Industry Trends and a Professor at Fordham University confirmed that in his research there have been "no significant changes in wholesale prices" in recent years as a result of the introduction of used book markets on the Internet.<sup>18</sup> Because of this, we assume in our welfare calculations that publisher prices have not changed as a result of the introduction of Internet exchanges for used books.<sup>19</sup> Under this assumption, we can use our elasticity estimates and equation (8) above to estimate the loss in publisher profit from the presence of Amazon's used book market.

Brynjolfsson, Hu, and Smith (2003) calculate that, after the introduction of the used marketplace Amazon sold approximately 99.4 million books per year. However, this represents sales after the

---

<sup>17</sup> For example, Brynjolfsson, Dick, and Smith (2004) find elasticities of between -6.75 and -9.77 for a shopbot listing new books; Baye et al. (2005) find elasticity of approximately -6 in a market for PDAs listed at a shopbot and Ellison and Ellison (2001) estimate the elasticity of -50 for DRAM memory modules.

<sup>18</sup> Source: Al Greco, Fordham University, May 12, 2005, conversation.

<sup>19</sup> This is a conservative assumption. If it is incorrect, we will overstate the publisher loss, and our resulting total welfare estimates will represent a lower bound on the true welfare gains.

introduction of the used book marketplace ( $D_N^U$ ) not before ( $D_N$ ). Thus, we must rewrite (8) in terms of observed variables (specifically  $D_N^U$ ), which gives us:

$$\Delta Q = (D_N^U - D_N) = D_N^U \frac{\eta \cdot \Delta P\%}{(1 - \eta \cdot \Delta P\%)} \text{ where } \Delta P\% = \left( \frac{P_N^U - P_u}{P_N^U} \right) \quad (14)$$

From our calculations above, the cross price elasticity of new book sales to used books prices ( $\eta$ ) is approximately 0.088. Finally, used books in our sample are sold at an average discount of 50.6% off Amazon's new price ( $\Delta P\% = .506$ ). From these figures, equation (14) shows that Amazon lost 4.63 million sales ( $\Delta Q$ ) due to the presence of used book markets.

We can use (6) to characterize the loss in publisher profit from this change in sales by noting that according to Brynjolfsson, Hu, and Smith (2003) the wholesale cost of adult trade books is between 43-51% off the book's list price and publisher gross margins on sales are typically 56-64%. Taking 60% as the typical margin, 47% as the typical discount off list price for wholesale prices, and noting the average list price in our sample is \$30.60, we calculate that publisher lost gross profit from Amazon.com's used book market is approximately \$45.05 million.<sup>20</sup>

### 6.3.2. *Retailer Welfare*

With regard to retailer welfare, Amazon.com incurs a loss from the decrease in the quantity of new books sold, which may be mitigated by an increase in revenue from used book marketplace sales that otherwise would not have occurred at the new book prices. Since we observe  $Q_n$ ,  $P_n$ , and the relevant cross-price elasticities, we can measure the net change in retailer welfare from these two effects. To do this, we first calculate Amazon's dollar contribution margins on new and used books. For new books, Wingfield (2003) places Amazon's gross margins on new book

---

<sup>20</sup> \$45.05 million = 4.63 million \* \$30.60 \* (1-.47) \* .60

sales at approximately 22% which, given Amazon's average price \$24.04 for new books, gives a dollar contribution margin of \$5.29.

For used books, Amazon also earns revenue of \$0.99 plus 15% of the sale price on their used book sales.<sup>21</sup> In addition, Amazon charges buyers \$3.49 for shipping and reimburses sellers \$2.26 for their shipping costs, and thus earns \$1.23 on shipping per unit sold. Given that the average price of used books in our sample that are sold is \$8.76,<sup>22</sup> the dollar contribution margin on used book sales is approximately \$3.04 (35%).<sup>23</sup> Thus, Amazon's losses on cannibalized new book sales net their gains on the corresponding used book sale work out to approximately (\$2.25 = \$5.29 - \$3.04) for each cannibalized new book: or a total of \$10.42 million for the 4.63 million cannibalized new book sales.

However, as noted above, Amazon also gains incremental customers from the presence of their used marketplaces. Our results suggest these incremental customers could be quite substantial. Milliot (2002) notes that across all product categories sold at Amazon.com, used products accounted for 23% of Amazon's sales. Moreover, used sales in the book category were one of the strongest of any product category according to Jeffrey Bezos. Thus, 23% may be an underestimate of the actual proportion of used sales in the book category.

If Amazon sells 99.4 million new books annually (Brynjolfsson, Hu, and Smith 2003) and used book sales made up 23% of total book sales (both new and used), then approximately 29.69 mil-

---

<sup>21</sup> Amazon.com waives the \$0.99 fee for "Pro Merchant Subscribers." Pro Merchant Subscribers are charged \$19.99 per month for membership. While we have no way of knowing how many of Amazon's sales come from these Pro Merchant Subscribers, we assume it to be 50%. We believe this is a conservative assumption, and it also ignores any additional revenue gains to Amazon from Pro Merchant membership fees.

<sup>22</sup> Note that the average used price that leads to a sale (\$8.76) is less than the average price of all used books listed in the marketplace (Table 1) as one would expect. Also note that Amazon never takes possession of the product, and thus revenue is approximately equal to gross profit.

<sup>23</sup>  $\$3.04 = \$8.76 * .15 + \$0.99 * .5 + \$1.23$

lion used books are sold through Amazon's marketplace annually. Recalling that only 4.63 million of these sales cannibalized new book sales, we estimate that Amazon sold approximately 25.06 million used book copies that would not otherwise have been sold new on the site. Said another way, only 16% (4.63 million / 29.69 million) of Amazon's used book sales directly cannibalize new book sales. The remaining 84% of used book sales apparently would not have occurred at the new book prices on Amazon's site. Using our figures above, these additional 25.06 million used book sales add approximately \$76.18 million to Amazon's profitability (25.06 million \* \$3.04 / used sale). Thus, on balance, the presence of Amazon's used book market added \$65.76 million to the company's profitability.<sup>24</sup>

### **6.3.3. Consumer Surplus**

To calculate the consumer surplus gain from the introduction of Amazon's used book markets, we apply Hausman and Leonard's (2002) methodology with respect to the consumer surplus gain from the introduction of new goods. Prior research has shown that income elasticity can be ignored can be ignored for consumer products that represent a small proportion of overall consumer expenditures (e.g., Brynjolfsson (1995) in the context of computer purchases, Hausman (1997a) in the context of telecommunications services, and Brynjolfsson, Hu, and Smith (2003) in the context of Internet book sales). Thus, ignoring income elasticity, the consumer surplus gain from the introduction of the used book market at Amazon.com should be given by:<sup>25</sup>

---

<sup>24</sup> It is important to note that this calculation of retailer surplus does not include surplus that may accrue to used booksellers. While Amazon takes a commission of 15% of the sale price on products sold through their marketplace, the remaining 85% of the sale goes to the seller. Under the assumptions that 50% of marketplace sellers are "Pro Merchant Subscribers" and thus do not pay \$0.99 to Amazon per sale), and that Amazon's \$2.26 payment to sellers exactly covers their shipping costs, marketplace sellers make approximately \$206.4 million (29.69 million used books \* (\$8.76 \* 85% - \$0.99 \* 50%)) on the sale of used books through Amazon's site. This almost certainly overestimates of the true marketplace seller surplus since many sellers would have a residual value of retaining the product. However, whatever the actual seller surplus might be, it likely represents another large source of social surplus from these transactions. We thank Paul Kattuman for making this observation.

<sup>25</sup> Note that, if income elasticity were positive for books, as seems likely, including income elasticity would increase our consumer surplus estimates.

$$CV = -\frac{P_u q_u}{(1 + \eta_u)} \quad (15)$$

where  $p_u$  is the average price of used book sold,  $q_u$  is the number of used books sold and  $\eta_u$  is the own price elasticity of used book demand.

Given the used book own price elasticity of -4.87, the average sale price of used books (\$8.76), and the quantity of used books sold (29.69 million), we estimate that the consumer surplus gain from the introduction of Amazon.com's used book market is \$67.21 million.

## 7. Discussion

While many papers in the IT, economics, and marketing literatures have analyzed the characteristics of new books sold in electronic markets (e.g., Brynjolfsson and Smith 2000; Clay, Krishnan, and Wolff 2001; Pan, Ratchford, and Shankar 2002; Baye, Morgan, and Scholten 2004), used books — and other used products — sold in IT-enabled exchanges may have an even larger impact on both electronic and physical markets.

IT-enabled markets for used products are able to aggregate supply and demand over a global marketplace, making it easier for buyers to find sellers and for sellers to find buyers. Because of this, these markets have significant advantages in terms of price, search costs, and selection over physical markets. As noted above, while Amazon's used book marketplace features at least one used book for almost every book in print and many out of print books, a typical physical used bookstore carries only between 5,000 to 30,000 unique titles. Likewise, prices of used books sold on the Internet are 38-75% lower than comparable prices in physical stores.<sup>26</sup> Because of

---

<sup>26</sup> As a point of comparison Brynjolfsson and Smith (2000) find that new book prices are *only* 15.5% lower online than in physical stores.

these advantages, the Internet makes up an estimated 67% of all used book sales (Wyatt 2005), versus *only* 12.7% of new book sales in the same channel (Rappaport 2004).

The question remains: how will these IT-enabled exchanges impact new product sales and resulting social welfare? The increased viability of used product sales in electronic markets may pose a significant threat for many categories of information goods — such as books, music, and movies — where there is relatively little degradation in the quality of the good over time and where artists and publishers are only compensated for the initial sale of the product. In these product categories, the increased variety, low prices, and low search costs available in online used product markets may attract customers who would have otherwise purchased a new copy of the product. If cannibalization of new product sales were to become widespread, it could undermine the profitability of the publishing business and reduce authors' and artists' creative incentives. Because of this, the Association of American Publishers and the Authors' Guild have asked Amazon to create artificial search costs for new book shoppers to locate used book copies by separating the two markets on Amazon's site.

In this research, we analyze the impact of used book markets on new book sales at Amazon.com. Using a unique dataset, we find that the cross-price elasticity of new book sales with respect to used book prices is rather low (0.088). This means that only 16% of Amazon's used book sales directly cannibalize new book purchases; the remaining 84% of sales represent purchases that otherwise would not have occurred at new book prices. Thus, while cannibalized sales result in an estimated \$45.05 million loss to publishers annually, the total welfare *gain* to society from this IT-enabled market is \$87.92 million annually after considering the \$65.76 million gain in Amazon.com's gross profits and the \$67.21 million gain in consumer surplus.

The implication of this finding for publishers is that, at least at present, used books do not appear to be a strong substitute to new book purchases for most consumers. Further, any lost publisher revenue from these used markets must be viewed in the context of the many ways the Internet has helped to increase new book sales: by lowering retail prices (while holding wholesale prices constant) (Brynjolfsson and Smith 2000), increasing product variety (Brynjolfsson, Hu, and Smith 2003), and providing a sales channel to customers who do not have local access to large bookstores (Brynjolfsson and Smith 2000). Finally, it is significant to note that the estimated \$45.05 million loss in annual gross profits, represents only 0.3% of total publisher gross profits.

For authors, while the observed levels cannibalization will lead to somewhat lower royalty payments, authors may experience an additional, indirect, gain through the 84% of used book sales that otherwise would not have occurred. For example, authors may accrue some added income from these additional readers through speaking fees, licensing deals, or advances on future books. Similarly, these new readers may buy new versions of subsequent releases by the same author(s). Further, used book purchases could also spur new book sales due to an increase in valuation from the possibility of resale (Ghose, Telang and Krishnan 2005).

However, book publishers, and producers of other comparable information goods, should remain attentive to potential changes in customer sensitivity toward used products. In the book category, customer sensitivity to used books may change over time as customers gain comfort with the quality and reliability of products sold in used markets. Likewise, producers of products with a stronger digital component than books should analyze the impact of cannibalization by used products. We speculate that cannibalization may be particularly acute for digital products, such as CDs and DVDs. Higher cannibalization levels might arise on the demand side because digital content typically does not degrade from use, reducing the importance of quality differentiation.

Further, on the supply-side most digital content (including CDs and DVDs) can be easily copied (and thus effectively retained) before they are resold, potentially making them more likely to be introduced for resale by (unscrupulous) sellers. In this regard, we believe that another key contribution of our work is in providing an easily executed methodology for product sellers and academic researchers to analyze the impact of cannibalization in other product categories.

It is also important to note that while cannibalization can reduce revenue to publishers and content creators, Information Technology can also provide new tools for controlling or eliminating used product markets. For example, licensing restrictions can prohibit the resale of some products purchased digitally (e.g., eBooks and music purchased through online music stores) and can create rental markets under more direct industry control (e.g., the older Divx and more recent Flexplay movie formats). Issues surrounding the viability and effectiveness of technology-enabled control over used product markets would make an important area for future research.

Finally, it is important to note several limitations of our study. First, our data comes from a single online retailer. Future research could include data from other well known used booksellers as well such as half.com. Second, we have focused on books only. In order to test the generalizability of our methodology and implications, future studies may wish to look at other product categories like CDs and DVDs are needed. Finally, our study represents only a snapshot in the evolution of the online used book marketplace. Consumer sensitivity to used books may change over time as consumers gain more familiarity with used products sold through online markets.

## References

- Anderson, S., V. Ginsburgh. 1994. Price Discrimination by Second-hand Markets. *European Economic Review* **38** 23-44.
- Bakos, J. Y. 1997. Reducing buyer search costs: Implications for electronic marketplaces. *Management Science* **43**(12) 1613-1630.
- Bapna, R. P. Goes, A. Gupta. 2003. Replicating Online Yankee Auctions to Analyze Auctioneers' and Bidders' Strategies *Information Systems Research* **14**(3) 244-268.
- Bapna, R., W. Gank, G. Shmueli. 2005. Consumer Surplus in Online Auctions. Working Paper, University of Connecticut
- Baye, M. R., J. Morgan, P. Scholten. 2004. Price Dispersion in the Small and in the Large: Evidence from an Internet Price Comparison Site. *Journal of Industrial Economics* **52**(4) 463-496.
- Ben-Akiva, M., S. Lerman. 1985. *Discrete Choice Analysis: Theory and Application to Travel Demand*. The MIT Press, Cambridge, MA.
- Benjamin, D., R. Kormendi. 1974. The interrelationship between the markets for used and new durable goods. *Journal of Law and Economics* **17** 381-401.
- Bond, E., L. Samuelson. 1984. Durable good monopolies with rational expectations and replacement sales. *RAND Journal of Economics* **15**(3) 336-345.
- Book Industry Study Group. 2004. Book Industry Trends. Book Industry Study Group, New York, NY. (ISBN 0-940-016-842.)
- Brynjolfsson, Erik. 1995. The contribution of information technology to consumer welfare. *Information Systems Research* **7**(3) 281-300.
- Brynjolfsson, E., M. Smith, 2000. Frictionless Commerce? A Comparison of Internet and Conventional Retailers, *Management Science* **46**, 563-585.
- Brynjolfsson, E., Y. Hu, M. Smith. 2003. Consumer Surplus in the Digital Economy: Estimating the Value of Increased Product Variety. *Management Science* **49**(11) 1580-1596.
- Brynjolfsson, Erik, Astrid A. Dick, Michael D. Smith. 2004. Search and Product Differentiation at an Internet Shopbot. Working Paper, Carnegie Mellon University, Pittsburgh, PA.
- Bulow, J. 1982. Durable-goods monopolist. *Journal of Political Economy* **90**(2) 314-322.
- Chevalier, J., A. Goolsbee. 2003. Measuring Prices and Price Competition Online: Amazon and Barnes and Noble. *Quantitative Marketing and Economics* **1**(2) 203-222.
- Chevalier, J., A. Goolsbee. 2005. Are Durable Goods Consumers Forward Looking? Evidence from College Textbooks. Working Paper, Yale School of Management, New Haven, CT.

- Clay, K., R. Krishnan, E. Wolff. 2001. Price Strategies on The Web: Evidence From The Online Book Industry. *Journal of Industrial Economics* **49**(4) 521-540.
- Coase, R. 1972. Durability and monopoly. *Journal of Law Economics* 15 143–149.
- Ellison, G., S. Ellison. 2004. Search, Obfuscation, and Price Elasticities on the Internet. Working paper, Massachusetts Institute of Technology, Cambridge, MA.
- Ghose, A., R. Telang, R. Krishnan. 2005. Effect of Electronic Secondary Markets on Supply Chain., *Journal of Management Information Systems*, 22(2), 91-120.
- Greco, A. 2005. *The Book Publishing Industry*. Lawrence Erlbaum Associates, Mahwah, N.J.
- Greco, A, R.M. Wharton, H. Estalemi. 2005. *Journal of Scholarly Publishing*, **36**(2), July.
- Guadagni, P., J. Little. 1983. A Logit Model of Brand Choice Calibrated on Scanner Data. *Marketing Science* **2**(3) 203-238.
- Hausman, J.A. 1997a. Valuing the effect of regulation on new services in telecommunications. *Brookings Papers on Economic Activity: Microeconomics* Vol. 1997 1-38.
- Hausman, J.A. 1997b. Valuation of new goods under perfect and imperfect competition. Bresnahan, Timothy F., Robert J. Gordon, eds. *The Economics of New Goods*. The University of Chicago Press, Chicago, IL, 209-237.
- Hausman, J.A., G. Leonard. 2002. The Competitive Effects of a New Product Introduction: A Case Study. *Journal of Industrial Economics* **50**(3) 237-263.
- Hendel, I., A. Lizzeri. 1999. Interfering with Secondary Markets. *Rand Journal of Economics* **30**(1) 1-21.
- Hicks, J.R. 1942. Consumers' surplus and index numbers. *Review of Economic Studies* **9**(2) 126-137.
- Liebowitz, S. J. 1982. Durability, Market Structure and New-used Goods Models. *American Economic Review* **72**(9) 816–824.
- Malone, Thomas, JoAnne Yates and Robert Benjamin. 1987. Electronic Markets and Electronic Hierarchies. *Communications of the ACM* **30**(6) 484-497.
- Miller, L.H. 1974. On Killing off the Market for Used Textbooks and the Relationship Between Markets for New and Secondhand Goods. *The Journal of Political Economy* 82(3) 612-619.
- Milliot, J. 2002. Amazon focused on unit growth. *Publishers Weekly* **249**(44) 15.
- Poynter, D. 2000. Publishing Poynters. April-June 2000, (Available at <http://parapub.com/getpage.cfm?file=newsletter/News0400.html>), accessed March 21, 2005.

Publishing Trends. 2004. Used Blues: Used Books Become Newer Everyday, To Many Publishers Dismay." *Publishing Trends Newsletter* **11**(7), Market Partners International, New York, NY (July 2004).

Rappaport, B. 2004. *Ipsos BookTrends 2004*. Ipsos-Insight, Chicago, IL.

Siegel, S., D. Siegel. 2004. *A Portrait of the U.S. Used Book Market*. Book Hunter Press, Yorktown Heights, NY.

Swan, P.L. 1980. Alcoa: The Influence of Recycling on Monopoly Power. *The Journal of Political Economy* **88**(1) 76–99.

Tedeschi, B. 2004. Online Battle of Low-Cost Books. *New York Times* (July 12).

Waldman, M. 1997. Eliminating the market for secondhand goods: An alternative explanation for leasing. *Journal of Law and Economics* **40**(1) 61-92.

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

Wingfield, N. 2003. New Chapter: In Latest Strategy Shift, Amazon Is Offering a Home to Retailers. *The Wall Street Journal* (September 24) A1.

Wyatt, Edward. 2005. Internet Grows as a Factor in Used-Book Business. *New York Times* (September 29).



Figure 2: Sample New Book Listing at Amazon.com

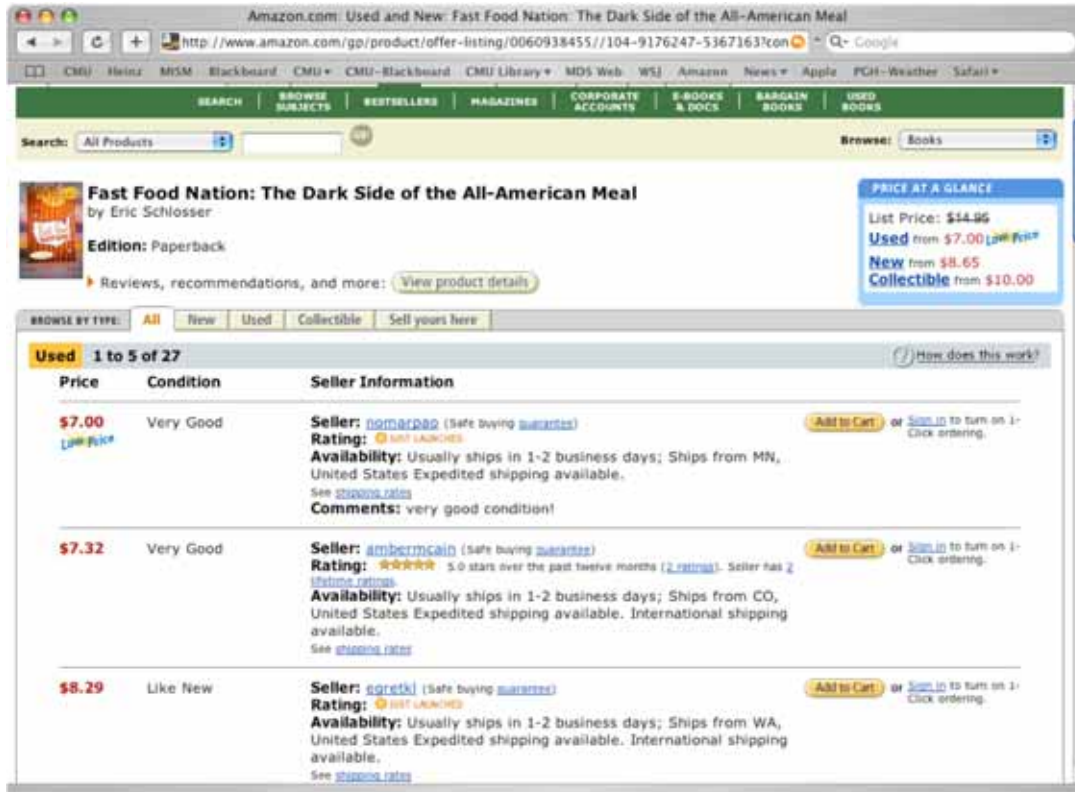


Figure 3: Sample Used Book Listing at Amazon.com

## Appendix

Recall the consumer utility functions on Page 11. By equating (i), (ii) and (ii), (iii), we derive the two indifference points  $\theta_1 = \frac{P_N - (1 + \alpha(1-k))P_U}{1-q}$  and  $\theta_2 = \frac{P_U}{q}$  which define the consumer market segments. By equating the demand of used goods ( $\theta_1 - \theta_2$ ) with the supply of used goods  $\alpha(1 - \theta_1)$ , we get the market clearing used-good price,

$$P_U = \frac{q(P_N^U(1+\alpha) - \alpha(1-q))}{1 - \alpha q(k(1+\alpha) - \alpha - 2)}.$$

The retailer's profit is then

$$\pi_R = (1 - \theta_1)(P_N^U - w^U) + (\theta_1 - \theta_2)k.$$

Substituting  $P_U$  in the retailer's profit equation and optimizing with respect to  $P_N$ , we get the optimal new good price,

$$P_N = \frac{(1 + w^U)(1 + \alpha q(\alpha - k)) + \alpha q[3 + \alpha^2 q(1 - k) + 2w^U + \alpha(2q - k(2q + w^U - 1))]}{2 + (6 + 4\alpha)q}$$

and the used good price

$$P_U = \frac{q((1 + \alpha(1 - k) + k)q + w^U)}{1 + (1 + 2\alpha(1 - k) + 2k)q}.$$

These expressions enable us to derive the relevant profits expressions as follows:

$$\pi_S^U = (1 - \theta_1)w = \frac{1}{2} \frac{(1 + \alpha q - w^U)w^U}{(1 + \alpha q(2 + \alpha))}$$

$$\text{and } \pi_R^U = \frac{1}{4} \frac{(1 + \alpha q - w^U)^2}{(1 + \alpha q(2 + \alpha))}.$$

The publisher can optimize its profit with respect to  $w^U$  and can set  $w^U = \frac{1}{2}(1 + \alpha q)$ .

Without the used book market, the publisher makes a profit of  $\pi_S = \frac{w(1-w)}{2}$  and the retailer makes a profit of  $\pi_R = \frac{(1-w)^2}{4}$ . Comparing these respective profits it is easy to see that  $\pi_S^U - \pi_S < 0$  for all values of  $w^U$ ,  $\alpha$  and  $q$  and  $w$  is set at optimal  $w = 1/2$ . On the other hand,  $\pi_R^U - \pi_R > 0$  for all  $w^U < 1 + aq - \frac{1}{2}\sqrt{1 + a(2+a)q}$ . Thus, as long as the publisher does not increase its wholesale price, the retailer always benefits and the publisher always loses.