# On the Efficiency of Internet Markets for Consumer Goods

Brian T. Ratchford, Xing Pan, and Venkatesh Shankar

*The authors examine the consumer welfare implications of changes in the structure of e-commerce markets employing comprehensive data collected from BizRate.com in November 2000 and November 2001. They find that price dispersion decreased substantially between these two periods and that measured differences in e-tailer services bear little relation to e-tailer prices.*

The Internet is an electronic marketplace, which Bakos (1997, p. 1676) defines as an "interorganizational information system that allows the participating buyers and sellers ... to exchange information about prices and product offerings." Bakos argues that electronic marketplaces allow buyers easier access to price and other information, which thereby decreases search costs. In turn, lower search costs lead to increased competition and a better allocation of resources (Bakos 1997).

On the basis of this analysis, it might be expected that online competition would force prices charged by e-tailers to a uniformly low level and that consumers would benefit greatly from the presence of Internet markets. The reality of this, however, seems quite different on the surface. Brynjolfsson and Smith (2000) find that Internet prices are lower than conventional retailer prices for books and compact discs (CDs), but they still find substantial dispersion, comparable to that found in conventional markets, in the prices posted by different e-tailers. Using data from a broader sample of items collected in November 2000, Pan, Ratchford, and Shankar (2001) document that the range of prices posted by e-tailers for a given item often exceeds the average price of the item. Pan, Ratchford, and Shankar (2002) show that only a small proportion of this large degree of price dispersion is explained by differences in services provided by e-tailers. Clay, Krishnan, and Wolff (2001) find evidence of large and persistent price dispersion in the Internet market for books during the period from August 1999 to January 2000. Similarly, Baye, Morgan, and Scholten (2001) find evidence of large and persistent price dispersion in a sample of 1000 items during the period from August 2000 to March 2001. This evidence indicates that Internet markets are not as frictionless as Bakos (1997) and others would have predicted, which raises the following questions: (1) Does the Internet actually enhance allocative efficiency, thereby providing substantial benefits to consumers? (2) How can large degrees of dispersion in posted prices exist in a market that is supposedly frictionless or approximately so? and (3) Is the observed dispersion in offerings symptomatic of an immature market, and has that dispersion disappeared more recently as Internet markets have matured? The demise of many e-tailers in the economic downturn of 2001 might have enhanced the functioning of e-tail markets.

To shed light on these questions and provide guidance to our empirical analysis, we present a framework based on Ratchford and colleagues' (1996) work for measuring the relationship between consumer surplus and the dispersion of retail prices. To provide further guidance, we summarize what is known about pricing when search costs and imperfect information are present and what is known about the impact of services offered by e-tailers on prices when information is imperfect. We apply our theoretical analysis to interpret the comparisons between prices at the e-tailers and conventional retailers that Brynjolfsson and Smith (2000) present. We also analyze two comprehensive data sets on e-tailer prices and services that we collected mainly from BizRate.com in November 2000 and November 2001. We use these data sets to draw conclusions about changes in price dispersion and consequent changes in consumer welfare resulting from the maturation of e-tail markets between these two periods. We find that price dispersion decreased substantially between these two periods and that measured differences in e-tailer services bear little relation to e-tailer prices.

## Related Literature

Since the pioneering work of Alba and colleagues (1997) and Bakos (1997), there has been a great deal of theoretical and empirical work on the impact of e-commerce on markets and consumer welfare. Several theoretical models explore subtleties associated with the emergence of Internet shopping. Bakos (1997) postulates that the Internet enhances competition, but Lal and Sarvary (1999) and Lynch and Ariely (2000) show that this is not necessarily the case. Lal and Sarvary present a theoretical model in which Internet usage can decrease search by lowering the cost of buying a preferred item compared with searching in the store. Welfare, defined as the sum of consumer and pro-

BRIAN T. RATCHFORD is the PepsiCo Chair Professor of Marketing, XING PAN is a doctoral student, and VENKATESH SHANKAR is Ralph J. Tyser Fellow and Associate Professor of Marketing and Entrepreneurship, Robert H. Smith School of Business, University of Maryland, College Park. The authors gratefully acknowledge the helpful comments of the special issue editors and two anonymous *JPP&M* reviewers.

ducer surplus, increases, but sellers are able to appropriate the cost savings to consumers through higher prices. Lynch and Ariely present a simulated Web search environment for wine. They show that designing the Web environment to facilitate quality comparison decreases price sensitivity for unique items, giving them more monopoly power. Conversely, designing the Web environment to facilitate price comparison makes common items more price sensitive but lowers demand for unique items.

Some studies point out that the Internet lowers search costs for both price information and nonprice information, such as product and quality information (Degeratu, Rangaswamy, and Wu 2000; Shankar, Rangaswamy, and Pusateri 2001). Whereas lower search costs on price information may lead to lower prices, lower search costs on nonprice information could lead to lower price sensitivity and, consequently, to higher prices.

Ellison and Ellison (2001) and Baye and Morgan (2001) study the role of price search engines. Both articles note the following paradox, which Ellison and Ellison succinctly state (p. 4): "If search engines create Bertrand competition then there will be no price dispersion and consumers will be unwilling to pay for the information the search engine provides." Ellison and Ellison propose that the problem can be circumvented if firms adopt strategies to make searches more difficult either by obfuscating prices or by some other means. Baye and Morgan propose that the shopping service will solve the problem posed in the paradox when the services are priced in a way that preserves ex ante price dispersion. Although the shopping service benefits consumers, it will be costly to firms because it increases competition and, as Baye and Morgan show, will not always increase overall welfare.

There is evidence that information available on the Internet can substantially affect prices in conventional markets. Brown and Goolsbee (2002) estimate that the presence of Internet referral services led to a reduction of 8%–15% in term life insurance policies. They also find that these reductions are accompanied by an increase in price dispersion when the share using online referrals is low. Brown and Goolsbee point out that this is consistent with the search model of Stahl (1989). Morton, Zettelmeyer, and Silva-Risso (2001) find that Autobytel customers in California on average saved approximately $450 per car, and Ratchford, Lee, and Talukdar (2003) document that the Internet has become a major source of information about automobiles. Their estimates indicate that the Internet leads to both time savings and better buys. A few studies have compared prices and price dispersion at pure-play e-tailers and multichannel retailers (e.g., Ancarani and Shankar 2002; Pan, Shankar, and Ratchford 2002; Tang and Xing 2001). We discuss the results of those studies in the "Comparison of Conventional and Internet Retailers" section.

As noted previously, Brynjolfsson and Smith (2000), Pan, Ratchford, and Shankar (2001), Clay, Krishnan, and Wolff (2001), and Baye, Morgan, and Scholten (2001) find evidence of substantial and persistent price dispersion in Internet markets. Along the same lines, Clemons, Hann, and Hitt (2002) find evidence of substantial differences in the quality of online travel agent recommendations. Clay, Krishnan, and Wolff (2001), who document several strategies that online booksellers appear to follow, find that price dispersion and propensity to discount are much greater for *New York Times* best-sellers than for other books. Smith and Brynjolfsson (2001) show that there is a substantial amount of brand preference for established book retailers, which may help explain price dispersion in that market.

Although studies show persistent price dispersion, the most recent study we reviewed ends in March 2001, before the economic downturn may have had its full impact. In this study, with data collected in November 2001, we address whether price dispersion declined as economic conditions became less favorable for e-tailers.

## Consumer Surplus and Price Dispersion

In general, allocative efficiency, a standard measure of consumer welfare, can be defined as the sum of consumer and producer surplus, where consumer surplus is the difference between willingness to pay and price and producer surplus is the difference between price and cost. We concentrate first on the consumer surplus part of this equation, taking prices as given. Our objective is to develop measures of consumer surplus that can be operationalized and related to price dispersion present in a market at any time.

Consider that the consumer's problem is acquiring one unit of an item such as a book, CD, or personal computer, at retail.[1] To simplify the analysis, assume that the consumer values the item enough to make a purchase at any price offered. At the highest level, the consumer is faced with the decision of whether to buy this item in a conventional store or on the Internet at an e-tailer. After the consumer makes this decision, he or she searches among retailers or e-tailers for a good deal and, after putting some effort into searching, makes a choice. We consider the case in which the search effort optimizes the trade-off between a better buy and the search cost.

Assume that consumer i wishes to allocate his or her budget between the focal item m and another composite commodity, which represents all other goods, and that the utility function is separable between the focal item class and all other goods. If the price of the focal item at retailer r from channel j (bricks-and-mortar or Internet) is $p_{mrj}$, the amount available to spend on other goods can be expressed as $y_i - p_{mrj} - S_i - c_{imrj}$, where y is income, S is expenditure on search for the best retailer or e-tailer at which to buy the focal item, and $c_{imrj}$ is transaction costs of purchasing the focal item. Transaction costs may include travel, waiting at the checkout, entering credit card information online, or any other costs of completing the transaction. These costs are reduced by services provided by the e-tailer or retailer. The sum of $p_{mrj}$ and $c_{imrj}$ is the full price of the item (Ehrlich and Fisher 1982). Given these assumptions, the consumer's indirect utility function, conditional on the purchase of m, can be written as

$$(1) \qquad U_{imrj} = z_{im} + \alpha_i\left(y_i - p_{mrj} - S_i - c_{imrj}\right).$$

[1]Alternatively, the consumer could purchase a predetermined market basket of items.

where $z_{im}$ is the utility of the focal item to consumer i. Without loss of generality, we can rescale Equation 1 so that it is measured in monetary units by dividing through by $\alpha$. Define $V_{imrj} = U_{imrj}/\alpha_i$ and $\phi_{im} = z_{im}/\alpha_i$. In addition, noting that expenditure on search mainly comprises time spent on search, t, × a unit time cost, w, we can write $S_i = w_i t_i$. This leads to the following expression for the monetary value of the consumer's purchase of the focal item:

$$(2) \qquad V_{imrj} = \phi_{im} + \left(y_i - p_{mrj} - w_i t_i - c_{imrj}\right).$$

For simplicity, assume that the consumer specializes his or her search on the channel that provides the highest expected utility. Let $\pi_{irj}$ be the probability that the consumer buys at retailer r conditional on buying in channel j. This can be interpreted as the probability that a consumer who has stopped searching after a given amount of time will buy at that store. Then the expected consumer surplus (without considering search costs) from the purchase in that channel is (DePalma, Myers, and Papageorgiou 1994; Small and Rosen 1981)

$$(3) \qquad CS_{ij} = \sum_{r=1}^{n_j} \pi_{irj}\left(\phi_{im} - p_{mrj} - c_{imrj}\right)$$
$$= \phi_{im} - \sum_{r=1}^{n_j} \pi_{irj}\left(p_{mrj} + c_{imrj}\right),$$

where $n_j$ is the number of retailers in channel j. We would obtain consumer surplus net of search costs by subtracting search costs from Equation 3. There are two polar cases: (1) when the consumer does not search at all (or has no prior knowledge) and just buys at random, in which case $\pi_{irj} = (1/n_j)$, and (2) when the consumer searches until locating the best retailer offering the best buy, b, so that $\pi_{irj} = 1$ if r = b, 0 otherwise. If we abstract from search costs, the difference between the maximum consumer surplus, b, and consumer surplus from a random choice, a, is

$$(4) \qquad CS_{ij}^{b} - CS_{ij}^{a} = \left(\bar{p}_{mj} + \bar{c}_{imj}\right) - \left(p_{mbj} + c_{imbj}\right),$$

that is, the difference between the average full price and the lowest full price. Equation 4 provides a measure of the potential gains of searching for a consumer facing a given set of full prices. Equation 4 becomes greater as the amount of price dispersion in channel j increases.

The consumer searches in channel j to maximize expected utility, which is equivalent to optimizing the trade-off between expected consumer surplus and the cost of search. This trade-off is optimized when

$$(5) \qquad \frac{dCS_{ij}}{dt_{ij}} = -\sum_{r=1}^{n_j} \frac{d\pi_{irj}}{dt_{ij}}\left(p_{mrj} + c_{imrj}\right) = w_i,$$

where $t_{ij}$ is time spent by consumer i at search in channel j. If the time is spent wisely, the search leads to decreases in $d\pi_{irj}/dt_{ij}$ for stores with a high full price and to increases in $d\pi_{irj}/dt_{ij}$ for stores with a low full price, thereby increasing consumer surplus. Eventually, diminishing returns to time will set in (e.g., if the consumer finds the lowest full price, returns to spending more time become zero), and the consumer will stop searching.

The solution to Equation 5 leads to a set of values $\pi^*_{irj}$ and $t^*_{ij}$ for each channel. Suppose that the consumer solves Equation 4 for each channel before deciding in which channel to buy. If we denote bricks-and-mortar as retailer type B and the Internet as retailer type I, we can define the expected full price at which the consumer is indifferent between buying online and going to the bricks-and-mortar retailer as

$$(6) \qquad \sum_{r=1}^{n_I} \pi^*_{irI}\left(p_{mrI} + c_{imrI}\right) = \sum_{r=1}^{n_B} \pi^*_{irB}\left(p_{mrB} + c_{imrB}\right)$$
$$+ w_i\left(t^*_{iB} - t^*_{iI}\right),$$

where higher online full prices make buying from the bricks-and-mortar retailer a better deal, and lower online full prices make buying from the e-tailer a better deal.

We draw two basic conclusions from Equation 6. First, any saving in search time by using the Internet is more valuable to those with high time costs. Consumers with high search costs are willing to pay a premium over standard retailers and benefit from the Internet even if they must pay higher prices for the channel. These are also the consumers who search less, as is shown in Equation 5, and are more willing to accept a risk of paying a high price. Because of time savings, consumers can benefit from the Internet even though they may pay a relatively high price (Lal and Sarvary [1999] reach a similar conclusion in a somewhat different context). The benefits of time savings increase with search costs.

Thus, one explanation for large observed price dispersion on the Internet is that its timesaving properties make it valuable to a group of consumers with high time costs that is willing to accept high prices rather than incur additional search costs. In effect, the high prices are a payment to the e-tailer for the privilege of using the timesaving medium. At the same time, other Internet users who do not have such high time costs can afford to search and can search efficiently with the medium, which creates a demand for low prices. Price dispersion results from the differences in incentives to search.[2]

Second, the channel in which the consumer can search most efficiently, as indicated by values of $\pi^*_{irj}$ and $t^*_{ij}$ obtained from solving Equation 5, gains an advantage. Thus, consumers who are relatively efficient at using the Internet are willing to pay more to use this channel. However, these consumers are also more likely to be able to locate a relatively good buy using this medium.

A general conclusion we draw from this section is that if choice probabilities and full prices associated with any outlet can be computed for a given state of information, the consumer surplus expression presented in Equation 3 can be used to make statements about potential gains to information. Equation 4 presents an example of such a statement for potential gains when consumers go from no information to complete information. It is important to note that Equations 3 and 4 are quite general and do not depend on any assumption of optimizing behavior on the consumers' part.

[2]We pursue this more formally in the next section. Brynjolfsson and Smith (2000) make a similar argument but do not provide a formal demonstration of this point.

Although Equation 4 presents a measure of gains to search, we can also say something about the highest search cost in the market if we impose more structure. Consider a model in which consumers search sequentially among retailers (or products) for the lowest price, as in Carlson and McAfee's (1983) study.[3] Prior to search, consumers are assumed to know the general distribution of prices but not the exact price charged by any seller.[4] Consumers are assumed to use a stopping rule in which they search if the expected gain relative to the best price they have observed so far is greater than the cost of another search, but they stop otherwise. The stopping rule can be used to place a lower bound on the search cost of the consumer who buys the highest-priced item in the market. If a consumer encounters a price of $p_h$, where $p_h$ is the highest price, the expected gain from searching further is equal to $\Sigma_{r \neq h} (1/n_r)(p_h - p_r)$. If the consumer is content to buy at the highest-price rather than search further, search cost must exceed this expected gain. So if T is the highest search cost, we know that

$$(7) \qquad T > \sum_{r \neq h}\left(\frac{1}{n_r}\right)(p_h - p_r) = \left(\frac{n_r + 1}{n_r}\right)p_h - \bar{p},$$

where the overbar denotes mean and $n_r$ is the number of firms with a price lower than $p_h$. Thus, knowledge of the highest price and the average price in the market allows the conclusion that the highest search cost is at least T.

Carlson and McAfee (1983) show that if there is a uniform distribution of search costs between zero and the upper bound T, and if consumers apply the reservation price rule, a downward-sloping demand curve, in which quantity is a linear function of the difference between price and average price, results. If they choose randomly, consumers with the highest range of search costs will spread their purchases evenly among all stores. Consumers with the next highest range, however, will avoid the highest-priced offering but will spread their business evenly among all others. Therefore, the highest-priced retailer (e-tailer) receives only the business of consumers with the highest search costs. The next highest-priced seller receives the business of only the two groups with the highest search costs, and so on.

### Summary

The following general conclusions can be drawn from the analyses in this section:

1. Differences in consumer surplus gross of search costs between channels, periods, or consumers can be computed as a weighted sum of full prices, where the weights are choice probabilities.
2. The difference between the unweighted average price and the lowest price reveals the difference in expected surplus (gross of search costs) between a completely uninformed and a fully informed consumer. The difference between the highest price and the average price reveals a lower bound on the highest search cost in the market. Given their interpretability, these measures can be used in empirical analyses of price dispersion.
3. Because of the timesaving properties of the Internet, the Internet appeals to consumers with high time costs who are willing to pay a high price. Despite paying a high price, these consumers benefit from the time savings afforded by the Internet and will be better off because of its existence.

In the next section, we discuss alternative explanations for price dispersion and use this discussion to develop hypotheses about drivers of price levels and price dispersion and about how these variables might change as Internet markets mature.

## Market Behavior

### Search Costs As an Explanation of Market Behavior

There is a large body of literature (e.g., Bakos 1997; Burdett and Coles 1997; Carlson and McAfee 1983; Salop and Stiglitz 1982; Stahl 1989; Varian 1980) that shows that price dispersion can be an equilibrium outcome when some consumers find that it is too expensive to search for complete information. To keep all prices from converging to the monopoly level, these models typically assume that some consumers either are perfectly informed or have zero search costs. On the supply side, equilibrium price dispersion can be driven by differences in firm costs (e.g., Carlson and McAfee 1983) or by firms with identical costs that randomize their prices to capture some mixture of informed and uninformed consumers in a competitive game (e.g., Stahl 1989; Varian 1980). A general result of these models is that prices and price dispersion fall and welfare increases as search costs decline.[5] To the extent that consumers learn to use the Internet more efficiently, their costs of searching among e-tailers should decline over time, which leads to the following hypothesis:

> $H_1$: As a result of declining search costs and improved consumer information, price levels and price dispersion in e-tail markets should decline over time.

Because there are sites that provide complete lists of prices (e.g., Shopper.com), we might conclude that the Internet already provides low search costs and complete information. However, search costs might still be present for at least some consumers if awareness of these sites is imperfect or if nonprice attributes of e-tailers are important. Even if these qualifications do not hold, price dispersion may still exist. Baye and Morgan (2001) show that providers of price lists preserve ex ante price dispersion (and demand for their services) through the imposition of fees on firms and consumers. Alternatively, Ellison and Ellison (2001) argue that firms using Internet price-listing services attempt to preserve price dispersion by obfuscating their prices. These arguments suggest that Internet price dispersion will persist because sellers and service providers will take steps to maintain it.

[3]To keep notation simple, we assume that c = 0 for this argument and that search is for price. Alternatively, the objective could be to find the lowest full price. We drop the consumer subscript i to simplify notation.

[4]Carlson and McAfee (1983) point out that their results can also be derived from a more complex model that relaxes the assumption that consumers know the distribution of prices.

[5]Price dispersion can increase when the number of informed consumers is low before it eventually decreases (Stahl 1989).

At least one search model also suggests that price dispersion will persist if consumers enter and exit the market at a steady rate. Burdett and Coles (1997) construct a model in which entrants compete aggressively for the business of consumers while established firms take advantage of their customers' switching costs by continually increasing their prices. The result is ongoing price dispersion, with the most recent entrants at the low end of the distribution and the most established entrants at the upper end. Comparative statics show that prices and price dispersion still decrease as search costs decline in this model. Still, the model indicates that entry and aging of firms and consumers, which alter search and pricing incentives, create another explanation of persistent price dispersion. In summary, service provider and e-tailer incentives, and dynamics of entry and exit, suggest an alternative to $H_1$.

> $H_2$: Price levels and price dispersion in e-tail markets will persist over time.

Another consideration in predicting the behavior of prices over time is changes in the number of firms selling an item, a subject investigated in detail by Baye, Morgan, and Scholten (2001). Over time, the number of e-tailers increased because of the dot-com boom and more bricks-and-mortar retailers going online but subsequently declined when the Internet bubble burst. Baye, Morgan, and Scholten show that existing theoretical models make conflicting predictions about the effect of number of firms on price dispersion; some predicted an increase in price dispersion with number of firms, and others predicted a decrease.[6] Their empirical results tend to show an inverted U-shaped relationship, which leads to the following hypothesis.

> $H_3$: Price dispersion will increase at a decreasing rate with number of firms selling an item, then decline after reaching a maximum.

### Differences in E-Tailer Service As an Alternative Explanation for Price Dispersion

The consumer's maximum utility in Equation 1 would result if he or she minimized the full price of the focal item defined as $FP_{imr} = p_{mr} + c_{imr}$. The transaction costs are affected by services offered by the e-tailer: $c_{imr} = c_{imr}(\mathbf{RS_{mr}})$, where $\mathbf{RS_{mr}}$ is a vector of services. Full price would be unchanged if $\partial p_{mr}/\partial RS_{kr} = -\partial c_{imr}/\partial RS_{kr}$, that is, if the increase in price due to a marginal increase in any service k equaled the reduction in transaction costs due to a marginal increase in service k. This relationship can be used to trace out indifference or willingness to pay curves for services (Ehrlich and Fisher 1982; Rosen 1974). Similarly, the e-tailer would be willing to offer more of the service if it could obtain a large enough price increase to cover its cost. Under perfect competition and perfect information, the simultaneous decisions of consumers and sellers leads to a regression relationship of the form (Rosen 1974)

$$p_{mr} = h(\mathbf{RS_{mr}}). \tag{8}$$

In this model, often termed the "hedonic regression model," the dispersion in prices of an item is completely determined by the dispersion in services. Thus, variation in services across e-tailers offers an alternative explanation of price dispersion to high search cost. This hypothesis would be supported empirically if regressions of prices on attributes could explain variation in prices up to random error due to omitted attributes. Thus, we can state the following hypothesis:

> $H_4$: Variation in prices across e-tailers is largely explained by variation in services offered by e-tailers, and measured price dispersion after correcting for differences in services is negligible over time.

The empirical strategy for testing $H_4$ is to estimate Equation 8, compute quality-adjusted prices for each item, and compute the dispersion of these quality-adjusted prices. For example, if we estimate the linear relation $p_{mr} = a + \mathbf{B}_m(\mathbf{RS_{mr}})$, the quality-adjusted price for any seller is

$$p^a_{mr} = p_{mr} - \hat{\mathbf{B}}_m(\mathbf{RS_{mr}} - \overline{\mathbf{RS}}_m), \tag{9}$$

where the quality-adjusted price is expressed as the price of that item at an average level of service. Adjusted price dispersion is then the dispersion of $p^a_{mr}$, which should approach zero if $H_4$ is correct. Under the maintained hypothesis of perfect information and perfect competition, this adjustment process is valid (Rosen 1974). Pakes (2001) shows that when information and competition are imperfect, the function in Equation 8 still exists, but the coefficients are a complex function of services offered by competing retailers and the distribution of consumer preferences. Therefore, the coefficients no longer have a clear interpretation or even an expected sign. We employ the adjustment outlined in Equation 9 as an empirical approximation; however, it can be inaccurate under significant departures from the conditions underlying $H_4$.

In the following sections, we perform two empirical analyses that apply the models and hypotheses developed thus far. In our first analysis, we attempt to make statements about the efficiency of the Internet compared with that of conventional retailers by interpreting Brynjolfsson and Smith's (2000) results in terms of our formulas for consumer surplus. In our second analysis, we employ data on prices and service levels for a large number of categories to study changes in the efficiency of the Internet as a retail medium between November 2000 and November 2001. This period coincides roughly with the periods just before and after the shakeout of dot-com firms in the economic downturn of 2001.

## Comparison of Conventional and Internet Retailers

Brynjolfsson and Smith (2000) tracked the prices of 20 book titles and 20 CD titles at eight Internet outlets and eight conventional outlets on a monthly basis from February 1998 to

[6]The authors contend that the best measure of price dispersion is the gap between the lowest and the second-lowest prices, because other ex ante prices could be irrelevant to consumer choices. However, because it is evident from our data that the largest firms in many Internet markets are somewhere in the middle of the price distribution and do not have either the lowest or the second-lowest prices, we decided to use more-conventional measures of price dispersion in this study.

Table 1. Summary of Relevant Results in Brynjolfsson and Smith (2000)

| Type | Conventional | Internet | Difference |
|---|---|---|---|
| **Mean Prices, Internet Share Weighted** | | | |
| Book | $13.90 | $11.74 | $2.16 |
| CD | $16.07 | $13.49 | $2.58 |
| **Mean Prices, Unweighted** | | | |
| Book | $13.90 | $12.68 | $1.22 |
| CD | $16.07 | $13.78 | $2.29 |
| **Minimum Prices** | | | |
| Book | n.a. | n.a. | $1.29 |
| CD | n.a. | n.a. | $1.40 |
| **Mean Full Prices, Internet Share Weighted** | | | |
| Book | $15.04 | $13.69 | $1.35 |
| CD | $17.41 | $15.15 | $2.26 |
| **Minimum Full Prices** | | | |
| Book | n.a. | n.a. | $1.09 |
| CD | n.a. | n.a. | $1.23 |
| **Range of Full Prices** | | | |
| Book | n.a. | $ 5.98 | n.a. |
| CD | n.a. | $ 4.45 | n.a. |

Notes: n.a. = not available.

May 1999. Their basic findings were that prices are generally lower on the Internet than in conventional outlets, but price dispersion is roughly comparable across the two channels. Price dispersion was found to be higher on the Internet for books and lower for CDs. In an effort to make welfare statements about the value of the Internet compared with conventional channels, we interpret Brynjolfsson and Smith's results in light of our models.

Table 1 summarizes Brynjolfsson and Smith's (2000) results that are relevant to this comparison. The first panel compares an unweighted average of conventional channel prices with a weighted average of Internet prices, where the weights are Web traffic shares, a (possibly crude) proxy for market shares. If we abstract from the differences in time costs and transaction costs across the two channels, these estimates provide a rough estimate of the difference in expected cost across the channels for a representative consumer of the type outlined in Equation 6. The results show that the Internet provides lower costs to consumers.[7]

The second panel compares unweighted means. These can be interpreted as expected prices for a consumer who does not search. Again, the Internet appears to provide better buys on average for such a consumer. The differences in minimum prices in the third panel indicate that a fully informed consumer would also pay a lower price on the Internet.

[7]The authors were understandably unable to obtain share weights for conventional retailers. To the extent that consumers in conventional channels are able to locate the best buys, Table 1 may overstate expected retail prices for conventional retailers.

Brynjolfsson and Smith (2000) estimated the full prices (p + c) for items in both channels, accounting for taxes, shipping and handling, and transportation costs consumers are likely to incur. Although this comparison tends to narrow the advantage of the Internet, the Internet still has lower prices for both an uninformed and a fully informed consumer. This comparison does not consider intangible services, such as the ability to buy a book or a CD right away at a conventional retailer. It also does not consider the consumer's opportunity cost of time, which should favor the Internet.

The final panel of Table 1 lists the range of full prices on the Internet, which averages $5.98 for books and $4.45 for CDs. This may be interpreted as the difference between the price that an uninformed consumer will settle for and the price paid by a fully informed consumer. Because our earlier discussion indicates that the expected gain from searching for one more price for such an uninformed consumer is roughly the difference between the maximum and the average price, which is roughly half the range, consumers who accept the highest price rather than search further should have a search cost of $5.98/2 = $2.99 for books and $4.45/2 = $2.225 for CDs.

Other studies comparing prices and price dispersion at different types of retailers show interesting differences. Tang and Xing (2001) find that the prices of pure-play Internet retailers are significantly (about 14%) lower than those of online multichannel retailers, which is consistent with Zettelmeyer's (2000) analytical result. Pan, Ratchford, and Shankar (2002) find that prices are lower for pure-play e-tailers than for bricks-and-clicks e-tailers for CDs, digital video discs (DVDs), and desktop and laptop computers; prices are similar for personal data assistants (PDAs) and electronics and higher for pure-play e-tailers for books and software. Pan, Shankar, and Ratchford (2002) analytically and empirically show that prices at pure-play e-tailers are lower than those at multichannel retailers in eight categories: apparel, computer hardware, consumer electronics, gifts and flowers, health and beauty, home and garden, office supply, and sports and outdoors. Ancarani and Shankar (2002) show that when list prices are considered for books and CDs, traditional retailers have the highest prices, followed first by multichannel retailers and then by pure-play e-tailers. However, when shipping costs are included, multichannel retailers have the highest prices, followed first by pure-play e-tailers and then by traditional retailers. With regard to price dispersion, pure-play e-tailers have the highest range of prices, but the lowest variability (standard deviation); multichannel retailers have the highest standard deviation in prices with or without shipping costs. These findings suggest that online pricing is complex.

In general, these results indicate that even if there is a large degree of dispersion of Internet prices, the price savings available on the Internet can provide benefits to consumers who can efficiently use this medium. However, those who do not have access to the Internet cannot obtain these benefits without incurring the costs of obtaining access and learning how to use the Internet. If the advantage of the Internet persists as this medium matures (and conventional channels decline), efforts to enhance accessibility may be justified. To determine what happens as Internet markets mature, we study a wide variety of Internet prices during two periods in the next section.

## Comparison of Internet Prices—November 2000 and November 2001

The data for this part of the study derive mainly from BizRate.com, a well-known price comparison Web site. This site searches and updates the product, price, and deal information daily for a large number of e-tailers. To overcome the potential shopbot participation effect, we also tried to search and collect prices of those e-tailers that are not listed at BizRate.com, though the site's list is generally quite complete. Moreover, by comparing e-tailers' prices at BizRate.com with those on their own Web sites, we verified that prices are identical for most e-tailers, except for a few that offer lower prices at Bizrate.com than at their Web sites.

In November 2000, we collected 6739 price quotes for 581 identical items sold by 105 e-tailers; in November 2001, we collected 6762 price quotes for 826 identical items sold by 89 e-tailers. Because these are posted prices, a critical assumption is that some transactions took place at each observed price. Because an average item had 11.60 sellers in 2000 and 8.17 in 2001, there appeared to be some attrition of sellers between the two periods. In addition to prices, we collected ratings of various e-tailer services published by BizRate.com.

We purposely focus on identical items to avoid the potential problem of unmeasured product heterogeneity. Such products are found in the following categories: books, CDs, computer software and hardware, consumer electronics, and DVDs. For example, the Toshiba Satellite 2775XDVD laptop computer (part number PS277U-6M9J0K) with a PIII 650 MHz processor, 64 MB memory, 12 GB hard disk, 8x DVD, 56 Kbps modem, and 14.1″ TFT screen is the same sold by any e-tailer. We compare the prices of such homogeneous items across the e-tailers in our sample selling them at any point in time. However, we were unable to track prices of identical items over time because the model numbers and identities of items for sale tend to change over the course of a year. Thus, our analysis is mainly useful for studying how the general dispersion in prices of identical items changed between the two periods. It is less useful for tracking changes in price levels.[8]

In our analysis, we work with basic prices as the dependent measure and do not directly add in shipping and handling costs. This is because there are usually several options for shipping and handling, which makes it problematic to construct shipping and handling costs that are consistent across e-tailers.[9] However, because we consider consumer ratings of shipping and handling costs in constructing our estimates of quality-adjusted prices, these costs are incorporated into our analysis.

Table 2 provides a summary of the means and standard deviations of the average item prices for both samples. Differences in average price levels between the two samples are partly due to differences in the mix of items sampled. For example, our 2001 desktop computer sample included a few relatively expensive servers, whereas our 2000 sample included more low-end items. Similarly, our software and consumer electronics samples included more and relatively

[8]The rapidly evolving technology for computer and electronic products and the quick changes in popularity of music, movies, and books led the prices of these products to change drastically over their short life cycles. Thus, comparing the changes of their price levels would confound the effect of market maturation with the effect of product maturation and therefore would be inappropriate.

[9]This may be a manifestation of the obfuscation referred to by Ellison and Ellison (2001).

**Table 2. Price Level by Category: Comparison of 2001 and 2000 E-Tailer Samples**

| Category | Mean and Standard Deviation (2001) | Observations (2001) | Mean and Standard Deviation (2001) | Observations (2001) | Difference in Mean | t-Value |
|---|---|---|---|---|---|---|
| Book | 20.65 (22.89) | 105 | 19.26 (26.04) | 134 | −1.39 | −.44 |
| CD | 13.51 (1.66) | 43 | 14.64 (6.53) | 120 | 1.13 | 1.74*** |
| DVD | 26.64 (18.73) | 96 | 22.53 (10.79) | 103 | −4.11 | −1.88*** |
| Desktop | 1209.70 (1077.70) | 105 | 2509.70 (2766.10) | 107 | 1300.00 | 4.52* |
| Laptop | 2391.60 (653.77) | 78 | 1981.30 (563.55) | 96 | −410.31 | −4.38* |
| PDA | 446.86 (317.97) | 37 | 350.70 (205.44) | 52 | −96.17 | −1.62 |
| Software | 281.42 (685.26) | 51 | 597.31 (1383.1) | 120 | 315.90 | 1.99** |
| Consumer electronics | 440.24 (498.74) | 66 | 671.94 (710.78) | 94 | 231.70 | 2.42** |

*Significant at $p < .01$.
**Significant at $p < .05$.
***Significant at $p < .10$.

**Table 3. Measures and Explanation of e-Tailers' Features by BizRate.com**

| Measure | Explanation |
|---|---|
| Ease of ordering | Convenience and speed of ordering |
| Product selection | Breadth/depth of products offered |
| Product information | Information quantity, quality, and relevance |
| Web site navigation and looks | Layout, links, pictures, images, and speed |
| On-time delivery | Expected versus actual delivery date |
| Product representation | Product description/depiction versus what is received |
| Customer support | Status updates and complaint/question handling |
| Tracking | Tracking order status |
| Shipping and handling | Shipping and handling charges and options |
| Certify 5 | Number of certifications from five agencies |
| Years certified | Number of years certified by BizRate.com |

more sophisticated items in 2001. The fall in average prices for DVDs, laptops, and PDAs may reflect general price trends for these categories. In general, prices are not directly comparable between the two samples because of differences in items sampled and general market trends. However, the dispersion of prices among sellers of physically identical items can be compared between the two periods.

## Hedonic Analysis

The first step in our analysis of the BizRate.com data is to examine the extent to which service differences account for measured prices and the hedonic model accounts for price dispersion. Our general approach is to employ hedonic regressions of price levels on services to develop measures of prices adjusted for the effects of service quality.

The first stage in our analysis of services is to define a set of measures of services on available data. BizRate.com presents the consumer evaluations of e-tailers on the first nine attributes described in Table 3. The items are scored on ten-point scales, where higher scores measure better performance. Although the first nine attributes in Table 3 capture functional dimensions of service, Brynjolfsson and Smith (2000) conjecture that trust is an important dimension of e-tailer service: a consumer would go to a trusted e-tailer to avoid spending the time to resolve problems that might crop up otherwise. To attempt to capture the trust dimension, we employ two variables, which are the final two measures listed in Table 3. The first variable, Certify 5, is a count of the number of certifications that an e-tailer receives from the following certifying agencies: Better Business Bureau, Gomez.com, BizRate.com, Truste.org, and VeriSign.org. The second variable, years certified, is the number of years the seller has been certified by BizRate.com. The rationale is that a high number of years indicates that the e-tailer has been active long enough to develop a reputation.

Because these 11 measures of e-tailer services are not independent, and some are likely to be measures of the same underlying construct, we subjected them to a factor analysis. Because we want our measures to be invariant to period, we performed the factor analysis on the pooled 2000 and 2001 data. The results of the factor analysis of the service measures indicate the existence of four underlying factors, which capture 84% of the variance in the original data. Table 4 provides the component matrix we obtained using Varimax rotation. The factors are labeled as follows: reliability, shopping convenience, certification, and shipping and handling. Because the factors explain a high proportion of the variance in the data, we employ factor scores as our measure of e-tailer services. The service measures employed in our analysis are related to the dimensions of retail services specified by Betancourt and Gautschi (1993). Reliability corresponds to Betancourt and Gautschi's assurance of product delivery dimension; shopping convenience is related to their assortment, accessibility, and ambiance dimensions; product information is related to their availability of information dimension; and certification is related to their assurance of product delivery dimension and, as pointed out previously, to the trust dimension specified by Brynjolfsson and Smith (2000). As the statement about

**Table 4. Factor Analysis of E-Tailer Services: Rotated Component Matrix[a]**

| | Component | | | |
|---|---|---|---|---|
| Variable | Reliability | Shopping Convenience | Certification | Shipping and Handling |
| Ease of ordering | .165 | **.891** | .064 | .248 |
| Product selection | .212 | **.833** | .197 | .083 |
| Product information | .453 | **.629** | .068 | –.099 |
| Web site navigation | .160 | **.914** | .144 | .138 |
| On-time delivery | **.921** | .166 | .109 | .116 |
| Product representation | **.729** | .392 | .325 | .060 |
| Customer support | **.908** | .152 | –.015 | .263 |
| Tracking | **.908** | .232 | .051 | .109 |
| Shipping and handling | .314 | .222 | .116 | **.853** |
| Certify 5 | .136 | .124 | **.898** | –.114 |
| Years certified | .046 | .177 | **.810** | .352 |

[a]Rotation method is Varimax.
Notes: Numbers in bold are the variables that loaded most heavily on a given factor.

Table 5. **Results of Pooled Regressions of Normalized Prices on Attributes by Category[a]**

| | | Independent Variable | | | | | |
|---|---|---|---|---|---|---|---|
| **Category** | **Parameter** | **Reliability** | **Shopping Convenience** | **Certification** | **Shipping and Handling** | **$R^2$** | **N** |
| Book | Estimate | .054 | –.009 | .008 | .058 | .222 | 2172 |
| | t-value | 11.93 | –3.68 | 1.83 | 21.58 | | |
| CD | Estimate | .014 | .008 | –.069 | .008 | .176 | 1275 |
| | t-value | 3.42 | 1.73 | –15.52 | 1.36 | | |
| Desktop | Estimate | .002 | –.010 | –.016 | .027 | .082 | 1721 |
| | t-value | .60 | –2.83 | –5.03 | 10.19 | | |
| DVD | Estimate | .002 | –.015 | –.023 | –.026 | .084 | 2309 |
| | t-value | .74 | –5.86 | –7.78 | –7.25 | | |
| Electronics | Estimate | .011 | .002 | .006 | .017 | .077 | 1478 |
| | t-value | 5.07 | .70 | 2.18 | 8.39 | | |
| Laptop | Estimate | .000 | .007 | –.002 | .008 | .021 | 1871 |
| | t-value | –.11 | 3.51 | –.71 | 3.94 | | |
| PDA | Estimate | .007 | –.011 | –.004 | .016 | .040 | 1039 |
| | t-value | 2.01 | –2.19 | –.85 | 4.83 | | |
| Software | Estimate | –.012 | –.012 | .012 | .025 | .105 | 1636 |
| | t-value | –4.65 | –3.88 | 4.00 | 12.40 | | |

[a]Regressions have the functional form outlined in Equation 10.

shipping and handling is worded in BizRate.com, this dimension relates mainly to shipping and handling charges and options.

Using scores on the service factors and the measures of trust as independent variables, we ran hedonic regressions of the form outlined in Equation 10 on the pooled data for 2000 and 2001.

$$(10) \qquad \frac{p_{mrt}}{\bar{p}_{mt-1}} = \sum_{k=1}^{6} b_k\left(RS_{mrkt} - \overline{RS}_{mkt}\right) + v_{mrt},$$

where m is the item, r is the e-tailer, k is the attribute, t is the time period, and v is an error term. Because effects become magnified as service levels increase, we divide by the mean of that item's price in the corresponding period to stabilize the error variance.[10] Measuring all effects as deviations from item means within a given period eliminates item effects due to generally high or low levels of attributes for sellers of that item. It has the same general effect as including a set of item dummy variables and creates a zero intercept. We pool across periods to make our quality adjustment consistent over time.

We ran regressions for each major category. The results are presented in Table 5. Although all are statistically significant, none of the regressions in Table 5 has a high $R^2$ value. Differences in e-tailer services, at least the ones measured in our data, do not explain a great degree of the variation in e-tailer prices, contrary to the hedonic hypothesis that services explain price dispersion.[11] In addition, we would generally expect positive signs on the various coefficients if this hypothesis is true, because these coefficients would measure marginal willingness to pay under this hypothesis, and we would expect consumers to be willing to pay nonnegative amounts for each attribute.[12] This phenomenon of wrong signs in regressions of prices on e-tailer service characteristics was also noted by Brynjolfsson and Smith (2000), who rejected using hedonic regressions partly for this reason. Whereas the negative signs provide evidence that Rosen's (1974) model of hedonic prices under perfect competition does not hold for our data, Pakes (2001) shows that they are possible in more general settings. Thus, the estimates in Table 5 are not necessarily biased or otherwise problematic. Consequently, we use these estimates to develop our estimates of quality-adjusted prices.

Among the variables in our regressions, shipping and handling tends to have the largest effect, which is positive in seven of eight cases. Favorable shipping and handling charges tend to be accompanied by higher prices. Even when significant, however, effects are generally small relative to the observed variation in prices. For example, an increase of one unit (standard deviation) in the factor score for shipping and handling increases the ratio of price to its mean by .058 (increases price by approximately 5.8 percentage points) for books. Among items, prices of books are explained best by the four service factors; prices of CDs are explained second best. However, most of the explanatory power for CDs comes from the negative effect of certification, which is difficult to interpret. A general conclusion is that the observed dispersion in e-tailer prices is not explained to any great degree by the variation in e-tailer services. The hedonic explanation for price dispersion can be rejected.[13]

Although the hedonic explanation appears not to hold, we still make empirical comparisons between quality-adjusted

[10]The form in Equation 10 gave better results than deviations from average prices and the ratio of log of price to its mean. However, different functional forms gave similar results.

[11]Pan, Ratchford, and Shankar (2002) make the same point from our 2000 data.

[12]Although the within-category nature of our analysis creates some correlations between the four attributes, these are small so that multicollinearity does not appear to be a serious problem in our regressions.

[13]Running regressions on raw attributes rather than factor scores did not lead to substantial improvements in fit.

**Table 6. Change in Price Dispersion Relative to Average Price Across Categories**

| | Sample Size | | Dispersion in Prices | | | | Dispersion in Adjusted Prices | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 2000 | 2001 | 2000 | 2001 | Difference | t | 2000 | 2001 | Difference | t |
| **Percentage of Price Difference** | | | | | | | | | | |
| Book | 105 | 134 | 48.90 | 48.08 | –.82 | –.47 | 49.16 | 34.38 | –14.77 | –8.31* |
| CD | 43 | 120 | 51.04 | 39.30 | –11.74 | –3.62* | 49.61 | 31.31 | –18.30 | –5.13* |
| Desktop | 105 | 107 | 34.39 | 15.01 | –19.38 | –6.83* | 36.19 | 15.83 | –20.36 | –7.68* |
| DVD | 96 | 103 | 43.67 | 32.28 | –11.39 | –4.57* | 38.34 | 33.61 | –4.73 | –2.02** |
| Electronics | 66 | 94 | 30.99 | 22.12 | –8.87 | –4.38* | 31.18 | 22.31 | –8.87 | –4.36* |
| Laptop | 78 | 96 | 25.70 | 17.87 | –7.82 | –3.44* | 25.91 | 17.78 | –8.13 | –3.68* |
| PDA | 37 | 52 | 37.10 | 30.26 | –6.84 | –1.47 | 36.37 | 30.88 | –5.49 | –1.23 |
| Software | 51 | 120 | 35.58 | 18.95 | –16.63 | –4.44* | 36.09 | 17.15 | –18.94 | –4.36* |
| **Price Coefficient of Variation** | | | | | | | | | | |
| Book | 105 | 134 | 15.29 | 16.63 | 1.34 | 2.47** | 15.47 | 11.93 | –3.54 | –7.00* |
| CD | 43 | 120 | 15.46 | 13.02 | –2.45 | –2.61* | 14.93 | 10.96 | –3.97 | –3.98* |
| Desktop | 105 | 107 | 10.78 | 5.46 | –5.32 | –6.36* | 10.56 | 5.71 | –4.85 | –6.42* |
| DVD | 96 | 103 | 13.05 | 10.22 | –2.84 | –3.25* | 11.94 | 10.42 | –1.52 | –1.79*** |
| Electronics | 66 | 94 | 9.65 | 8.22 | –1.44 | –2.33** | 9.33 | 7.81 | –1.51 | –2.50** |
| Laptop | 78 | 96 | 7.55 | 6.11 | –1.44 | –1.97** | 7.54 | 6.05 | –1.48 | –2.08** |
| PDA | 37 | 52 | 10.49 | 9.86 | –.62 | –.48 | 10.22 | 9.77 | –.46 | –.35 |
| Software | 51 | 120 | 10.55 | 6.51 | –4.04 | –3.90* | 10.10 | 5.79 | –4.31 | –4.35* |

*Significant at $p < .01$.
**Significant at $p < .05$.
***Significant at $p < .10$.

prices and unadjusted prices. Using the regression coefficients from Table 5, we calculate a quality-adjusted price for each item according to the previous formula.

$$\text{(11)} \qquad p^{a}_{mrt} = \overline{p}_{mt}(1 + v_{mrt}) = \overline{p}_{mt}\left[\frac{p_{mrt}}{\overline{p}_{mt}} - \sum_{k=1}^{6} \hat{b}_k\left(RS_{mrkt} - \overline{RS}_{mkt}\right)\right].$$

On the basis of the logic outlined in Equation 9, Equation 11 expresses the item's price adjusted for service quality at time t as price less the effect of deviations from the average level of attributes on price. It is therefore an estimate of what the price of any item would be if it had an average level of attributes.

## Analysis of Prices

In this section, we compare changes in dispersion in prices and quality-adjusted prices between the 2000 and the 2001 samples. The comparison in Table 6 employs two general measures of price dispersion that are commonly used in studies of this phenomenon. The first is percentage difference, where percentage price difference is defined as 100 × (range of item prices/mean item price). For example, across the 104 books in the 2000 sample, the average book has a price range of 48.9% of its mean price. The second measure is the coefficient of variation. Because both measures are expressed relative to price, they have the advantage of controlling for price differences across categories and years.

Table 6 presents estimates for both unadjusted and quality-adjusted prices as defined in Equation 11. Except for books and CDs in 2001, the quality-adjustment procedure does not have much effect on the dispersion measures, which is not surprising given the low $R^2$ values in the corresponding regressions. The dispersion in prices reported in Table 6 appears to be quite large, and our estimates of price percentage difference for books and CDs are somewhat larger than those of Brynjolfsson and Smith (2000), possibly because our sample contains more e-tailers. However, with the exception of books and PDAs, dispersion declined significantly for all measures between 2000 and 2001, consistent with $H_1$. Incorporating the quality adjustment also led to a decline in dispersion for books. The decline in measured dispersion between 2000 and 2001 is especially large for items in the desktop computer and software categories, with the average dispersion approximately cut in half between those periods.

We previously showed that the return to search for an uninformed consumer is equal to the difference between the average price and the lowest price (Equation 4), whereas an upper bound on search costs is related to the difference between the maximum price and the average price (Equation 7). We calculated these measures for the items in our sample using both unadjusted and quality-adjusted prices. We present results for the 2000 and 2001 samples in Table 7. As shown in Table 7, differences between mean price and minimum price show no clear pattern of change between the two periods. However, Table 7 shows a clear pattern of decrease in all categories for differences between maximum and mean price, which are related to the highest cost of search. These decreases are significant at the .05 level or better in five of the eight cases for both unadjusted and adjusted prices.[14] Further inspection of Table 7 and Table 5

[14]Weighting the highest price by (n + 1/n), as derived in our theoretical section, gave qualitatively similar results. We preferred to present the unweighted results in Table 6 because the distributions are easier to interpret.

**Table 7. Change in Dispersion Above and Below Average Price Across Categories**

| | Sample Size | | Mean Price Difference | | | | Mean Adjusted Price Difference | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 2000 | 2001 | 2000 | 2001 | Difference | t | 2000 | 2001 | Difference | t |
| **Difference Between Average and Minimum Price** | | | | | | | | | | |
| Book | 105 | 134 | 4.56 | 4.18 | –.39 | –.62 | 4.76 | 2.84 | –1.92 | –3.23* |
| CD | 43 | 120 | 2.43 | 2.44 | .02 | .07 | 2.41 | 2.26 | –0.16 | –.70 |
| Desktop | 105 | 107 | 94.32 | 171.64 | 77.32 | 3.01* | 143.70 | 183.10 | 39.40 | 1.40 |
| DVD | 96 | 103 | 4.35 | 3.19 | –1.16 | –2.35** | 4.26 | 3.43 | –.83 | –1.77*** |
| Electronics | 66 | 94 | 63.19 | 79.74 | 16.55 | 1.16 | 57.28 | 67.33 | 10.05 | .87 |
| Laptop | 78 | 96 | 300.35 | 157.21 | –143.13 | –4.10* | 299.37 | 156.67 | –142.70 | –4.26* |
| PDA | 37 | 52 | 39.68 | 52.70 | 13.02 | 1.35 | 40.53 | 50.55 | 10.01 | 1.03 |
| Software | 51 | 120 | 34.87 | 52.89 | 18.02 | 1.02 | 39.47 | 52.09 | 12.61 | .69 |
| **Difference Between Maximum and Average Price** | | | | | | | | | | |
| Book | 105 | 134 | 5.69 | 3.80 | –1.89 | –2.50** | 5.76 | 3.63 | –2.13 | –2.32** |
| CD | 43 | 120 | 4.57 | 3.29 | –1.28 | –2.91* | 4.40 | 2.32 | –2.08 | –4.61* |
| Desktop | 105 | 107 | 213.83 | 188.71 | –25.13 | –.87 | 209.88 | 196.24 | –13.64 | –.46 |
| DVD | 96 | 103 | 7.31 | 3.79 | –3.52 | –4.88* | 5.90 | 3.85 | –2.05 | –3.36* |
| Electronics | 66 | 94 | 80.30 | 61.41 | –18.90 | –1.25 | 82.35 | 70.10 | –12.25 | –.80 |
| Laptop | 78 | 96 | 336.16 | 186.67 | –149.49 | –4.60* | 340.53 | 185.81 | –154.72 | –4.80* |
| PDA | 37 | 52 | 95.73 | 49.41 | –46.32 | –2.50** | 92.20 | 53.59 | –38.61 | –2.55** |
| Software | 51 | 120 | 102.82 | 54.78 | –48.04 | –.81 | 99.96 | 44.53 | –55.43 | –.99 |

*Significant at $p < .01$.
**Significant at $p < .05$.
***Significant at $p < .10$.

shows that the distribution of prices was skewed above the mean in 2000 but is roughly symmetric in 2001. Evidently, there were fewer extremely high prices on the market in 2001, which is consistent with a decrease in the highest search costs.

To examine further whether there were general changes in price dispersion between the samples, we regressed the measures of dispersion in Table 6 on the following variables:

- Ln (item average price)—a control for the possibility that dispersion relative to price might decline with the price level because search costs are unlikely to increase proportionally with prices. The natural log gave slightly better results than a linear term (overall results were insensitive to this choice).
- Linear and quadratic terms in number of firms—to capture the effect of number of firms on dispersion, which Baye, Morgan, and Scholten (2001) and Pan, Ratchford, and Shankar (2002) find to be nonlinear.
- Category dummies—to control for effects that are idiosyncratic to category.
- A dummy = one in 2001 and zero otherwise.

The results of this analysis are presented in Table 8. With all other factors held constant, the coefficients of the 2001 dummy in Table 8 indicate a significant decline in price dispersion. For unadjusted prices, the decline is approximately 18% relative to the 2000 mean for both dispersion measures. For adjusted prices, the corresponding decline is approximately 24% for both measures. The general results in Table 6 hold up across categories when we control for price levels and numbers of firms. Consistent with $H_3$, the coefficients of the number of firm variables imply that price dispersion increases with number of e-tailers until it hits its maximum at about 15 for the percentage price difference measure and about 10 for the coefficient of variation measure.[15] Finally, the results in Table 8 indicate that relative dispersion declines with price.

## Conclusions

One of the results in this article is that price dispersion in the Internet markets studied declined substantially between November 2000 and November 2001. This reflects a maturation of these markets, and though other explanations are possible (e.g., increased collusion), these results are consistent with improvements in information and consequent gains in consumer welfare. Our finding of decreased price dispersion is contrary to findings of no trend by Clay, Krishnan, and Wolff (2001) and Baye, Morgan, and Scholten (2001). Both of these studies used data from an earlier period (until March 2001, beyond which the Internet markets started witnessing significant shakeouts), which may account for the difference in results.

Our data reject differences in e-tailer services as a major driver of observed price dispersion over time. Although it is possible that this was due to shortcomings of our measures of services, our measures were consistent with existing the-

[15]Our results for the relation between number of firms and coefficient of variation are similar to those of Baye, Morgan, and Scholten (2001); however, they found that dispersion increased with number of firms for a percentage of price difference measure.

**Table 8. Determinants of Price Dispersion Relative to Average Price**

| | Percentage Price Difference Dependent | | | | Price Coefficient of Variation Dependent | | | |
|---|---|---|---|---|---|---|---|---|
| | Unadjusted | | Adjusted | | Unadjusted | | Adjusted | |
| Variable | Estimate | t-Value | Estimate | t-Value | Estimate | t-Value | Estimate | t-Value |
| Intercept | 44.005 | 8.65 | 27.221 | 5.57 | 18.742 | 11.73 | 13.874 | 9.13 |
| Ln (average price) | −3.229 | −5.41 | −2.221 | −3.87 | −1.123 | −5.99 | −.777 | −4.36 |
| Number of firms | 2.685 | 3.50 | 4.052 | 5.49 | .345 | 1.43 | .662 | 2.89 |
| Firms squared | −.085 | −2.67 | −.140 | −4.58 | −.019 | −1.87 | −.032 | −3.32 |
| CD | −3.617 | −2.10 | −1.445 | −.88 | −2.084 | −3.86 | −.932 | −1.82 |
| Desktop | −8.732 | −2.83 | −4.043 | −1.36 | −3.062 | −3.16 | −1.976 | −2.14 |
| DVD | −11.255 | −6.71 | −6.138 | −3.81 | −3.832 | −7.27 | −1.885 | −3.76 |
| Electronics | −11.641 | −4.53 | −6.777 | −2.74 | −3.422 | −4.24 | −2.263 | −2.95 |
| Laptop | −11.978 | −3.52 | −9.266 | −2.84 | −3.522 | −3.30 | −2.692 | −2.65 |
| PDA | −6.546 | −2.34 | −1.722 | −.64 | −2.081 | −2.37 | −.668 | −.80 |
| Software | −15.565 | −7.01 | −11.034 | −5.17 | −5.144 | −7.38 | −3.913 | −5.91 |
| 2001 dummy | −6.998 | −6.25 | −9.008 | −8.38 | −2.126 | −6.05 | −2.816 | −8.43 |
| $R^2$ | .325 | | .279 | | .311 | | .254 | |
| N | 1407 | | 1407 | | 1407 | | 1407 | |

ories of retail services and were obtained from the best source of this information of which we are aware. It is much more likely that explanations for price dispersion that rest on cost information are correct. With the exception of our results for the book category, it did not make much difference to our results whether unadjusted or adjusted prices for measured differences in services were used. An implication is that analysts are generally safe in working with unadjusted prices, which has generally been done in the theoretical and empirical literature on price dispersion.

Our model of consumer surplus may provide a useful insight into why there is a large degree of price dispersion in Internet markets even though these markets allow information to be gathered relatively quickly without traveling to retailers. Because it saves time, the Internet should appeal to consumers with high time costs who do not find it cost effective to search. These consumers are willing to accept high prices. At the same time, if the Internet allows relatively efficient search, consumers who do not have such high time costs might be able to locate attractive selling prices expeditiously. The existence of groups with radically different search costs may help drive Internet price dispersion.

To the extent that the existing evidence indicates that Internet prices are generally lower than prices for comparable items at bricks-and-mortar retailers, the Internet improves consumer welfare. We also should point out that price dispersion in Internet markets does not indicate allocative inefficiency per se. Although consumers who pay high prices may lose, producers may capture corresponding gains, and gains and losses may cancel each other out.

Even though active intervention in markets is currently unfashionable and is not something we would advocate, there has been a long-standing interest in intervening to eliminate inefficiencies in retail markets (Maynes and Assum 1982). Given this history, this issue of active intervention could again surface for Internet markets. Our finding that online price dispersion declined over a one-year period suggests that interest in this issue may be premature. Further study of trends in the behavior of prices in e-tail markets would be helpful for monitoring the markets' efficiency.

It would also be useful to learn if our conjecture about the identities of consumers paying high prices on the Internet is correct. Consumers with a high value for time savings are likely to be wealthy and therefore are quite different from less wealthy people who must pay high prices in conventional retail markets because of their lack of mobility. This leads to another potential policy issue of making the benefits of the Internet more accessible to those who currently lack either access to the medium or knowledge of how to use it, that is, the issue of the "digital divide," which has been receiving a lot of attention. If the Internet leads to lower prices and is able to overcome mobility constraints, steps to promote its use may be warranted.

A key missing piece of data limits the applicability of this and other studies of pricing behavior in Internet markets. We generally observe only posted prices and do not know how many sales take place at each price. Actual sales data, similar in scope to store-level scanner data available in conventional markets, are needed for Internet markets. Another need is more-comprehensive data on prices in conventional retail markets. This would allow more-comprehensive comparisons of price levels between markets and more-general statements about price levels on the Internet versus in the conventional markets than can be drawn from the limited number of product categories that have been compared to date.

## References

Alba, Joseph, John Lynch, Barton Weitz, Chris Janiszewski, Richard Lutz, Alan Sawyer, and Stacy Wood (1997), "Interactive Home Shopping: Consumer, Retailer, and Manufacturer Incentives to Participate in Electronic Marketplaces," *Journal of Marketing*, 61 (July), 38–53.

Ancarani, Fabio and Venkatesh Shankar (2002), "Price Levels and Price Dispersion: A Comparison of Pure Play vs. Bricks-and-

Mortar vs. Bricks-and-Clicks e-Tailers," working paper, SDA Bocconi Graduate School of Management, Universita Commercialé Luigi Bocconi, Milan.

Bakos, Yannis (1997), "Reducing Buyer Search Costs: Implications for Electronic Marketplaces," *Management Science*, 43 (December), 1676–92.

Baye, Michael R. and John Morgan (2001), "Information Gatekeepers on the Internet and the Competitiveness of Homogeneous Product Markets," *American Economic Review*, 91 (June), 454–74.

———, ———, and Patrick Scholten (2001), "Price Dispersion in the Small and in the Large: Evidence from an Internet Comparison Site," working paper, Kelley School of Business, Indiana University.

Betancourt, Roger R. and David Gautschi (1993), "Two Essential Characteristics of Retail Markets and Their Economic Consequences," *Journal of Economic Behavior and Organization*, 21 (August), 277–94.

Brown, Jeffrey R. and Austan Goolsbee (2002), "Does the Internet Make Markets More Competitive? Evidence from the Life Insurance Industry," *Journal of Political Economy*, 110 (June), 481–507.

Brynjolfsson, Erik and Michael Smith (2000), "Frictionless Commerce? A Comparison of Internet and Conventional Retailers," *Management Science*, 46 (April), 563–85.

Burdett, Kenneth and Melvyn G. Coles (1997), "Steady State Price Distributions in a Noisy Search Equilibrium," *Journal of Economic Theory*, 72 (1), 1–32.

Carlson, John and Preston McAfee (1983), "Discrete Equilibrium Price Dispersion," *Journal of Political Economy*, 91 (June), 480–93.

Clay, Karen, Ramayya Krishnan, and Eric Wolff (2001), "Prices and Price Dispersion on the Web: Evidence from the Online Book Industry," *Journal of Industrial Economics*, 49 (4), 521–39.

Clemons, Erik, Il-Horn Hann, and Lorin Hitt (2002), "Price Dispersion and Differentiation in Online Travel: An Empirical Investigation," *Management Science*, 48 (April), 534–49.

Degeratu, Alexandru, Arvind Rangaswamy, and Jianan Wu (2000), "Consumer Choice Behavior in Online and Traditional Supermarkets: The Effect of Brand Name, Price, and Other Search Attributes," *International Journal of Research in Marketing*, 17 (March), 55–78.

DePalma, Andre, Gordon M. Myers, and Yorgos Y. Papageorgiou (1994), "Rational Choice Under Imperfect Ability to Choose," *American Economic Review*, 83 (June), 419–40.

Ehrlich, Isaac and Lawrence Fisher (1982), "The Derived Demand for Advertising: A Theoretical and Empirical Investigation," *American Economic Review*, 72 (June), 366–88.

Ellison, Glenn and Sara Fisher Ellison (2001), "Search, Obfuscation and Price Elasticities on the Internet," working paper, Department of Economics, Massachusetts Institute of Technology.

Lal, Rajiv and Miklos Sarvary (1999), "When and How Is the Internet Likely to Decrease Price Competition?" *Marketing Science*, 18 (4), 485–503.

Lynch, John and Dan Ariely (2000), "Wine Online: Search Costs Affect Competition on Price, Quality and Distribution," *Marketing Science*, 19 (Winter), 83–103.

Maynes, E. Scott and Terje Assum (1982), "Informationally Imperfect Consumer Markets: Empirical Findings and Policy Implications," *Journal of Consumer Affairs*, 16 (Summer), 62–87.

Morton, Fiona Scott, Florian Zettelmeyer, and Jorge Silva-Risso (2001), "Internet Car Retailing," *Journal of Industrial Economics*, 49 (4), 501–19.

Pakes, Ariel (2001), "A Reconsideration of Hedonic Price Indices with an Application to PCs," working paper, Department of Economics, Harvard University.

Pan, Xing, Brian T. Ratchford, and Venkatesh Shankar (2001), "Why Aren't the Prices of the Same Item the Same at Me.com and You.com? Drivers of Price Dispersion Among E-Tailers," working paper, R.H. Smith School of Business, University of Maryland, College Park.

———, ———, and ——— (2002), "Can Price Dispersion in Online Markets be Explained by Differences in e-Tailer Service Quality?" *Journal of the Academy of Marketing Science*, 30 (4), 429–41.

———, Venkatesh Shankar, and Brian T. Ratchford (2002), "Price Competition Between Pure Play vs. Bricks-and-Clicks E-Tailers: Analytical Model and Empirical Analysis," in *Advances in Applied Microeconomics, The Economics of the Internet and E-Commerce*, Vol. 11, Michael Baye, ed. Greenwich, CT: JAI Press.

Ratchford, Brian T., Jagdish Agrawal, Pamela E. Grimm, and Narasimhan Srinivasan (1996), "Toward Understanding the Measurement of Market Efficiency," *Journal of Public Policy & Marketing*, 15 (Fall), 167–84.

———, Myung-Soo Lee, and Debabrata Talukdar (2003), "The Impact of the Internet on Information Search for Automobiles," *Journal of Marketing Research*, 40 (May), forthcoming.

Rosen, Sherwin (1974), "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *The Journal of Political Economy*, 82 (1), 34–55.

Salop, Steven and Joseph Stiglitz (1982), "The Theory of Sales: A Simple Model of Equilibrium Price Dispersion with Identical Agents," *The American Economic Review*, 72 (December), 1121–30.

Shankar, Venkatesh, Arvind Rangaswamy, and Michael Pusateri (2001), "The Online Medium and Customer Price Sensitivity," working paper, R.H. Smith School of Business, University of Maryland, College Park.

Small, Kenneth A. and Harvey S. Rosen (1981), "Applied Welfare Economics with Discrete Choice Models," *Econometrica*, 49 (January), 105–30.

Smith, Michael and Erik Brynjolfsson (2001), "Consumer Decision Making at an Internet Shopbot," *Journal of Industrial Economics*, 49 (4), 541–58.

Stahl, Dale (1989), "Oligopolistic Pricing and Sequential Consumer Search," *American Economic Review*, 79 (September), 700–12.

Tang, Fang-Fang and Xiaolin Xing (2001), "Will the Growth of Multi-Channel Retailing Diminish the Pricing Efficiency of the Web?" *Journal of Retailing*, 77 (Fall), 319–33.

Varian, Hal (1980), "A Model of Sales," *American Economic Review*, 70 (4), 651–59.

Zettelmeyer, Florian (2000), "Expanding to the Internet: Pricing and Communication Strategies When Firms Compete on Multiple Channels," *Journal of Marketing Research*, 37 (3), 292–308.

Copyright of Journal of Public Policy & Marketing is the property of American Marketing Association and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.