Inf Syst E-Bus Manage
DOI 10.1007/s10257-017-0346-y

ORIGINAL ARTICLE

Pricing strategies in online book industry: a comparative study

Jifeng Luo[1] · **Han Zhang**[2] · **Haizheng Li**[3]

Received: 3 January 2015 / Revised: 5 March 2015 / Accepted: 16 March 2015

Abstract In this paper, we examine pricing strategies in the U.S. online book industry over two time periods, with an aim to understand whether and how the driving factors of price dispersion change over time. Our empirical results show that dispersion in prices has remained substantial over the period of 2001–2006, but the driving factors of these variations in price have evolved. In 2001 online book retailers generally engaged in obfuscation, frustrating consumer search by manipulating shipping options. As documented by prior literature and revealed in our 2001 data, higher prices charged by retailers were positively related with longer shipping time. This strategy has been abandoned, as shown by our results of a 2006 sample. Online retailers are now competing to ship items quicker than rivals and to pass fewer or no shipping costs on to consumers. The impact of trust assurance seals (e.g., seals of online security and privacy) on price has materialized over the period of 2001–2006. This is because as more consumers become security conscious, the effects of assurance seals on the price becomes better recognized. Moreover, although retailers are roughly clustered into three cohorts, they strategize prices across different product items within each cohort.

✉ Han Zhang
han.zhang@scheller.gatech.edu

Jifeng Luo
luojf@sjtu.edu.cn

Haizheng Li
haizheng.li@econ.gatech.edu

[1] Antai College of Economics and Management, Shanghai Jiao Tong University, Shanghai 20030, China

[2] Scheller College of Business, Georgia Institute of Technology, Atlanta, GA 30308, USA

[3] School of Economics, Georgia Institute of Technology, Atlanta, GA 30332, USA

Published online: 30 June 2017

Keywords Pricing strategies · Price dispersion · Search costs · Obfuscation · Trust assurance seals · Electronic markets · Book industry

1 Introduction

Price dispersion, the variation in prices across sellers of the same item, is just as prevalent on the Internet as it is for conventional retailers, and it persists even after accounting for differences in product and service quality (cf. Ancarani and Shankar 2004; Ba et al. 2008; Baye et al. 2004; Brynjolfsson and Smith 2000; Clay et al. 2001, 2002; Morton et al. 2001; Pan et al. 2002, 2004; Tang and Xing 2001; Xing 2010). Using laboratory data of prices, Garcia-Gallego et al. (2014) examines the evolution of prices in markets with Internet price-comparison search engines. They find that price distributions are bimodal: monopoly price and interior price dispersion. In fact, the assertion made more than thirty years ago that "…the 'law of one price' is no law at all" (Varian 1980, p. 651) still holds true on the Internet.

The Internet has provided a natural field in which to re-examine price dispersion. Scholars have extensively studied price dispersion and offer various explanations. According to the disequilibrium perspective, online price dispersion, as a disequilibrium phenomenon, does not persist as the market matures. However, extant research has suggested that deviations from the "law of one price" are the norm rather than the exception. If dispersed price is an equilibrium, it is important to know the source of price dispersion. If the source is mainly differences in service quality as proposed by retailer heterogeneity models (Hotelling 1929; Shaked and Sutton 1982; Ba et al. 2008), then firms offering high service levels should be able to charge premium prices. In this case, firms should focus more on creating and leveraging internal technological, organizational, and managerial processes in order to provide better service. If the equilibrium dispersed price is a consequence of search costs for price (e.g., Salop and Stiglitz 1977, 1982; Varian 1980; Burdett and Judd 1983; Stahl 1989), dispersion exists even after controlling for quality difference. For example, Liu et al. (2012) theoretically prove that due to the presence of informed and uninformed buyers, it is possible for a high-reputation seller charges a lower price than a low-reputation seller. Baye et al. (2004) argue that prices would be strategically unpredictable, therefore preventing rivals from systematically undercutting a firm's price and acquiring the mass of informed customers.

More empirical investigation is needed to reveal which model or combination of models of price dispersion best explains the phenomenon and its evolution. More specifically, what are the driving factors behind online price dispersion? Do the driving factors evolve over time? The answers to these questions will offer online retailers a deeper understanding of pricing strategy. To address this issue, we utilize data collected online in both November 2001 and May 2006 to investigate online pricing strategies and price dispersion. This large time window allows us to examine online pricing strategies over the consolidation stage of e-commerce development. The development of electronic commerce is usually classified into different stages: 1995–2000 was the first stage of e-commerce (innovation stage); 2001–2006 was

the second stage of e-commerce (consolidation stage); and 2006-present is the third stage of e-commerce (reinvention stage).

Our results suggest that while the magnitude of price dispersion remains quite stable, the driving factors of price dispersion have evolved over time. In the 2001 sample, online book retailers generally engaged in obfuscation (frustrating consumer search) practice and took advantage of asymmetric information regarding shipping options. This strategy was later abandoned. In the 2006 data, firms with fast standard shipping options charged higher total prices. The existence of trust assurance seals did not seem to have any association with price levels in the 2001 sample. In 2006 the economic effect of seals materialized and online bookstores with seals were able to charge higher prices.

This paper is organized as follows: Sect. 2 presents the literature review. Section 3 formulates the testable hypotheses, and Sect. 4 describes the data. Finally, Sect. 5 discusses the empirical results. Section 6 concludes.

2 Literature review

Through an extensive literature review, we identified four major theories that explain price dispersion. First, the disequilibrium perspective assumes that Internet markets were inefficient during the initial years of the Internet, and that price dispersion may be a disequilibrium phenomenon that reflects the random noise of an immature market (Brynjolfsson and Smith 2000). The second perspective proposes that the difference among online prices is the result of quality differences among retailers (Ba et al. 2008; Hotelling 1929; Pan et al. 2002). Firms offering better services can charge premium prices. While the Internet provides consumers with an easy access to price, search costs for information on price or quality are not eliminated. Some retailers may take advantage of high search costs and charge high prices while still offering low-quality service (Baylis and Perloff 2002). This leads to a third perspective, which asserts that search costs resulting from consumers' imperfect information allow firms to take advantage of ignorant consumers (Salop and Stiglitz 1977; Varian 1980). The fourth theory takes a non-search equilibrium angle. For example, Baye and Morgan (2004) suggest that bounded rationality among sellers will result in price dispersion.

2.1 Online price dispersion as a disequilibrium phenomenon

Inefficiency of an immature market has been proposed as one of the reasons for online price dispersion (Brynjolfsson and Smith 2000). But recent longitudinal analyses on price dispersion suggest that online price dispersion is a persistent phenomenon, and that the maturity of electronic markets has not yielded a frictionless market. For example, Baylis and Perloff (2002) find that the shape of the price distribution by week remains relatively constant through an 11–14 week sample period. Ratchford et al. (2003) compare levels of price dispersion for over 500 product items in November 2000, 2001, and February 2003 and find that the magnitude of price dispersion continues to be substantial. Baye et al. (2004)

examine four million daily price observations for consumer electronics products listed at Shopper.com from August 2000 to March 2001. The three measures of price dispersion in their paper are found to be quite stable over the sample period. Hence Baye and Morgan (2004) conclude that price dispersion for a variety of products sold in online markets is ubiquitous and persistent.

2.2 Online price dispersion from retailer characteristics

A rational consumer with enough information generally will not choose an online retailer with a higher price and bad service quality. Varian (2000) predicts that over time two groups of e-retailers would emerge: one providing little service at a low price and the other offering good service at a high price. Indeed, Pan et al. (2003) argue that although empirical studies have documented significant and persistent price dispersion and have investigated homogeneous products to avoid potential confounding effects of unmeasured product heterogeneity, one critical aspect of heterogeneity, namely, retailer service offering, has been generally overlooked. The observed price dispersion could well be the result of retailer service differentiation. According to the models of retailer heterogeneity (e.g., Ba et al. 2008), the relationship between price and service quality is positive: stores with higher service offerings charge higher prices.

However, the results from empirical studies regarding service quality are mixed. Pan et al. (2002) find that post-purchase service was not significantly related to price for some markets and was negatively related to price levels in other markets. They speculate that such results may come from a lack of differentiation in post-purchase service between various retailers. Cao and Gruca (2004), on the contrary, find a positive relationship between price levels and post-purchase performance.

One plausible reason for the mixed findings between service quality and price is the existence of uncertainties online on the quality of the product or retailer. Without being able to physically examine the product or the retailer, online consumers face uncertainties about both (Dimoka et al. 2012; Luo et al. 2012; Pavlou et al. 2007). It is possible that some retailers may take advantage of information asymmetry about retailer characteristics. For example, Baylis and Perloff (2002) show that Internet retail markets for digital cameras and scanners consisted of *good* firms with low prices and superior service and *bad* firms that charged high prices but offered poor service. They attribute this phenomenon to firms' discrimination among consumers with different knowledge, search costs, or patience. Ba et al. (2012) call this phenomenon adverse price effect and Liu et al. (2012) name it negative price premium effect. Jin and Kato (2006) find that higher winning prices in eBay auctions for baseball cards do not necessarily indicate that the product delivered by the seller is of higher true physical quality. Buyers are willing to pay 33–51 percent more for cards with outrageous (exaggerated) quality claims, even though such seller claims represent nothing except for an elevated risk of default or counterfeit and no better card quality. They conclude that some buyers are not only uninformed about the quality of a specific seller or the quality of the goods sold by a specific seller, but these buyers also hold an incorrect belief about how true quality varies by seller claims. They argue that assuming price to be a

monotone function of quality is misleading, especially in a market where quality is important but where no verifiable information about quality is available to consumers.

2.3 Online price dispersion from search costs

The growth and rapid spread of the Internet has significantly reduced search costs on price, but it does not necessarily mean that search costs on price will be trivial. Internet retailers could utilize seemingly random prices (Varian 1980; Stahl 1989, 1996) or engage in obfuscation (frustrating consumer search) when interacting with price-search technologies (Ellison and Ellison 2009). Ellison and Ellison (2009), for example, argue that Internet retailers have an incentive to engage in obfuscation by putting some friction back in the market and by making price search more difficult or less of a threat to profitability. Both strategies can easily make the search for price information harder. Therefore, online price dispersion from asymmetric information of price is not as trivial as one would expect.

Salop and Stiglitz (1977) consider a market to have two kinds of consumers: the "informed" consumer and the "uninformed" consumer. For some parameter regions, the equilibrium takes a form where some firms sell at a competitive price and others sell at a higher price. The high-price firms' consumers consist only of uninformed consumers. This is so called "spatial" price dispersion. Varian (1980) proposes "temporal" price dispersion, in which consumers can only be informed for short periods of time, because firms continually change their prices. Firms engage in randomizing prices in an attempt to keep consumers uninformed about their prices and consequently price discriminate between informed and uninformed customers. In models of spatial price dispersion, firms have to make an important decision when setting their prices: they can choose to simply charge a high price and only sell to a fraction of their product to uninformed consumers, or they can set lower prices to target both informed and uninformed consumers. But the temporal price dispersion models predict that firms randomize prices to keep consumers ignorant. Baye et al. (2004) propose a "hit and run" model and predict that firms charging high prices consistently are at a competitive disadvantage relative to their rivals who can exploit the predictability of their pricing strategy. Recently, Chellappa et al. (2010) empirically examine how price format, as a pricing strategy, helps to explain price dispersion in the domestic U.S. airline markets. Oh and Lucas (2006) find that online sellers change their price strategies frequently, which makes it difficult for consumers to respond appropriately.

2.4 Other price dispersion models

Some other work on online markets emphasizes non-search explanations for price dispersion in homogeneous "good" markets. For instance, Baye and Morgan (2001) present a theoretical model for understanding the economic mechanism behind information gatekeepers. It demonstrates that the gatekeeper charges firms advertising fees above socially optimal levels, therefore inducing only partial participation by firms. Although all consumers purchase from the firm offering the

lowest price, price dispersion persists because advertised prices are lower than unadvertised prices. Faced with the puzzle that price dispersion occurs in a laboratory setting with perfect information, Baye and Morgan (2004) propose bounded rationality-based theories. They use bounded rationality to derive equilibrium price dispersion in a homogeneous product market. Their analysis of data sets from two independent laboratory experiments and a leading Internet price comparison site provided statistical results consistent with the bounded rationality-based explanation of price dispersion.

Each of the models described above facilitates our understanding of price dispersion. In this study, we highlight the importance of understanding price dispersion from a comprehensive perspective by integrating different perspectives of price dispersion in the research model. More specifically, this study focuses on how the driving factors behind online price dispersion evolve over time, with the purpose of offering a deeper understanding about pricing strategy.

3 Hypotheses development

It has been proposed that price dispersion may be a disequilibrium phenomenon that reflects the random noise of an immature market (Brynjolfsson and Smith 2000). But recent longitudinal analyses on price dispersion suggest that online price dispersion is a persistent phenomenon and that the maturity of electronic markets has not yielded a frictionless market. Over the past several decades, two major views have emerged to account for the dispersion of prices on the Internet for the same item: asymmetric information and retailer heterogeneity. Modern economic theory suggests that price dispersion will arise when there is a positive probability that a randomly arriving consumer knows only one price due to high search costs (see, e.g., Salop and Stiglitz 1977, 1982; Varian 1980; Burdett and Judd 1983; Stahl 1989). Meanwhile, Pan et al. (2004) argue that the observed price dispersion could well be the result of retailer service differentiation. Baye and Morgan (2004) conclude that price dispersion for a variety of products sold in online markets is ubiquitous and persistent. But their empirical conclusion is limited to a short time window, and it is unknown whether price dispersion on the Internet will persist over a long time window. Therefore, we propose the following testable hypothesis:

H1 Online price dispersion exists over a large time window.

3.1 Search costs

There are varying implications within the models of price dispersion from the perspective of search costs for price information. In models of spatial price dispersion (Salop and Stiglitz 1977), some firms will persistently sell their product at a lower price than other firms. But Varian (1980) argues that firms run sales to keep consumers uninformed about prices. He proves that the equilibrium that is sustainable in a homogeneous good market is one in which firms charge seemingly random prices. That is, the identity of the firm offering the lowest price will vary

unpredictably over time. Stahl (1989, 1996) considered a model in which some consumers incur a search cost every time they receive a price quote. This model has a mixed strategy equilibrium: retailers randomize prices. Baye et al. (2004) predict that firms charging high prices consistently are at a competitive disadvantage relative to their rivals who can exploit the predictability of its pricing strategy. By undercutting a rival's price by a small amount, the small reduction in margins to existing consumers is more than offset by the increased volume stemming from the demand of price-motivated consumers. The "hit and run" pricing strategies, as named by Baye et al. (2004), may be particularly effective in highly competitive online sectors. Their results suggest considerable variation in the identity of the firm offering the lowest price over time. Running infrequent sales can make the price unpredictable. Each of the above models assumes one product per firm. We expect that random pricing strategy can be extended to include multiple products per firm. Therefore, we hypothesize that:

H2 Across different product items, there are considerable variations in the identity of the firms offering the lowest price.

3.2 Service quality

Shipping and handling are usually treated as one aspect of retail service (Pan et al. 2002, 2003; Baylis and Perloff 2002). Because of spatial separation between buyers and sellers, shipment of goods is an indispensable part of Internet retailing. The reliance on shipping makes delivery an essential component of competition among Internet retailers (Dinlersoz and Li 2006). From the consumer's point of view, longer shipping time means additional waiting time to get the goods and should be offset by a discounted price. When consumers are perfectly informed, controlling for other firm characteristics, firms with shorter shipping time offer higher total prices.

However, if some consumers have imperfect information and learn about shipping options only after spending additional effort and time at a firm's website, manipulating shipping options becomes possible to the extent that some retailers with longer shipping times or higher shipping fees may actually charge higher prices. According to Ellison and Ellison (2009), one of the most visible search versus obfuscation battles was fought over shipping costs and shipping times. Tedeschi (2001) reports that CDNow made a profit on shipping in the early days of e-commerce by charging customers $3 for the first item and $1 for each additional item, prices that were then far above the average cost of shipping a CD. Dinlersoz and Li (2006) find that firms that charge lower base prices tend to offer lower shipping fees and higher shipping quality, as measured by average delivery time.

As more search engines are displaying shipping charges and more consumers are well informed, online retailers with better shipping options charge higher prices. This obfuscation strategy with respect to shipping is not sustainable in the long run. Therefore we propose that:

H3 Shipping quality can be negatively related to price levels when obfuscation is possible, but will be positively related to price levels in the long run.

3.3 Trust assurance seals

To take advantage of online transactions, consumers need to provide valuable information, such as their demographic information and credit card number. It may be impossible to conduct transactions on the Internet without revealing private information. Firms can also collect information on customer online behavior by using cookies and click-through data. Therefore, consumers are concerned about online vendors' trustworthiness. According to Van den Poel and Leunis (1999), consumers perceive that the Internet is more risky than traditional channels. Prior literature (Bhimani 1996; Griffin et al. 1997) also suggests that online shopping risks are amplified by issues such as the security and disclosure of information during and after the transaction process. Stewart (2003) suggests that consumers' willingness to buy from an online vendor is a function of both trust in that store and of perceived Internet-related risks.

To promote consumers' trust, many business-to-consumer (B2C) online merchants are currently displaying trust assurance seals on their websites. Displaying more seals generally signals a stronger commitment of the retailer to security and privacy policies. The trust assurance seals displayed by online retailers could alleviate the risk faced by consumers and could provide ways for smaller, less-established stores to attract online shoppers. Such effects, if they materialize, should be transferred to pricing behaviors of online vendors and lead to differentiated prices. Consumers may not directly pay out of pocket for such services, but they may pay indirectly through the premium for the goods and services that they purchase from these stores (Rust and Kannan 2003). Stores with seals charge a higher price yet still attract a profitable number of consumers, while stores with no seal have to lower their price in order to allure consumers to buy from their shops. However, McKnight et al. (2004) do not find any significant impact and suggest that the ineffectiveness of assurance seals could be because many consumers simply do not understand their function.

With the proliferation of viruses, phishing attacks, and other forms of malice online, more consumers have become security and privacy-conscious. According to an IBM survey (January 25 2006), more Americans anticipate falling victim to a cyber attack than to a physical crime. This increased anxiety about the possibility of a cyber attack changes consumer behavior. Seventy percent of Americans now only use Internet shopping sites that display a security protection seal. In fact, using a data set of 9098 shopping sessions collected from a randomized field experiment conducted by a large seal in 2008, Özpolat et al. (2013) find that the presence of the assurance seal increases purchase conversion rate. We want to check whether trust assurance seal can also lead to higher price. Therefore, we hypothesize that:

H4 Having trust assurance seals on the storefronts is positively related to price, and the relationship will materialize in the long run.

3.4 Product assortment

A wide selection of products offers consumers the convenience to locate and purchase obscure products. Consumers may be willing to pay a price premium for such conveniences if they wish to reduce their search costs or enjoy one-stop shopping. Extensive research on product assortment Boatwright and Nunes (2004) has been conducted but provides conflicting results. Several papers find that reducing the selection of less popular items a store carries within a category results in unchanged or even increased category sales (e.g., Drèze et al. 1994; Broniarczyk et al. 1998; Boatwright and Nunes 2001, 2004). In contrast, Borle et al. (2005) find that cutting product assortment may erode consumer retention and reduce overall store sales. Reducing product selection can drive consumers to a rival with a broader product selection. Brynjolfsson et al. (2003) find that increased product variety significantly enhances consumer welfare. Their study reveals a $0.7 to $1 billion increase in consumer welfare that comes from the increased product variety of online bookstores. If the convenience, which results from broad product assortment, could build consumer retention and loyalty, retailers may charge a price premium. Therefore, we hypothesize that:

H5 E-retailers with broader product assortment charge higher prices in the long run.

4 The data

To control for possible product heterogeneity across stores, we collected price data from the online book market. One reason we choose the online book market is that books are a simple, physical good that can be cheaply shipped to consumers over a large geographical area. The homogeneous nature of new books enables us to investigate price dispersion without worrying about difficulties rendered by any product quality differences. Moreover, book publishers charge the same wholesale prices on an individual book across retailers, regardless of the size of the retailer or the channel in which the retailer operates (e.g., Clay et al. 2001; Brynjolfsson et al. 2003).

Two-period data (2001 and 2006) was collected in this study. The development of electronic commerce is usually classified into different stages: 1995–2000 was the first stage of e-commerce (innovation stage); 2001–2006 was the second stage of e-commerce (consolidation stage); and 2006-present is the third stage of e-commerce (reinvention stage). Our two data samples matched the e-commerce second stage (consolidation stage). In 2001, some e-commerce companies may still adopt some strategies which were popular in early stage of e-commerce. In 2006, e-commerce reached the end of the consolidation stage, so a lot of new strategies have been adopted. Our empirical analysis can shed light on the issue of pricing strategies over time.

The sample includes two categories of books: a random sample of bestsellers and a random sample of books in print. This method is consistent with that used by

Brynjolfsson and Smith (2000). In November 2001, we randomly selected five books from Yahoo! book top sellers and another five books from Yahoo! Shopping (Books). In order to get the most complete list of online bookstores possible, we first used several shopbots such as pricescan.com and pricegrabber.com to search these 10 books, which gave us a list of online bookstores. Second, we added some bookstores that were not found by the shopbots but were mentioned by prior literature (e.g., Brynjolfsson and Smith 2000). By doing so, we compiled a list of all available online bookstores at the data collection time. We found 20 different online bookstores available in 2001, with 148 book-retailer observations overall. It is worth noting that the focus of the research is about the pricing strategies of online bookstores rather than books. Therefore, in this study, the number of bookstores is more important in revealing different pricing behaviors among stores, while books only represent homogenous product. In May 2006, we randomly selected 13 bestsellers on The New York Times bestseller list and another 13 books from Yahoo! Book top sellers. Moreover, we randomly selected 80 books from Yahoo! Shopping (Books). Using the before mentioned methodology, we acquired a list of 22 different bookstores, of which only eight stores had been in business throughout the entire period between 2001 and 2006.

We collected all the prices and shipping and handling fees for the standard shipping option. Unit price is the list price that retailers charged for each book, and total price is measured by the list price plus a standard shipping and handling fee. Several stores differentiate shipping rates based on shipping location. We collected the standard shipping charges using zip code GA 30075. Only one store in the 2006 sample, Wal-Mart, charged tax for books shipped to all U.S. states. We excluded tax from the total price, because the tax is affected by other factors that cannot be captured in this study. Stores selling only second-hand books were eliminated. Some stores use "bait and switch," that is, they strategically advertise a low price online but then refrain from honoring that price. By collecting price data directly from the book stores instead of from price search engines, we were able to alleviate the confounding impact of bait and switch.

For each book we also collected data on "Acquisition time," "Shipping time," and whether it was "In stock." "Acquisition time" is the time from order placement until the item ships, while "Shipping time" is the time from shipment until delivery, measured as the average of the reported minimum and maximum expected shipping time by a retailer for its standard shipping option (Dinlersoz and Li 2006).

"In stock" notifies consumers of the availability of the book. To collect the information on "in stock," we put each book into the shopping cart and checked whether the store had it in stock and could immediately ship it to consumers. We calculated the percentage of books, which were offered and were in stock, out of the sample at a particular store, and we named the new variable "Product assortment." Broad product assortment means there is a high probability that a customer can acquire the product she or he wants from an online bookstore.

We browsed each bookstore's website and recorded any trust assurance seals that were displayed. In total, there were four distinct seals represented in the sample in 2001 and nine in 2006. One store in 2006, Textbookx.com, displayed four seals on its website at the time the data was collected, including seals of online security,

privacy, and corporate code. In our 2001 sample, about 25 percent of online retailers had seals, of which about 15 percent had more than one seal. These numbers increased to 59 percent and 37 percent in 2006, respectively. A dummy variable named "Seal" was constructed, which equals 1 for the stores with at least one seal and 0 otherwise. We also included the number of seals ("NSeal") to study the impact of having more than one seal on price.

A number of control variables were collected to control for store-specific factors that may influence pricing strategies. Multi-channel retailers could charge higher prices than pure-play Internet retailers. This is because they provide better pick-up and return service, more convenient product inspection, and greater consumer trust (Venkatesan et al. 2007). Ancarani and Shankar (2004) compare the levels of price of books and CDs between multi-channel and pure Internet retailers. Their results show that multi-channel retailers had higher average prices than pure Internet retailers, regardless of whether the listed price or total price was considered. Tang and Xing (2001) find that multi-channel retailers had significantly higher prices than pure play e-retailers using a data set containing 4896 price observations for 51 DVD titles sold at pure play e-retailers and top multi-channel retailers. Pan et al. (2002) also find evidence that multi-channel retailers have higher prices after controlling for retailer characteristics. Therefore, we added a dummy variable "Channel" with 0 denoting "pure e-commerce" stores and 1 for "brick-and-click" stores. Another dummy variable, "Public," was included, with 1 for websites owned by public companies and 0 otherwise. We checked each store against the COMPUSTAT database. If the store was listed in the COMPUSTAT database, it was considered to be a store owned by a public company. We constructed a dummy, "Brand Name," with 1 for Amazon.com and Barnes & Noble.com and 0 otherwise. Variable description is reported in Table 1, and descriptive statistics are shown in Tables 2 and 3. There is a significant variation in shipping time and shipping charge, as shown in Tables 2 and 3. For the same product, some vendors offer longer delivery

Table 1 Description of variables used in the empirical analysis

Variable	Variable description
List price	Unit list price of each book charged by retailers ($)
Shipping charge	Shipping and handling fee charged by a retailer for each book (standard shipping)
Total price	Sum of the list price and shipping and handling charge
Acquisition time	The time from placing an order until it ships (in business days)
Shipping time	Average reported shipping time for the standard shipping option (in business days)
Product assortment	The percentage of books that were offered and in stock for each store (%)
Seal	Dummy variable: 1 for stores with seals, 0 for stores without
NSeal	The number of seals displayed on the store's website
In stock	Dummy variable: 1 if the book is in stock, 0 otherwise
Channel	Dummy variable: 1 for "brick and click" stores and 0 for pure online stores
Public	Dummy variable: 1 for publicly traded companies and 0 otherwise
Brand name	Dummy variable: 1 for Amazon.com and Barnes & Noble.com, 0 otherwise

Table 2 Descriptive statistics and correlations 2001

Constructs	Mean	SD	1	2	3	4	5	6	7	8	9	10	11
List price	18.1	11.4											
Shipping charge	5.01	1.03	0.04										
Total price	23.2	11.5	0.99	0.14									
Acquisition time	2.61	2.76	0.11	0.26	0.14								
Shipping time	4.45	0.77	0.04	0.30	0.07	0.06							
Product assortment	91.3	17.5	0.07	0.11	0.08	0.17	−0.18						
Seal	0.25	0.44	0.05	0.15	0.07	0.43	0.02	0.16					
NSeal	0.53	1.00	0.02	0.11	0.03	0.15	−0.10	0.16	0.92				
In stock	0.94	0.24	0.06	0.12	0.07	0.05	−0.00	0.03	0.05	0.05			
Channel	0.17	0.38	0.08	−0.49	0.03	−0.18	0.13	0.02	0.07	0.18	−0.16		
Public	0.22	0.41	0.06	−0.64	−0.01	−0.30	0.13	0.04	−0.36	−0.33	−0.14	0.44	
Brand name	0.10	0.31	0.04	−0.47	−0.01	−0.22	0.14	0.02	−0.24	−0.22	0.21	0.28	0.66

Table 3 Descriptive Statistics and Correlations 2006

Constructs	Mean	SD	1	2	3	4	5	6	7	8	9	10	11
List price	24.5	30.4											
Shipping charge	4.04	1.11	0.03										
Total price	28.5	30.5	0.99	0.07									
Acquisition time	6.21	10.7	0.01	0.30	0.02								
Shipping time	5.73	2.64	−0.03	−0.32	−0.05	0.22							
Product assortment	65.1	28.4	0.05	−0.19	0.04	−0.51	−0.53						
Seal	0.59	0.49	0.02	−0.16	0.01	−0.02	0.31	−0.01					
NSeal	1.14	1.18	−0.01	−0.10	−0.01	−0.16	0.04	0.06	0.73				
In stock	0.56	0.50	−0.02	−0.23	−0.03	−0.83	−0.27	0.51	−0.04	0.15			
Channel	0.41	0.49	0.03	0.09	0.04	0.06	−0.08	0.11	−0.35	−0.36	0.03		
Public	0.27	0.45	−0.01	−0.32	−0.02	−0.27	−0.15	0.33	−0.33	−0.40	0.21	0.29	
Brand name	0.09	0.29	0.03	−0.02	0.03	−0.19	−0.33	0.40	−0.60	−0.45	0.16	0.13	0.52

times, while others offer immediate delivery for standard shipping. Shipping charges also differ significantly.

5 Empirical analysis and results

5.1 Market efficiency

Table 4 provides descriptive statistics of price dispersion for the 2001 and 2006 samples. Overall, our descriptive statistics suggest that price dispersion is persistent, supporting H1. Large variations of prices were found across online book retailers in both 2001 and 2006. For example, price range, measured as the difference between the maximum and minimum unit price of each book, are around 11.06 and 11.55, respectively. To control for price levels, we normalized maximal and minimal price by average price and calculated the percentage difference. The results are consistent.

Price range and percentage difference only consider the two extreme observations and ignore all other prices. They may not capture the competitive structure of the market. Therefore, we calculate the standard deviation and coefficient of variation to take into consideration every price observation. Both measures indicate substantial price dispersion.

Baye and Morgan (2001) make a compelling argument that these measures (coefficient of variation and price range) may not be an ideal way to measure price dispersion, because they do not truly capture the degree of competitiveness in the market. Given that the goods themselves are perfectly homogeneous, the vast majority of users that visit shopbots purchase the product at comparatively low prices. Thus, the only real competitors in the market are the firms that sell at the lower end of the price spectrum. Therefore, it makes sense to define price dispersion as the difference between the lowest prices and the second lowest prices charged. Accordingly, Baye et al. (2004) define the price gap as the percentage difference between the lowest two prices. Following their work, we also measured the price dispersion using the price gap and reported the results in Table 4. The mean price gap for total price across books is 12 percent in the 2001 sample and 11 percent in

Table 4 Price dispersion across online book retailers in 2001 and 2006

Price dispersions	Range	Percentage difference	Standard deviation	Coefficient of variation	Price gap	Information value
List price						
2001	11.06	0.61	3.23	0.18	0.25	0.34
2006	17.55	0.76	4.76	0.21	0.18	0.23
Total price						
2001	11.55	0.49	3.30	0.14	0.12	0.20
2006	19.29	0.68	5.10	0.18	0.11	0.22

$N_{2001} = 10$ and $N_{2006} = 106$

the 2006 sample, meaning that a consumer can expect to save more than ten percent of the total price by choosing the cheapest retailer. The price gap is much larger for list price than for total price in both samples.

In our data sample, 90 percent of the time Amazon.com and Barnes & Noble.com fail to offer the cheapest product in the 2001 sample and 96 percent of the time in the 2006 sample. Yet these two firms account for more than 85 percent of online book sales, and Amazon.com alone accounted for 60–75 percent of online book sales around 2001 (Goolsbee and Chevalier 2002). In 2009, Amazon.com's market share of online book sales was still 70–75 percent. Because of Amazon.com's overwhelming control of the market for online books, we define a measure of price dispersion as the percentage difference between the lowest price and the price charged by Amazon.com. It could be the value of brand, i.e., the amount that a loyal customer of Amazon.com is willing to pay rather than purchasing at the lowest price. Following Leiter and Warin (2007), we call this type of price dispersion information value, i.e., the amount that an informed consumer can expect to save by purchasing at the lowest price. The average value of information is 23 percent in the 2006 sample, meaning that an average consumer can expect to save nearly one quarter of the total price just by choosing the cheapest retailer. It is worth noting that in 2001, the price gap of the total price is much smaller than that of the list price; the same is true for information value. This result seems to suggest that the stores with the lowest list prices ask for higher shipping and handling charges.

5.2 Oligopolistic pricing

In order to examine the pricing behavior of online book stores across different items, we count the times that each store offers the lowest or the highest price for a given book. The results are shown in Tables 5 and 6. Our results provide mixed support for H2. The identity of the firm offering the lowest price did not vary much in our 2001 sample. The store, Doublediscounts.com, offered the lowest price for nine of the 10 books surveyed. But in 2006, the identity of the firm offering the lowest price changed hands frequently between five stores. These are AllDirect, Buy.com, Walmart, Biggerbooks, and SuperBookDeals. This suggests considerable variation in the identity of the low-price firm even at a snapshot of the price distribution at one point in time.

Interestingly, three cohorts of online bookstores emerged in both samples, as suggested by Tables 5 and 6. The first cohort, composed of price leaders, frequently charged lower prices than others. Within this cohort, stores undercut each other aggressively in the 2006 sample. The second cohort, mainly opportunists, utilized a high-end pricing strategy. These stores, which are not well-known, charge higher prices to prey on uninformed consumers. The last cohort priced somewhere in between: most of the time, stores in this cohort charged the most common prices (e.g., mode of the price data). Amazon.com and Barnes & Noble.com are representatives of the third cohort. In a similar vein, Ellison and Snyder (2014) also find three strategic groups in their working paper, where group 1 generally occupy the lower price ranks and change prices relatively frequently, group 2 occupy the middle price ranks and tend to change prices infrequently, and group 3 charge the

Table 5 Times with Minimal and Maximal Price—Store Level 2001

Online book store	W/minimal price		W/maximal price	
	Times	Frequencies (%)	Times	Frequencies (%)
Doublediscount	**9**	90	0	0
Varsity BOOKS	0	0	**10**	100
Powells	0	0	**8**	80
1bookstreet	0	0	**6**	60
Wordsworth	0	0	**3**	30
Bookcloseouts	1	10	0	0
BOOKVARIETY	0	0	1	10
Amazon	0	0	1	10
Barnes & noble	0	0	1	10
A1books	0	0	1	10
TextbookX	0	0	0	0
AllDirect	0	0	0	0
eCampus	0	0	0	0
Bookamillion	0	0	0	0
AlphaCraze	0	0	0	0
Page1Books	0	0	0	0
ElGrande	0	0	0	0
Bcybookloft	0	0	0	0
Totalcampus	0	0	0	0
Textbooksource	0	0	0	0

The sum of the times with maximal price is larger than 10. This is due to the fact that more than one book store charged the same maximal price for one book

Table 6 Times with minimal and maximal price—store level 2006

Online book store	W/minimal price		W/maximal price	
	Times	Frequencies (%)	Times	Frequencies (%)
AllDirect	**37**	35	0	0
Buy.com	**14**	13	1	1
Walmart	**11**	10	0	0
BiggerBooks	**11**	10	0	0
Super book deals	**10**	9	0	0
Papamedia	1	1	**55**	52
Blackwell US	1	1	**26**	25
Powells	1	1	**20**	19
Page1books	4	4	**12**	11
Textbookrus	1	1	**10**	9
Allbooks4less	5	5	1	1

Table 6 continued

Online book store	W/minimal price		W/maximal price	
	Times	Frequencies (%)	Times	Frequencies (%)
Overstock	4	4	0	0
Strandbook	2	2	0	0
Hamiltonbook	2	2	0	0
Amazon	1	1	4	4
Barnes & noble.com	0	0	3	3
Bookbyte	0	0	2	2
Books a million	0	0	2	2
Textbookx	1	1	1	1
SanDiego technical books	1	1	1	1
eCampus	0	0	1	1
bestprices	0	0	0	0

The sum of the times with minimal (maximal) price is larger than 106. This is due to the fact that more than one book store charged the same minimal (maximal) price for one book

highest prices. We conducted pair wise t-tests to see whether there are significant price differences between retailers within the three cohorts in both the 2001 and 2006 samples. Table 7 shows that unit prices offered by each cohort indeed differ from one another significantly. Of course, we acknowledge that the result fails to account for the confounding effects of firm characteristics.

5.3 Bookstore heterogeneity

To help disentangle the confounding effects, we use a hedonic price function, where the logarithm of total price (P) charged by store i for book j is given by.

$$\begin{aligned}\ln(P_{ij}) = {} & \beta_0 + \beta_1 SHIP_{ij} + \beta_2 SEAL_i + \beta_3 ASSORTMENT_i \\ & + \beta_4 ACQUISITION_{ij} + \beta_5 INSTOCK_{ij} \\ & + \beta_6 PUBLIC_i + \beta_7 CHANNEL_i + \beta_8 BRAND_i + \varepsilon_{ij}\end{aligned} \quad (1)$$

where SHIP is the delivery time, SEAL is a dummy denoting if the book retailer has trust assurance seals presented on the website, ASSORTMENT denotes product assortment, ACQUISITION is the acquisition time, INSTOCK denotes if the book is in stock, PUBLIC denotes if the book retailer is a public company, CHANNEL denotes whether the store is a brick and click, and BRAND denotes if the store is a brand book retailer. We take the natural logarithm of total price to alleviate the heteroskedastic problem. Following Baylis and Perloff (2002), we do not include firm-specific dummies, because firm dummies would be perfectly collinear with variables representing firm characteristics.

We examined the possibility of multicollinearity. A pairwise correlation analysis ensured that no two regressors were highly correlated (see Tables 2 and 3). The VIF statistics (Belsley et al. 1980) in the preliminary estimations suggested that

Table 7 Price differences across cohorts of e-retailers in 2001 and 2006

Pair wise *t* tests	1–2	1–3	2–3
2001	9.37**	4.96*	4.40**
2006	8.62***	2.18	6.44***

1–2, 1–3, and 2–3 indicate the mean price difference between cohorts 1 and 2, between cohorts 1 and 3, and between cohorts 2 and 3, respectively

*, **, *** = significant at 10, 5 and 1% level. The sample sizes for 2001 are 49 for 1–2, 145 for 1–3, and 174 for 2–3, respectively. The sample sizes for 2006 are 911, 1216, and 1209 respectively

multicollinearity was not a concern. The confounding impact of book heterogeneity is controlled by book fixed effects. In this case, the price differences caused by the quality and production cost of a book are differenced out so that we could focus on price differentials caused by retailers. To address heteroscedasticity and the correlation of errors within books, we ran fixed effects models with robust standard errors clustered by book. The results are shown in Specifications I and II. In addition, because the number of clusters is small (10 books) in the 2001 sample, the critical values for the tests of significance are drawn from a t-distribution with eight (10−2) degrees of freedom (Cameron et al. 2008; Cohen and Dupas 2010). The critical values for the 1, 5, and 10% significance levels are thus 3.35, 2.31, and 1.86, respectively.

While one-way cluster-robust standard errors allow for correlation of unknown form within cluster, it is assumed that errors are uncorrelated across clusters. Cameron et al. (2011) propose an extension of cluster-robust standard errors that allows for clustering along more than one dimension. In contrast to one-way clustering, non-nested, two-way clustering allows for both time-series and cross-sectional dependence (in our case, book and retailer). Gow et al. (2009) find two-way clustering produces unequivocally better inferences than any of these methods even with as few as 10 clusters. Specifications III and IV show the results of two-way cluster.

We ran regressions using 2001 and 2006 data separately and compare the results, focusing on the evolution of retailers' online pricing strategies. Regression results are reported in Table 8.

As shown in Table 8, shipping time has a positive and significant association with total price for the 2001 sample. In particular, an additional day of delivery time increases the price by about $1.10 in 2001, evaluated at the mean total price. It seems that online book retailers with longer delivery times generally charge higher prices. The same pattern holds for acquisition time. Such a pricing strategy suggested by the empirical results does not seem to favor the consumer, indicating that online retailers engage in obfuscation, as suggested by Ellison and Ellison (2009). This result is also consistent with the finding of "good firms and bad firms" by Baylis and Perloff (2002): firms with desirable attributes would charge less to attract informed consumers, while firms with undesirable attributes, such as low guarantees and high shipping fees, would charge a high price. Dinlersoz and Li (2006) also find that shipping time is positively related with price. One reason lies in the fact that shipping service is almost always separately priced on the Internet. If

Table 8 Estimation results of pricing strategies in the online book industry

	2001				2006			
	I	II	III	IV	I	II	III	IV
Intercept	2.80***	2.79***	2.35***	2.79***	3.26***	3.25***	3.18***	3.17***
	(0.16)	(0.17)	(0.30)	(0.17)	(0.03)	(0.03)	(0.20)	(0.22)
Seal	0.036	–	0.036	–	0.18***	–	0.20***	–
	(0.022)		(0.10)		(0.01)		(0.01)	
NSeal	–	0.005	–	0.005	–	0.019***	–	0.015
		(0.013)		(0.046)		(0.003)		(0.018)
Shipping time	0.047***	0.048***	0.047	0.048*	−0.033***	−0.019***	−0.034***	−0.019*
	(0.009)	(0.009)	(0.032)	(0.028)	(0.003)	(0.003)	(0.007)	(0.012)
Product Assortment	0.001	0.001	0.001	0.001	−0.002***	−0.001***	−0.001	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.0003)	(0.0003)	(0.001)	(0.001)
Acquisition Time	0.009**	0.011***	0.009	0.011*	0.00	−0.00	0.001	−0.00
	(0.003)	(0.002)	(0.008)	(0.007)	(0.001)	(0.001)	(0.005)	(0.006)
In stock	−0.098**	−0.096**	−0.098*	−0.096	−0.079***	−0.117***	−0.15	−0.19
	(0.038)	(0.040)	(0.059)	(0.062)	(0.019)	(0.020)	(0.14)	(0.15)
Public	0.038	0.029	0.038	0.029	−0.14***	−0.14***	−0.16***	−0.17***
	(0.021)	(0.021)	(0.14)	(0.021)	(0.015)	(0.016)	(0.029)	(0.058)
Channel	−0.023	−0.017	−0.023	−0.017	0.17***	0.13***	0.17***	0.12***
	(0.033)	(0.035)	(0.11)	(0.12)	(0.015)	(0.008)	(0.019)	(0.023)
Brand name	−0.076**	−0.075**	−0.076	−0.075	0.28***	0.16***	0.30***	0.16***
	(0.028)	(0.028)	(0.12)	(0.13)	(0.017)	(0.014)	(0.022)	(0.052)
Book fixed effects	Y	Y	Y	Y	Y	Y	Y	Y
Adj. R^2	0.91	0.91	–	–	0.96	0.96	–	–
Observations	148	148	148	148	1261	1261	1261	1261

Standard errors are reported in parentheses

*, **, *** = significant at 10, 5 and 1% level

some consumers have imperfect information about prices and shipping service, and if they learn about shipping options only after spending additional effort and time at a retailer's website, firms can charge a higher price with lower shipping quality by discriminating against consumer with higher search costs.

As more consumers become informed on shipping options, obfuscation with respect to shipping is less likely to provide a benefit to retailers. Indeed, the list price no longer responds to an additional day of acquisition, and shipping time now has a negative effect on the price as expected in the 2006 sample. Specifically, an additional shipping day decreases the total price by an amount ranging from 2 to 3%. The dramatic change between 2001 and 2006 implies that the market became more competitive. In other words, when shipping time is longer, retailers charge a lower price in 2006, but not in 2001. Online retailers are now competing to ship items more quickly than rivals and are passing fewer or no costs on to consumers, suggesting a more competitive online book market. These results provide support for H3.

In 2006 online bookstores with seals could generally charge a higher price after controlling for other factors (Columns 6 and 8 Table 8), which is in support of H4. We also ran regressions with "*NSeal*" included, as shown in Column 7 Table 8. As expected, displaying more seals on storefronts helps. Interestingly, the existence of seals did not seem to have significant effects on pricing in 2001, although the coefficient before "*Seal*" has the right sign. In 2001, online retailers with seals did not, in general, charge a premium price for having seals, even though purchasing seals costs additional funds and efforts. It may be that most consumers do not recognize the value of online seals, which makes it difficult for stores to charge a higher price. The impact of displaying trust assurance seals materialized in 2006, as more consumers became security conscious. The increasing number of book retailers with seals in 2006, to some extent, indicates the increasing perceived usefulness of seals and supports our finding.

We did not find strong evidence in support of H5. As suggested by our results, product assortment had a positive though statistically insignificant influence on the price level charged by online bookstores in 2001. The effect became negative in 2006, but it is economically insignificant (ranging from -0.001 to -0.002). It could be that because the cost of holding large selection is negligible and entry barrier is low, wide selection became a necessity instead of a competitive differentiator.

The results for most control variables are generally consistent with expectations, although some may not always be statistically significant. For example, when a book is in stock, it takes retailers no effort to acquire it, and therefore its price is usually lower. The coefficients in 2001 and 2006 are not significantly different from each other, suggesting the impacts of "In stock" remained stable over the years. Retailers mainly look at the problem from a cost perspective. In 2006, retailers with dual channels were able to charge higher prices, consistent with Pan et al. (2002) and Ancarani and Shankar (2004). Yet in 2001, having dual channels did not give stores an advantage to charge higher prices. This result is attributed to better integration of channels in recent years.

In 2001, all other things equal, online book retailers with brand names charged lower prices. It seems that bookstores with brand names tried to drive other stores

Table 9 Summary of the key findings

Driving factors	β_{2001}	β_{2006}	Implications
Shipping time	0.047***	−0.033***	In early days price comparison websites did not collect information on shipping costs. Retailers took advantage by charging a higher price for longer shipping time. The obfuscation strategy became unsustainable in later years
Seal	0.036	0.18***	The economic effect of trust assurance seals materialized in later years
Product assortment	0.001	−0.002***	Broader product assortment became a necessity, instead of a competitive differentiator
Brand name	−0.076**	0.28***	Bookstores with brand names tried to drive other stores out of business by lowering their prices, willing to give up premiums in the short run in order to dominate the market in the long run

Coefficients in Columns 2 and 3 are directly taken from Table 8 Columns 2 and 6

*, **, *** = significant at 10, 5 and 1% level

out of business by lowering their price, willing to give up premiums in the short run in order to dominate the market in the long run. In the 2006 sample, the impact of brand names on price was positive. The difference in results between 2001 and 2006 shows that retailers with brand names charged lower prices in 2001 to grab market share but then charged higher prices in 2006. We summarize our key findings in Table 9.

5.4 Search costs and others

The evidence presented above strongly suggests that the price differentials in book prices can be partially explained by bookstore heterogeneity. However, the pricing patterns illustrated in Tables 5 and 6 reveal that the observed price differences cannot result entirely from online bookstore and book characteristics. In both 2001 and 2006, the stores that appear less attractive on the basis of non-price characteristics (e.g., product selection, brand name, etc.) seem to charge a higher price than their competitors. In our 2006 data, Papamedia.com (Miami Florida), for instance, charged the maximum prices for 55 titles, near 82.1% of the 67 books it offered. Service differences do not appear to justify the observed high price; instead, the store is relatively lesser known, with average product selection, and it only provides an online channel. In addition, the fact that the factors driving price dispersion have evolved over time also suggests that some stores engaged in obfuscation to take advantage of higher search costs. As Leiter and Warin (2007) note, when talking about online markets, it is important to know that there are two types of search costs: those linked to obtaining price information and those associated with obtaining product and seller information. Although the pricing behaviors in 2006 became beneficial to consumers (that is, stores with better services on average charged higher), it is interesting to observe how large the magnitude of price dispersion is after accounting for retailer heterogeneity.

Table 10 Price dispersion after controlling for retailer heterogeneity

Residual price dispersions	Range	Standard deviation
List price		
2001	5.09	1.65
2006	7.10	2.31
Total price		
2001	4.86	1.60
2006	6.33	2.04

$N_{2001} = 10$ and $N_{2006} = 106$

We only calculate range and standard deviation as other measures are either sensitive to the zero mean residual or do not retain original sense

To understand the magnitude of price dispersion after controlling for retailer characteristics, a relevant dispersion measure may be one that measures price dispersion after purging the differential arising from store heterogeneity (Sorensen 2000). Following Sorensen (2000), we calculate the dispersion in the residuals from a regression of price on book and store fixed effects. The new measures of price dispersion are reported in Table 10. The important result here is that price dispersion remains substantial even after retailer and book characteristics are taken into account. Moreover, the price dispersions after purging retailer heterogeneity are quantitatively similar in 2001 and 2006. Even though the driving factors within retailer heterogeneity changed over the years 2001–2006, the magnitude of residual price dispersion remains constant.

6 Discussion and conclusion

Price dispersion is a readily observed feature of online markets. Extant research has suggested that deviations from the "law of one price" are the norm rather than the exception. Though price dispersion has been investigated extensively, this study makes several important contributions to the research literature on online price dispersion and pricing decisions. First, this study empirically investigates the evolution of online pricing strategies in B2C electronic markets over time, using data collected online in November 2001 and May 2006. The unique data set gives this study a historic view of the e-commerce evolution, and this study finds evidence in support of its conceptual model. Price dispersion is found to be a persistent phenomenon, but the factors driving price dispersion have evolved over time.

Online book retailers generally engaged in obfuscation practice with respect to service quality (Ellison and Ellison 2009). For example, retailers take advantage of shipping options and charge higher prices for longer delivery time. However, with the increasing use of search engines and third party ratings, consumers could easily access information on a store's delivery quality. If there are enough perfectly informed consumers, the potential search keeps the market competitive (Salop and Stiglitz 1977). The informed exert a positive pecuniary externality on the

uninformed, and obfuscation practices with respect to shipping are not sustainable. Indeed, in 2006 retailers with longer shipping times could only charge lower book prices, consistent with the prediction of models with informed consumers. Of course, one can expect the emergence of other new obfuscation strategies from online retailers.

Second, our results also indicate that trust assurance seals play an important role in pricing decisions in recent years. Online bookstores with trust assurance seals charged higher prices in 2006, after controlling for other factors. The empirical results also show support for a hypothesis that websites displaying more seals can charge higher prices. Interestingly, the existence of seals did not seem to have a significant effect on price in 2001. One plausible explanation could be that many consumers began to recognize the value of these seals. While prior research has examined the effect of seals on online customer trust (e.g., Hu et al. 2010; Wang et al. 2004; Noteberg et al. 2003; Odom et al. 2002), our research is one of the first to investigate the economic impacts of trust assurance seals with real market data, which has been largely neglected in the literature. Extension of this analysis to other industries is desirable for robustness and further insight.

Third, by integrating different perspectives of price dispersion in the research model, this study highlights the importance of understanding price dispersion from a comprehensive perspective. In this research, all six measures of price dispersion suggest that online price dispersion is a persistent phenomenon even in a homogenous market. The price does not converge to either the competitive price or the monopoly price. Our results suggest the role of bookstore heterogeneity in explaining price differentials in book prices. At the same time, online retailers price their product items strategically to retard buyer learning and to combat increased market transparency by putting friction back into the market. For instance, there is considerable turnover in the identity of the firms offering the lowest price across different product items. Further, even after controlling for retailer heterogeneity, the magnitude of residual price dispersion remains large. Our results also reveal three cohorts of pricing strategies: low, medium, and high. It is noteworthy that the market leaders, Amazon.com and Barnes and Noble.com, have taken a mid-range pricing strategy.

As with any research, this paper comes with a number of limitations, which provide opportunities for further exploration in future research. First, although our data was collected over two time periods, the data for each time period is merely a snapshot of the market. This research would provide more insight on price adjustment if a longer time window were used for each time period. Second, some retailers offer bundling service, providing free shipping for larger purchases. Since we do not have real consumer transaction data, we were not able to disentangle the pricing strategy from this bundling service. Third, we collected the price information directly from each store's website. These prices are posted prices and may not reflect the prices that consumers actually pay. Ghose and Yao (2011) find that once actual transaction prices, rather than posted prices, are used, price dispersion in some electronic markets is substantially reduced. Therefore, empirical tests using transaction price data would improve this research.

Acknowledgements This research is partially sponsored by the National Natural Science Foundation of China (Project No. 71201101) and "Shu Guang" project supported by Shanghai Education Development Foundation and Shanghai Municipal Education Commission (Project No. 16SG11).

References

Ancarani F, Shankar V (2004) Price levels and price dispersion within and across multiple retailer types: further evidence and extension. J Acad Mark Sci 32(2):176–187

Ba S, Stallaert J, Zhang Z (2008) Oligopolistic price competition and adverse price effect in online retailing markets. Decis Support Sci 45(4):858–869

Ba S, Stallaert J, Zhang Z (2012) Online price dispersion: a game-theoretic perspective and empirical evidence. Inf Syst Res 23(2):575–592

Baye MR, Morgan J (2001) Information gatekeepers on the internet and the competitiveness of homogeneous product markets. Am Econ Rev 91(3):454–474

Baye MR, Morgan J (2004) Price dispersion in the lab and on the Internet: theory and evidence. Rand J Econ 35(3):449–466

Baye MR, Morgan J, Scholten P (2004) Temporal price dispersion: evidence from an online consumer electronics market. J Interactive Marketing 18(4):101–115

Baylis K, Perloff JM (2002) Price dispersion on the internet: good firms and bad firms. Rev Ind Organ 21:305–324

Belsley DA, Kuh E, Welsch RE (1980) Regression diagnostics: identifying influential data and sources of collinearity. Wiley, New York

Bhimani A (1996) Securing the commercial internet. Commun ACM 39(6):29–31

Boatwright P, Nunes JC (2001) Reducing Assortment: an attribute-based approach. J Mark 65:50–63

Boatwright P, Nunes JC (2004) Correction Note for 'reducing assortment: an attribute-based approach'. J Mark 68:iv

Borle S, Boatwright P, Kadane JB, Nunes JC, Shmueli G (2005) The effect of product assortment changes on customer retention. Mark Sci 24(4):616–622

Broniarczyk SM, Hoyer WD, McAlister L (1998) Consumers' perceptions of the assortment offered in a grocery category: the impact of item reduction. J Mark Res 35(May):166–176

Brynjolfsson E, Smith MD (2000) Frictionless commerce? a comparison of internet and conventional retailers. Manag Sci 46(4):563–585

Brynjolfsson E, Hu Y, Smith MD (2003) Consumer surplus in the digital economy: estimating the value of increased product variety at online booksellers. Manag Sci 49(11):1580–1596

Burdett K, Judd KL (1983) Equilibrium price dispersion. Econometrica 51:955–969

Cameron AC, Gelbach JB, Miller DL (2008) Bootstrap-based improvements for inference with clustered errors. Rev Econ Stat 90(3):414–427

Cameron AC, Gelbach JB, Miller DL (2011) Robust inference with multi-way clustering. J Bus Econ Stat 29(2):238–249

Cao Y, Gruca TS (2004) The influence of pre- and post-purchase service on prices in the online book market. J Interact Mark 18(4):51–62

Chellappa RK, Sin RG, Siddarth S (2010) Price formats as a source of price dispersion: a study of online and offline prices in the domestic U.S. Airline markets. Inf Syst Res, published online in Articles in Advance

Clay K, Krishnan R, Wolff E (2001) Prices and price dispersion on the web: evidence from the online book industry. J Ind Econ 49(4):521–539

Clay K, Krishnan R, Wolff E, Fernandes D (2002) Retail strategies on the web: price and non-price competition in the online book industry. J Ind Econ 50(3):351–367

Cohen J, Dupas P (2010) Free distribution or cost-sharing? evidence from a randomized malaria prevention experiment. Q J Econ 125(1):1–45

Dimoka A, Hong Y, Pavlou PA (2012) On product uncertainty in online markets: theory and evidence. MIS Q 36(2):395–426

Dinlersoz EM, Li H (2006) The shipping strategies of internet retailers: evidence from internet book retailing. Quant Mark Econ 4(4):407–438

Drèze X, Hoch SJ, Purk ME (1994) Shelf management and space elasticity. J Retail 70(4):301–326

Ellison G, Ellison SF (2009) Search, obfuscation, and price elasticities on the internet. Econometrica 77(2):427–452
Ellison SF, Snyder CM (2014) An empirical study of pricing strategies in an online market with high-frequency price information. CESifo working paper, No. 4655, http://hdl.handle.net/10419/93414
Garcia-Gallego A, Georgantzis N, Jaramillo-Gutierrez A, Pereira P, Pernias-Cerrillo JC (2014) On the evolution of monopoly pricing in internet-assisted search markets. J Bus Res 67(5):795–801
Ghose A, Yao Y (2011) Using transaction prices to re-examine price dispersion in electronic markets. Inf Syst Res 22(2):269–288
Goolsbee A, Chevalier J (2002) Measuring prices and price competition online: amazon and barnes and noble. Yale ICF working paper No. 0–23
Gow ID, Ormazabal G, Taylor DJ (2009) Correcting for cross-sectional and time-series dependence in accounting research. Account Rev, Forthcoming. http://ssrn.com/abstract=1175614
Griffin K, Ladd P, Whitehead R (1997) Internet commerce: security is still a concern. Rev Account Inf Syst 2(1):67–71
Hotelling H (1929) Stability in competition. Econ J 39:41–57
Hu X, Wu G, Wu Y, Zhang H (2010) The effects of web assurance seals on consumers' initial trust in an online vendor: a functional perspective. Decis Support Syst 48(2):407–418
Jin GZ, Kato A (2006) Price quality and reputation: evidence from an online field experiment. Rand J Econ 37(4):983–1005
Leiter DB, Warin T (2007) An empirical study of price dispersion in homogenous goods markets. working paper, Middlebury College
Liu Y, Feng J, Wei KK (2012) Negative price premium effect in online market: the impact of competition and buyer informativeness on the pricing strategies of sellers with different reputation levels. Deci Support Syst 54(1):681–690
Luo J, Ba S, Zhang H (2012) The effectiveness of online shopping characteristics and well-designed websites on satisfaction. MIS Q 36(4):1131–1144
McKnight DH, Kacmar CJ, Choudhury V (2004) Shifting factors and the ineffectiveness of third party assurance seals: a two-stage model of initial trust in a web business. Electron Mark 14(3):252–266
Morton FS, Zettelmeyer F, Silva-Risso J (2001) Internet car retailing. J Ind Econ 49:501–520
Noteberg A, Christiaanse E, Wallage P (2003) Consumer trust in electronic channels: the impact of electronic commerce assurance on consumers' purchasing likelihood and risk perceptions. e-Ser J 2(2):46–67
Odom MD, Kumar A, Saunders L (2002) Web assurance seals: how and why they influence consumers' decisions. J Inf Syst 16(2):231–250
Oh W, Lucas HC (2006) Information technology and pricing decisions: price adjustments in online computer markets. MIS Q 30(3):755–775
Özpolat K, Gao G, Jank W, Viswanathan S (2013) The value of third-party assurance seals in online retailing: an empirical investigation. Inf Syst Res 24(4):1100–1111
Pan X, Ratchford BT, Shankar V (2002) Can price dispersion in online markets be explained by difference in e-tailer service quality? J Acad Mark Sci 30(4):433–445
Pan X, Shankar V, Ratchford BT (2003) The evolution of price dispersion in internet retail markets. Adv Appl Microecon Organ New Ind Econ 12:85–105
Pan X, Ratchford BT, Shankar V (2004) Price dispersion on the internet: a review and directions for future research. J Interact Mark 18(4):116–135
Pavlou PA, Liang H, Xue Y (2007) Understanding and mitigating uncertainty in online exchange relationships: a principal-agent perspective. MIS Q 31(1):105–136
Ratchford BT, Pan X, Shankar V (2003) On the efficiency of internet markets for consumer goods. J Pub Policy Mark 22(1):4–16
Rust RT, Kannan PK (2003) E-service: a new paradigm for business in the electronic environment. Commun ACM 46(6):36–42
Salop S, Stiglitz J (1977) Bargains and ripoffs: a model of monopolistically competitive prices. Rev Econ Stud 44:493–510
Salop S, Stiglitz J (1982) The theory of sales: a simple model of equilibrium price dispersion with identical agents. Am Econ Rev 72:1121–1130
Sorensen AT (2000) Equilibrium price dispersion in retail markets for prescription drugs. J Politi Econ 108(4):833–850
Shaked A, Sutton J (1982) Relaxing price competition through product differentiation. Rev Econ Stud 49:3–13

Stahl DO (1989) Oligopolistic pricing with sequential consumer search. Am Econ Rev 79:700–712
Stahl DO (1996) Oligopolistic pricing with sequential consumer search and heterogeneous search costs. Int J Ind Organ 14:243–268
Stewart K (2003) Trust transfer on the world wide web. Organ Sci 14(1):5–17
IBM Survey (2006) Consumers Think Cybercrime Now Three Times More Likely Than Physical Crime. January 25
Tang F, Xing X (2001) Will the growth of multi-channel retailing diminish the pricing efficiency of the web? J Retail 77(3):319–333
Tedeschi B (2001) E-commerce Report; shipping fees: some scrimp, some profit. The New York Times, New York, p 11
Van den Poel D, Leunis J (1999) Consumer acceptance of the internet as a channel of distribution. J Bus Res 45(3):249–256
Varian HR (1980) A model of sales. Am Econ Rev 70:651–659
Varian HR (2000) Market structure in the network age. In: Bryjolfsson E, Kahin B (eds) Understanding the digital economy. MIT Press, Cambridge
Venkatesan R, Kumar V, Ravishanker N (2007) Multichannel shopping: causes and consequences. J Mark 71(2):114–132
Wang S, Beatty SE, Foxx W (2004) Signaling the trustworthiness of small online retailers. J Interact Mark 18(1):53–69
Xing X (2010) Can price dispersion be persistent in the internet markets. Appl Econ 42(15):1927–1940