Management Science

Price Dispersion and Loss-Leader Pricing: Evidence from the Online Book Industry

Xinxin Li, Bin Gu, Hongju Liu,

Please scroll down for article—it is on subsequent pages

INFORMS is the largest professional society in the world for professionals in the fields of operations research, management science, and analytics.
For more information on INFORMS, its publications, membership, or meetings visit http://www.informs.org

Price Dispersion and Loss-Leader Pricing:
Evidence from the Online Book Industry

Xinxin Li

Department of Operations and Information Management, School of Business, University of Connecticut,
Storrs, Connecticut 06269, xinxin.li@business.uconn.edu

Bin Gu

Department of Information Systems, W. P. Carey School of Business, Arizona State University,
Tempe, Arizona 85287, bingu@asu.edu

Hongju Liu

Department of Marketing, School of Business, University of Connecticut,
Storrs, Connecticut 06269, hongju.liu@business.uconn.edu

In this paper, we develop a theoretical model to analyze the pricing strategies of competing retailers with asymmetric cross-selling capabilities when product demand changes. Our results suggest that retailers with better opportunities for cross-selling have higher incentives to adopt loss-leader pricing on high-demand products than retailers with low cross-selling capabilities. As a result, price dispersion of a product across retailers rises when its demand increases. The predictions of our model are consistent with the empirical evidence from the online book retailing industry. Using product breadth as a proxy for cross-selling capability, we find that retailers with high cross-selling capabilities reduce prices on best sellers more aggressively than retailers with low cross-selling capabilities. As a result, price dispersion increases when a book makes it to the best-seller list, and the increase is mainly driven by the difference in pricing behavior between retailers with different cross-selling capabilities. Our empirical results are robust against a number of alternative explanations.

Key words: price dispersion; loss-leader strategy; competitive pricing; cross-selling capability
History: Received April 4, 2010; accepted August 6, 2012, by Sandra Slaughter, information systems. Published online in *Articles in Advance* November 9, 2012.

1. Introduction

Price dispersion for homogeneous products has received considerable attention from researchers due to its wide-range implications for retailers, consumers, and markets in general (Pan et al. 2004). The emergence of the Internet has spawned a large number of online retailers that compete for consumers' attention on the same products. Moreover, the growth of price comparison websites has significantly reduced search costs for consumers. Researchers thus expect a more efficient market that reduces price dispersion for homogeneous goods (Baye et al. 2004).

Previous empirical studies on online price dispersion provide some support (albeit limited) to this view of improved market efficiency. Although research has documented significant online price dispersion in a wide range of product categories (see, e.g., Brynjolfsson and Smith 2000, Clay et al. 2001, Clemons et al. 2002), several studies find confirming evidence that price dispersion declines as the Internet market matures and more online retailers join

the competition (Pan et al. 2004, Baye et al. 2004, Venkatesan et al. 2007, Ghose and Yao 2011).

There exists, however, a well-known anomaly to this view. In one of the earliest studies on this topic, Clay et al. (2001) found that price dispersion observed for best-selling books is significantly higher than that for average books, despite the fact that best sellers are carried by more online bookstores and price information is widely available for these books. This finding contradicts the common wisdom that increased competition and lower search costs lead to lower price dispersion.

In this study, we revisit this anomaly and offer an in-depth analysis on a plausible cause of higher price dispersion observed for popular products. We construct an analytical model to examine competing retailers' pricing strategies in a market characterized by the following two features: (1) retailers have asymmetric cross-selling capabilities, and (2) product demand changes in response to demand shocks. Our results suggest that the higher price dispersion often observed for popular products is related to retailers'

practice of loss-leader pricing, i.e., retailers reduce prices of popular products to attract consumers to their online stores, in hopes that these consumers purchase other more profitable products.[1] Retailers' motivation to engage in loss-leader pricing varies with their cross-selling capabilities. When product demand increases, retailers with high cross-selling capabilities have higher incentives to set loss-leading prices than retailers with low cross-selling capabilities. This enlarges the price difference between the two types of retailers, which contributes to the observed higher price dispersion for popular products.

Using a panel data set from 22 online retailers for 1,614 hardcover books in 2004, we empirically assess the predictions of our model by analyzing changes in book price and price dispersion when product demand increases, namely, when books hit the *Publishers Weekly* best-seller list (Sorensen 2007). The results of our empirical analysis are consistent with the predictions of our model. Specifically, using a retailer's product breadth as a proxy for cross-selling capability, we show that retailers with higher cross-selling capabilities reduce a book's price more aggressively after it hits the best-seller list than retailers with lower cross-selling capabilities. We also show that price dispersion increases when a book gets on the best-seller list. We decompose price dispersion into three components: the price difference between the group of retailers with high cross-selling capabilities and the group with low cross-selling capabilities, and the price dispersion within each group. We find that the main source of the increased price dispersion observed for best-seller books is the price difference between the two groups of retailers. These findings are robust against a number of alternative explanations.

Our model predictions and empirical findings are echoed by a recent *Wall Street Journal* article (Bustillo and Trachtenberg 2009) highlighting a book price war that broke out between Walmart, Amazon, and Target in October 2009. These retailers cut their online prices for 10 best sellers from $25 to $9. A publishing analyst, Michael Norris, commented that "Non-bookstore retailers can sell books at a loss, and they can sell books for less than the price they paid, because they usually like to treat them as loss leaders" (Hobson 2009). Meanwhile, independent bookstores decided not to cut their prices as much to match their mega competitors, as predicted in this study.

---

[1] Our definition of loss-leader pricing follows Hess and Gerstner (1987, p. 358), who define it as "a pricing strategy in which retailers set very low prices, sometimes below cost, for some products to lure customers into stores." We do not emphasize whether prices are below cost, although our analytical model does show that retailers with high cross-selling capabilities may charge a price lower than cost to attract consumers and profit from cross-selling.

This paper contributes to two research streams. First, we develop an analytical model to provide theoretical grounds for the higher price dispersion often observed for high-demand products. We further test the predictions of the analytical model with data from the online book retailing industry. Our model complements previous studies on price dispersion by providing a new explanation that is consistent with the observed price pattern in the industry. Second, this study also complements existing studies on loss-leader pricing, because these studies have not focused on how the incentive for loss-leader pricing varies across different retailers as product demand changes. We provide a framework to understand such incentives given retailers' asymmetric cross-selling capabilities and changes in product demand condition.

The rest of this paper is organized as follows. We review the related literature in §2. Section 3 presents the analytical model on competing retailers' pricing strategies when they have different cross-selling capabilities. We conduct the empirical analysis in §4 using panel data collected from a major price comparison site, followed by robustness analyses against a number of alternative explanations. Section 5 concludes with a discussion on limitations and future research opportunities.

2. Literature Review

This study stems from two research streams. First, we extend the online price dispersion literature by developing an analytical model on retailers' pricing strategies to explain the increased price dispersion often observed for popular products. Second, we extend the literature of loss-leader pricing by showing that retailers have asymmetric incentives to engage in loss-leader pricing as product demand changes.

2.1. Literature on Price Dispersion

Past studies have examined price dispersion in various industries and have proposed several explanations. Some of these studies analyze price dispersion within a category but across products, whereas others focus on the price dispersion for homogeneous products across retailers. Our study focuses on the latter.

Price dispersion across retailers can arise from heterogeneity in consumer loyalty, and results in a mixed-strategy equilibrium in pricing. For example, Shilony (1977) and Varian (1980) presented models in which oligopolistic sellers use mixed strategies in prices. These models predict that the relative position of a retailer within the price distribution changes over time in a random fashion. Lach (2002) presented evidence consistent with this prediction using monthly prices of refrigerators, chicken, coffee, and flour. However, by examining the prices of a digital camera and a flatbed scanner, Baylis and Perloff (2002)

found that the store ranking in the cross-sectional price distribution was persistent over time.

Price dispersion can also be a result of limited consumer awareness of competing retailers, or heterogeneity in consumer preference for retailers (Chen and Hitt 2001). Branded retailers tend to charge higher prices on average due to brand loyalty. However, when there are both limited consumer awareness of competing retailers and heterogeneity in brand loyalty, the market may exhibit a mixed-strategy equilibrium in which branded retailers randomize prices such that sometimes branded retailers charge lower prices than unbranded retailers. Similarly, Smith (2001) presented an analytical model of online price dispersion that focuses on consumer awareness of Internet retailers. He pointed out that consumers' search cost in electronic markets is primarily a function of consumers' mental awareness of different retailers. In equilibrium, the price dispersion is high because a few well-known retailers cooperate to set high prices, whereas the remaining fringe retailers randomize over prices. Smith (2001) examined prices set by 24 Internet retailers for 23,744 books in late 1999 and found that well-known retailers have very similar prices, whereas fringe retailers do not.

Another common explanation for price dispersion on the Internet is service heterogeneity. Firms that provide good services may charge premium prices. Pan et al. (2002) examined the prices for 581 items in eight categories and service quality ratings for 105 online retailers. They showed that the service quality difference is not the main driver of the observed price dispersion. Furthermore, Baylis and Perloff (2002) found that low prices are often associated with superior services, and vice versa. Retailers' pricing format also has a significant impact on price dispersion. Chellappa et al. (2011) empirically analyzed the effect of vendors' price format, everyday low price (EDLP) versus non-EDLP, on price dispersion of domestic airfares. They showed that price format remains an important source of price dispersion after accounting for other factors.

None of the aforementioned studies, however, consider the role of cross-selling capability in retailers' pricing strategies or analyze how different types of retailers respond to changes in product demand. In contrast, we focus on the pricing strategies of asymmetric retailers with different cross-selling capabilities in response to demand changes; that is, we use demand change as an identification strategy to assess the influence of loss-leader strategy on price dispersion. Our model suggests that online bookstores with higher cross-selling capabilities tend to make deeper cuts on prices when books hit the best-seller list, which leads to higher overall price dispersion on best sellers. This pattern is consistent with the observations from our data.

Extant studies suggest that price dispersion is likely to decrease as competition among online retailers intensifies. For example, Ghose and Yao (2011) found substantially lower price dispersion using actual transaction prices collected from an online government purchase market where retailers compete side by side under strict service requirements. The 0.22% price dispersion observed in this market suggests that the "law of one price" can prevail in a highly competitive online market. Similarly, Tang et al. (2010) found that increased use of shopbots among consumers is associated with lower product price and lower price dispersion. Sorensen (2000) suggested that consumers' increased propensity to price shop for repeatedly purchased products will constrain the prices of such products to be lower and less dispersed. Applying a similar argument to the online book retailing industry, one would expect best sellers to have smaller price variations across vendors because these books tend to be more heavily advertised, which reduces search costs. However, both our data and prior studies (e.g., Clay et al. 2001) indicate that the observed price dispersion for popular products is often higher than that for less popular products, despite lower search costs. Our study provides a plausible explanation for this intriguing phenomenon.

This study is also related to prior research on the book retailing industry (e.g., Raff 2000, Clay et al. 2002). Most closely related to our study, Clay et al. (2001) examined price dispersion in the online book retailing industry using data collected between August 1999 and January 2000 covering 399 books from 32 online bookstores. They found that best sellers have lower prices but larger price dispersion than other books. However, they did not formally model or explain the higher price dispersion. In contrast, we rationalize the observed pattern using an analytical model of loss-leader pricing and show that the predictions of the model are consistent with the price patterns observed in our data.

In a research commentary paper, Pan et al. (2004) pointed out that the Clay et al. (2001) result regarding price dispersion for best sellers is contrary to the expectation that more advertised books should exhibit lower price dispersion. They conjectured that different retailers may use different popular products as loss leaders, resulting in a higher degree of price dispersion. They did not formally model or empirically test the conjecture. Although we agree that the observed price dispersion is related to loss-leader pricing, our analysis indicates that rather than all retailers adopting loss-leader pricing, the main driving force is the heterogeneity in retailers' cross-selling capabilities and in their motivations in adopting loss-leader pricing. Specifically, we find that retailers with high cross-selling capabilities lower prices for

best sellers significantly more than retailers with low cross-selling capabilities, and the increase in price dispersion is mainly driven by the systematic price difference between the two groups of retailers. The predictions of our analytical model are consistent with the price patterns observed in our data.

2.2. Literature on Loss-Leader Pricing

Our study is also related to the literature on loss-leader pricing. Hess and Gerstner (1987) used a two-period model to study loss-leader pricing in combination with a rain-check policy. Each store sells a selection of "impulse goods" (products bought on sight without price comparisons across stores) and a "shopping good" (for which consumers compare prices and determine which store to visit). Consumers are assumed to visit only one store in each period. They studied the pricing behavior of a representative store and showed that it could be in the best interest of stores to price the shopping good below its marginal cost to attract consumers into the store and profit through the impulse goods. In their framework, the potential candidate for a loss leader is exogenously determined, and retailers are assumed to be symmetric.

Bliss (1988) developed a general model of monopolistic competition among multiproduct retailers and showed that the equilibrium markup for each product depended on the partial derivatives of the demand function. Because the inverse of the matrix of partial derivatives is unsigned, a retailer's optimal pricing could involve loss leading. In a more flexible framework, Lal and Matutes (1994) treated all decisions of retailers and consumers as endogenous. Each retailer carries two goods and decides what prices to charge and which prices to advertise. Consumers choose which store to visit based on the advertised prices and their rational expectations of the unadvertised prices. In equilibrium, retailers advertise the price of the same good, and the advertised price can be below marginal cost. Furthermore, they found that products with lower reservation prices are more natural candidates of loss leaders.

Rajiv et al. (2002) investigated the pricing and advertising decisions of asymmetrically positioned retailers with different service quality. They considered two retailers, each with a loyal segment of customers. It costs more for the loyal customers to shop at the other store. They found that a retailer's quality positioning impacts its frequency and depth of advertised price discounts—relative to the low-quality store, the high-quality retailer offers advertised sales more frequently but with smaller discounts. Their study focused on the frequency and depth of price discounts, while abstracting away from the multiproduct aspect of retailers.

Adding to this line of research, we investigate which retailers are more likely to adopt loss-leader pricing in a competitive environment, and what products are more likely to be selected as loss leaders. To address these questions, it is important to consider a retailer's cross-selling capability and product demand. We show that when the demand for a product increases, retailers with higher cross-selling capabilities are likely to offer a deeper discount.

3. Analytical Model

3.1. A Model of Loss-Leader Pricing

Consider two product markets (market for product A and market for product B) where four retailers (denoted as retailer 1, retailer 2, retailer 3, and retailer 4) compete. Each product attracts a separate set of consumers, but a consumer who visits a retailer to buy one product may become interested in purchasing the other product with probability b. (Whether the consumer indeed purchases the other product and where the consumer purchases it will depend on its prices offered by different retailers, as discussed later in this section.) This creates cross-selling opportunities because attracting a buyer for one product may enable the retailer to generate profit from that buyer for both products. This copurchasing situation may arise because of impulse purchase or proactive recommendations that retailers use to inform consumers of products of interest.

Because our goal is to examine the difference in pricing behavior of the retailers endowed with different cross-selling opportunities, we concentrate our analysis on the scenario in which retailers 1 and 3 carry both products, whereas retailers 2 and 4 carry only one product (product A). In this setting, consumers who come to retailer 1 or 3 for either product may become interested in purchasing the other product, whereas this cross-selling opportunity is not present for retailers 2 and 4. Therefore, under this assumption, retailers 1 and 3 possess higher cross-selling capabilities than retailers 2 and 4, whose cross-selling capabilities are normalized to zero. Thus, our model captures two types of retailers based on their cross-selling capabilities and considers the competition both across the two types of retailers and within each type of retailer.

For simplicity and without affecting the main insights from the model, we assume that product value is v and normalize product cost to zero for both products. Each consumer is assumed to purchase at most one unit of each product. Because we are interested in examining how retailers with different cross-selling capabilities differ in their pricing when product demand changes, we normalize the market size for product B to be one and assume the market

Figure 1 The Spatial Competition Model
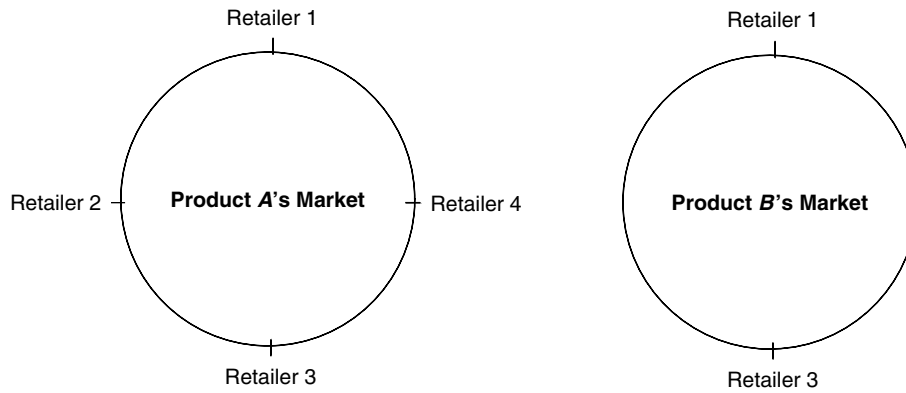
Note. Both consumers and retailers are evenly distributed along the preference circle.

size for product A to be d. Our analysis will focus on how product A's prices offered by different types of retailers change when d changes.

Consumers have different preferences toward the retailers. Following the standard Salop (1979) circle model, we assume that in both markets, consumers are uniformly distributed along a unit circle in the preference space with the four retailers located equidistantly on the circle as shown in Figure 1. For a consumer located at x on the circle, her utility from purchasing either product from retailer j is $v - p_{jk} - t|x - x_j|$, where v denotes the maximum value of the product, p_{jk} represents retailer j's price for product k $(j = 1, 2, 3, 4; k = A, B)$,[2] x_j represents retailer j's location, and t is the disutility for a unit distance in the preference space incurred due to the mismatching of the consumer's location and the retailer's location in the preference space.

To derive retailers' optimal prices, we first derive each retailer's demand and profit function by analyzing consumers' purchase decisions. Consumers in the market for product k $(k = A, B)$ will purchase product k from the retailer that provides the highest nonnegative utility. We assume v is sufficiently large so that all of the consumers in each market will purchase from one of the retailers. Therefore, without cross-selling opportunities, retailer j's profit, denoted by π_j^D (superscript D indicates "direct" because cross-selling profit is not included yet), is given as follows:

$$\pi_1^D(p_{1A}, p_{1B}) = p_{1A}\left(\frac{t/4 - p_{1A} + p_{2A}}{2t} + \frac{t/4 - p_{1A} + p_{4A}}{2t}\right)d$$
$$+ p_{1B}\left(\frac{t/2 - p_{1B} + p_{3B}}{t}\right), \quad (1)$$

$$\pi_2^D(p_{2A}) = p_{2A}\left(\frac{t/4 - p_{2A} + p_{1A}}{2t} + \frac{t/4 - p_{2A} + p_{3A}}{2t}\right)d, \quad (2)$$

[2] Because retailers 2 and 4 do not offer product B, we can consider p_{2B} and p_{4B} to be infinity.

$$\pi_3^D(p_{3A}, p_{3B}) = p_{3A}\left(\frac{t/4 - p_{3A} + p_{2A}}{2t} + \frac{t/4 - p_{3A} + p_{4A}}{2t}\right)d$$
$$+ p_{3B}\left(\frac{t/2 - p_{3B} + p_{1B}}{t}\right), \quad (3)$$

$$\pi_4^D(p_{4A}) = p_{4A}\left(\frac{t/4 - p_{4A} + p_{3A}}{2t} + \frac{t/4 - p_{4A} + p_{1A}}{2t}\right)d, \quad (4)$$

where $|p_{1A} - p_{2A}| < t/4$, $|p_{1A} - p_{4A}| < t/4$, $|p_{2A} - p_{3A}| < t/4$, $|p_{3A} - p_{4A}| < t/4$, and $|p_{1B} - p_{3B}| < t/2$.

To take cross-selling into consideration, we note that the consumers purchasing product k $(k = A, B)$ from retailer 1 or 3 could become interested in the other product (product h, where $h = A, B; h \neq k$) with a probability of b. Thus, in expectation, fraction b of the consumers purchasing product k from retailer 1 or 3 become interested in product h. Out of these consumers, some are willing to take time and effort to compare prices for product h from all retailers offering the product, whereas others are not, possibly due to a higher opportunity cost of time. The net effect of these behaviors, similar in spirit to search cost models, is that some fraction of these copurchasing consumers will simply purchase product h from the same retailer from which they purchase product k as long as it provides nonnegative utility, whereas the remainder of these consumers will purchase product h from the retailer offering the highest nonnegative utility for product h. For simplicity, we assume that these two types of consumers are of equal size, but our results are robust to alternative size assumptions.

Note that this cross-selling opportunity applies to both products; that is, the consumers who purchase product A from retailer 1 or 3 may become interested in and copurchase product B, and likewise, the consumers who purchase product B from retailer 1 or 3 may also become interested in and copurchase product A. Thus, any price decision made by retailer 1 or 3 needs to take both cross-selling opportunities into account.

After taking into account the cross-selling opportunities, we can derive retailer j's total profit function, denoted by π_j, as follows:

$$\pi_1(p_{1A}, p_{1B})$$

$$= (p_{1A} + bp_{1B})\left(\frac{t/4 - p_{1A} + p_{2A}}{2t} + \frac{t/4 - p_{1A} + p_{4A}}{2t}\right)d$$

$$+ \left(p_{1B} + \frac{1}{2}bp_{1A}\right)\left(\frac{t/2 - p_{1B} + p_{3B}}{t}\right)$$

$$+ \frac{1}{2}bp_{1A}\left(\frac{t/4 - p_{1A} + p_{2A}}{2t} + \frac{t/4 - p_{1A} + p_{4A}}{2t}\right), \quad (5)$$

$$\pi_2(p_{2A})$$

$$= p_{2A}\left(\frac{t/4 - p_{2A} + p_{1A}}{2t} + \frac{t/4 - p_{2A} + p_{3A}}{2t}\right)d$$

$$+ \frac{1}{2}bp_{2A}\left(\frac{t/4 - p_{2A} + p_{1A}}{2t} + \frac{t/4 - p_{2A} + p_{3A}}{2t}\right), \quad (6)$$

$$\pi_3(p_{3A}, p_{3B})$$

$$= (p_{3A} + bp_{3B})\left(\frac{t/4 - p_{3A} + p_{2A}}{2t} + \frac{t/4 - p_{3A} + p_{4A}}{2t}\right)d$$

$$+ \left(p_{3B} + \frac{1}{2}bp_{3A}\right)\left(\frac{t/2 - p_{3B} + p_{1B}}{t}\right)$$

$$+ \frac{1}{2}bp_{3A}\left(\frac{t/4 - p_{3A} + p_{2A}}{2t} + \frac{t/4 - p_{3A} + p_{4A}}{2t}\right), \quad (7)$$

$$\pi_4(p_{4A})$$

$$= p_{4A}\left(\frac{t/4 - p_{4A} + p_{3A}}{2t} + \frac{t/4 - p_{4A} + p_{1A}}{2t}\right)d$$

$$+ \frac{1}{2}bp_{4A}\left(\frac{t/4 - p_{4A} + p_{3A}}{2t} + \frac{t/4 - p_{4A} + p_{1A}}{2t}\right), \quad (8)$$

where $|p_{1A} - p_{2A}| < t/4$, $|p_{1A} - p_{4A}| < t/4$, $|p_{2A} - p_{3A}| < t/4$, $|p_{3A} - p_{4A}| < t/4$, and $|p_{1B} - p_{3B}| < t/2$.

When deriving the profit functions in (5)–(8), we focus on the interior scenario where all the retailers have positive demand in the market of product A and focus on the symmetric equilibrium in this scenario. This equilibrium holds as long as the copurchasing probability b is not so large that retailers 1 and 3 would be able to drive retailers 2 and 4 out of product A's market, or that retailers 1 and 3 would charge competitive prices only for one product and sell the other product purely through cross-selling.[3] The exact condition for b is given in the appendix.

---

[3] For example, retailer 1 or 3 may charge a very high price for product A. In this case, even though it cannot attract any consumer from product A's market directly, it may sell product A to the consumers in product B's market who buy product B from the retailer and become interested in product A if the consumers face high opportunity cost to search for lower prices from other retailers.

3.2. Model Predictions and Hypotheses

Each retailer selects the optimal price(s) to maximize its profit given its competitors' pricing strategies. In symmetric equilibrium, the retailers' optimal prices are

$$p_{1A}^* = p_{3A}^* = \frac{(6d - b(6bd^2 + 8d - 7))t}{4(6d - b(2bd(1 + d) - 3))}, \quad (9)$$

$$p_{1B}^* = p_{3B}^* = \frac{(24d + b(12 - 7b + 2(-3 + b)d + 12d^2))t}{8(6d - b(2bd(1 + d) - 3))}, \quad (10)$$

$$p_{2A}^* = p_{4A}^* = \frac{(6d - b(d(4 + b + 4bd) - 5))t}{4(6d - b(2bd(1 + d) - 3))}. \quad (11)$$

The equilibrium is derived in the appendix. A detailed examination of the optimal prices in (9)–(11) yields the following results.

Result 1. When product A's demand increases, retailers with high cross-selling capabilities reduce prices more than retailers with low cross-selling capabilities: $\partial p_{1A}^*/\partial d = \partial p_{3A}^*/\partial d < \partial p_{2A}^*/\partial d = \partial p_{4A}^*/\partial d < 0$.

Result 2. When product A's demand increases, the price dispersion on product A increases as long as the demand for product A is not too low: $(\partial \text{Variance}(p_{1A}^*, p_{2A}^*, p_{3A}^*, p_{4A}^*))/\partial d > 0$ if $d > 1/2$.

Result 3. Retailers with high cross-selling capabilities set prices lower than retailers with low cross-selling capabilities as long as the demand for product A is not too low: $p_{1A}^* = p_{3A}^* < p_{2A}^* = p_{4A}^*$ if $d > 1/2$. In addition, retailers with high cross-selling capabilities charge a price below cost for product A if the demand for product A is sufficiently high, whereas this strategic behavior is not present for retailers with low cross-selling capabilities: $p_{1A}^* = p_{3A}^* < 0$ if $d > (3 - 4b + \sqrt{9 + 2b(-12 + b(8 + 21b))})/(6b^2)$, and $p_{2A}^* = p_{4A}^* \geq 0$.

These results reveal the different pricing strategies followed by retailers with high cross-selling capabilities (represented by retailers 1 and 3 in our model) and those with low cross-selling capabilities (represented by retailers 2 and 4 in our model). The first result suggests that when product A's demand increases, all of the retailers have an incentive to lower prices, but this incentive is especially strong for retailers with high cross-selling capabilities because they can benefit more from the increased demand through cross-selling.

Result 2 shows the impact of a demand increase on price dispersion. As long as the demand for product A is not too low, retailers with high cross-selling capabilities will charge a lower price than retailers with low cross-selling capabilities, as suggested by Result 3. This happens because retailers with high cross-selling capabilities can obtain additional profits through cross-selling product B, which enables

them to afford charging lower prices for product A to attract more consumers to their websites. Combined with Result 1, this suggests that an increase in demand will enlarge the price gap between retailers. Therefore product A will exhibit higher price dispersion when its demand increases.

Result 3 suggests that in equilibrium, retailers 2 and 4 are never willing to lower their prices of product A below cost (because they cannot recoup the loss through cross-selling), but retailers 1 and 3 may choose to sell product A below cost and to profit from cross-selling product B. This tactic of lowering the price for one product, sometimes below cost, while benefiting from the sales of other products resembles the loss-leader pricing strategy in the marketing literature (e.g., Lal and Matutes 1994).

This set of results has two direct implications that can be empirically tested using actual market data. First, retailers with high cross-selling capabilities are likely to reduce prices more than retailers with low cross-selling capabilities after the demand for a product increases. This directly follows Result 1 and suggests a moderating role of a retailer's cross-selling capability to its pricing behavior when product demand changes. We thus propose the following:

HYPOTHESIS 1 (H1). *Retailers with high cross-selling capabilities are likely to lower prices more than retailers with low cross-selling capabilities when the demand for a product increases.*

Second, by understanding the pricing behavior of different retailers as product demand changes, our results also have implications for the pattern of price dispersion. Result 2 suggests that a demand increase would increase price dispersion across retailers, as long as the product demand is not too low. Furthermore, our model also indicates that the increase in the price dispersion is mainly driven by the price difference between retailers with high cross-selling capabilities and those with low cross-selling capabilities, because the price variation within each group is zero given our symmetric model setup. In reality, other factors may introduce variation in prices even among retailers with the same cross-selling capability. But the prediction that the increase in the overall price dispersion will be mainly driven by the price difference between the two retailer groups should be relatively robust. We thus propose the following:

HYPOTHESIS 2A (H2A). *Demand increases lead to higher observed price dispersion across all retailers.*

HYPOTHESIS 2B (H2B). *The increase in overall observed price dispersion when product demand increases is mainly driven by the price difference between retailers with high cross-selling capabilities and those with low cross-selling capabilities.*

4. Empirical Analysis

Our theoretical model provides an analytical framework to predict retailers' pricing behavior in response to demand change given their different cross-selling capabilities. In this section, we empirically test our hypotheses using data collected in the online book retailing industry and investigate the changes in retailers' pricing strategies when a book makes it to the best-seller list.

4.1. Data Description

To test our model predictions, we need to observe retailers' pricing behavior under different demand conditions. In this study, we use best-seller status to identify demand change (Sorensen 2007). We started with all of the hardcover books in *Books in Print* that meet the following criteria: (1) the book is in an active status, (2) the book is written in English, and (3) the book was published between January 2000 and February 2004. To capture future best sellers, we focused on active books that received at least one review indicated by *Books in Print* and had an active profile on Amazon.com,[4] or that appeared at least once in the *Publishers Weekly* best-seller list between January 1, 2000, and February 9, 2004. This produced a sample of 2,651 books. We focus on hardcover books because they play an important role in book retailers' pricing strategies to generate traffic and attract customers (Raff 2000).

For each book in our sample, we collected its price data from all of the online bookstores listed on pricescan.com, one of the largest price comparison sites during the study period. It covered 22 online bookstores ranging from large retailers such as Amazon and Barnes & Noble, to less known stores such as HamiltonBook and Reiter's. We used a software agent to collect price data from pricescan.com for the 2,651 books in each week between March 5, 2004, and October 25, 2004. The information we collected includes book prices, shipping costs, and availability from each of the 22 online bookstores. To prevent our analysis from being affected by the presence of obsolete or out-of-stock books—prices that may not be actively monitored by these retailers—our final sample for price analysis consists of only 1,614 books with the following features: (1) the book's shipping status is "ships within 24 hours" on both Amazon and Barnes & Noble during the entire data collection period, and (2) the book appears on pricescan.com during the entire data collection period.

---

[4] We required each book's sales rank at Amazon.com to be lower than 100,000. For books published before 2004, we required (1) an average of at least 10 customer reviews on Amazon.com per year and (2) at least one review on Amazon.com by the end of first month after release.

Table 1   Summary Statistics across Bookstores

Bookstore	(1) Average price ($)	(2) Average shipping cost ($)	(3) Average price as a percentage of list price (%)	(4) Multicategory retailer	(5) Book coverage (%)
A1Books	12.99	2.99	52.40	No	99.92
AllDirect	15.42	3.45	62.29	No	96.87
Amazon	17.19	4.48	69.72	Yes	100.00
Barnes & Noble	20.39	3.54	82.43	No	99.51
BCY	16.52	4.00	66.67	No	91.21
BiggerBooks	18.15	3.97	73.36	No	100.00
Blackwell	24.70	0.00	99.78	No	99.06
BookCloseouts	6.68	4.00	27.83	No	69.37
Books-A-Million	18.36	3.95	74.18	No	100.00
Buy	16.43	2.40	66.48	Yes	96.94
eCampus	18.52	3.97	74.88	No	99.81
HamiltonBook	10.92	3.90	41.63	No	55.60
Overstock	14.45	1.40	59.57	Yes	94.42
Page 1	17.89	3.95	72.74	No	97.10
Powell's	20.08	3.50	81.36	No	98.57
Reiter's	27.74	5.00	100.10	No	21.46
San Diego Technical	20.81	5.00	84.85	No	91.85
Textbookx	16.53	3.58	66.72	No	97.70
TotalCampus	16.25	3.49	66.79	No	65.30
VarsityBooks	24.74	4.95	100.03	No	99.81
Walmart	15.67	2.48	63.57	Yes	98.00
WordsWorth	22.56	4.95	91.88	No	97.66
All bookstores	17.86	3.59	71.79	N/A	89.55

Columns (1)–(3) of Table 1 provide the average book price, average shipping cost, and average price as a percentage of list price for each of the 22 online bookstores. In our sample, the average book price is $17.86, which reflects a 28.21% discount off of the list price. There are, however, significant price variations across bookstores. Table 1 also shows significant variations in shipping costs, ranging from free shipping offered by Blackwell to as high as a $5.00 charge by Reiter's and San Diego Technical.

To examine the impact of a retailer's cross-selling capability on its choice of loss-leader strategy, we need to measure each retailer's cross-selling capability. Because it is not directly observable, we use product breadth to approximate a retailer's cross-selling capability. In particular, we construct two variables to assess the breadth of product offering at each retailer. The first variable is a dummy variable that identifies whether a retailer is a multicategory retailer that offers more product categories than what are typically available in bookstores (i.e., books, music CDs, and DVDs). This variable captures a retailer's product breadth at the category level. The second variable is the percentage of the 2,651 books selected for our study that are available at each retailer.[5] This variable measures product breadth within the book category.

Columns (4) and (5) of Table 1 show the multicategory retailer status and book coverage percentage for each bookstore. Four of the 22 bookstores in our study—Amazon, Buy.com, Overstock.com, and Walmart—are multicategory retailers. The summary statistics also show significant variations in book coverage. Whereas some bookstores (e.g., Books-A-Million and Bigger Books) offer all of the 2,651 books for sale, some other bookstores (e.g., HamiltonBook and Reiter's) carry less than 60% of them. The significant variation in book coverage across bookstores is consistent with what has been observed in prior studies (e.g., Clay et al. 2001).

Our analytical model predicts that retailers follow different pricing strategies when product demand increases. For our empirical analysis, we consider the changes in retailers' pricing behaviors when a book makes it to the best-seller list (Sorensen 2007). We collected the entire *Publishers Weekly* best-seller list for the study period. We then matched books in our sample with those in the best-seller list to create a dummy variable *BestSellerDummy* to identify whether a book is on the best-seller list for a given week.

We also collected book-specific characteristics from Amazon.com, including publication date, paperback availability and the publication dates of corresponding paperbacks (if available). We created two variables from the book characteristic data: *Age* measures the number of days passed since the release

---

[5] If we use the percentage of the 1,614 books used for analysis, our results are qualitatively the same.

Table 2   Summary Statistics and Correlation Matrix for Major Independent Variables

Variables	Mean	Std. dev.	Min	Max	(1)	(2)	(3)	(4)	(5)
(1) *BestSellerDummy*	0.01	0.11	0.00	1.00	1				
(2) log *Age*	5.72	0.80	1.95	7.47	−0.00	1			
(3) *PaperbackDummy*	0.35	0.48	0.00	1.00	−0.02	0.61	1		
(4) *Multicategory*	0.17	0.38	0.00	1.00	0.00	0.00	0.00	1	
(5) *BookCoverage*	0.90	0.19	0.21	1.00	0.00	0.00	0.00	0.23	1

of the hardcover book, and *PaperbackDummy* identifies whether a book faces competition from its paperback editions in a given week. Table 2 provides the summary statistics of the variables.

To corroborate the robustness of our empirical findings against a number of alternative explanations, we also collected additional supplementary data on retailer characteristics and conducted a series of robustness analyses. The detail is provided in §4.4.

4.2.   Empirical Models

To assess how changes in book demand influence retailers' pricing strategies, we first analyze H1, which suggests that retailers with high cross-selling capabilities are likely to reduce the prices of best sellers more than retailers with low cross-selling capabilities. To test the hypothesis empirically, we develop a fixed-effects log-linear model for book price as follows:

$$
\begin{aligned}
\log(Price_{ijt}) &= \beta_1 BestSellerDummy_{it} + \beta_2 BestSellerDummy_{it} \\
&\quad \times Multicategory_j + \beta_3 BestSellerDummy_{it} \\
&\quad \times BookCoverage_j + \beta_4 \log Age_{it} + \beta_5 PaperbackDummy_{it} \\
&\quad + \theta_i + \delta_j + \tau_t + \varepsilon_{ijt}.
\end{aligned}
\tag{12}
$$

In Equation (12), the dependent variable is the log transformation of book *i*'s price at retailer *j* in week *t*.[6] We use log transformation because there is significant variation in book prices, and it allows us to normalize the price distribution. The main independent variables in Equation (12) are $BestSellerDummy_{it}$, $BestSellerDummy_{it} \times Multicategory_j$, and $BestSellerDummy_{it} \times BookCoverage_j$. The dummy variable $BestSellerDummy_{it}$ identifies whether book *i* appears on the best-seller list during week *t*. The variables $Multicategory_j$ and $BookCoverage_j$ are proxies for retailer *j*'s cross-selling capability. H1 suggests that, when product demand increases, retailers with higher cross-selling capabilities will reduce price more

than retailers with lower cross-selling capabilities. We therefore expect the coefficients on the interaction terms between $BestSellerDummy_{it}$ and the two proxies for cross-selling capability, β_2 and β_3, to be negative.

Other factors may also influence book prices. Thus, in Equation (12) we include book fixed effects, time fixed effects, and retailer fixed effects to control for book-specific heterogeneity, seasonality, and retailer-specific heterogeneity that could influence book prices but are unobservable to researchers. These fixed effects control for many drivers of online price dispersion identified in previous literature (e.g., see a summary in Pan et al. 2004) and allow us to leverage demand changes to identify the impact of cross-selling capability on retailers' pricing strategies. These fixed effects also capture the main effects of $Multicategory_j$ and $BookCoverage$. Because $Multicategory_j$ and $BookCoverage_j$ are time invariant for each retailer, each of them can be expressed as a linear combination of retailer dummies and thus are collinear with retailer fixed effects. The issue of estimating interaction effects when main effects are collinear with fixed effects was discussed by Allison (2005). We follow the approach used in prior literature (e.g., Altonji and Dunn 1996, Bertrand et al. 2000, Rothstein 2006) to report the estimation results by including retailer fixed effects and omitting the collinear main effects.

In Equation (12), we include the log transformation of book age to control for any trend in prices over product life cycles, and the paperback dummy variable to control for possible price changes in response to the availability of paperback editions. We use Huber–White robust clustered standard errors to adjust for any potential heteroskedasticity and serial correlation within panel in all of the regressions in this paper (Wooldridge 2002). Because the sampling probabilities are different for best sellers and non–best sellers, we estimate the regression Equation (12) using both ordinary least squares (OLS) and weighted least squares (WLS).[7]

---

[6] We use a book's retail price as the dependent variable in the regression. Alternatively, we can use a book's retail price plus shipping cost as the dependent variable. Because shipping costs remain the same during the study period, our analysis of the price changes in response to demand shocks is not affected by the choice of the dependent variable.

[7] In our sample, the sampling probability for best sellers is approximately 18.05% and the sampling probability for non–best sellers is approximately 0.73%. We use the inverse of the sampling probability as weight for each observation in our WLS estimation.

To assess H2A and H2B on price dispersion, we develop a fixed effects model for price dispersion to examine the direction in which the overall price dispersion changes when product demand increases. This design allows us to control for unobserved book-specific characteristics that may influence price dispersion. We also note that a number of statistics such as standard deviation and price range have been used in prior studies to measure price variation (e.g., Clay et al. 2001, Baye et al. 2004). We choose price variance in this study because it is easily decomposable, and its decomposition allows us to better understand the underlying factors that contribute to the price variation in later analysis. Similar to Equation (12), we apply a log transformation on $PriceVariance_{it}$ to normalize its distribution. The key independent variable is the dummy variable for best sellers because our goal is to assess how demand increases influence price dispersion. Besides book fixed effects, we also include time fixed effects, book age, and the paperback dummy to control for product-specific factors and seasonality that may influence price dispersion over time:

$$\log(PriceVariance_{it})$$
$$= \beta_1\, BestSellerDummy_{it} + \beta_2 \log Age_{it}$$
$$+ \beta_3\, PaperbackDummy_{it} + \theta_i + \tau_t + \varepsilon_{it}. \quad (13)$$

Equation (13) does not include retailer fixed effects or retailer-specific characteristics because price variance is a market level measure. H2A suggests that the coefficient on the best-seller dummy, β_1, shall be positive.

To examine whether the difference in pricing behavior between retailers with high cross-selling capabilities and those with low cross-selling capabilities is the main contributing factor to the changes in price dispersion, we classify retailers into two groups of equal size based on their cross-selling capabilities measured by multicategory status and book coverage. Because multicategory retailers such as Amazon.com and Walmart offer a huge variety of products beyond the book category, they tend to have better opportunities for cross-selling. Therefore, the high cross-selling capability group includes all four of the multicategory retailers and the top seven book retailers ranked on their book coverage.

We then decompose the overall price dispersion into three components: the first two correspond to the price variation within each group of retailers with similar cross-selling, capabilities, and the third corresponds to the price difference between the two groups of retailers with different cross-selling capabilities. This decomposition allows us to develop a better understanding of what drives the observed increase in price dispersion:

$$PriceVariance_{it}$$
$$= a\, PriceVariance\_HIGH_{it} + b\, PriceVariance\_LOW_{it}$$
$$+ ab(AvgPrice\_HIGH_{it} - AvgPrice\_LOW_{it})^2,$$

where

$$a = \frac{NumStores\_HIGH}{NumStores\_HIGH + NumStores\_LOW} \quad \text{and}$$

$$b = \frac{NumStores\_LOW}{NumStores\_HIGH + NumStores\_LOW}.$$

Here, $PriceVariance\_HIGH_{it}$, $AvgPrice\_HIGH_{it}$, and $NumStores\_HIGH$ refer to the price variance, the average price, and the number of stores for the group of retailers with high cross-selling capabilities, respectively. Similarly, $PriceVariance\_LOW_{it}$, $AvgPrice\_LOW_{it}$, and $NumStores\_LOW$ represent the price variance, the average price, and the number of stores for the group of retailers with low cross-selling capability, respectively.

To understand what drives the changes in the overall price variation, we assess the three components of price variance separately:

$$\log(PriceVariance\_HIGH_{it})$$
$$= \beta_1^H\, BestSellerDummy_{it} + \beta_2^H \log Age_{it}$$
$$+ \beta_3^H\, PaperbackDummy_{it} + \theta_i^H + \tau_t^H + \varepsilon_{it}^H, \quad (14)$$
$$\log(PriceVariance\_LOW_{it})$$
$$= \beta_1^L\, BestSellerDummy_{it} + \beta_2^L \log Age_{it}$$
$$+ \beta_3^L\, PaperbackDummy_{it} + \theta_i^L + \tau_t^L + \varepsilon_{it}^L, \quad (15)$$
$$\log[(AvgPrice\_HIGH_{it} - AvgPrice\_LOW_{it})^2]$$
$$= \beta_1^P\, BestSellerDummy_{it} + \beta_2^P \log Age_{it}$$
$$+ \beta_3^P\, PaperbackDummy_{it} + \theta_i^P + \tau_t^P + \varepsilon_{it}^P. \quad (16)$$

Because the regression error terms may be correlated across the three equations, we utilize seemingly unrelated regressions to address this issue. We also use Huber–White robust clustered standard errors to adjust for any potential heteroskedasticity and serial correlation within panel. H2B suggests that the main driver of the price dispersion is the increase in price difference between retailers with different cross-selling capabilities. We thus expect coefficient β_1^P to be positive and significantly higher than coefficients β_1^H and β_1^L.

4.3. Regression Results
Table 3 reports the regression results for Equation (12), which describes retailers' pricing behavior in response to demand increases and the moderating effect of cross-selling capability. The first column

Table 3   Regression Analysis of Pricing Strategy

	OLS	WLS
BestSellerDummy	0.320***	0.326***
	(0.075)	(0.076)
BestSellerDummy × Multicategory	−0.073***	−0.074***
	(0.011)	(0.011)
BestSellerDummy × BookCoverage	−0.335***	−0.342***
	(0.078)	(0.079)
PaperbackDummy	−0.004	−0.004
	(0.003)	(0.003)
log *Age*	0.023***	0.024***
	(0.003)	(0.003)
R^2	73.14%	73.22%
Number of observations	904,307	904,307

Notes. Coefficients for product fixed effects, time fixed effects, and store fixed effects are omitted from the table. Huber–White robust clustered standard errors are in parentheses.

     ***$p < 0.01$.

presents the OLS regression results, whereas the second column presents the WLS regression results.

The coefficients of the interaction terms between the best-seller dummy and the two proxies for cross-selling capability are negative and significant in both OLS and WLS regressions. The negative coefficients suggest that retailers with high cross-selling capabilities decrease price significantly more than retailers with low cross-selling capabilities when a book hits the best-seller list, supporting H1.

The first column of Table 4 presents the regression result for Equation (13) that analyzes the changes in price variation across all of the online book retailers when a book makes it to the best-seller list. The result shows that the overall price dispersion increases about 6% on average when a book makes it to the best-seller list, supporting our H2A.

The second, third, and fourth columns of Table 4 report the regression results for Equations (14)–(16), which isolate the causes of this increase in price variation by assessing the impact of *BestSellerDummy* on each of the three components of the overall price variance. These results suggest that appearing on the best-seller list has a substantial impact on the price difference between the two groups of retailers (grouped by their cross-selling capabilities) with an increase of about 29%, whereas the price dispersion within each group does not change significantly when a book makes it to the best-seller list. A direct comparison on the coefficient of *BestSellerDummy* across the three equations suggests that $\beta_1^P > \beta_1^H$ and $\beta_1^P > \beta_1^L$, both significant at the 0.05 level.

Taken together, these results support H2B and indicate that the increased price dispersion observed on best sellers is mostly driven by the different pricing strategies employed by different types of retailers—retailers with high cross-selling capabilities are likely to use best sellers as loss leaders, whereas the remaining retailers have less incentive to do so.

4.4. Robustness Analysis

To corroborate the robustness of our results, we perform several robustness analyses on our results. First, considering the possibility that it takes time for retailers to adjust prices after demand shocks, we rerun our analysis using lagged *BestSellerDummy*. Using lagged *BestSellerDummy* also allows us to alleviate the concern of reverse causality. Second, product demand before a book makes it to the best-seller list could also influence retailers' pricing behavior. We thus include each book's sales rank at Amazon.com in the previous week and exclude Amazon.com price observations in our analysis to control for this effect. Third, in our main analysis, we include the *PaperbackDummy*

Table 4   Regression Analysis of Price Dispersion

	Overall price dispersion	Price dispersion among retailers with high cross-selling capability	Price dispersion among retailers with low cross-selling capability	Price difference between the two retailer groups
BestSellerDummy	0.055***	0.048	−0.015	0.503**
	(0.018)	(0.045)	(0.033)	(0.207)
PaperbackDummy	0.042***	−0.032	0.103***	−0.153
	(0.012)	(0.023)	(0.017)	(0.096)
log *Age*	0.078***	−0.166***	0.182***	1.333***
	(0.015)	(0.025)	(0.023)	(0.102)
$R^{2\text{a}}$	87.75%		72.95%	
Number of observations[b]	54,876		54,876	

Notes. Coefficients for product fixed effects and time fixed effects are omitted from the table. Huber–White robust clustered standard errors are in parentheses.

     [a]For seemingly unrelated regressions, the system weighted R^2 provided by SAS is reported.

     [b]We have fewer observations than Table 3 because the regressions for prices are conducted at the retailer level, whereas the regressions for price dispersion are conducted at the market level.

     **$p < 0.05$; ***$p < 0.01$.

to capture the potential competition between different book editions, but it cannot capture the effect of a corresponding paperback making it to the best-seller list. Therefore we test the robustness of our results by adding a dummy variable that indicates whether a corresponding paperback makes it to the best-seller list. Fourth, when we group retailers according to their cross-selling capabilities, we split the 22 online retailers naturally into two groups of equal size—the high group includes the 4 multicategory retailers and the top 7 book retailers ranked on book coverage, and the low group includes the remaining 11 book retailers. We check to make sure that our results are not sensitive to the way we group the retailers. For example, we can split the 18 book retailers equally according to their book coverage; thus the high group includes the 4 multicategory retailers and the top 9 book retailers ranked on book coverage, and the low group includes the bottom 9 book retailers. Our results are not affected. Finally, to make sure that our results are robust to different measures of price dispersion, we rerun the price dispersion analysis using the coefficient of variation as the measure of price dispersion.

In addition, although we use product breadth as a proxy of a retailer's cross-selling capability to examine the retailer's incentive to adopt loss-leader pricing, other retailer characteristics may also affect the retailer's pricing strategy on popular products. For example, previous literature suggests that retailers' reputation, scale, and physical store presence can potentially affect retailers' pricing strategies (Bliss 1988, Brynjolfsson et al. 2009, Forman et al. 2009, Raju et al. 1990, Smith and Brynjolfsson 2001, Tang and Xing 2001). We thus construct two variables to control for retailers' reputation and scale effects during our study period: (1) *Reach* measures the percentage of Internet users who visit a retailer and was collected as a one-time measure from Alexa.com, and (2) *Popularity* indicates the relative number of searches conducted on a retailer by U.S. consumers and was collected from Google Insights for Search based on average Web search interest for a retailer's URL (without the "www" prefix). We use non-time-varying measures because conceptually a retailer's reputation or scale should be stable in the short run. Our observations during data collection were also consistent with this notion. To control for the effect of physical store presence, we construct a dummy variable, *DualChannel*, that equals one for the retailers with both online and offline stores. When incorporating the three variables into Equation (12), we note that their main effects are captured by the retailer fixed effects; therefore, we include the interaction terms between *BestSellerDummy* and these three variables in Equation (12) to control for any moderating effect

that reputation, scale, or physical store presence could have on a retailer's incentive to engage in loss-leading prices.

Finally, it is important to investigate whether our results can be explained by cost differences between retailers due to variation in wholesale volume discount or retailer supply chain efficiency. After extensive research and additional robustness analysis, we are confident that the observed changes in retail prices and price dispersion when a book makes it to the best-seller list cannot be driven by cost differences because of the following five reasons.

(1) Publishers used to offer secret deals to large retailers or require a hefty volume to get their maximum discounts. In 1994, the American Booksellers Association (ABA) and independent bookstores filed lawsuits against major publishers to stop such practices (Clark 1998). To settle these lawsuits, publishers agreed to publish their discount schedules in ABA Book Buyer's Handbook,[8] that are actively monitored by ABA and independent bookstores to ensure a level playing field between large and small retailers. After the lawsuits, the required volume to achieve the maximum wholesale discounts has dropped significantly and should be easily reached by the retailers in our study, if not by all independent bookstores. For example, the maximum discount by Clear Light Publishing can be achieved as long as a retailer buys at least 10 books, across titles.[9] Even the small retailers in our study are really not that small—they offered an extensive selection of books on their websites. Furthermore, wholesale discounts typically apply to the total number of books across titles, and supply chain efficiency is also reflected across book titles. Therefore, if a retailer has cost advantage, this advantage should be reflected in its prices of all books. We thus calculate the average price of all books for each retailer and then compare the average price among the group of retailers with high cross-selling capabilities with that among the group with low cross-selling capabilities. The two-sample *t*-test suggests that the price difference between the two groups is not statistically significant (*p*-value = 0.33), further confirming that the retailers in our study do not differ significantly in their costs.

---

[8] See, e.g., *American Booksellers Association, Inc. v. Houghton Mifflin Company, Inc.*, 95-2 Trade Case (CCH) P71, 217 (S.D.N.Y. 1995).

[9] Clear Light Publishing offers a discount of 40% for one to nine mixed titles and 43% for more (http://www.clearlightbooks.com/promotions.htm, accessed October 25, 2012). Similar discounts are offered by Macmillan (http://us.macmillan.com/CMS400/uploadedFiles/MacmillanSite/Non-Menu_Items/bookseller_services/US_Retail_Discount_Schedule_%28Non-Returnable%29_2011.pdf, accessed October 25, 2012) and Workman (http://www.workman.com/resources/WorkmanRetailBookDiscount.pdf, accessed October 25, 2012).

(2) Our empirical strategy of focusing on the changes in retailers' pricing strategies when a book gets on the best-seller list also helps alleviate the concern of cost differences. Even if some retailers receive larger wholesale discounts than others or if some retailers are more efficient than others in supply chain management, such advantage does not change when an individual book makes it to the best-seller list. In particular, considering that large stores, such as Amazon.com, Walmart, and Books-A-Million, should easily reach the required volume to achieve the maximum wholesale discount regardless of whether or not a book is on the best-seller list, a book making it to the best-seller list should not lead to any cost change. Therefore, cost difference hypothesis cannot explain their significant price decreases observed after a book makes it to the best-seller list. These retailers' high cross-selling capabilities and their motives to use popular products as loss leaders, however, provide a viable explanation. We also analyze each retailer's pricing behavior individually when books make it to the best-seller list. Specifically, for each retailer, we calculate its percentage price change when books get on the best-seller list. We then compare the percentage price change between the two groups of retailers with different cross-selling capabilities. The two-sample t-test suggests that when a book makes it to the best-seller list, the retailers in the high cross-selling group lower prices significantly more than the retailers in the low cross-selling group (p-value = 0.04). Given that in this analysis we focus on the same retailer's price change for the same set of books when these books get on the best-seller list, wholesale cost or supply chain efficiency should stay constant and therefore cannot be the driver behind the observed price change.

(3) The magnitude of retail price discounts observed for best sellers cannot be explained by the volume discount on cost alone. The maximum wholesale discount is typically less than 50%,[10] but we often observe that best sellers are sold at a discount of 60% or more, i.e., below cost even if we ignore other overhead costs. These deep discounts offered on best sellers also should not be the result of inventory management where retailers dump books at the later stage of product life cycles, because of the following two observations. First, we observe that while retailers tended to lower prices when books made it to the best-seller list, they usually raised their prices

back after books dropped off the best-seller list.[11] Second, we observe that former best sellers were still in active sales for a long period after they dropped off the best-seller list.[12] These observations provide direct evidence against the inventory management argument. The practice of offering retail prices lower than wholesale prices (and thus incurring losses) for best sellers, however, is supportive of the loss-leader strategy mentioned repeatedly in the industry.[13]

(4) As discussed earlier, our results are not affected after incorporating the moderating effects of *Reach* and *Popularity*. Because these two variables provide reasonable controls for retailer size, it suggests that our results are robust after controlling for the scale effect on costs.

(5) To further alleviate this concern, we conduct an additional robustness check in which we use a different measure of cross-selling capability while explicitly controlling for the scale effect on costs. This robustness check is based on the 2004 comScore transaction data. This data set provides detailed transaction information for 50,000 Internet users across the United States for a subset of the retailers in our sample. Removing the retailers with too few transactions, seven retailers remain—Amazon.com, Barnes & Noble, Books-A-Million, Buy.com, eCampus, Overstock.com, and Walmart. Although this data set only provides transactions for a subset of the retailers, it works well for this robustness check because we can thus focus on large retailers with minimal cost differences. Using the comScore data, we construct a more refined measure of cross-selling capability that is less confounded with retailer size—the percentage of book orders that have multiple items. We also use the two measures on retailer reputation and scale to control for the scale effect on costs in this analysis.

---

[10] From various sources we have seen a typical range of 40%–50% given by the industry insiders we interviewed, 45%–50% provided by the American Booksellers Association (2009), 40%–44% published by Clear Light, and 40%–46% published by Macmillan. Even if retailers order books nonreturnable and pay freight, Workman discounts no more than 54%.

[11] Among the large retailers used for our later robustness analysis discussed in (5), we find significant price decreases when books make it to the best-seller list and significant price increases after books drop off the list. There is no significant difference between the prices before books make the list and the prices after books drop off of the list.

[12] After books dropped off of the best-seller list, their sales ranks on Amazon.com remained much higher than the sales ranks of average books until the end of our study period (which corresponds to an average of 20 weeks after they dropped off of the list).

[13] See the comments by Michael Norris mentioned earlier (§1). Also, Steve Riggio, then vice chairman of Barnes & Noble, once remarked: "best sellers, which make up only about 3% of sales, have long been treated as loss leaders" (Rose and Quick 2000, B1). After a recent price war between Amazon.com, Walmart, and Target, the ABA believed that publishers did not offer any special terms to the big-box retailers, and complained, "They're using our most important products—mega bestsellers, which, ironically, are the most expensive books for publishers to bring to market—as a loss leader to attract customers to buy other, more profitable merchandise" (American Booksellers Association 2009).

Table 5   Robustness Analysis of Pricing Strategy

	Use lagged *BestSellerDummy*	Control for lagged *AmazonSalesRank*	Control for *Paperback BestSellerDummy*	Control for reputation and dual channel presence
BestSellerDummy	0.319*** (0.074)	0.314*** (0.079)	0.320*** (0.075)	0.203*** (0.067)
BestSellerDummy × *Multicategory*	−0.072*** (0.012)	−0.071*** (0.014)	−0.073*** (0.011)	−0.045*** (0.014)
BestSellerDummy × *BookCoverage*	−0.337*** (0.078)	−0.336*** (0.082)	−0.335*** (0.078)	−0.326*** (0.079)
BestSellerDummy × log *Reach*				0.026*** (0.009)
BestSellerDummy × log *Popularity*				−0.041*** (0.010)
BestSellerDummy × *DualChannel*				−0.020* (0.011)
PaperbackDummy	−0.003 (0.003)	−0.004 (0.003)	−0.004 (0.003)	−0.003 (0.004)
PaperbackBestSellerDummy			−0.005 (0.010)	
log *Age*	0.026*** (0.004)	0.029*** (0.004)	0.023*** (0.003)	0.027*** (0.004)
log *AmazonSalesRank* (lagged)		0.001* (0.001)		
R^2	73.05%	73.15%	73.14%	71.53%
Number of observations[a]	874,064	820,934	904,307	660,943

Notes. Coefficients for product fixed effects, time fixed effects, and store fixed effects are omitted from the table. Huber–White robust clustered standard errors are in parentheses.

[a]We have fewer observations than in Table 3 because the observations that do not have lagged *BestSellerDummy*, lagged *Amazon-SalesRank*, *Reach*, or *Popularity* data are excluded from analysis. In addition, Amazon prices data are excluded from analysis in the second column, where the lagged *AmazonSalesRank* is added to control for the prior period demand.

$^*p < 0.1$; $^{***}p < 0.01$.

The results of all the robustness analyses are reported in Tables 5–10. Our main results hold qualitatively: (1) a retailer's cross-selling capability moderates its pricing strategies on best sellers so that retailers with high cross-selling capabilities reduce prices significantly more than retailers with low cross-selling capabilities when a book makes it to the best-seller list, and (2) the overall price variation increases when a book makes it to the best-seller list and the increase is mainly driven by the significant increase in the price difference between the two groups of retailers.

5. Conclusion
This paper contributes to the price dispersion literature by providing a new explanation for the increased price dispersion often observed when product demand increases. We show that the heterogeneity in retailers' cross-selling capabilities can be a major driving force behind this increased price dispersion. Specifically, our model indicates that, in equilibrium, retailers with higher cross-selling capabilities are more likely to adopt loss-leader pricing on high demand products, whereas retailers with low cross-selling capabilities have less incentive to do so.

Consequently, when demand jumps, the price difference between the two groups of retailers widens, resulting in higher price dispersion.

The prediction of our model is consistent with empirical evidence from the online book retailing industry. By examining book prices from 22 online retailers, we observe higher price dispersion when books make the *Publishers Weekly* best-seller list. Using product breadth as a proxy for cross-selling capability, we find that retailers with high cross-selling capabilities reduce prices when a book makes it to the best-seller list significantly more than retailers with low cross-selling capabilities. Furthermore, the increased price dispersion for best sellers is mainly driven by the price difference between the two groups of retailers divided based on their cross-selling capabilities. Our findings are robust against a number of alternative explanations.

Our analysis also contributes to the research stream on loss-leader pricing. Although loss-leader pricing has been examined before, previous studies have not focused on how the incentive for loss-leader pricing varies across different retailers when product demand changes. This paper provides a framework to understand such incentive given retailers' asymmetric

Table 6   **Robustness Analysis of Price Dispersion Using Prior Period Best-Seller Status**

	Overall price dispersion	Price dispersion among retailers with high cross-selling capability	Price dispersion among retailers with low cross-selling capability	Price difference between the two retailer groups
BestSellerDummy	0.042** (0.018)	0.045 (0.043)	−0.038 (0.037)	0.468** (0.233)
PaperbackDummy	0.037*** (0.012)	−0.034 (0.024)	0.093*** (0.017)	−0.143 (0.100)
log *Age*	0.108*** (0.017)	−0.208*** (0.029)	0.225*** (0.026)	1.898*** (0.124)
R^{2a}	87.98%		73.39%	
Number of observations[b]	53,262		53,262	

Notes. Coefficients for product fixed effects time fixed effects, and store fixed effects, are omitted from the table. Huber–White robust clustered standard errors are in parentheses.

   [a]For seemingly unrelated regressions, the systems weighted R^2 provided by SAS is reported.

   [b]We have fewer observations than in Table 4 because the observations that do not have lagged data are excluded from analysis.

   **$p < 0.05$; ***$p < 0.01$.

cross-selling capabilities and uses changes in product demand condition to identify retailers' pricing strategies.

Our analysis has significant practical implications. Consumers often purchase multiple products together, and different retailers carry different product lines. Therefore, the "law of one price" cannot easily prevail in this environment. A retailer's pricing strategy shall be based on its ability to cross-sell other products, as we show in this paper. While reducing prices on popular products to attract customers can be an effective strategy for some retailers, it may not be applicable to all retailers. Our model and empirical analysis indicate that the optimal strategy for a retailer depends on its cross-selling capability relative to its competitors.

The intuition from this analysis applies to other product categories as well. For example, online electronic retailers such as Newegg.com and TigerDirect.com periodically offer popular personal computer components at deep discounts, in hopes of attracting customers to their stores to purchase other (more profitable) products on the same shopping trip. Such practice can increase price dispersion on the products used as loss leaders.

Our study has a number of limitations that provide opportunities for future research. First, our empirical results are drawn from observations of the online book retailing industry in 2004. Although anecdotal evidence suggests that significant price dispersion still persists in today's online book industry as does the practice of loss-leader pricing, whether these results

Table 7   **Robustness Analysis of Price Dispersion Controlled for Prior Period Demand**

	Overall price dispersion	Price dispersion among retailers with high cross-selling capability	Price dispersion among retailers with low cross-selling capability	Price difference between the two retailer groups
BestSellerDummy	0.047** (0.019)	0.029 (0.049)	−0.001 (0.030)	0.382* (0.228)
PaperbackDummy	0.037*** (0.012)	−0.027 (0.023)	0.089*** (0.017)	−0.157 (0.100)
log *Age*	0.107*** (0.017)	−0.151*** (0.029)	0.192*** (0.026)	1.766*** (0.129)
log *AmazonSalesRank* (lagged)	0.001 (0.003)	−0.041*** (0.005)	0.024*** (0.004)	0.093*** (0.023)
R^{2a}	87.98%		73.43%	
Number of observations[b]	53,262		53,262	

Notes. Coefficients for product fixed effects and time fixed effects are omitted from the table. Huber–White robust clustered standard errors are in parentheses.

   [a]For seemingly unrelated regressions, the system weighted R^2 provided by SAS is reported.

   [b]We have fewer observations than in Table 4 because the observations that do not have lagged *AmazonSalesRank* data are excluded from analysis.

   *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

Table 8   Robustness Analysis of Price Dispersion Under Alternative Grouping of Retailers

	Overall price dispersion	Price dispersion among retailers with high cross-selling capability	Price dispersion among retailers with low cross-selling capability	Price difference between the two retailer groups
BestSellerDummy	0.055***	0.081***	−0.050	0.461***
	(0.018)	(0.031)	(0.042)	(0.151)
PaperbackDummy	0.042***	−0.033*	0.133***	−0.089
	(0.012)	(0.019)	(0.018)	(0.096)
log *Age*	0.078***	−0.020	0.175***	1.415***
	(0.015)	(0.021)	(0.025)	(0.103)
R^2	87.75%		73.35%	
Number of observations	54,876		54,876	

Notes. Coefficients for product fixed effects and time fixed effects are omitted from the table. Huber–White robust clustered standard errors are in parentheses.

[a]For seemingly unrelated regressions, the system weighted R^2 provided by SAS is reported.

*$p < 0.1$; ***$p < 0.01$.

can be generalized to other industries or more recent observation periods needs to be evaluated in future studies. Second, we present a stylized model where retailers within each group are symmetric. In reality, each retailer has its own unique characteristics that could also affect its pricing strategy. Third, for our empirical analysis, we use multicategory status and product coverage as proxy variables for a retailer's cross-selling capability. As a robustness check, we also use the percentage of book orders containing multiple items as a proxy (but only for a subset of the retailers). These are admittedly crude measures. One issue with these measures is that they can be correlated with other retailer characteristics. For example, because of the dynamics of the online book industry, retailers with higher reach or popularity may have incentives to take advantage of their market power and increase their product breadth. As a result, our measures of product breath, reach, and popularity are likely to be correlated. This potentially gives rise to the issue

of multicollinearity when all of these variables are included in a regression model. Although OLS estimators are unbiased in the presence of multicollinearity, multicollinearity increases standard errors and thus works against finding statistically significant results. In our case, the coefficients of these variables are significant in spite of the increased standard errors, so we are confident that our results are robust to this issue. But future studies could benefit from a more refined measure of cross-selling capability or data from a different industry where product breath is uncorrelated with reach and other retailer characteristics. Fourth, our results show that the increased price dispersion often observed for best sellers is mainly driven by the heterogeneity in retailers' cross-selling capabilities. This is consistent with our analytical results, but we acknowledge that other factors may also contribute to this increased price dispersion, and needs to be investigated in future research. Finally, we observe only retailers' prices and not consumers'

Table 9   Robustness Analysis of Price Dispersion Using Coefficient of Variation as the Dependent Variable

	Overall price dispersion	Price dispersion among retailers with high cross-selling capability	Price dispersion among retailers with low cross-selling capability	Price difference between the two retailer groups
BestSellerDummy	0.007***	0.006	0.002	0.018***
	(0.002)	(0.004)	(0.003)	(0.005)
PaperbackDummy	0.005***	−0.003	0.014***	−0.005*
	(0.002)	(0.003)	(0.002)	(0.003)
log *Age*	−0.000	−0.018***	0.007**	0.043***
	(0.002)	(0.003)	(0.003)	(0.003)
R^{2a}	79.21%		66.94%	
Number of observations	54,876		54,876	

Notes. Coefficients for product fixed effects and time fixed effects are omitted from the table. Huber–White robust clustered standard errors are in parentheses.

[a]For seemingly unrelated regressions, the system weighted R^2 provided by SAS is reported.

*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

Table 10 **Robustness Analysis Using New Cross-Selling Measure for Large Retailers**

	New cross-selling measure for large retailers
BestSellerDummy	0.101**
	(0.042)
BestSellerDummy × *Multi-Item*	−0.161***
	(0.061)
BestSellerDummy × log *Reach*	0.005
	(0.007)
BestSellerDummy × log *Popularity*	−0.024***
	(0.006)
BestSellerDummy × *DualChannel*	−0.005
	(0.005)
PaperbackDummy	0.004**
	(0.002)
log *Age*	−0.006***
	(0.002)
R^2	87.35%
Number of observations[2]	350,564

Notes. Coefficients for product fixed effects, time fixed effects, and store fixed effects are omitted from the table. Huber–White robust clustered standard errors are in parentheses.

[a]We have fewer observations than in Table 3 because the observations that do not have *Multi-Item*, *Reach*, or *Popularity* data are excluded from analysis.

**p < 0.05; ***p < 0.01.

purchasing behavior at the retailers. In-depth information on consumer purchases at each of the retailers will be able to provide more insights into retailers' pricing strategies and competitive behavior.

Acknowledgments

Appendix. Derivation of the Symmetric Equilibrium

Given that $\partial^2 \pi_j(p_{jk})/\partial p_{jk}^2 < 0$ for all j and k ($j = 1, 2, 3, 4$; $k = A, B$), each firm's optimal price can be derived by solving $\partial \pi_j(p_{jk})/\partial p_{jk} = 0$ simultaneously for all j and k. This leads to

$$p_{1A}^* = p_{3A}^* = \frac{(6d - b(6bd^2 + 8d - 7))t}{4(6d - b(2bd(1+d) - 3))},$$

$$p_{1B}^* = p_{3B}^* = \frac{(24d + b(12 - 7b + 2(-3+b)d + 12d^2))t}{8(6d - b(2bd(1+d) - 3))}, \quad \text{and}$$

$$p_{2A}^* = p_{4A}^* = \frac{(6d - b(d(4+b+4bd) - 5))t}{4(6d - b(2bd(1+d) - 3))}.$$

We then check whether all the retailers have positive demand in product A's market by verifying the following conditions: $|p_{1A} - p_{2A}| < t/4$, $|p_{1A} - p_{4A}| < t/4$, $|p_{2A} - p_{3A}| < t/4$, $|p_{3A} - p_{4A}| < t/4$, and $|p_{1B} - p_{3B}| < t/2$. These conditions require $b < (5 - 4d + \sqrt{25 - 40d + 40d^2 + 96d^3})/(2d + 8d^2)$. We assume v is sufficiently large so that all consumers receive positive utility from purchasing from any retailer in equilibrium.

Next, we check to make sure that neither retailer 1 nor retailer 3 has an incentive to deviate from this equilibrium by lowering its price of product B significantly to steal away some copurchasing consumers from the other retailer also selling product B. Given the symmetric structure, it is sufficient to conduct this analysis just for retailer 1. In equilibrium, because retailers 1 and 3 charge the same price of product B, we have $(t/4 - p_{3A}^* + p_{2A}^*)/2t < (t/2 - p_{3B}^* + p_{1B}^*)/t$ and $(t/4 - p_{3A}^* + p_{4A}^*)/2t < (t/2 - p_{3B}^* + p_{1B}^*)/t$, so the consumers who visit retailer 3 to buy product A and become interested in product B will buy product B from retailer 3 as well. If retailer 1 deviates by lowering its price of product B significantly so that $(t/4 - p_{3A}^* + p_{2A}^*)/2t > (t/2 - p_{3B}^* + p_{1B})/t$ and $(t/4 - p_{3A}^* + p_{4A}^*)/2t > (t/2 - p_{3B}^* + p_{1B})/t$, retailer 1 can attract some copurchasing consumers from retailer 3 and earn

$$\pi_1^{\mathrm{Dev1}}(p_{1A}, p_{1B})$$
$$= (p_{1A} + bp_{1B})\left(\frac{t/4 - p_{1A} + p_{2A}^*}{2t} + \frac{t/4 - p_{1A} + p_{4A}^*}{2t} \right)d$$
$$+ \frac{1}{2}bp_{1B}\left(\frac{t/4 - p_{3A}^* + p_{2A}^*}{2t} + \frac{t/4 - p_{3A}^* + p_{4A}^*}{2t} \right.$$
$$\left. -2\frac{t/2 - p_{3B}^* + p_{1B}}{t} \right)d$$
$$+ \left(p_{1B} + \frac{1}{2}bp_{1A} \right)\left(\frac{t/2 - p_{1B} + p_{3B}^*}{t} \right)$$
$$+ \frac{1}{2}bp_{1A}\left(\frac{t/4 - p_{1A} + p_{2A}^*}{2t} + \frac{t/4 - p_{1A} + p_{4A}^*}{2t} \right). \quad (17)$$

If retailer 1 deviates, it chooses p_{1A} and p_{1B} to maximize $\pi_1^{\mathrm{Dev1}}(p_{1A}, p_{1B})$ subject to the conditions that $(t/4 - p_{3A}^* + p_{2A}^*)/2t > (t/2 - p_{3B}^* + p_{1B})/t$ and $(t/4 - p_{3A}^* + p_{4A}^*)/2t > (t/2 - p_{3B}^* + p_{1B})/t$. It can be derived that retailer 1 never finds it profitable to deviate to attract copurchasing consumers from retailer 3 under this sufficient condition:

$$(96d^2 - 8bd(4+d)(-4+3d) - b^3(1+4d)(5+6d+4d^3)$$
$$+4b^4d(2+d+3d^2+4d^3) - 2b^2(-20+d(19+2d(14+d(23+6d)))))$$
$$\cdot ((-16d+b(-8+b+4(-2+b)d^2))(-6d+b(-3+2bd(1+d))))^{-1} > 0.$$

Finally, we check to make sure that neither retailer 1 nor retailer 3 has an incentive to deviate from this equilibrium by charging a very high price for either product so that it earns profit from this product purely through cross-selling. Again, given the symmetric structure, it is sufficient to conduct this analysis just for retailer 1. If retailer 1 deviates by charging a very high price to product A, its maximum deviation profit is

$$\pi_1^{\mathrm{Dev2}}(p_{1A}, p_{1B}) = \left(p_{1B} + \frac{1}{2}bv \right)\left(\frac{t/2 - p_{1B} + p_{3B}^*}{t} \right). \quad (18)$$

If retailer 1 deviates by charging a very high price to product B, its maximum deviation profit is

$$\pi_1^{\mathrm{Dev3}}(p_{1A}, p_{1B})$$
$$= \left(p_{1A} + \frac{1}{2}bv \right)\left(\frac{t/4 - p_{1A} + p_{2A}^*}{2t} + \frac{t/4 - p_{1A} + p_{4A}^*}{2t} \right)d. \quad (19)$$

If retailer 1 deviates, it either chooses p_{1B} to maximize $\pi_1^{\text{Dev2}}(p_{1A}, p_{1B})$ or chooses p_{1A} to maximize $\pi_1^{\text{Dev3}}(p_{1A}, p_{1B})$. It can be derived that the maximum profit that retailer 1 can achieve if it deviates is

$$\max\Big\{\big((b(-24+7b)+6(-8+b+b^2)d+4b(-3+2b)d^2)t$$
$$+4b(6d+b(3-2bd(1+d)))v\big)$$
$$\cdot\big(-96d+16b(-3+2bd(1+d))\big)^{-1},$$
$$\tfrac{1}{8}\big(((-12d+b(-8+d(4+b(3+6d))))t\big)$$
$$\cdot(-6d+b(-3+2bd(1+d)))^{-1}-2bv\big)\Big\}.$$

If retailer 1 does not deviate, its profit is

$$\big((-b+(-2+b^2)d)\big(-36d(4+d)+3b^3d(7+4d(3+2d))$$
$$+b^2(-7+10d(-2+5d))-24b(3+d(2+5d)))t\big)$$
$$\cdot\big(32(-6d+b(-3+2bd(1+d)))^2\big)^{-1}.$$

Thus, the equilibrium holds as long as b is sufficiently small so that it satisfies the following three conditions:

$$\max\Big\{\big((b(-24+7b)+6(-8+b+b^2)d+4b(-3+2b)d^2)t$$
$$+4b(6d+b(3-2bd(1+d)))v\big)$$
$$\cdot\big(-96d+16b(-3+2bd(1+d))\big)^{-1},$$
$$\tfrac{1}{8}\big(((-12d+b(-8+d(4+b(3+6d))))t\big)$$
$$\cdot(-6d+b(-3+2bd(1+d)))^{-1}-2bv\big)\Big\}$$
$$<\big((-b+(-2+b^2)d)\big(-36d(4+d)+3b^3d(7+4d(3+2d))$$
$$+b^2(-7+10d(-2+5d))-24b(3+d(2+5d)))t\big)$$
$$\cdot\big(32(-6d+b(-3+2bd(1+d)))^2\big)^{-1};$$

$$(96d^2-8bd(4+d)(-4+3d)-b^3(1+4d)(5+6d+4d^3)$$
$$+4b^4d(2+d+3d^2+4d^3)-2b^2(-20+d(19+2d(14+d(23+6d)))))$$
$$\cdot\big((-16d+b(-8+b+4(-2+b)d^2))(-6d+b(-3+2bd(1+d)))\big)^{-1}$$
$$>0; \quad \text{and} \quad b<\frac{5-4d+\sqrt{25-40d+40d^2+96d^3}}{2d+8d^2}.$$

References

Allison PD (2005) *Fixed Effect Regression Methods for Longitudinal Data Using SAS* (SAS Institute, Cary, NC).

Altonji JG, Dunn TA (1996) The effects of family characteristics on the return to education. *Rev. Econom. Statist.* 78(4):692–704.

American Booksellers Association (2009) ABA asks Department of Justice to investigate bestseller price wars. *American Booksellers Association News* (October 22), http://news.bookweb.org/news/aba-asks-department-justice-investigate-bestseller-price-wars.

Baye M, Morgan J, Scholten P (2004) Price dispersion in the small and in the large: Evidence from an Internet price comparison site. *J. Indust. Econom.* 52(4):463–496.

Baylis K, Perloff JM (2002) Price dispersion on the Internet: Good firms and bad firms. *Rev. Indust. Organ.* 21(3):305–324.

Bertrand M, Luttmer EFP, Mullainathan S (2000) Network effects and welfare cultures. *Quart. J. Econom.* 115(3):1019–1055.

Bliss C (1988) A theory of retail pricing. *J. Indust. Econom.* 36(4):375–391.

Brynjolfsson E, Smith MD (2000) Frictionless commerce? A comparison of Internet and conventional retailers. *Management Sci.* 46(4):563–585.

Brynjolfsson E, Hu YJ, Rahman MS (2009) Battle of the retail channels: How product selection and geography drive cross-channel competition. *Management Sci.* 55(11):1755–1765.

Bustillo M, Trachtenberg JA (2009) Wal-Mart strafes Amazon in book war. *Wall Street Journal* (October 16) A1.

Chellappa RK, Sin RG, Siddarth S (2011) Price-formats as a source of price dispersion: A study of online and offline prices in the domestic U.S. airline markets. *Inform. Systems Rev.* 22(1):83–98.

Chen PY, Hitt L (2001) Brand awareness and price dispersion in electronic markets. *Proc. 22nd Internat. Conf. Inform. Systems* (Association for Information Systems, Atlanta), 233–246.

Clark DS (1998) The Robinson-Patman Act: Annual update. Accessed October 25, 2012, http://www.ftc.gov/speeches/other/spring98.shtm.

Clay K, Krishnan R, Wolff E (2001) Prices and price dispersion on the Web: Evidence from the online book industry. *J. Indust. Econom.* 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. Indust. Econom.* 50(3):351–367.

Clemons EK, Hann I-H, Hitt LM (2002) Price dispersion and differentiation in online travel: An empirical investigation. *Management Sci.* 48(4):534–549.

Forman C, Ghose A, Goldfarb A (2009) Competition between local and electronic markets: How the benefit of buying online depends on where you live. *Management Sci.* 55(1):47–57.

Ghose A, Yao Y (2011) Using transaction prices to re-examine price dispersion in electronic markets. *Inform. Systems Rev.* 22(2):1–17.

Hess JD, Gerstner E (1987) Loss leader pricing and rain check policy. *Marketing Sci.* 6(4):358–374.

Hobson J (2009) Bookstores profit off online price wars. *Marketplace* (October 30), http://www.marketplace.org/topics/tech/bookstores-profit-online-price-wars.

Lach S (2002) Existence and persistence of price dispersion: An empirical analysis. *Rev. Econom. Statist.* 84(3):433–444.

Lal R, Matutes C (1994) Retail pricing and advertising strategies. *J. Bus.* 67(3):345–370.

Pan X, Ratchford BT, Shankar V (2002) Can price dispersion in online markets be explained by differences in e-tailer service quality? *J. Acad. Marketing Sci.* 30(4):443–456.

Pan X, Ratchford BT, Shankar V (2004) Price dispersion on the Internet: A review and directions for future research. *J. Interactive Marketing* 18(4):116–135.

Raff DMG (2000) Superstores and the evolution of firm capabilities in American bookselling. *Strategic Management J.* 21(10/11):1043–1059.

Rajiv S, Dutta S, Dhar SK (2002) Asymmetric store positioning and promotional advertising strategies: Theory and evidence. *Marketing Sci.* 21(1):74–96.

Raju JS, Srinivasan V, Lal R (1990) The effects of brand loyalty on competitive price promotional strategies. *Management Sci.* 36(3):276–304.

Rose M, Quick R (2000) Web booksellers give Potter fans rush delivery. *Wall Street Journal* (June 22) B1.

Rothstein JM (2006) Good principals or good peers? Parental valuation of school characteristics, Tiebout equilibrium, and the incentive effects of competition among jurisdictions. *Amer. Econom. Rev.* 96(4):1333–1350.

Salop SC (1979) Monopolistic competition with outside goods. *Bell J. Econom.* 10(1):141–156.

Shilony Y (1977) Mixed pricing in oligopoly. *J. Econom. Theory* 14(2):373–388.

Smith M (2001) The law of one price? The impact of IT-enabled markets on consumer search and retailer pricing. Working paper, Carnegie Mellon University, Pittsburgh.

Smith M, Brynjolfsson E (2001) Customer decision making at an Internet ShopBot: Brand still matters. *J. Indust. Econom.* 49(4):541–558.

Sorensen AT (2000) Equilibrium price dispersion in retail markets for prescription drugs. *J. Political Econom.* 108(4):833–850.

Sorensen AT (2007) Bestseller lists and product variety. *J. Indust. Econom.* 55(4):715–738.

Tang F, Xing X (2001) Will the growth of multi-channel retailing diminish the pricing efficiency of the Web? *J. Retailing* 77(3):319–333.

Tang Z, Smith MD, Montgomery A (2010) The impact of ShopBot use on prices and price dispersion: Evidence from online book retailing. *Internat. J. Indust. Organ.* 28(6):579–590.

Varian HR (1980) A model of sales. *Amer. Econom. Rev.* 70(4):651–659.

Venkatesan R, Mehta K, Bapna R (2007) Do market characteristics impact the influence of retailer characteristics on online prices? *J. Retailing* 83(3):309–324.

Wooldridge JM (2002) *Econometric Analysis of Cross Section and Panel Data* (MIT Press, Cambridge, MA).