Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Frictionless Commerce? A Comparison of Internet and Conventional Retailers

Erik Brynjolfsson, Michael D. Smith,

To cite this article:
Erik Brynjolfsson, Michael D. Smith, (2000) Frictionless Commerce? A Comparison of Internet and Conventional Retailers. Management Science 46(4):563-585. https://doi.org/10.1287/mnsc.46.4.563.12061

Frictionless Commerce? A Comparison of Internet and Conventional Retailers

Erik Brynjolfsson • Michael D. Smith
MIT Sloan School of Management, 50 Memorial Drive, Cambridge, Massachusetts 02139
erikb@mit.edu • mds@mit.edu

There have been many claims that the Internet represents a new nearly "frictionless market." Our research empirically analyzes the characteristics of the Internet as a channel for two categories of homogeneous products—books and CDs. Using a data set of over 8,500 price observations collected over a period of 15 months, we compare pricing behavior at 41 Internet and conventional retail outlets.

We find that prices on the Internet are 9–16% lower than prices in conventional outlets, depending on whether taxes, shipping, and shopping costs are included in the price. Additionally, we find that Internet retailers' price adjustments over time are up to 100 times smaller than conventional retailers' price adjustments—presumably reflecting lower menu costs in Internet channels. We also find that levels of price dispersion depend importantly on the measures employed. When we compare the prices posted by different Internet retailers we find substantial dispersion. Internet retailer prices differ by an average of 33% for books and 25% for CDs. However, when we weight these prices by proxies for market share, we find dispersion is lower in Internet channels than in conventional channels, reflecting the dominance of certain heavily branded retailers.

We conclude that while there is lower friction in many dimensions of Internet competition, branding, awareness, and trust remain important sources of heterogeneity among Internet retailers.

(*Search*; *Competition*; *Internet*; *Price Dispersion*; *Menu Costs*; *Pricing*; *Intermediaries*)

1. Introduction

> The Internet is a nearly perfect market because information is instantaneous and buyers can compare the offerings of sellers worldwide. The result is fierce price competition, dwindling product differentiation, and vanishing brand loyalty.
>
> Robert Kuttner in *Business Week*, May 11, 1998

> . . . industry titans such as Bill Gates, the boss of Microsoft, regale the world's leaders with the promise of "friction-free capitalism."
>
> *The Economist*, May 10, 1997

> All of this brings you closer and closer to the efficient market.
>
> Robert MacAvoy, President Eastman Consulting[1]

The conventional wisdom regarding Internet competition, expressed in the preceding quotes, is that the unique characteristics of the Internet will bring about a nearly perfect market. In the extreme version of this "Internet efficiency" view, the characteristics of the Internet will lead to a market where retailer "location" is irrelevant, consumers are fully informed of prices and product offerings, and all retailers make zero economic profit.

At the same time, there is anecdotal evidence that the Internet may not be completely efficient. For

[1] Quoted in "Good-Bye to Fixed Pricing," Business Week, May 4, 1998.

0025-1909/00/4604/0563$05.00
1526-5501 electronic ISSN

example, if price competition is inevitable, why do investors place such high valuations on Internet retailers such as Amazon.com and CDnow that sell undifferentiated products—the products most likely to experience fierce competition. If the Internet makes "location" irrelevant, why are Internet retailers making million-dollar deals for the right to showcase their products on major Internet portals and content sites.[2] While there may be answers to these questions consistent with the efficiency hypothesis, the degree of efficiency on the Internet deserves empirical verification.

As sales of consumer goods on the Internet grow from $7.8 billion in 1998 to an estimated $108 billion in 2003,[3] questions surrounding the level of efficiency on the Internet take on heightened importance for consumers, businesses, and stockholders: Will competition on the Internet lead to lower and more homogeneous prices. Can Internet retailers adjust their prices more readily than conventional retailers in response to structural changes in supply or demand. Do brand name and trust matter for homogeneous goods sold on the Internet.

Ultimately, the effects of the Internet on commerce are likely to be varied and occasionally unpredictable. Even the best theorizing will need to be based on empirical observations. Accordingly, our work seeks to address these questions by comparing actual prices charged by Internet and conventional retailers of books and compact discs (CDs). Our data include over 8,500 individual price observations collected from February 1998 to May 1999. We run a variety of statistical and econometric tests on our data to determine pricing characteristics both statically and across time.

In addressing these questions, our work follows several other studies of the effect of electronic commerce on differentiated goods markets. Degeratu et al. (1998) study online grocery sales and find that price sensitivity can sometimes be lower online than in conventional channels. Shankar et al. (1998) use survey data for travelers to show that prior positive experience with a brand in the physical world can decrease price sensitivity online. Lynch and Ariely (1998) use laboratory experiments regarding the sale of wine through electronic channels to show that the amount of product information provided to customers can affect price competition and increase customer loyalty. Lee (1998) finds that prices for used cars sold via electronic auction markets tend to be higher than prices for used cars sold via conventional auction markets. Clemons et al. (1998) find that prices for airline tickets offered by online travel agents vary by as much as 20%. Related studies are reviewed in Smith et al. (2000).

This paper extends these results by studying prices for homogeneous physical goods (CDs and books) matched across conventional and Internet channels. We focus on homogeneous goods as the product categories most likely to experience strong price competition given the characteristics of Internet channels (Bakos 1998). We focus on a sample of goods carefully matched across conventional and Internet channels to eliminate much of the unobserved heterogeneity in the sample, leaving us with a clearer measure of the difference between the channels.

Our study is especially influenced by a pioneering work by Bailey (1998). Bailey analyzed prices for books, CDs, and software in Internet and conventional outlets from 1996 to 1997. He found evidence that prices on the Internet were, on average, higher than prices in conventional outlets. In addition to analyzing a more recent time period, we refine his methodology to better account for "typical" prices in both channels. We also extend his empirical tests to better analyze differences in price levels, menu costs, and price dispersion between Internet and conventional channels. While we rely on relatively simple econometrics for this study, a number of interesting characteristics of Internet retailing are readily apparent:

1) Prices for books and CDs sold on the Internet average 9–16% less than the identical items sold via

[2] For example, Barnes and Noble's $40 million dollar, 4-year deal to be the exclusive bookseller for America Online and K-Tel's "seven-figure" 2-year deal to sell music on Playboy's Internet site. Amazon.com, CDnow, and Barnes & Noble have announced similar deals with sites such as Yahoo, CNN, New York Times, and Disney.

[3] Forrester Research, "Retail's Growth Spiral," On-line Retail Strategies, November 1998.

Table 1 Summary of Data Gathered

	Product	Retailers	Titles	Observations*	Source
Feb. 1998–May 1999	Books	4 (Internet)	20	1,144	Internet
		4 (Internet–Hybrid)	20	1,124	Internet
		4 (Conventional–Hybrid)	20	1,111	Conventional
		4 (Conventional)	20	1,037	Conventional
	CDs	4 (Internet)	20	1,115	Internet
		4 (Internet–Hybrid)	20	1,102	Internet
		4 (Conventional–Hybrid)	20	1,109	Conventional
		4 (Conventional)	20	978	Conventional

* The number of observations varies because some retailers were unable to provide prices for some out-of-stock items.

conventional channels. The mean price for books was \$2.16 less and \$2.58 less for CDs.

2) Internet retailers change prices in smaller increments than do conventional retailers. The smallest observed price change on the Internet was \$0.01 while the smallest observed price change by a conventional retailer was \$0.35.

3) There are substantial and systematic differences in prices across retailers on the Internet. Prices posted on the Internet differ by an average of 33% for books and 25% for CDs. At the same time, the dispersion of prices weighted by retailer popularity reveals that Internet markets are highly concentrated, but the retailers with the lowest prices do not receive the most sales.

These results provide support for the hypothesis that the Internet is a more efficient channel in terms of price levels and menu costs. However, the price dispersion results suggest that retailer heterogeneity with respect to factors such as branding, awareness, and trust remain important factors to understanding Internet markets.

The paper proceeds in three sections. In §2 we discuss our data collection methodology. In §3 we review the theoretical basis and empirical results of a variety of tests on our data. In §4 we present conclusions from our research and identify areas for future research.

2. Data Collection Methodology

Our study gathered price data, product characteristics, and retailer characteristics for CDs and books sold through Internet and conventional retail outlets. Books and CDs were chosen because the physical products themselves are homogeneous: Books can be uniquely identified by their ISBN number, and CDs can be uniquely identified by a record label catalog number. This enabled easier comparison of prices across channels. Books and CDs are also useful because each are sold through numerous Internet and conventional outlets, facilitating comparison across a wide variety of firms. Our methodology is summarized in Table 1, and detailed below.

2.1. Retailers

Our methodology tracks two types of retailers: those that sell over the Internet and those that sell through conventional outlets. For each product category, we selected eight Internet and eight conventional retailers. Half of these retailers (four Internet retailers and a matched set of four conventional retailers) are "hybrid" retailers: They maintain operations both on the Internet and in conventional outlets.[4,5]

To select our Internet retailers, we compiled a list of all U.S.-based book and CD retailers that were listed in Yahoo and that sold a general selection of titles. We used Yahoo because it was believed to be both comprehensive and unbiased (retailers are listed at no

[4] For example, Barnes and Noble is a hybrid retailer. It has an Internet outlet (barnesandnoble.com) and multiple conventional-world stores.

[5] A complete list of the specific retailers we tracked is available from http://ebusiness.mit.edu/papers/friction/.

charge).[6] We excluded retailers outside the United States and niche retailers to focus our attention on competition in a particular market—general selection bookstores in the United States.

From the list of general selection retailers provided by Yahoo, we attempted to identify a set of retailers with fairly comprehensive coverage in each product category. To estimate coverage, we used the number of hits for each retailer collected by the Web21 web ranking service. Web21 estimates website popularity by sampling traffic at selected points within the Internet and aggregating the total number of "hits" for each Internet address.[7] While this provides an imperfect measure of market share, we judged Web21 to be the most rigorous and widely recognized web ranking service at the time this study was initiated.[8]

Using the Web21 hit data we selected the four largest Internet-only retailers and the four largest hybrid retailers in both product categories. Our Internet retailers comprised 99.8% of the total number of hits for book retailers and 96.5% of the total number of hits for CD retailers, and thus comprise a fairly comprehensive sample.

It would not be feasible to collect data from an equally comprehensive set of the 22,000 conventional book and 12,000 conventional CD retailers in the United States. Instead, we selected our eight conventional retailers to be "representative" of sales made in brick and mortar stores for each product category. Four of the retailers comprise conventional outlets of our hybrid Internet retailers. The remaining four retailers were selected so that the eight conventional retailers matched national sales patterns. For book retailers we were able to obtain a profile of book sales by retail outlet type from the American Booksellers Association, and we attempted to match our selections to these sales proportions. For CD retailers, we attempted to obtain a mix of national chains and local stores based on data in the 1997 Market Share Reporter.

To verify that the conventional prices in our sample did not contain systematic biases, we used census data and yellow page directories to select 50 book and 50 CD retailers at random from all book and CD retailers operating in the United States. We made phone calls to each retailer requesting prices for two titles from our sample.[9] We then compared the prices charged by the random sample of retailers to prices charged by retailers in our study for the same week. Our empirical findings from this comparison suggest that the prices charged by conventional retailers in our sample are not biased in a way that would weaken our main results when comparing conventional prices to Internet prices. Specifically, we find that prices for conventional retailers in our study are no higher than prices in a random sample of conventional retailers. We also find that price dispersion for the conventional retailers in our sample is no smaller than price dispersion in a random sample of conventional retailers.

There are several other aspects of our retailer selection that deserve mention. First, we selected our conventional retailers from a geographically dispersed set of locations: California, Georgia, Maryland, Massachusetts, Oregon, Texas, and Virginia. This was in response to comments on the representativeness of earlier findings by Bailey (1998), who gathered data only from Boston-area retailers. We added a geographically dispersed set of conventional and hybrid retailers to gain a more representative measure of national average price levels for conventional retailer prices. Because of differences in income levels across geographical areas, obtaining retailers from different geographical markets may lead to greater variance in prices than if all conventional stores were selected from the same city. We discuss the impact of this in more detail below.

[6] We compared the list of retailers derived from Yahoo to retailers listed by several other directory services and price intermediaries (e.g., www.evenbetter.com/, www.simplymusic.com/, www.infospace.com/) and found that we had not excluded any relevant retailers from our list.

[7] More detail on Web21's methodology is provided on their website (www.web21.com/).

[8] We validated our hit data with data regarding the number of hypertext links to a web page. Analyzing the number of links reported by Altavista (www.altavista.net) shows them to be broadly consistent with the Web21 statistics.

[9] The methodology for conducting this comparison is available from http://ebusiness.mit.edu/papers/friction.

Second, we adapted our retailer selection as firms entered or exited the market. In early June 1998, Amazon.com started selling CDs in their Internet outlet and Borders.com opened an Internet outlet selling books and CDs. Both sites easily qualified as a "top 4" site according to the methodology outlined above. Thus, after they entered, we added Amazon.com as an online CD retailer, and dropped the retailer with the lowest number of hits (Tunes.com). Likewise, we added Borders.com as an Internet book and CD retailer, dropping Cody's Books and Camelot Records. Similarly, in May 1999, altbookstore.com ceased Internet operations and was replaced in our sample by buy.com, the next most popular store.

Third, while our strategy of sampling retailers from a broad geographic scope has many advantages, it also introduced some logistical problems. Several of the individuals who gathered remote prices for us were unavailable during the summer months (June through August). Therefore, during these months we shifted the location of price gathering for several stores to the Boston area. Additionally, because of a staffing shortage, we did not gather any prices from conventional outlets in September 1998, and for this reason we have dropped all September (Internet) observations from our data set.

Fourth, in October 1998 and again in February 1999 we replicated our selection of "top 4" retailers using the same methodology described above. In October 1998, we found that no changes were necessary to the retailers we were already tracking. In February 1999, we dropped BooksNow in favor of spree.com on the basis of popularity.

2.2. Titles

Our study tracks 20 book titles and 20 CD titles. Half of these titles were selected from current best-seller lists and the other half from a random selection of titles generally available in conventional outlets. We refer to the first category of titles as "popular" and the second as "miscellaneous."

Popular CD titles were selected from the top 10 albums in Billboard's best-selling album list at the time the study started. Our popular book titles were selected from the top 1 or 2 titles from each of the six *New York Times* best-seller categories. During the study period, we replaced titles that dropped off their respective best-seller list with the highest ranked title not already included in the sample.

The selection of miscellaneous titles was complicated by our dual goals of obtaining a suitably random selection of titles and obtaining titles that are available in both Internet and conventional outlets. To balance these goals, we selected a random set of titles from one conventional CD and one conventional book retailer that were not included in our study. From these titles we eliminated titles that were not carried by a majority of our conventional retailers.[10]

It is worth noting that Internet retailers typically have a much larger selection of books and CDs than conventional retailers (e.g., over 3 million books at BarnesandNoble.com versus about 175,000 at a Barnes & Noble superstore[11]). Since we track only titles that are available at most conventional and Internet stores, we implicitly put zero weight on titles that are carried *only* by Internet stores. By definition, the price for such titles is higher in conventional stores than at Internet stores, since the effective price (a.k.a. the "virtual price") for a product that is not available is higher than any price with positive demand (See Hicks 1940 and Hausman 1997). Therefore, our finding of lower prices on the Internet would be unambiguously strengthened if a "complete" selection of books and CDs could have been tracked and priced.

2.3. Source

The last column in Table 1 lists the source for our data. All of the data for our Internet retailers were gathered from the retailer's Internet web page. Likewise, the data for all but two of our conventional retailers were gathered directly from the brick and mortar stores in personal visits by people employed by our project. We found that the remaining two stores, Cody's Books and Powell's Books, simply posted the prices for their conventional stores on the Internet. For these two retailers, we first sent e-mail to the retailer and con-

[10] This was done to ensure that we would not have to discard large numbers of observations for titles not carried by our retailers. We only had to drop 3 books and 1 CD from the original list due to this constraint.

[11] *Business Week*, "Amazon.com: The Wild World of E-commerce," December 14, 1998, pp. 110.

firmed that the prices posted on the Internet were identical to those charged in the conventional outlet. We then made phone calls in February 1998 to both stores to verify the prices for all 20 books in our study. After finding that all prices were in fact the same as those posted on the web, we recorded the remaining observations based on the information contained on the Internet. In May 1998, December 1998, and again in May 1999 we verified that neither store had changed its pricing policy.

3. Empirical Results

We analyze our data by examining price levels, price changes over time, and price dispersion across stores. With regard to price levels and price changes, our goal is to compare the characteristics of Internet channels for books and CDs to the characteristics of existing conventional channels for the same products. For price dispersion we also discuss the characteristics of Internet channels alone in an effort to identify the sources of price dispersion in Internet channels.

This section is divided into four parts. The first three parts present our analysis of price levels, price changes, and price dispersion in Internet and conventional markets. For each test we first present the theoretical basis for the test and discuss possible sources of structural differences between Internet and conventional channels. We then review the empirical results and finally discuss potential interpretations of our findings. The fourth part discusses potential explanations for the larger than expected price dispersion we observe on the Internet.

3.1. Relative Price Levels

> Lower buyer search costs in electronic marketplaces promote price competition among sellers. This effect will be most dramatic in commodity markets, where intensive price competition can eliminate all seller profits.
>
> Yannis Bakos (1998, p. 40)

From an economic perspective, price levels are a particularly useful measure of efficiency. Within the classic economic model of social welfare, setting prices above marginal cost leads to deadweight loss as some consumers are forced to forgo socially efficient trades. In this setting, lower prices lead to an increase in social welfare as more welfare-enhancing trades are allowed to occur.

There are a variety of factors that can lead to lower prices in equilibrium. One important factor in our setting is lower search costs. Several economic models (e.g., Salop 1979) have argued that positive consumer search costs can lead to prices above marginal cost in equilibrium, even for otherwise homogeneous products. In a setting more closely tied to the Internet, Bakos (1997) uses Salop's circular city model to examine the effects of lower search costs on equilibrium prices in electronic markets. In Bakos' setting, consumers incur search costs to discover the prices and characteristics of products. The market then forms a Perfect Bayesian Equilibrium where consumers and producers choose actions based on beliefs about their opponents' actions. The result is that, in the presence of search costs, producers will set prices above marginal cost. Lower search costs, however, will drive Internet prices for homogeneous goods toward the Bertrand marginal cost pricing result.

The extent of search costs in Internet book and CD channels remains an open question. On one hand, it seems clear that the technical characteristics of the Internet work to lower search costs compared to conventional channels (see Bakos 1997, 1998, for example). In particular, Internet price search intermediaries[12] provide customers with easy access to prices at a wide variety of retailers. Further, even without price intermediaries, customers can use Internet hypertext links and retailers' search tools to quickly compare prices between retailers. From a standpoint of casual empiricism, gathering prices on the Internet seems far easier than physically visiting stores or trying to obtain prices by telephoning stores and speaking with a salesperson.[13]

[12] For example, www.infospace.com, www.bottomdollar.com, and www.mysimon.com for books and CDs; and www.evenbetter.com, www.addall.com, www.bestbookbuys.com, and www.bookfinder.com for books.

[13] For example, we found that gathering book prices on the Internet took 20 seconds per store using a price search intermediary (www.evenbetter.com) and approximately 1 minute per store by visiting the retailer's web pages directly. Gathering the same prices by phone from 5 conventional stores took nearly 3 minutes per store.

On the other hand, we do not expect search costs in Internet channels to be trivially small in comparison to search costs in the physical world. It may be that the types of consumers who shop on the Internet have higher wages or are busier and therefore have systematically higher costs of time than those who shop in conventional stores. Additionally, some Internet consumers may not be aware of more than one or two Internet retailers for books or CDs and similarly may not be aware of the existence of price search intermediaries. These consumers may have a high search cost to find and qualify additional retailers or to find price search intermediaries, and therefore would be unlikely to take advantage of the lower search costs offered by Internet channels.[14]

In addition to low search costs, low entry costs or low operational costs could also lead to lower equilibrium price levels on the Internet. One widely discussed aspect of Internet channels is low entry costs for Internet retailers. According to this view, an Internet retailer needs only a web page and a relationship with book or CD distributors to effectively "enter" the market. Lower entry costs should lead to more entry in equilibrium (e.g., Salop 1979); and more entry, or even the mere threat of entry (e.g., Milgrom and Roberts 1982), should lead to lower prices in equilibrium. Similarly, Internet outlets may have lower operational costs than conventional outlets. These lower operational costs among Internet retailers could also lead to lower prices in a long-run equilibrium.[15]

Ultimately, whether prices on the Internet are lower than prices in conventional outlets is an empirical question. Accordingly, we run several econometric tests to analyze the relative levels of prices between Internet and conventional channels.

Table 2 ***t* tests on Mean Store Prices**

Product Market	Conventional Price Mean	Internet Price Mean	Alternative Hypothesis	*t* test Significance
Books	13.90	11.74	$P_{INET} < P_{PHYS}$	0.001
CDs	16.07	13.49	$P_{INET} < P_{PHYS}$	0.001

First, we run simple hypothesis tests on mean Internet and conventional prices after weighting for the popularity of Internet retailers.[16] The weights reflect the fact that some Internet retailers (e.g., Amazon.com, CDnow) receive far more traffic than others (e.g., Cody's, tunes.com). Thus, our measure of "typical" prices on the Internet takes into account the relative number of customer visits to each site. To establish weights, we create a measure of retailer "screen share." Our screen share measure is the average of the relative number of hits (reported by Web21) and links (reported by Altavista) to each Internet retailer.[17,18]

The results of these tests are displayed in Table 2.[19] Our results suggest that the prices charged for books and CDs are lower on the Internet than in conven-

Obviously, physically visiting the stores would require far more time and effort.

[14] We also note that the Internet may lower search costs for shoppers at conventional stores if, for example, the stores post the same prices online as they charge in their conventional outlets. However, since only a fraction of hybrid book and CD retailers set prices in this way, we expect that in our setting easy access to prices over the Internet will affect Internet retailers more than conventional retailers.

[15] Some Internet retailers may price below marginal cost for some goods. In this case, of course, lower prices would be associated with less economic efficiency. However, discussions with several senior executives at Internet retailers suggests that such extremes of pricing are rare.

[16] Unless otherwise stated, our tests of Internet prices include prices for retailers who only operate on the Internet (e.g., Amazon.com, CDnow) and the Internet outlets of hybrid retailers (e.g., Barnesand-Noble.com, Borders.com). Conventional retailer prices include prices charged in conventional retailers (e.g., Crown, Sam Goody) and prices in hybrid retailers' physical outlets.

[17] Physical prices are left unweighted because, as noted above, our physical prices are fairly representative of prices charged at a random sample of U.S. retailers. Weighting by estimates of conventional stores' market share would slightly enhance our finding of lower prices on the Internet.

[18] Weighting Internet observations by links or hits alone would result in only a $0.02 (plus or minus) change in relative book prices and a $0.06 change in relative CD prices. Neither change effects our conclusions.

[19] *t* test results are based on the significance of the weighted regression of price onto an Internet dummy variable. To control for serial correlation, the *t* test significance levels are quoted using Newey-West standard errors.

Table 3 Proportion of the Time the Minimum Internet Price (of Eight Sample Stores) Is Less Than or Equal to the Minimum Conventional Price (of Eight Sample Stores)

Product Market	Min. Internet Price < Min. Conventional Price	Min. Internet Price = Min. Conventional Price	Alternative Hypothesis	*P*-Value Significance*
Books	92.0%	4.5%	$P_{INET} < P_{PHYS}$	0.05
CDs	84.6%	5.1%	$P_{INET} < P_{PHYS}$	0.05

* Under the null hypothesis that minimum prices are the same across channels, this statistic should follow a Bernoulli distribution—half the time the lowest price is found on the Internet. We test this null hypothesis (after discarding ties) against an alternative hypothesis that it is more likely to find minimum prices on the Internet. Since a Bernoulli distribution assumes independent draws, we run our tests separately for each time period (to control for serial correlation in the minimum price statistic). We reject the null in *all* time periods at the reported *p*-values.

tional stores: by \$2.16 (15.5%) for books and \$2.58 (16.1%) for CDs. *t* tests reject the null hypothesis that mean prices on the Internet are equal to mean prices in conventional stores in favor of the hypothesis that Internet prices are lower than conventional prices.[20]

This is, of course, only one possible test of differences in Internet and conventional prices. Measuring unweighted Internet prices, we still find that average prices are statistically significantly lower ($p < 0.001$) on the Internet than in conventional outlets: by \$1.22 for books and \$2.29 for CDs. Another test of price differences compares the lowest price found among all eight Internet retailers in our sample for a given book in a given time period with the lowest price found among all eight conventional retailers sampled. We find the minimum price on the Internet is, on average, \$1.29 lower for books and \$1.40 lower for CDs than the minimum price charged by conventional retailers. Further, Table 3 shows that over the study period, the lowest price for a given title in a given week across all retailers sampled is found on the Internet 92.0% of the time for books and 84.6% of the time for CDs. Each of these comparisons reflects the expected outcome which would result if a shopper compared prices at an equal number of retailers in each channel (one for the mean price comparison, all eight for the minimum price comparison). Since it takes less time to compare prices on the Internet, it is plausible that the average Internet shopper will compare prices at more retailers than the average conventional shopper. Any comparison that took this fact into account would be more likely to find the lower price on the Internet, strengthening our basic finding.

One potential problem in interpreting these statistics is that posted prices alone do not adequately reflect the price paid by consumers to purchase an item. A more appropriate measure of the actual price paid by consumers might include the amount consumers must spend on taxes, shipping and handling, and transportation to obtain the product. In other words, we are interested in a measure of the cost for consumers to have the product in their living room. Including sales tax in these calculations represents customers' prices instead of channel efficiency. As noted below, excluding sales tax charges from our tests does not effect our conclusion that prices are lower on the Internet.

We need to make several assumptions in order to include these factors in our calculations. For taxes charged by conventional retailers, we use the local tax rate where the store is located. For Internet retailers, tax calculations are more complex. Current tax law dictates that retailers must charge tax only in states where they have local operations—a.k.a. nexus. For example, Books.com (headquartered in Cleveland) must charge tax on all orders shipped to addresses in the state of Ohio. This puts hybrid stores with operations in a large number of states at a disadvantage. For example, Blockbuster music's Internet outlet must charge tax in all 50 states. To combat this disadvantage, some hybrid retailers have split off their Internet businesses and are therefore only required to charge taxes in states where they have Internet operations. Borders Books operates conventional stores in many states, but Borders.com, a separate business entity, charges tax only in Tennessee and Michigan.

Taking these factors into account, we calculate taxes

[20] The empirical results reported in this section hold whether one compares all prices (as reported here) or compares prices separately for the groups of retailers tracked from February to May 1998, June to August 1998, October 1998 to January 1999, and February to May 1999.

for Internet retailers based on a population-weighted average of the tax rates for states where they charge taxes.[21] This calculation corresponds to the average tax rate U.S.-based consumers would receive if they chose stores without regard to sales tax considerations. An alternative would be to assume that Internet consumers are able to avoid paying local sales tax by choosing retailers who do not have nexus in their state (the assumption adopted by Goolsbee (forthcoming)). If this situation were to hold, our calculations will overestimate the actual price of Internet goods, and our results are strengthened.[22]

Shipping and handling charges are sensitive to the number of items shipped and to the speed of delivery. Shipping and handling charges for Amazon.com (which are typical of other stores) illustrate this fact. Standard (3 to 7 day) shipping from Amazon.com costs $3.95 for the first book and $0.95 for each additional book, while next-day shipping costs $10.95 for the first book and $2.95 for each additional book. Thus, the "per book" cost of standard shipping for five books is $1.55 compared to $10.95 for one book using next-day shipping. Since our analysis focuses on "typical" prices charged to Internet consumers, we assume that customers order an average of three titles per transaction and request standard shipping for each of the Internet retailers in our sample.[23]

For transportation charges we estimate that consumers shopping at conventional stores must drive an average of five miles round trip to make their purchase. Consistent with our Internet calculations, we also assume that conventional shoppers purchase three books or CDs per trip.

We use the government reimbursement rate of $0.32 per mile to include this expense in our calculations.[24] We note that this estimate of transportation charges is likely to be conservative for two reasons. First, five miles per round trip is less than half our calculation of the average distance a consumer in the United States must travel to reach the nearest book or CD retailer.[25] Second, and more significantly, our estimate does not include any value consumers may place on their time to get to and from the store.[26]

Using these assumptions, we run the same tests as above on "typical full price" (i.e., price including tax, shipping and handling, and mileage charges after weighting our Internet prices by screen share). Making these adjustments (Table 4) we again find lower prices on the Internet. For books we find that Internet prices are $1.36 (9%) less than prices in conventional outlets. For CDs we find that Internet prices are $2.26 (13%) less than prices in conventional outlets.[27]

Table 4 *t* **tests on Mean Full Prices**

Product Market	Conventional Price Mean	Internet Price Mean	Alternative Hypothesis	*t* test Significance
Books	15.04	13.69	$P_{INET} < P_{PHYS}$	0.001
CDs	17.41	15.15	$P_{INET} < P_{PHYS}$	0.001

[21] Using Books.com as an example, Ohio has a 5.5% tax rate and makes up 4.2% of the U.S. population. Therefore, we use an effective tax rate of 0.23% (5.5% × 4.2%) for Books.com.

[22] Many states ask that buyers report and pay a "use tax" on goods purchased from out of state retailers. However, very few consumers are aware of these laws and essentially no one pays them on books or CDs (Goolsbee, forthcoming).

[23] These assumptions are consistent with available industry information. Forbes ASAP (April 6, 1998, p. 47) reports that customers at Tower Records Internet site order 2.7 recordings per transaction. Ghemawat and Baird (1998) report that Amazon.com's customers order an average of $50 in books per transaction. The average price of a book in the 100 most frequently ordered books at Amazon.com is $16.14, which, after adding shipping and handling charges, works out to 2.8 books per order.

[24] The result is 5 miles × $0.32 per mile/3 items per trip = $0.53.

[25] We used a national database of addresses to calculate that the average consumer in the United States must actually travel about 10.9 miles to reach the nearest book retailer, and 10.4 miles to reach the nearest CD retailer. Our methodology for this analysis can be found at http://ebusiness.mit.edu/papers/friction.

[26] We note that even if consumers place a very low value on their time, taking into account differences in shopping convenience would almost certainly enhance our results. For example, as part of a classroom exercise, we asked 70 MBA students to compare "shopping experiences" when shopping for CDs at Internet and conventional retailers. Including travel, search, and purchase times, our results showed that it took 35 minutes longer to shop in conventional outlets than it did in Internet outlets.

[27] Considering prices without including mileage charges at all (but still including shipping and handling and taxes) Internet prices are $0.82 lower for books and $1.75 lower for CDs. Assuming that consumers order only one title per transaction, Internet prices are $0.40 lower for

Table 5 **Proportion of the Time the Minimum Internet Price Is Less than the Minimum Conventional Price (Prices Include Shipping and Handling, Tax, and Mileage)**

Product Market	Min. Internet Price < Min. Conventional Price	Alternative Hypothesis	*P*-Value Significance*
Books	83.4%	$P_{INET} < P_{PHYS}$	0.01
CDs	82.5%	$P_{INET} < P_{PHYS}$	0.05

* As above, to control for serial correlation, we conduct separate tests in each time period under the null hypothesis of a Bernoulli distribution (equal probability of finding lower prices on the Internet). We reject this null hypothesis in *all* time periods at the reported *p*-values.

As above, examining the minimum price found in each channel shows that Internet prices including shipping and handling and taxes are lower than conventional prices by \$1.09 for books and \$1.23 for CDs. Likewise, the lowest price for any given item for any given time period is found on the Internet 83.4% of the time for books and 82.5% of the time for CDs (Table 5).

We conclude that prices for books and CDs are lower on the Internet whether one examines prices alone or prices including tax, shipping and handling, and mileage charges.[28] This finding supports the hypothesis that the Internet provides a more efficient channel for the products we track.

An implication of this finding is that an increasing share of sales should be expected to take place through the Internet channel. Indeed, anecdotal evidence suggests that competition between Internet and conventional book retailers is evolving in precisely this way. A major university bookstore has attributed "double digit decreases in trade book sales" to competition from Internet retailers.[29] Because university bookstores cater to an audience with unusually high levels of Internet access compared to the rest of the population, they may function as the "canaries in the mineshaft" whose demise previews the coming threats to bookstores in other locations. In fact, Gerry Masteller, co-owner of an independent bookseller in Palo Alto, California, cited competition from Amazon.com as one reason he decided to shut down his store: "I feel like Amazon.com is the straw that is potentially breaking the camel's back."[30]

3.2. Price Changes and Menu Costs

> Even small menu costs may be sufficient to generate substantial aggregate nominal rigidity and large business cycles.
>
> Daniel Levy et al. (1997, p. 792)

Menu costs—the costs a retailer incurs when changing a posted price—have long been studied by macroeconomists as a potential source of economic fluctuations (Sheshinski and Weiss 1993, Stigler and Kindahl 1970, for example). In the standard view, optimally retailers will make small price adjustments in response to shifts in supply and demand conditions. However, in practice, the cost to change prices is nonzero, and in some cases can be quite high (see Levy et al. 1997 for evidence of menu costs in grocery stores averaging \$0.52 per price change, \$105,887 per store, and consuming 35.2% of net margins). Because of this, retailers will only change prices when the benefit of a price change exceeds the associated cost. The resulting "price stickiness" can amplify even small menu costs by leading to an inefficient assignment of resources, potentially creating large business cycles (Levy et al. 1997) as noted above.

In conventional outlets, menu costs are driven by the cost of physically relabeling the prices of goods on store shelves. On the Internet, we hypothesize that menu costs should be much lower—comprised primarily of the cost to change a single entry (per title) in a database. If this hypothesis were true, we would expect to see less "price stickiness" among Internet retailers.

To examine this hypothesis we analyze the size and

books and \$1.96 lower for CDs. Excluding sales taxes, Internet prices are \$0.87 lower for books and \$1.69 lower for CDs. Each of these results is significant at the 0.05 confidence level or better.

[28] Our findings would be strengthened if we excluded hybrid retailers from our comparisons of price levels.

[29] The source of this quotation is an email message sent to one of the authors requesting advice on whether the bookstore should respond by setting up its own website.

[30] "Bookstore closing is same old story—with a twist" Reuters, December 29, 1998.

Table 6 Descriptive Data on Price Changes

	Books		CDs	
Statistic	Conventional Prices	Internet Prices	Conventional Prices	Internet Prices
Range	\$0.35–\$8.00	\$0.05–\$7.50	\$1.00–\$7.00	\$0.01–\$10.00
Mean Value	\$2.37	\$1.84	\$2.98	\$1.47

number of price changes between Internet and conventional channels (Table 6). The first and second rows in Table 6 present the range and mean value of observed price changes.

We note that these data support the hypothesis of lower menu costs in Internet channels in the sense that Internet retailers are willing to make smaller price changes than conventional retailers. As noted in the first row of Table 6, the smallest Internet price change we observed, for books, is \$0.05 compared to \$0.35 for conventional prices. For CDs, the smallest price change observed on the Internet is \$0.01, compared to among conventional retailers \$1.00.

A second way to view these data is through a histogram showing the number of price changes in Internet and conventional outlets. Figures A.1 and A.2 in the Appendix show histograms for our data using \$0.05 bins. These histograms demonstrate that, far from being an isolated incident, retailers on the Internet regularly make price changes that are smaller than the smallest price change we observe in conventional stores. For books, 25 of the Internet price changes (13% of the observed Internet price changes) were smaller than \$0.35 (the smallest price change observed among conventional outlets). Similarly for CDs, 90 Internet price changes (28%) were smaller than \$1.00.

A third way to view the data is through a cumulative distribution function showing the probability of observing a price change smaller than a particular level within each channel. Figures A.3 and A.4 show cumulative distribution plots of price changes for books and CDs respectively. For both products, the cumulative distribution function is larger for the Internet channel than the physical channel for all price changes up to \$6.00.

A fourth way to view the data by calculating the standard deviation of price changes between channels (Table 7). This calculation takes into account all price changes: positive, negative, and zero. Here again we find that the standard deviation of price changes on the Internet is lower than the standard deviation of price changes in conventional outlets for both books and CDs.

There are a variety of ways to evaluate the statistical significance of these statistics under the null hypothesis that the price changes are the same. A t-test on the size of price changes rejects the null, for both books and CDs, in favor of an alternative hypothesis that price changes on the Internet are smaller than price changes in conventional outlets ($p < 0.001$). One can also examine whether the observed number of small price changes in each channel could have arisen by chance. Under the null hypothesis that Internet and conventional price changes are drawn from the same underlying distribution, we would expect our distributions to follow a binomial distribution with probabilities corresponding to those observed in the combined set of price changes. Tests under this assumption reject the null, for both books and CDs, in favor of the stated alternative hypothesis whether one considers all price changes smaller than \$1.00 or price changes smaller than the smallest amount observed in conventional outlets ($p < 0.001$).[31]

3.3. Price Dispersion

> Price dispersion is a manifestation—and indeed it is the measure—of ignorance in the market.
>
> George Stigler (1961, p. 214)

In the classic Bertrand model of price competition,

[31] As with the previous Bernoulli distribution tests, these tests assume independence across observations. Controlling for serial correlation (by removing observations that have the same magnitude but opposite sign in consecutive time periods) does not change our results or the reported p-value.

Table 7 Standard Deviation of Price Changes

	Internet	Conventional
Books	0.7539	0.9792
CDs	0.7374	1.2356

products are perfectly homogeneous, retailers are afforded no spatial advantages in attracting consumers, and consumers are informed of all prices. The result is that competition occurs only in price, consumers buy from the lowest priced retailer and retailers all set the same price—a price equal to marginal cost. This result reflects the "law of one price" commonly taught in basic economics texts (e.g., Mankiw 1998, Chapter 7). Unfortunately, real-world markets rarely operate so smoothly. In fact, the existence of price dispersion is one of economics' most replicated findings (see Pratt et al. 1979, Dahlby and West 1986, or Sorensen 2000, for example).

Not surprisingly, price dispersion is usually explained as a violation of one of the Bertrand assumptions: product homogeneity, zero search costs, or perfectly informed consumers. Early work analyzed the role of product heterogeneity in explaining price dispersion (e.g., Griliches 1961, Chow 1967). This research views products as bundles of characteristics, with price dispersion arising from the presence or absence of characteristics in a particular product.

Price dispersion arising from differences in search cost has been analyzed by a variety of researchers. Burdett and Judd (1983) and Stahl (1989, 1996) model the role of search cost in explaining price dispersion in a setting where consumers engage in a costly search for price quotes. As noted above, Bakos (1997) analyzes the role of electronically mediated markets in lowering search costs, and finds that lower search costs should lead to lower and more homogeneous prices. Empirically, search costs have been found to explain price dispersion in auto insurance markets (Dahlby and West 1986) and, more recently, in prescription drug markets (Sorensen 2000).

Salop and Stiglitz (1977, 1982) and Varian (1980) have analyzed price dispersion arising from consumers who are differentially informed of prices. In these models some consumers are informed of all prices and other consumers know only one price (and do not look for other prices). The informed consumers purchase from the retailer with the lowest price; the uninformed consumers purchase if the price they are aware of is lower than their reservation value. The typical result is that some stores charge low prices in an attempt to attract informed consumers while other stores charge high prices to sell to uninformed consumers. In related empirical work, Van Hoomissen (1988) examines price data during periods of high inflation and finds that "price dispersion is strongly influenced by the presence of differentially informed consumers." (p. 1304)

Considering these factors alone, we expect to see only a small degree of price dispersion on the Internet. With regard to product heterogeneity, we have intentionally selected products—books and CDs—whose physical characteristics are *entirely* homogeneous. Considering search costs, as noted above, we expect lower search costs on the Internet than in conventional channels. Similarly, we expect the role of informed and uninformed consumers to be less of a factor in dispersion among Internet prices than it is among conventional prices.

Because the Internet is a multifaceted market, it is worth looking at the question of dispersion from a variety of perspectives. We first analyze several aspects of price dispersion by looking at posted prices. We then repeat analysis after weighting all the price observations by a proxy for market share in each channel. Each of these measures highlight different aspects of Internet commerce, and are both useful in characterizing Internet markets. Dispersion in posted prices corresponds to the price differences consumers would find if they were equally likely to observe prices from any store, e.g., after using a price comparison intermediary or some other listing of retailers, or if they searched among all the retailers in our sample without revisiting the same store repeatedly. Dispersion in weighted prices corresponds roughly to the prices one would observe by recording the prices consumers actually pay for goods on the Internet and in conventional outlets. We analyze dispersion resulting from posted and weighted prices below.

We create both absolute and relative measures to analyze dispersion in posted prices. Both measures reveal higher than expected dispersion in Internet prices. Absolute dispersion statistics for our data show a substantial range of prices available on the Internet for the same book or CD in the same time period. The range of prices across our eight Internet retailers (i.e., the largest price minus the smallest price for a partic-

Table 8 Proportion of Times Price Dispersion Is Larger on the Internet

Channel	Range	Trimmed Range	Standard Deviation
Books	87%	73%	84%
CDs	37%	43%	33%

ular title within a particular time period) averages \$5.98 for books and \$4.45 for CDs.[32] This corresponds to an average price range of 33% for books and 25% for CDs. Moreover, prices for some book and CD titles in our survey differ by as much as 47% across Internet retailers.[33]

We consider two tests of relative dispersion in posted prices across channels. First, following Sorensen (2000), we compare measures of price range, trimmed range, and standard deviation between the Internet and conventional channels. Second, we use kernel density plots to graphically analyze the dispersion of prices around their mean.

In our first set of tests, price range is defined as the difference between the largest and smallest price charged by our eight Internet and eight conventional retailers for a given title and date. Similarly, trimmed range is the difference between the second largest and second smallest prices charged by our retailers and is used to control for possible outliers in the data. Standard deviation, our third measure of dispersion, is calculated using the standard formula. Using these measures, for each date we count the number of titles (out of 20) where a particular measure of dispersion is larger on the Internet.

Table 8 summarizes our results for posted prices including tax and shipping and handling charges. The table presents the proportion of times our measures of price range, trimmed price range, and standard deviation are larger for the Internet prices in our study than for conventional prices. For books, we find that measures of range, trimmed range, and standard deviation are, in general, larger for Internet prices than for conventional prices. For example, measured across all time periods, the range of prices on the Internet is larger than the range in conventional outlets 87% of the time.

Within time periods, one can test for differences in price dispersion under the null hypothesis that the dispersion statistics are drawn from the same underlying distribution. As above, to control for potential serial correlation in the data, we conduct tests for each test statistic and for each month. Using this test, for 8 out of 12 months, we reject the null hypothesis of equal dispersion in favor of an alternative hypothesis that dispersion is larger among the Internet prices we tracked ($p < 0.01$) for each of the three dispersion statistics employed. In the remaining 4 months, for at least one dispersion measure, we fail to reject the null hypothesis of equal distributions.

For CDs, our data suggest that dispersion on the Internet is approximately equal to, or slightly smaller than, dispersion among conventional prices. Employing the same hypothesis test as above, in 7 of the 12 monthly observations, we fail to reject the null hypothesis of equal distributions for all three dispersion statistics. In the remaining 5 months, we reject the null hypothesis of equal dispersion in favor of less dispersion in Internet prices for at least one of the dispersion measures.

We can view these results graphically by plotting the full set of prices and observing how far these prices are from their means. To examine our data in this way, we calculate separate Internet and conventional mean prices for each date/title observation.[34] We then de-mean our data using the appropriate mean price. The result measures the differences in mean prices across channels after subtracting title and date effects.

[32] All dispersion measures are quoted for prices including shipping and handling and taxes. Similar results are obtained using store prices before adding tax and shipping charges.

[33] Furthermore, even the highest priced Internet retailers appear to make positive sales. Powell's Books is the most expensive Internet retailer in our sample—with average prices \$3.75 higher than prices at Amazon.com. Yet Powell's reports significant sales volume over the Internet (see, for example, "Powell's Books Pushes Back Its Boundaries, Both Physical and Virtual," *Publishers Weekly*, August 10, 1998, p. 243.)

[34] For example, for our February prices, we calculate 40 different mean prices (20 titles × 2 separate means for Internet and physical channels). We perform similar calculations for the other 8 monthly observations.

We plot our de-meaned data using a kernel density plot.[35] Our kernel density results, using an Epanechnikov kernel with optimal weights, are plotted in Figures A.5 and A.6 in the Appendix for books and CDs, respectively. We find that our graphical results are consistent with the descriptive results above. The graph for de-meaned book prices shows more weight in the tails of the plot for Internet prices compared to conventional prices. This suggests, again, that Internet prices have slightly more dispersion than comparable conventional prices. For CDs, neither channel seems to have obviously more dispersion.[36]

As noted above, it is also possible to measure the dispersion in share-weighted prices.[37] To do this we weight the posted prices by a proxy for market share in each channel. For Internet retailers, the weights are based on the screen share measure described earlier. For conventional retailers, the weighting is based on market share data provided by the Market Share Reporter (CDs) and the American Booksellers Association (books) for national chains, and estimates of market share in local stores. While these measures are only estimates of the true market share in a particular geographical market, our results are not particularly sensitive to changing assumptions of the conventional retailer's market share. In the extreme case, that of leaving the conventional outlets' prices unweighted, would lead to the same results quoted here.

After weighting the prices, we de-mean the data and create a kernel density plot as described previously. The resulting plots are presented in Figures A.7 and A.8. We note that these plots suggest that there is *less* dispersion in weighted prices on the Internet than in conventional outlets for both books and CDs.[38]

Taken at face value, this finding is consistent with a more efficient Internet market. However, the result should be interpreted with some caution for several reasons. First, we note that the results reflect the dominance of a few heavily branded retailers in each market. For example, Amazon.com has over 80% of the Internet market for books according to our screen share data and available market share data. By comparison, with only a 25.4% share, Barnes & Noble has the highest market share of any of the conventional retailers we track. Furthermore, the concentration around the market leader's price is enhanced by the fact that the next largest retailers in each product category seem to set prices very close to the market leader's price. In the Internet book market, the second and third most popular book retailers—Borders.com and Barnesandnoble.com—set prices that are nearly identical to prices at Amazon.com. Over all time periods, prices at these three retailers are on average −$0.19 and $0.09 different than Amazon.com's price. Looking at individual monthly observations, we find that the retailers' prices are typically within $0.01 of each other. In May 1999, for example, these retailers' prices were within $0.01 of each other for 16 of the 20 books we tracked. Similar results hold for CDs, where Amazon.com and CDnow (the two market leaders) set prices within $0.18 of each other on average.

Another reason our weighted dispersion tests may not necessarily suggest fully efficient markets is that several firms set prices well below the market leaders without getting significant market share. For example, during our sample period, Books.com consistently set prices that averaged $1.60 below Amazon.com's—but even with these low prices Books.com received relatively few sales (2.2% screen share). Indeed, the higher

[35] Kernel density plots can be thought of as a continuous version of a histogram. The kernel density plot attempts to estimate the "true" distribution $f(x)$ from a discrete set of observed x values. Like a histogram, it does this by selecting "bins" along the x-axis. However, unlike a histogram, kernel density bins can overlap—providing for smoothing of the function. The kernel density algorithm estimates the density of observed x-values within the bin, placing more weight on observations closer to the bin's center point.

[36] One may note that these graphs seem to correspond roughly to normal distributions. In this spirit we conducted F-tests on the de-meaned data displayed in these graphs. The F-test results are the same as a graphical intuition. For books we reject the null hypothesis that the standard deviations of prices are the same in favor of an alternative hypothesis that the standard deviation of prices is larger on the Internet. For CDs, an F-test fails to reject the null hypothesis that the standard deviation of Internet and conventional prices are the same.

[37] We thank Haim Mendelson for suggesting this approach.

[38] A parametric test finds that the Internet distribution has a smaller standard deviation than the conventional distribution in both markets ($p < 0.01$). This test is suggestive, not conclusive, because our conventional retailer's dispersion data is potentially biased to be too large, as noted above.

heterogeneity in market shares on the Internet may be evidence of important unobserved differences across retailers. Ironically, as we discuss below, a high market share may, in and of itself, be an important "feature" that supports a price premium because of the importance of network externalities in word-of-mouth marketing and trust-building.

Taken as a whole, the price dispersion results suggest that factors other than price influence the behavior in Internet markets. In the next section, we provide several potential explanations for the price dispersion observed in Internet markets.

3.4. Analysis of Internet Price Dispersion

As noted above, the economics literature has traditionally focused on three explanations for price dispersion in markets: search costs, asymmetrically informed consumers, and product heterogeneity. In this section, as a starting point for future empirical and theoretical research on this topic, we revisit each of these three factors to reassess how they may explain the observed price dispersion on the Internet.

Asymmetrically Informed Consumers and Search Costs. Models of asymmetrically informed consumers (e.g., Salop and Stiglitz 1977) and models of search cost (e.g., Salop 1979) are related in that both posit that (1) consumers are imperfectly informed, (2) information is costly to obtain, and (3) retailers set prices to leverage consumer heterogeneity in information and search cost. While it is plausible that these factors play a role on the Internet, a straightforward application of these models to our data does not adequately explain the variance we observe.

For example, in the models of Salop and Stiglitz (1977) and Varian (1980), informed consumers purchase from the lowest priced store(s) while purchases from uninformed consumers are evenly distributed among stores. A direct prediction of these models then is the retailer with the lowest prices should have the highest proportion of sales since it will get sales from all the informed consumers in addition to its "share" of the uninformed consumers.

However, this prediction is not supported by our data. Amazon.com is the undisputed leader in online book sales, and yet is far from the leader in having lower prices. Three of the eight online book retailers in our study have lower prices, on average, than Amazon.com; and the lowest priced retailer, Books.com, has prices that average $1.60 less than Amazon's prices. What's more, Books.com's price advantage was consistent across books and across time. Books.com's prices were lower than Amazon's price 99% of the time and all but two of the books tracked were cheaper at Books.com in each and every time period.

A similar situation holds for CDs. Amazon.com's CD sales surpassed the sales of the previous sales leader, CDnow, only three months after Amazon entered the market.[39] Yet both Amazon.com and CDnow have prices that average approximately $1.00 more than prices at Newbury Comics and CD Universe, the CD retailers with the lowest prices in our study.

Similarly, current models of search costs do not adequately explain the price dispersion in our data. For example, Bakos' (1997) model of search costs predicts that lower search costs in electronic markets should lead to lower price dispersion on the Internet than in comparable conventional outlets. As noted above, this prediction is not supported by our data for several measures of dispersion in posted prices.

Heterogeneity in consumer information and nonzero search costs may explain some of the price dispersion in our data, however. For example, it may be that consumers are aware of only one or two Internet retailers for books and CDs and find that the cost of searching for additional retailers exceeds the expected price benefit. This may explain why well-known retailers like Amazon.com and CDnow can charge price premiums compared to lesser known rivals such as Books.com or CD Universe. However, it does not explain why retailers like Powell's Books and Tower Records can successfully charge prices greater than better-known stores such as Amazon.com and CDnow. To examine this effect, we explore sources of product and retailer heterogeneity.

Product and Retailer Heterogeneity. Early work analyzed the role of product heterogeneity in explaining price dispersion (e.g., Griliches 1961, Chow 1967).

[39] "Amazon.com: The Wild World of Ecommerce," *Business Week*, December 14, 1998, p. 108.

This research views products as bundles of characteristics, with price dispersion arising from the presence or absence of characteristics in a particular product. In this context, it is important to note that the physical products in our sample were selected to be a matched set of identical commodities, in an attempt to eliminate any effects of product heterogeneity a priori. Nonetheless, one may also wish to consider the relevant "product" to be more than its purely physical characteristics. For example, retailers that provide additional services may be able to charge a price premium for the corresponding products they sell, and heterogeneity in the services offered by retailers may explain some of the price dispersion we observe.

However, there are two problems in using heterogeneity in service characteristics to explain the majority of price dispersion in our data. First, we find that the observed services offered by our Internet retailers (return policies, title reviews, artificial-intelligence-based suggestion tools, and audio clips for CDs) either do not vary significantly across retailers or are *negatively* correlated with price. For example, all but three of our CD retailers allow customers to return unopened CDs—and two of the retailers that do *not* provide this service are among the highest priced retailers in our survey. Similarly for books, the four retailers that offer suggestion tools have the four lowest prices in our survey. The lack of variance in services and the frequent negative correlation with prices explain why standard hedonic analysis of our data is unproductive. Regressions of price onto service characteristics yield shadow prices for services that are frequently negative and whose magnitude is often substantially larger than would be expected.

A second problem with using product heterogeneity to explain the differences in Internet prices is many of the service characteristics (suggestion tools, customer reviews, and the availability of audio clips) are purely informational in nature and are thus separable from the product. There is no technical or legal reason why a shopper could not go to one site for help in selecting a product and then simply click to another site to buy it. Apart from another explanation (e.g., search costs, externalities, or switching costs), providing better information about the characteristics of homogeneous products should not provide retailers with a strategic advantage over their rivals.

However, our data suggest that heterogeneity in other (unobserved) retailer characteristics may account for some of the observed dispersion in Internet prices. If heterogeneity in unobserved retailer characteristics accounted for the observed dispersion, we would expect that these characteristics would not vary significantly across titles. This would suggest that the price rank of retailers should be relatively consistent across titles—a prediction supported by our data. For example, in October, Books.com had the lowest price for 16 of the 20 titles tracked, and Powell's had the highest price for 17 of the 20 titles tracked ($p < 0.01$). Similarly, for CDs, either Newbury or CD Universe had the lowest price for 15 of the 20 titles and either Blockbuster or Tower had the highest prices for 18 of the 20 titles ($p < 0.01$).[40] The large differences in market shares across retailers are also consistent with important perceived differences in retailer characteristics.

While there are a variety of potential unobserved retailer characteristics, one promising candidate is heterogeneity in the "trust" consumers have for the various Internet retailers and the associated value of branding. Recent scholars have argued that trust is among the most important components of any effective Internet marketing program (see, e.g., Urban 1998). Indeed, we note that the importance of trust may arise directly from the characteristics of the Internet. Specifically, while the importance of factors such as search costs may be reduced on the Internet, factors such as trust may play an enhanced role because of the spatial and temporal separation between buyer, seller, and product on the Internet. Most consumers have little history or physical contact with Internet retailers and they must be wary of a falling prey to a site that posts low prices but is proficient only in charging credit cards, not delivering the goods.

If trust does provide a source of strategic advantage for Internet retailers, we would expect this to be

[40] The ranks of other retailers and ranks in other time periods also show consistency in retailer price ranks across titles and to an extent time. The test statistics are versus the null hypothesis that there are no firm-specific effects.

reflected in their actions. With this in mind, we review several different ways Internet retailers may be able to signal trust.

One obvious way Internet retailers can signal trust is by developing a reputation among customers for reliability. Thus, consumers may be willing to pay a slightly higher price to retailers they have dealt with before compared to retailers with which they have no prior relationship. This factor may explain why Internet retailers who also have conventional outlets are able to command a price premium from loyal physical-world consumers.[41] Thus, a high market share may be, to a degree, self-perpetuating.

Similarly, trust may be conveyed by word of mouth. Word of mouth may take a variety of forms in an electronic communication medium. For example, the presence of many active participants in an Internet retailer's online community may communicate that the retailer has many satisfied customers. Trusted individuals who operate "partner" or "associate" bookstores for retailers may also implicitly vouch for a retailer's trustworthiness.[42] We also note that there may be important network externalities to the conveyance of trust through word of mouth—more customers create a stronger signal of trust and strong signals of trust may lead to more customers.

Trust may also be signaled by the retailer through advertising (e.g., Milgrom and Roberts 1986, Wernerfelt 1988) or indirectly by being mentioned (favorably) in the popular press. This may explain the large advertising spending conducted by Internet retailers. Amazon.com, for example, advertises very heavily, with projected spending of over $200 million in 1999—amounting to $29 per customer or 24% of total revenue. In comparison, Barnes & Noble spends 4% of its revenues on marketing for its conventional stores. We also note that a Lexis-Nexis search of all news articles in 1998 shows that Amazon.com was mentioned in twice as many articles as Barnesandnoble.com and in 50 times as many articles as Books.com.

Finally, prominent links from other trusted sites on the web may vouch for a retailer's trustworthiness. For example, consumers who trust the *New York Times* book section may be more inclined to trust Barnesandnoble.com's service because the *New York Times* was willing (albeit for a fee) to allow Barnesandnoble.com to have a link on the *New York Times* Internet site.

Thus, while we agree with George Stigler (1961, p. 214) that "it would be metaphysical, and fruitless, to assert that all dispersion is due to heterogeneity," retailer heterogeneity in factors such as trust may offer a fruitful starting point for future research into sources of price advantage on the Internet. Indeed, one of the ironies suggested by our data is that, far from being a great equalizer of retailers and eliminating the need for branding as is often claimed,[43] the Internet may heighten the importance of differences among retailers in dimensions such as trust and branding.

4. Conclusions

It has been widely speculated that electronically mediated markets will have less friction than comparable conventional markets. We examine this hypothesis empirically using the prices charged by Internet and conventional retailers for homogeneous products—books and CDs.

Our analysis indicates that Internet retailers charge lower prices than conventional retailers—whether one considers prices alone or "prices" including the costs of getting the item to the users' homes (e.g., shipping and handling, taxes, mileage). It is easy to predict that

[41] For example, our data show that Internet retailers who also have conventional outlets are able to charge price premiums of 8.7% for books and 8.6% for CDs compared to retailers who only operate on the Internet.

[42] "Partner" programs (also referred to as "associate" and "affiliate" programs) give "partners" commissions for product sales generated through their site. In a typical program, the partner sets up a site advertising a particular book or books. This site contains a link to the bookstore sponsoring the partner program (Amazon.com for example). The partner gets a commission (typically 5% to 15%) on any sales generated at Amazon.com through her site. See http://www.bookwurms.com/ for an example of an Amazon.com "associate" bookstore.

[43] For example, Jim Borland of the Knight Ridder news service observes "[t]he Internet is a great equalizer, allowing the smallest of businesses to access markets and have a presence that allows them to compete against the giants of their industry." ("Move Over Megamalls, Cyberspace Is The Great Retailing Equalizer" April 13, 1998.)

a substantially larger fraction of the U.S. (and world) population will gain access to the Internet in the next decade. Thus, an implication of our findings is that conventional retailers will find it increasingly difficult to compete on price so long as the substantial differences between channels persist.

We also find that Internet retailers make price changes in smaller increments than comparable conventional retailers. Smaller menu costs may allow Internet retailers to more efficiently adjust their prices to structural changes in the market. We also note that if interfirm competition is stronger on the Internet, Internet retailers may have a higher incentive to make small price changes than conventional retailers.

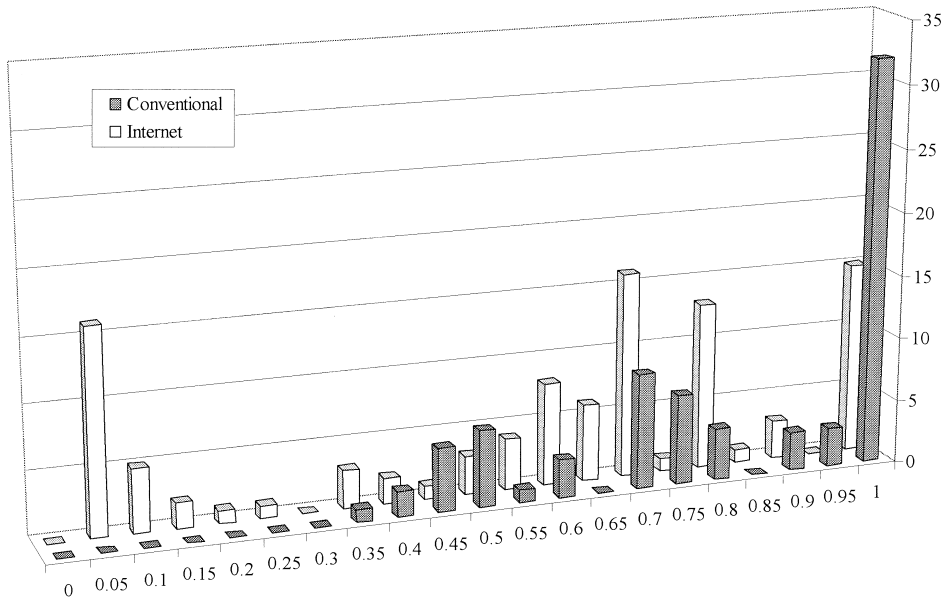
Lastly, we find that the level of price dispersion on the Internet depends importantly on the measures employed. In light of both existing theory and the earlier results on price levels and price changes, the dispersion in posted prices is surprisingly high. Posted prices vary by as much as 47% across Internet retailers. Furthermore, the retailers with the lowest prices do not make the most sales. At the same time dispersion in weighted prices is lower on the Internet than in conventional outlets—reflecting a dominance among certain heavily branded retailers.

Given these findings, we analyze potential sources for the high degree of price dispersion on the Internet. We conclude the Internet price dispersion may arise from two different sources of retailer heterogeneity: heterogeneity in customer awareness, and heterogeneity in retailer branding and trust. We also note that, far from being equalized, these differences among sellers may be amplified on the Internet as compared to conventional channels. The question of whether these differences are a symptom of an immature market or reflect more permanent characteristics of Internet retailing may provide a fruitful starting point for future research into sources of price advantage on the Internet.[44]

[44] This research has benefited from valuable comments by Joseph Bailey, Yannis Bakos, Rajiv Banker, Steven Berry, Il-Horn Hann, Lorin Hitt, Haim Mendelson, Alan Sorensen, two anonymous referees, and seminar participants at MIT, Harvard, the National Bureau of Economic Research, and the 1998 Workshop on Information Systems and Economics. We thank Stephane Bressan, Thomas Cheng, Mikey Fradin, Min Huang, Wallace Newton, Alice Ng, Matt Norvell, Virginia O'Connell, Glorimar Ripoli, Julie Shin, Kristin Smith, and Lou Smith and MIT's Context Interchange Project for data collection support and Web21 for providing valuable data. We gratefully acknowledge generous funding from NASA, the OECD, the MIT Center for Coordination Science, the Center for eBusiness@MIT under a grant from Fleet Bank, and the Stanford Computer Industry Project.

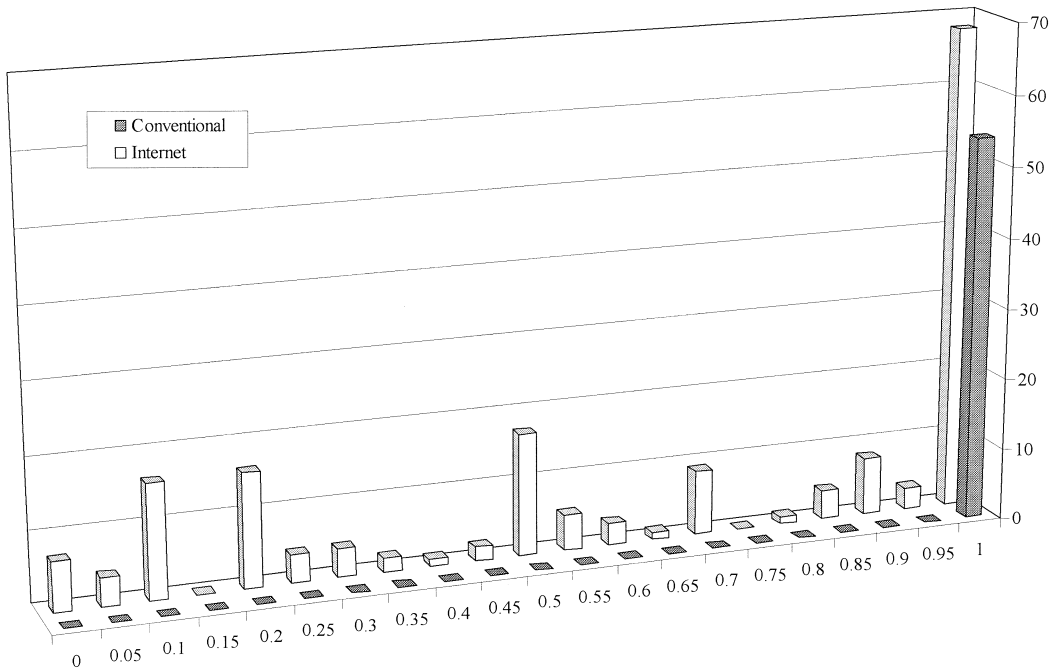
Appendix. Histograms and Kernel Density Plots for Price Changes and Price Variance

Figure A1 Histogram of Book Price Changes



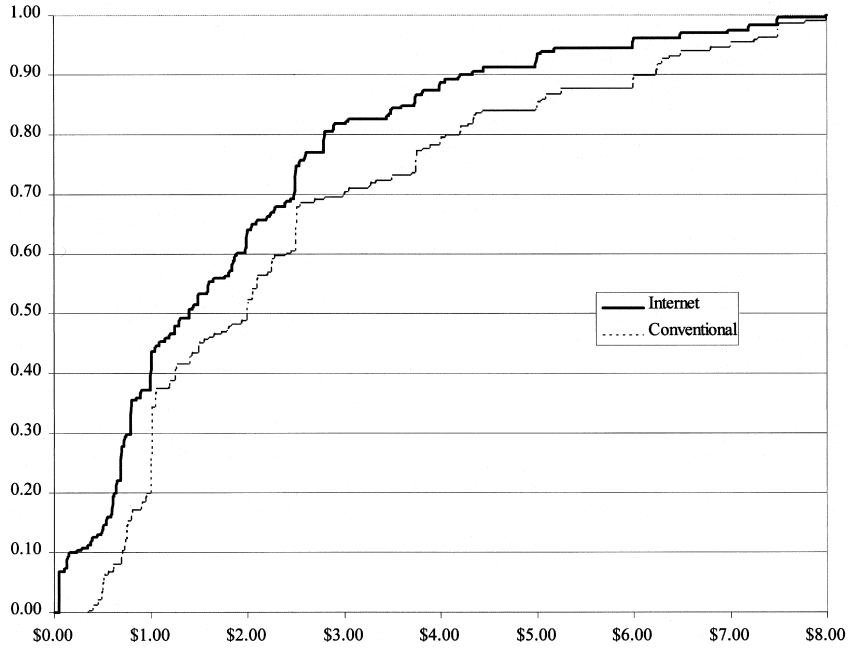
Note. Internet book retailers are substantially more likely to make small changes in their posted prices than are conventional book retailers. (The figure only displays price changes smaller than $1.00 in keeping with the focus on menus costs.)

Figure A2 Histogram of CD Price Changes



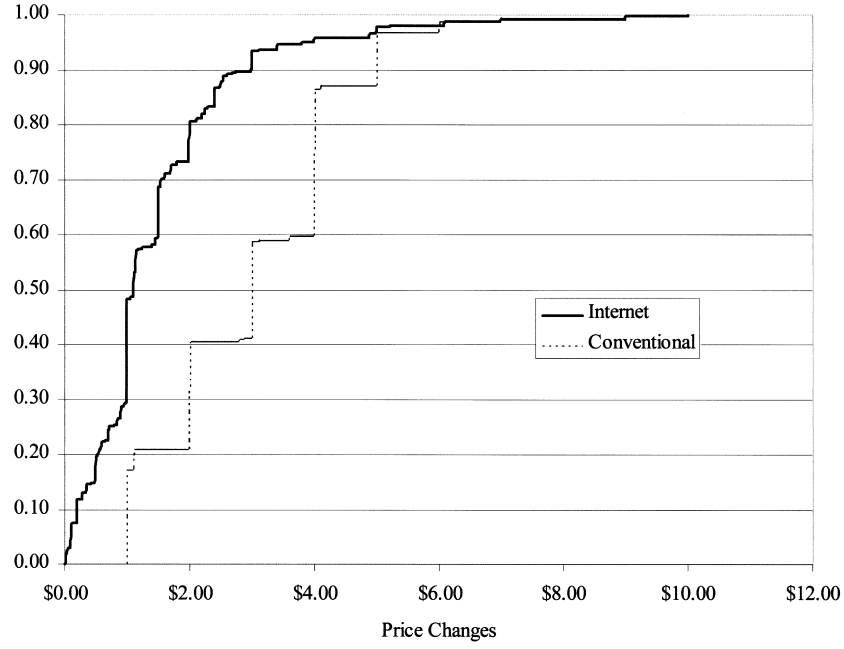
Note. Internet CD retailers are substantially more likely to make small changes in their posted prices than are conventional CD retailers. (The figure only displays price changes smaller than $1.00 in keeping with the focus on menus costs.)

Figure A3 Cumulative Distribution Function of Book Price Changes



Note. Internet book retailers are more likely to make small price changes over all observed price change values.

Figure A4 Cumulative Distribution Function of Book Price Changes



Note. Internet CD retailers are more likely to make small price changes over all observed price change values less than $6.00.

Figure A5 **Kernel Density for De-Meaned Full Prices for Books (Epanechnikov Kernel)**

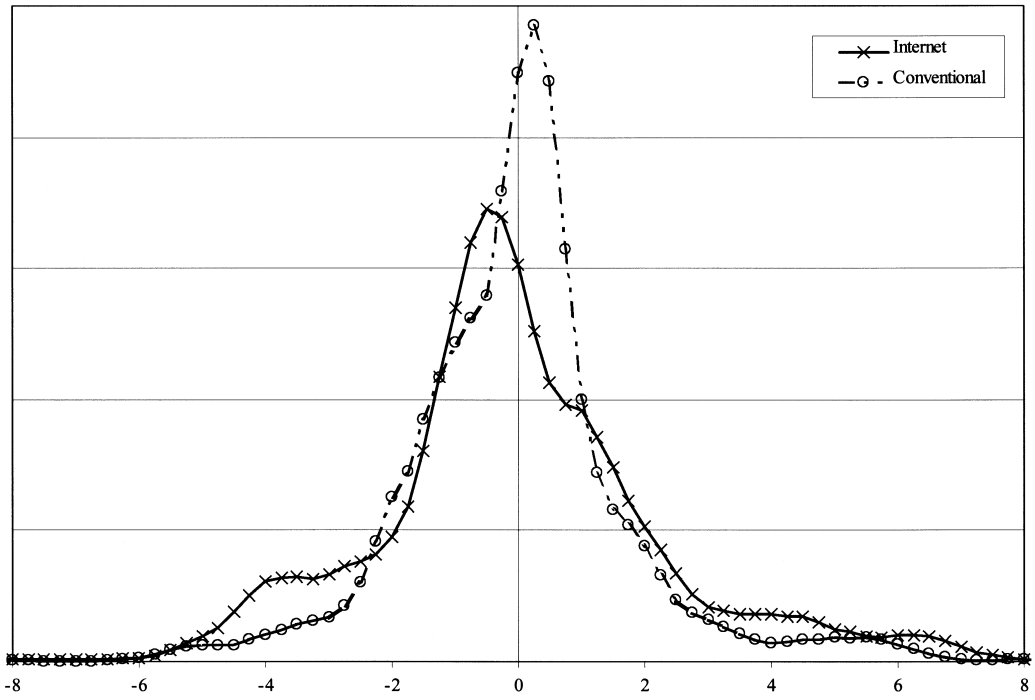


Figure A6 **Kernel Density for De-Meaned Full Prices for CDs (Epanechnikov Kernel)**

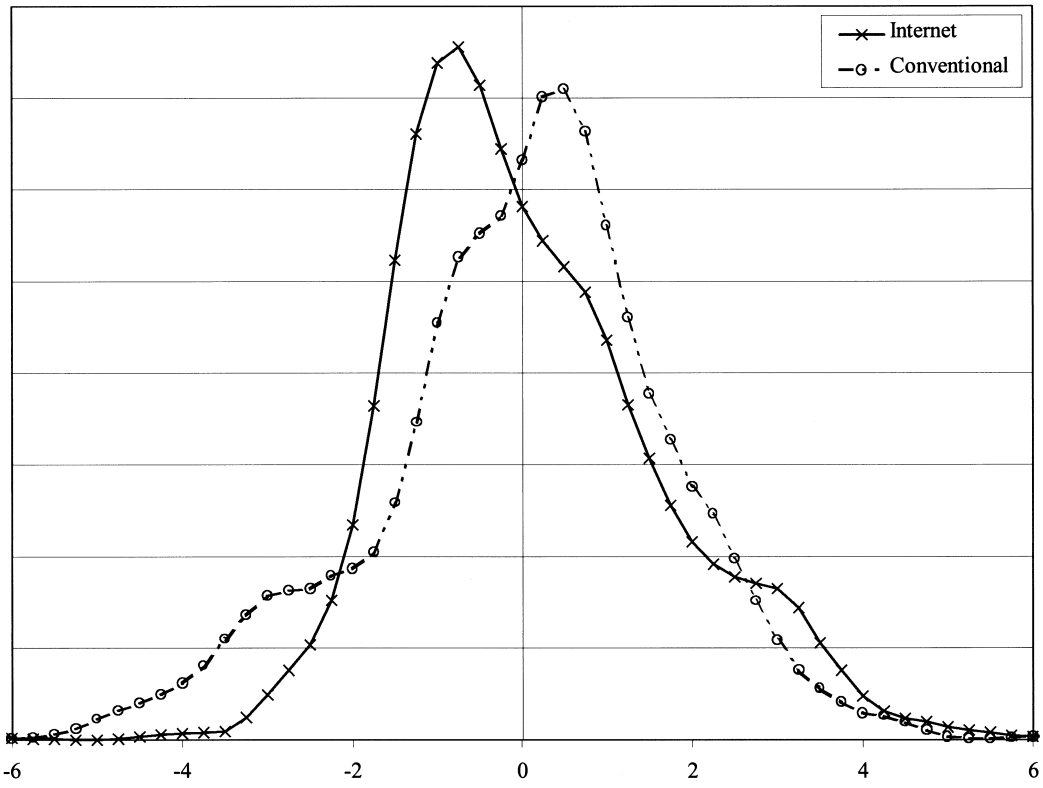


Figure A7 **Kernel Density for De-Meaned Full Prices for Books, Observations Weighted by Screen/Market Share (Epanechnikov Kernel)**

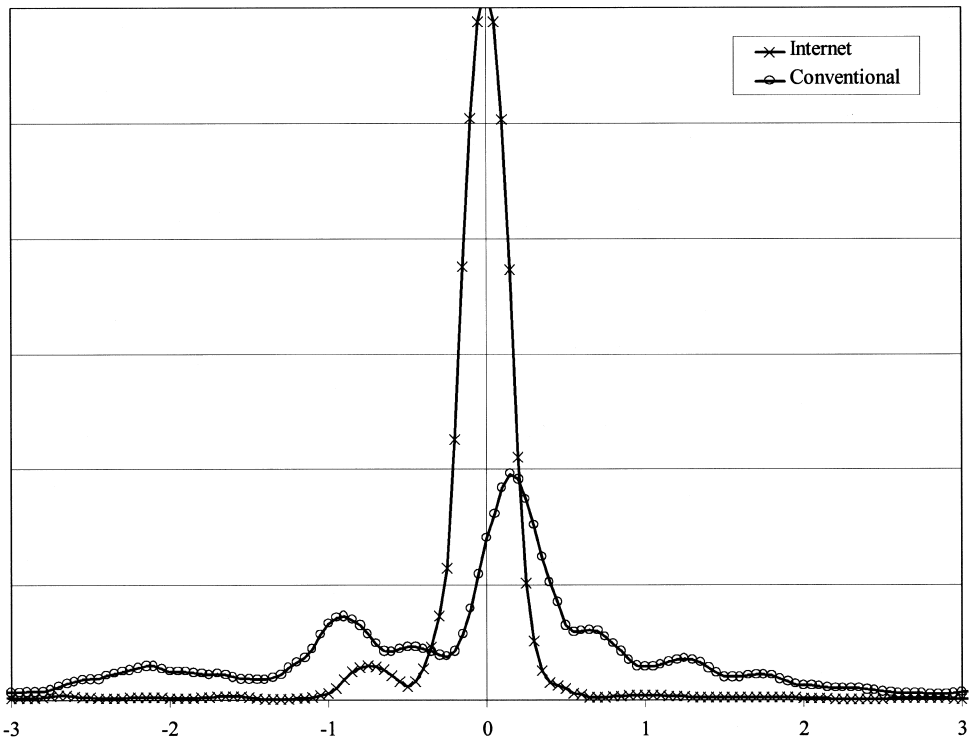
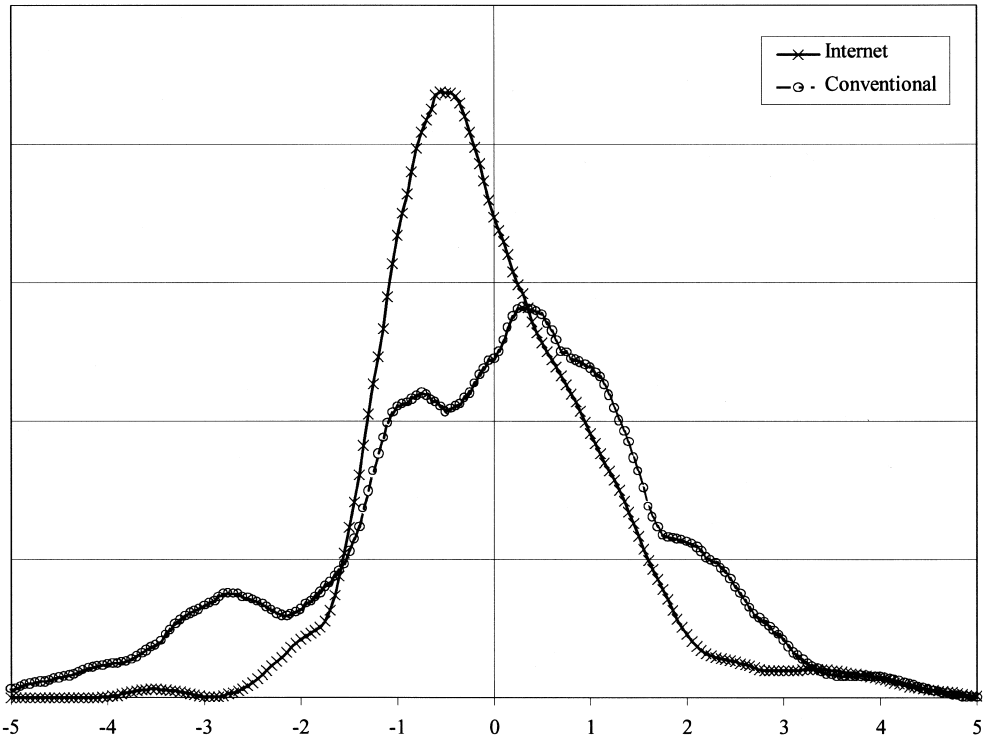


Figure A8 **Kernel Density for De-Meaned Full Prices for CDs, Observations Weighted by Screen/Market Share (Epanechnikov Kernel)**



References

Bailey, Joseph P. 1998. Intermediation and electronic markets: Aggregation and pricing in Internet commerce. Ph.D. Thesis, Technology, Management and Policy, Massachusetts Institute of Technology, Cambridge, MA.

Bakos, J. Yannis. 1997. Reducing buyer search costs: Implications for electronic marketplaces. *Management Sci.* **43** (12) 1613–1630.

——. 1998. The emerging role of electronic marketplaces on the Internet. *Comm. ACM* **41**(8) 35–42.

Burdett, Kenneth, Kenneth Judd. 1983. Equilibrium price dispersion. *Econometrica* **51**(4) 955–969.

Chow, Gregory C. 1967. Technological change and the demand for computers. *Amer. Econom. Rev.* **57**(5) 1117–1130.

Clemons, Eric K., Il-Horn Hann, Lorin M. Hitt. 1998. The nature of competition in electronic markets: An empirical investigation of online travel agent offerings. Working Paper, The Wharton School of the University of Pennsylvania, Philadelphia, PA.

Dahlby, Bev, Douglas S. West. 1986. Price dispersion in an automobile insurance market. *J. Political Econom.* **94**(2) 418–438.

Degeratu, Alexandru, Arvind Rangaswamy, Jeremy Wu. 1998. Consumer choice behavior in online and regular stores: The effects of brand name, price, and other search attributes. *Marketing Science and the Internet*, INFORMS College on Marketing Mini-Conference, Cambridge, MA.

Ghemawat, Pankaj, Bret Baird. 1998. Leadership online: Barnes & Noble vs. Amazon.com. Case N9-798-063, Harvard Business School, Boston, MA.

Goolsbee, Austan. 2000. In a world without borders: The impact of taxes on Internet commerce. *Quart. J. Econom.* Forthcoming.

Griliches, Zvi. 1961. Hedonic price indexes for automobiles: An econometric analysis of quality change. Griliches, ed. *Price Indexes and Quality Change*: *Studies in New Methods of Measurement*. Harvard University Press, Cambridge, MA, 55–87.

Hausman, Jerry. 1997. Valuation of new goods under perfect and imperfect competition, Bresnahan and Gordon, eds. *The Economics of New Goods*. The University of Chicago Press, Chicago, IL, 209–237.

Hicks, John R. 1940. The valuation of the social income. *Economica* **7** 105–124.

Hotelling, Harold. 1929. Stability in competition. *Econom. J.* **39** (153) 41–57.

Lee, Ho Geun. 1998. Do electronic marketplaces lower the price of goods. *Comm. ACM* **41** (1) 73–80.

Levy, Daniel, Mark Bergen, Shantanu Dutta, Robert Venable. 1997. The magnitude of menu costs: Direct evidence from large U.S. supermarket chains. *Quart. J. Econom.* **112** (3) 791–825.

Lynch, John G., Jr., Dan Ariely. 1998. Wine online: Search cost and competition on price, quality, and distribution. *Marketing Sci.* Forthcoming.

This paper was accepted by Haim Mendelson and Seungjin Whang.

Mankiw, N. Gregory. 1998. *Principles of Economics*. Dryden Press, Fort Worth, TX.

Milgrom, Paul, John Roberts. 1982. Limit pricing and entry under incomplete information. *Econometrica* **50** 443–460.

——, ——. 1986. Price and advertising signals of product quality. *J. Political Econom.* **94**(4) 796–821.

Pratt, John W., David A. Wise, Richard Zeckhauser. 1979. Price differences in almost competitive markets. *Quart. J. Econom.* **93**(2) 189–211.

Salop, S. 1979. Monopolistic competition with outside goods. *Bell J. Econom.* **10** 141–156.

——, J. E. Stiglitz. 1977. Bargains and ripoffs: A model of monopolistically competitive price dispersion. *Rev. Econom. Stud.* **44**(October) 493–510.

——, ——. 1982. The theory of sales: A simple model of equilibrium price dispersion with identical agents. *Amer. Econom. Rev.* **72**(5) 1121–1130.

Shankar, Venkatesh, Arvind Rangaswamy, Michael Pusateri. 1998. The impact of Internet marketing on price sensitivity and price competition. *Marketing Science and the Internet*, INFORM College on Marketing Mini-Conference. Cambridge, MA.

Sheshinski, Eytan, Yoram Weiss. 1993. *Optimal Pricing, Inflation, and the Cost of Price Adjustment*. MIT Press, Cambridge, MA.

Smith, Michael, Joseph Bailey, Erik Brynjolfsson. 2000. Understanding digital markets. E. Brynjolfsson, B. Kahin, eds. *Understanding the Digital Economy*. MIT Press, Cambridge, MA.

Snedecor, G. W., W. G. Cochran. 1989. *Statistical Methods*, 8th ed. Iowa State University Press, Ames, IA.

Sorensen, Alan T. 2000. Equilibrium price dispersion in retail market for prescription drugs. *J. Political Econom.* (August)

Stahl, Dale O. 1989. Oligopolistic pricing with sequential consumer search. *Amer. Econom. Rev.* **79**(4) 700–712.

——. 1996. Oligopolistic pricing with heterogeneous consumer search. *Internat. J. Indust. Organ.* **14**(2) 243–268.

Stigler, George. 1961. The economics of information. *J. Political Econom.* **69**(3) 213–225.

——, James K. Kindahl. 1970. *The Behavior of Industrial Prices*. Columbia University Press, New York.

Urban, Glen L., Fareena Sultan, William Qualls. 1998. Trust-based marketing on the Internet. Working Paper #4035-98, MIT Sloan School of Management, Cambridge, MA.

Van Hoomissen, Theresa. 1988. Price dispersion and inflation: Evidence from Israel. *J. Political Econom.* **96**(6) 1303–1314.

Varian, Hal R. 1980. A model of sales. *Amer. Econom. Rev.* **70**(4) 651–659.

Wernerfelt, Birger. 1988. Umbrella branding as a signal of new product quality: An example of signaling by posting a bond. *RAND J. Econom.* **19**(3) 458–466.