INTMAR-00237; No. of pages: 1; 4C:

Available online at www.sciencedirect.com

ScienceDirect

Journal of Interactive Marketing xx (2017) xxx

JOURNAL OF INTERACTIVE MARKETING

www.elsevier.com/locate/intmar

Highlights

Journal of Interactive Marketing xxx (2017) xxx–xxx

Pricing Best Sellers and Traffic Generators: The Role of Asymmetric Cross-selling

Cenk Kocas[a] & Koen Pauwels[b] & Jonathan D. Bohlmann[c,*]

[a] *Sabanci University, Faculty of Management, YBF 1076 Orhanlı, 34956 Tuzla, Istanbul, Turkey*
[b] *D'Amore-McKim School of Business, Northeastern University, 205E Hayden Hall, 360 Huntington Avenue, Boston, MA 02118, USA*
[c] *North Carolina State University, Dept. of Business Management, Campus Box 7229, Raleigh, NC 27695, USA*

- Asymmetric retailer price strategies of best sellers are modeled under cross-selling.
- Price discount strategies are examined under cross-selling conversion and inclusion.
- Cross-selling potential of products even far down a best seller list is demonstrated.
- Larger multicategory retailers offer deeper discounts on top best seller products.
- Empirical analysis of online pricing provides support for key findings of the model.

http://dx.doi.org/10.1016/j.intmar.2017.09.001
1094-9968/© 2017

ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

Journal of Interactive Marketing xx (2017) xxx – xxx

JOURNAL OF INTERACTIVE MARKETING

www.elsevier.com/locate/intmar

Pricing Best Sellers and Traffic Generators: The Role of Asymmetric Cross-selling

Cenk Kocas [a] & Koen Pauwels [b] & Jonathan D. Bohlmann [c,*]

[a] *Sabanci University, Faculty of Management, YBF 1076 Orhanlı, 34956 Tuzla, Istanbul, Turkey*

[b] *D'Amore-McKim School of Business, Northeastern University, 205E Hayden Hall, 360 Huntington Avenue, Boston, MA 02118, USA*

[c] *North Carolina State University, Dept. of Business Management, Campus Box 7229, Raleigh, NC 27695, USA*

Abstract

Among the many items online retailers sell, some stand out as best sellers and are often sold at considerable discounts. Best seller discounting can encourage customer traffic and the purchase of a basket of other products in the same transaction. Although most studies treat retailers as symmetric, the cross-selling potential is generally asymmetric across retailers, since some online retailers have more products to sell. In addition, the cross-selling effect works both ways — customers intending to buy a best seller may buy other items in their shopping basket, while other customers intending to buy a basket may buy a best seller while visiting the retailer. The authors model the pricing implications of this rich variety of asymmetric cross-selling, with both best sellers and typical baskets acting as traffic generators and cross-sold products. The common wisdom that loss leader pricing leads to neither a significant increase in store traffic nor an increase in profits does not apply in an asymmetric case where one retailer has more products to cross-sell. The cross-selling potential of products even far down the best seller list is demonstrated. Empirical analyses provide support for key findings of the theoretical model using book pricing and sales rank data from multiple online retailers.

Keywords: Pricing; Online retailing; Best seller; Cross-selling; Loss leader

Introduction

On October 22, 2009, the American Booksellers Association sent a letter to the U.S. Department of Justice (DOJ) accusing Amazon.com, Wal-Mart, and Target of illegal predatory pricing. These three retailers had sold ten hardcover new releases, including best sellers by James Patterson, John Grisham, and Stephen King, for less than $9, though such books typically retail between $25 and $35 (Trachtenberg 2009). The letter also reported that publishers were not offering special terms to these retailers, so the titles were being sold below cost. Taking issue with this claim, The Wall Street Journal Law Blog commented that retailers setting prices below profit-making levels was not a sign of predatory pricing but rather an indicator of healthy price competition (Jones 2009). Promoting and selling the top-ten titles below cost represented a loss leader strategy to draw in customers who might purchase other titles or merchandise.

The DOJ case focused on 10 best sellers, but we also observe strong price competition for many products with even far lower sales ranks. News reports in October 2009 suggested that Wal-Mart was already offering up to 200 best sellers for 50% off their list price (Reisinger 2009). Amazon.com typically lists 100 books at considerable discounts under its "Best Sellers in Books" list. In other product categories, more than 500 generic prescription drugs are offered either for free (e.g., antibiotics at Publix and Meijer) or for only $4 for a month's supply (e.g., Wal-Mart, Kmart, Target) (National Conference of State Legislatures 2011). Amazon.com even provides sales ranks of books up to 10 million, similar to buy.com and other sites that track and report the sales ranks of almost all products offered for sale online. Retailers recognize that many products are able to generate some degree of traffic and cross-selling opportunity.

* Corresponding author.
E-mail addresses: kocas@sabanciuniv.edu (C. Kocas), kpauwels@northeastern.edu (K. Pauwels), jon_bohlmann@ncsu.edu (J.D. Bohlmann).

http://dx.doi.org/10.1016/j.intmar.2017.09.001
1094-9968/© 2017

Given the observed richness of price discounting across hundreds of items, we aim to clarify the pricing implications of the traffic generation potential for products with diverse sales ranks. We model and empirically examine price discounting strategies for online retailers. Although our model has application to retail competition more generally, the online pricing issues are more pertinent for several reasons. First, although products at the top of best seller lists are clear traffic generators and prime candidates for loss leader pricing, many products with lower sales ranks also exhibit some traffic generation potential. In other words, "best seller" is not so much a category as it is a matter of degree. Considering that an online retailer can offer millions of items, the retail pricing decision is much more complex since even less popular items may generate at least some traffic and cross-selling potential, prompting an online retailer to consider how to best discount such items. A key question thus emerges: What is the price discounting implication of the diminishing but positive traffic generation potential of products farther down the best seller ranks? Second, if a best seller is meant to generate traffic and sales of other products, then retailer size may be an important variable. Some retailers are bigger than others in that they offer more products for customers to purchase. Such asymmetric competition means that some retailers can benefit more from best seller discounting since the opportunity for cross-selling is bigger. Online stores have achieved very large assortments, so consideration of shopping basket size is important for online retailing. How do price discounting strategies and cross-selling vary with a retailer's size of the typical shopping basket it sells? Third, the psychological and economic motivations to visit a retailer and be cross-sold can be more prevalent in an online setting. The large product assortment can impact traffic for the online retailer and be an important basis of differentiation (Pan, Shankar, and Ratchford 2002; Ratchford 2009). Online recommendations for other items to purchase during online shopping introduce prolific cross-selling opportunities, including instances where a best seller is the product being cross-sold. How are price discounting strategies affected when additional shopping items or a best seller may be cross-sold to different shoppers? Finally, offering lower prices may be more prevalent and important for online retailers compared to brick-and-mortar stores (Pan, Shankar, and Ratchford 2002). Ratchford (2009) suggest that online price dispersion deserves additional explanations, particularly in relation to "heterogeneity in services" such as the product variety offered by retailers. Our study of cross-selling with asymmetric retailer size adds new insights to online price discounting strategies.

Given these important online pricing issues, we pose several research questions:

1. How do competing, profit-maximizing retailers determine price discounts for best sellers?
2. How does the loss-leading price of best sellers depend on retailer size?
3. How do retailers price best sellers and traffic generators of varying ranks?
4. When does best seller pricing increase traffic and profits?

Current marketing literature is limited on the first two research questions and absent on the rest, even though these questions are crucial to understanding the retail dilemma of which items to price higher or lower and when. The 2009 case about best-selling books reveals that not all retailers can offer the same lowest price. If the optimal (loss leader) price of a best seller is not the same across retailers, on what does it depend? Can a retailer with relatively smaller basket sizes offer the same loss leader prices as a larger basket-size retailer?

To examine these questions, our model includes two main characteristics of realistic retailer cross-selling activity generally ignored in prior research. First, retailers are asymmetric in that they vary in how many products they sell, meaning that their cross-selling capabilities differ.[1] Second, cross-selling is not a one-way activity where a customer buys a single best seller and then buys another basket of items while visiting the retailer. Some customers intending to buy a typical shopping basket may be cross-sold a best seller.

We examine the price discounting strategies of multiproduct retailers that incorporate these cross-selling characteristics. We use the term "best seller" to refer to any product with a higher potential to generate traffic for the retailer than a product lower down the sales rank.[2] We analyze a model in which best sellers can lead to the cross-sale of a basket of goods, just as the sale of a shopping basket can lead to the cross-sale of a best seller. An online retailer might be willing to reduce the price of a best seller if it would lead to cross-selling opportunities, but it also wants to increase the price of the best seller to the degree that it is cross-sold to buyers of other items. We show that the loss leader prices of best sellers depend critically on the typical basket size of a retailer. This finding explains why big-box retailers, such as Amazon.com, can offer discounts that cannot be matched by smaller retailers. We examine the boundary conditions of this phenomenon, and provide empirical evidence with online book pricing data that supports key propositions from our model: price discounts positively correlate with sales rank (even far down the best seller list), best sellers with low list prices are discounted more, and large basket retailers offer deeper discounts on the top best sellers.

Best Seller Discounts and Loss Leaders

Best sellers are books for which demand vastly exceeds what is then considered to be large sales (Steinberg 1996). Recent research has uncovered three major content reasons a book becomes a best seller: (1) its main themes, (2) symmetric plot with 3-act structure, and (3) everyday language (Archer and Jockers 2016). Becoming a best seller is also driven by the reputation of the author, gatekeepers such as publishing houses and publishers of book reviews and bestseller lists,

[1] Li, Gu, and Liu (2013) analyze asymmetry in a retailer cross-selling, but the asymmetry is binary in that a retailer either cross-sells or it doesn't.

[2] While we use "best seller" to indicate a traffic-generating product, other research uses similar labels of "loss leader" or "shopping good." We use loss leader to reflect a best seller product priced below cost. A composite good "basket" in our study represents one or more items purchased in addition to a focal best seller item.

Please cite this article as: Cenk Kocas, et al., Pricing Best Sellers and Traffic Generators: The Role of Asymmetric Cross-selling, Journal of Interactive Marketing (2017), http://dx.doi.org/10.1016/j.intmar.2017.09.001

word-of-mouth networks, advertising, and a host of techniques to become included in best seller lists (Hill and Power 2005). Price is not considered a key driver of becoming a best seller because book prices are low compared to consumer budgets in mature markets. However, discounted or even loss leader pricing may influence shopping traffic.

Loss leader pricing has been the subject of considerable research in marketing. Hess and Gerstner (1987) were the first to employ a formal model of loss leaders. Lal and Matutes (1994) explain many facets of loss leader pricing, including Walters and MacKenzie's (1988) empirical finding that on average it leads to neither a significant increase in store traffic nor an increase in profits (in a supermarket setting). Our model findings are parallel to Lal and Matutes (1994) in some respects. However, our model captures asymmetries in both products and retailers, which enables us to show that the classic finding of no significant increases in traffic and profits holds only for the symmetric retailer case. When there is asymmetry among online retailers, the retailer with a marginal advantage in benefiting from cross-selling can increase both traffic and profits, compared with a smaller retailer with weaker cross-selling potential.

DeGraba (2006) considers loss leader pricing as a way to capture high-profit customers. He shows that by offering discounts on products that are more likely to be purchased by high-profit customers, loss leader pricing can price discriminate in a competitive setting. Our model approach is similar in that the profit potential of a shopping basket determines the pricing of traffic-generating best sellers. The competitive bundling literature also deals with a similar problem, such as Balachander, Ghosh, and Stock (2010) who combine bundle discounts and price promotions in a model of cross-category bundling.

In a brick and mortar setting, absent price communication, the consumer is at the risk of zero consumer surplus, because the retailer could price the products at the reservation price given the consumer has already incurred the sunk travel cost. Signaling low prices on some products is suggested to be a solution to this setback (Lal and Matutes 1994). Simester (1995) also argues advertised prices may signal the efficiency of the retailer and her low marginal costs and hence low prices on unadvertised products. Signaling with low prices can thus lead to increased store traffic. This rationale however, does not apply to an online setting, since sunk travel costs are minimal and price information is typically available for most items. For online retailing, other factors such as product variety (retailer size asymmetry) may be at play for cross-selling and loss leader pricing.

Researchers have also extensively examined online and offline price dispersion (Ancarani and Shankar 2004; Bakos 1997; Baye, Morgan, and Scholten 2004; Brynjolfsson and Smith 2000; Clay, Krishnan, and Wolff 2001; Pan, Ratchford, and Shankar 2004; Pan, Shankar, and Ratchford 2002, 2003; Ratchford, Pan, and Shankar 2003). Ambrus and Weinstein (2008) show that equilibrium loss leaders can occur with positive profits if there are certain demand complementarities among goods sold. Ratchford, Pan, and Shankar (2003), Ellison and Ellison (2005) and Ratchford (2009) provide thorough reviews of prices and price dispersions in electronic commerce. Our work concentrates on the lower bound of prices (loss leader pricing and discounts) which we claim to be a function of the traffic generation potential of products. Furthermore, our work is among a few (Chen and Hitt 2003; Kocas and Bohlmann 2008; Smith 2002) where retailers play asymmetric mixed strategies of temporary "randomized" price discounting that produce online price dispersion. The mixed strategy pricing equilibria of competing firms are reflected through observed temporal price discounting and dispersion (Narasimhan 1988; Ratchford 2009; Varian 1980) across multiple products and retailers (see also Iyer and Pazgal 2003; Raju, Srinivasan, and Lal 1990). Actual pricing data represent repeated observations of a mixed pricing strategy over time.

Our work shares the mixed strategy equilibrium interpretation of temporary price discounts with the preceding research. However, our study is unique in that it does not rely on the dynamics of loyal and switcher customer groups, but rather on cross-selling. We utilize DeGraba's (2006) perspective that the profit potential of a cross-sold composite good (basket) determines the pricing of the best seller, but we do so via the approach of probabilistic retailer pricing strategies advocated by Varian (1980) and Narasimhan (1988). We compare symmetric with asymmetric cases and show that the profit potential of basket sizes shapes the price discounting equilibria. Our work thereby bridges the research streams of loss leaders and competitive price promotions by examining cross-selling pricing strategies in a single framework. Our discounted pricing model allows us to also determine when loss leader pricing will apply to a best seller.

Online Cross-selling and Baskets of Goods

> Amazon's super saver free shipping is truly a piece of marketing genius. It works on the premise that people will buy more items in the same order just to achieve the free shipping. I can admit that I find myself doing just that on a constant basis. Every time I go to Amazon to buy a $15 DVD, I will likely buy another $10 item just to get to that $25. There is something endlessly satisfying about getting the items you want without having to pay those nasty extra fees.
>
> (May 20, 2008, anonymous Internet posting)

Consider a typical online shopping experience, in which a customer shops for a new or best seller product (book, DVD, CD, console game). The customer may visit her favorite retailer's site or visit a price comparison site first to view the range of prices available for the item. She could visit the online seller that offers the product at the lowest price, or consider just a short list of favorite retailers and choose the retailer that offers the lowest price. When the item enters the shopping basket on the online store's website, a variety of forces then push the customer to purchase other items. She may get free shipping if she spends just $5 more, remember a book she wanted to buy next time she was online, or receive

Please cite this article as: Cenk Kocas, et al., Pricing Best Sellers and Traffic Generators: The Role of Asymmetric Cross-selling, Journal of Interactive Marketing (2017), http://dx.doi.org/10.1016/j.intmar.2017.09.001

a suggestion for yet another book (or even an unrelated item) by a content or collaborative filter-based recommender system (Fleder and Hosanagar 2009).

Beyond arguments arising from total costs of shopping (e.g., processing and shipping), psychological factors may also lead to additional item purchases. Dhar, Huber, and Khan (2007) define the term "shopping momentum effect" as the inertia to continue purchasing unplanned items after an initial purchase, independent of the economies-of-scale arguments. Heilman, Nakamoto, and Rao (2002) and Stilley, Inman, and Wakefield (2010b) also show that unexpected savings on planned items can create a psychological windfall effect, leading to an increased purchase of unplanned items.

These psychological and economic effects of a sales promotion on the size and composition of the shopping basket are diverse; promotional items attract both cherry-pickers with very small baskets and customers who eventually purchase large baskets[3] (Dhar, Huber, and Khan 2007; McAlister, George, and Chien 2009; Stilley, Inman, and Wakefield 2010a). Mulhern and Padgett (1995) find that more than three-fourths of shoppers who based store choice on promoted items spent even more money on other regularly-priced items. The overall implication is that cross-selling can cut both ways. Shoppers who are mainly interested in a "best seller" may impulse buy one or more items (a basket). We label this successful cross-selling as "conversion." Also, a buyer not necessarily interested in a best seller may, in addition to purchasing the planned shopping basket, also buy a best seller. We label this cross-selling as an "inclusion."

Given the wide variety of items offered by online retailers, and the widespread occurrence of purchase recommendations and impulse buying, any model of online price discounting should consider the different types of realistic cross-selling opportunities. Unlike prior research, we consider both types of cross-selling in our pricing model. Further, given the role of the shopping basket in cross-selling, online price discounting for the best seller must consider that online retailers can differ greatly in the items they offer to sell. In other words, asymmetry in the shopping basket size among online retailers means some retailers will benefit from cross-selling more than others, with important implications for the optimum best seller price discounting strategy. Our model therefore focuses on retailer asymmetry under cross-selling, making a unique contribution to the online pricing literature.

We focus on the traffic generation potential of discounted best sellers and consolidate the diverse effects of cross-selling by introducing an "effective rate of average baskets sold." Suppose m customers are drawn to a retailer to purchase the best seller. Some of these customers will buy only the best seller, while others will be successfully cross-sold a basket of size x_i, i = 1 to m, where $x_i = 0$ if the customer buys only the best seller. Instead of modeling a large series of basket sizes purchased $(x_1, x_2, \ldots x_m)$, we can easily express an effective rate of cross-selling conversion relative to a retailer j's average basket size it sells s_j:

$$\alpha_j = \frac{\sum_{i=1}^{m} x_i}{m\, s_j} \tag{1}$$

The effective rate of average baskets sold α simply captures the degree of cross-selling conversion by the retailer, equivalently scaled by a measure of the retailer's average basket size (s). A retailer with higher α is more successful at cross-selling conversion sales. The effective rate of average baskets sold also allows us to convert the distribution of additional items sold into an effective Bernoulli distribution that has just two outcomes: an α probability that a customer is cross-sold an average basket, and a $(1-\alpha)$ probability that the customer buys the promoted best seller item but no additional items.

In the next section, we analyze and compare a symmetric and an asymmetric duopoly of retailers, in which retailers can sell a best seller and a composite good (average basket) to potential customers. We also provide empirical support for the model's findings, using book pricing and sales rank data from Amazon.com, as well as pricing data from 18 other retailers.

Model

We consider a 2 × 2 × 2 market with two retailers (R1 and R2) selling two products (A and B) to two customer segments (n shoppers of A, and N shoppers of B). Let good A be a best seller (book, CD, DVD, or console game) that creates traffic for retailers. Let good B be an average basket. Similar to other models (DeGraba 2006; Li, Gu, and Liu 2013) the two segments reflect two types of customers that differ in how they choose a retailer — n shoppers visit a retailer intending to purchase best seller A, and N shoppers visit a retailer intending to purchase product basket B. Cross-selling opportunities exist for both segments, whether conversion (best seller buyers also purchase basket B) or inclusion (basket buyers also purchase best seller A). Variables used in the model are defined and summarized in Table 1.

Retailers and Products

The retailers choose prices for the best seller product, strategically considering competitor prices. We assume a one-shot simultaneous-move game for the price choice of the best seller A to maximize profits, similar to Varian (1980) and Narasimhan (1988). R1 has the price couple (a_1, b_1) for products A and B, and R2 has the price couple (a_2, b_2). In determining prices of two goods, Lal and Matutes (1994) show that the non-loss leader good is priced at the exogenous reservation price. We therefore keep the price of the average basket exogenous to the model to better assess the price dependence of best seller A on the average basket.[4] Initially, we

[3] McAlister, George, and Chien (2009) report the basket size distribution of a supermarket; the range is 1 to 130, the distribution is skewed (60% of the baskets contain fewer than ten items) and average basket size is around ten items.

[4] The solution when both prices are endogenous is highly involved. Beard and Stern (2008) examine a similar 2 × 2 × 2 model and acknowledge that such models are complex and that general formulations are likely to be intractable.

Please cite this article as: Cenk Kocas, et al., Pricing Best Sellers and Traffic Generators: The Role of Asymmetric Cross-selling, Journal of Interactive Marketing (2017), http://dx.doi.org/10.1016/j.intmar.2017.09.001

Table 1
Variables and definitions used in the model.

α_i	Effective rate of average baskets sold as a result of the sale of product i, i = a or b
m	Number of customers in the calculation of the effective rate
x_i	Number of items in the basket of the ith customer, i = 1 to m,
s	Retailer's average basket size
Ri	Retailer i, i = 1 or 2
a_i	Price of the bestseller (product A) at Retailer i, i = 1 or 2
b_i	Price of the average basket (product B) at Retailer i, i = 1 or 2
n	Number of price comparison shoppers of the bestseller (product A)
N	Number of shoppers of an average basket (product B)
r	Reservation price of the bestseller (product A)
$E\pi_i$	The expected profit of Retailer i, i = 1 or 2
$F_i[a]$	Cumulative distribution function of price of the bestseller (product A) at Retailer i, i = 1 or 2
a_{min}	Lowest possible quoted price of the bestseller (product A) in the mixed strategy
α/β	The conversion-to-inclusion ratio ${}^{\alpha_a}/_{\alpha_b}$
b_{sym}	The average basket size under retailer symmetry
$F_{sym}[a]$	Cumulative distribution function of price of the bestseller (product A) in the symmetric case
$F_{iasym}[a]$	Cumulative distribution function of price of the bestseller (product A) in the asymmetric case at Retailer i, i = 1 or 2

assume a symmetric duopoly in which R1 and R2 are similar in terms of the price value of their average basket ($b_1=b_2=b$). We then relax this assumption to analyze an asymmetric case in which one retailer has an average basket larger than the other retailer. Without loss of generality, we assume no fixed costs and zero marginal cost.[5]

Customers

Two groups of buyers visit the retailers in this model.

1. There are n customers who price compare for product A, the best seller, and buy it at the retailer that offers it for less as long as the price is below their reservation price r. If prices a_1 and a_2 are equivalent, both retailers share n customers equally. To capture conversions among the n customers, we assume an effective rate of average baskets sold, as defined previously, equal to α_a. This is akin to assuming that there is an α_a probability ($0<\alpha_a\leq 1$) that any given customer in this segment will convert to purchasing an average basket. Thus, an (effective) α_a proportion of this segment, $n\alpha_a$, also buys an average basket. The α_a parameter captures the conversion incidence (Lam et al. 2001).
2. There are N other customers who shop for product basket B, and buy it at the retailer where it is available for the lower price. These customers do not have any preferences for either retailer, and because we assume under the symmetric case that the price of the average basket is identical, both retailers share the N customers equally (we later discuss the asymmetric case). To capture inclusions among the N customers, there is an α_b probability ($0<\alpha_b\leq 1$) that any given customer in this segment will also purchase the best seller.[6] Thus, $N\alpha_b$ customers also buy product A from the same retailer. The α_b parameter captures the inclusion incidence (McAlister, George, and Chien 2009).

For expositional simplicity and to establish the intuition behind our results, we present the case when $\alpha_a=\alpha_b=\alpha$ in the main part of the article. We then present the equilibria for $\alpha_a\neq\alpha_b$ and discuss them subsequently. Our equilibrium solutions for optimum price discounting follow standard mixed strategy mechanics (Kocas and Bohlmann 2008; Narasimhan 1988; Ratchford 2009; Varian 1980) under the absence of pure strategies. The mixed-strategy pricing solution is reflected as a probability distribution of the best seller discounted prices, which we term the "price promotion strategy" for the retailer. The highest price in the distribution represents no discount, while lower prices reflect a discount.

Symmetric Case

The general profit function of retailer Ri is given by:

$$E\pi_i = n(a_i+\alpha b)\mathrm{Prob}(a_i<a_j) + N(b+\alpha a_i)\frac{1}{2} \qquad (2)$$

The term $n(a_i+\alpha b)\mathrm{Prob}(a_i<a_j)$ is the sum of profits from the sale of the best seller to n customers and the profits from the sale of an average basket to αn customers when $a_i<a_j$. The term $N(b+\alpha a_i)\frac{1}{2}$ is the sum of the profit from the sale of the average basket to N/2 customers and the profit from the sale of the best seller to $\alpha N/2$ customers. Denoting $F_j[a]$ as the cumulative distribution function of Rj's prices for good A, we can rewrite the profit function for Ri as $E\pi_i = n(a+\alpha b)(1-F_j[a]) + N(b+\alpha a)\frac{1}{2}$.

P_1. *The symmetric retailers' profit-maximizing price promotion strategy is given by a mixed strategy equilibrium of price discounts. The best seller price distribution is given by* $[a] = 1-\frac{N\alpha(r-a)}{2n(a+b\alpha)}$*. The resulting symmetric equilibrium profit is* $E\pi = \frac{N(b+\alpha r)}{2}$*. (Proofs are provided in Appendix A).*

No pure strategy equilibrium exists in this one-shot game. The tension between the desire to lower prices of traffic generators and the desire to increase their prices when part of high-margin baskets leads to mixed strategy discounting of the best seller, in which lower prices are more likely; that is, $\frac{\partial f[a]}{\partial a}<0$ for all $a\in\{a_{min}, r\}$. For ease of interpretation we can set r = 1, such that the bestseller price (a) can be interpreted as a fraction of the highest "regular" price — any price that is less than one reflects a discount. Fig. 1 illustrates the distribution functions for the best seller price under specific parameter values, showing a considerable occurrence of loss leader prices (a is negative). Such tension under symmetric competition can

[5] Although larger retailers may enjoy cost efficiencies, our model shows that the larger retailer can be strategically motivated to lower prices even without any simplistic cost advantage.

[6] Because the cross-sold item is a single best seller, the sales distribution of the success (sale of the best seller) is already a Bernoulli distribution, and therefore the effective rate is equal to the nominal proportion α_b.

Please cite this article as: Cenk Kocas, et al., Pricing Best Sellers and Traffic Generators: The Role of Asymmetric Cross-selling, Journal of Interactive Marketing (2017), http://dx.doi.org/10.1016/j.intmar.2017.09.001

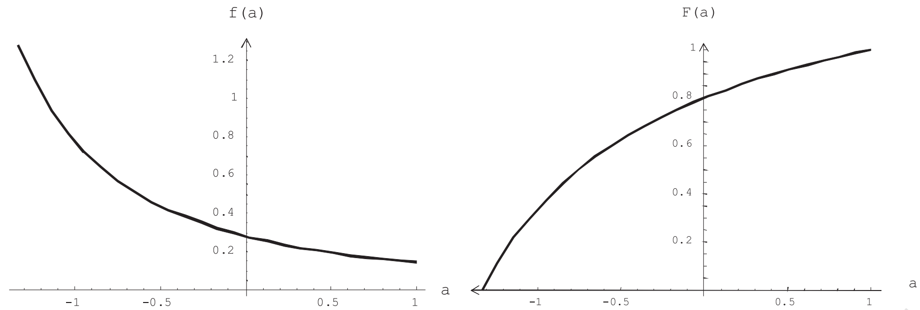


Fig. 1. The symmetric mixed strategy equilibria: probability and cumulative distributions for best seller price a under parameter values $\frac{b}{r}=5$, $\frac{n}{N}=.5$, r=1, AND α=.5.

be severe, and as P_1 indicates the expected profit is equivalent to the profit the retailer would make if it priced the best seller high at r and lost all n customers to the other retailer, selling the average basket to $\frac{N}{2}$ and the best seller to α $\frac{N}{2}$ customers. Therefore, we can conclude that in the case of a symmetric duopoly, the discounts offered on the best seller do not raise profits, consistent with the work of Walters and MacKenzie (1988) and Lal and Matutes (1994). Our analysis also provides support for these authors' insights into the lack of increase in traffic. Because of symmetry, the customer traffic remains the same at $\frac{n+N}{2}$.

P_2. *In the symmetric retailer equilibrium, loss leader pricing can occur if the retailer's incentives to discount, through larger basket size and traffic generation potential, are large enough. The lowest quotable price is given by* $a_{min}=\frac{Nr\alpha-2nb\alpha}{N\alpha+2n}$. *This lower bound,* a_{min}, *is negative (a loss-leading price) when* $\frac{b}{r}\frac{n}{N}>\frac{1}{2}$, *where* $\frac{b}{r}$ *is the relative average basket size compared with the reservation price of the best seller and* $\frac{n}{N}$ *is the relative traffic generation potential of the best seller compared with the average basket.*

Fig. 2 provides a visual analysis of the comparative statics of discounting in our model. Panel A depicts the cumulative distribution function of the mixed strategy best seller prices for different values of the traffic generation potential of the best seller, $\frac{n}{N}$. As the traffic generation potential of the best seller increases (a larger segment size n gives more opportunity to cross-sell the basket), the lower bound of the support shifts to the left and allows for deeper price cuts, while the percentage of prices below cost increases. We test this finding as H1 in the "Empirical Support" section. The relative average basket size compared with the reservation price of the best seller has a similar effect on the distribution of prices. As $\frac{b}{r}$ increases, so does the frequency of loss-leading prices and the depth of the discounts (see Fig. 2, Panel B). Markets with larger basket sizes experience deeper discounts. Given similar traffic generation potential, the lower-priced items are more likely to be loss leaders. This finding partially explains why staple items with relatively lower base prices are more likely to be loss leaders. We also test these finding as H2 and H3 in our empirical analyses. By putting the "loss" in loss leading, our model shows

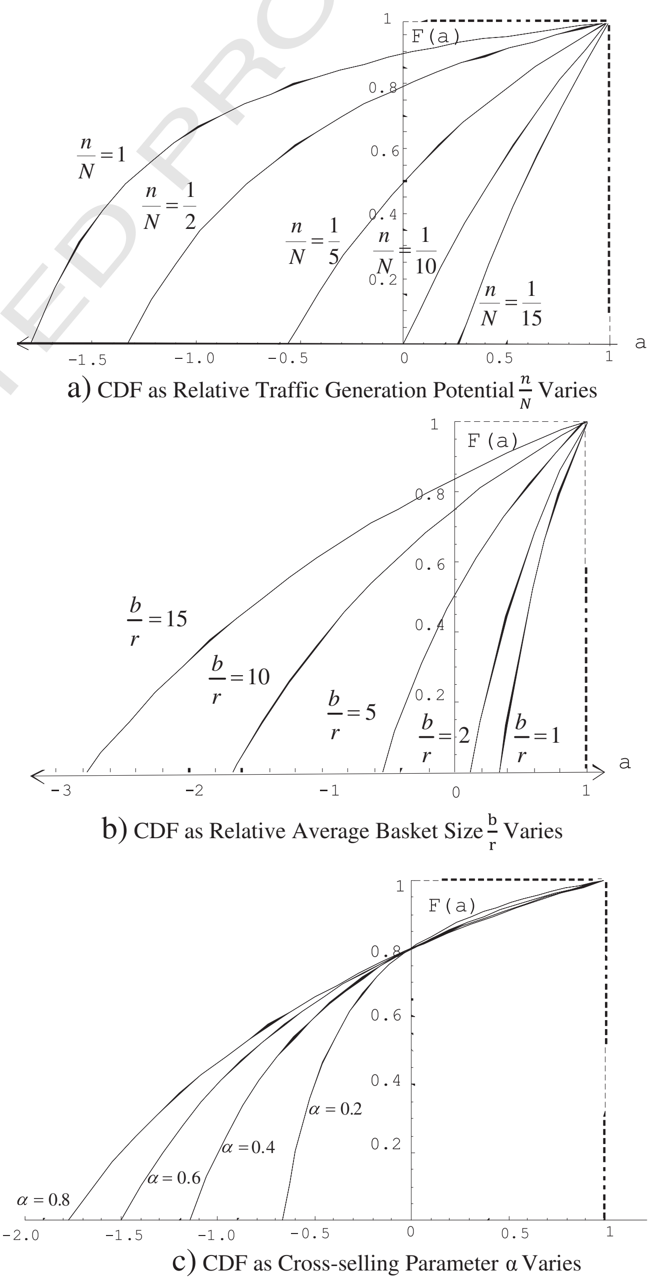


Fig. 2. Comparative statics of the cumulative distribution function.

Please cite this article as: Cenk Kocas, et al., Pricing Best Sellers and Traffic Generators: The Role of Asymmetric Cross-selling, Journal of Interactive Marketing (2017), http://dx.doi.org/10.1016/j.intmar.2017.09.001

that a best seller can be priced below cost if either its traffic generation potential is great enough or the revenue potential with respect to the average basket it could cross-sell is high enough.

What happens to the best seller price a as the cross-selling rate α changes? As Fig. 2 Panel C shows, α does not affect the frequency of below-cost (loss-leading) prices but rather the depth of loss-leading discounts. Only through considering probabilistic pricing strategies can we effectively distinguish the depth and frequency of pricing discounts. Greater cross-selling potential leads to deeper loss leader discounts. The frequency of below-cost discounts is given by $P(a<0) = F(0) = 1-\frac{N\,r}{2\,nb}$, which is independent of α. This finding, however, holds only when the inclusion incidence of an item is similar to its conversion incidence ($\alpha_a=\alpha_b=\alpha$).We now consider the case when the cross-selling incidences are different, providing additional insight into loss leader pricing.

When Inclusion Incidence Is Different from Conversion Incidence

The conversion and inclusion cross-selling rates may differ in some retail settings. For example, for seasonal items such as turkeys at Thanksgiving or eggs at Easter, the conversion rate is probably higher than the inclusion rate; that is, customers go shopping for these particular items rather than simply happening to buy these items on shopping trips initiated by other needs. The results, summarized in Appendix B, show that the optimal frequency of discounts should be higher for items with conversion rates higher than the inclusion rates. Formally, we use the notations $\alpha_a=\alpha$ and $\alpha_b=\beta$ and define α/β as the conversion-to-inclusion ratio. The frequency of discounts for the best seller price (a) when $\alpha\neq\beta$ is given by $F[a] = 1-\frac{N\beta(r-a)}{2n(a+b\alpha)}$. This frequency is higher for products with higher conversion-to-inclusion ratios, α/β, such as seasonal items. This finding provides an analytical explanation to the empirical generalization that seasonal items are discounted heavily (Chevalier, Kashyap, and Rossi 2003).

Asymmetric Case

Without loss of generality, assume that $b_1>b_2$, all else being equal, such that R1 has a larger average basket size. We expect that as a result of this asymmetry, R1 has potentially more to gain from cross-selling and is motivated to offer deeper price cuts on the best seller than R2.

To focus on basket size asymmetry ($b_1>b_2$) rather than on customer segment size asymmetry, we assume in our discussion that N customers are shared equally by both retailers; that is, $N_1 = N_2 = \frac{N}{2}$. We provide an analysis of the case when $N_1\neq N_2$ in Appendix C. We again assume that $\alpha_a=\alpha_b=\alpha$.

P_3. *The profit-maximizing distribution of best seller prices for the retailer with the larger average basket size, R1, is given by the mixed strategy* $F_1[a] = 1-\frac{N\alpha(r-a)}{2n(a+b_2\alpha)}$, *and the bounds are given by* $a_{min} = \frac{Nr\alpha-2nb_2\alpha}{N\alpha+2n}$ *and r.*

P_4. *The retailer with the smaller average basket size, R2, has a higher average price than R1. Although R2 has equal discount depths as R1, the frequency of discounts is lower for R2, with a mass* $M = \frac{\alpha(b_1-b_2)}{r+\alpha b_1}$ *at r. The profit-maximizing distribution of prices for R2 is given by the mixed strategy:*

$$F2[a] = 1-\frac{N\alpha(r-a)+2Mn(r+b_1\alpha)}{2n(a+b_1\alpha)} = \frac{2n(a+b_2\alpha)+N\alpha(r-a)}{2n(a+b_1\alpha)}. \tag{3}$$

The analysis of the asymmetric case helps explain the pricing dynamics of our opening vignette. Because they offer products in many categories and subcategories, mass merchandisers such as Amazon.com and Wal-Mart achieve average basket sizes larger than other sellers, whether online or offline. Their larger potential profit margin, due to their larger average basket sizes, motivates and allows them to offer deeper discounts on the most anticipated best sellers. From P_4 we indeed expect the larger retailers to engage in loss leader pricing more frequently for a given set of items (see Fig. 3). Although a retailer with a smaller average basket size can offer similarly deep discounts, it can do so only less frequently or on fewer items given it has less to gain in a cross-selling conversion of a smaller basket. Thus, as P_3 and P_4 demonstrate, the larger average basket size retailer R1 can grant (1) a deeper average discount on a given set of items than R2, (2) the same discounts on the same items as R2 but more frequently, and (3) the same discounts on more items than R2. We will also demonstrate (P_5) that such an advantage leads not only to lower prices but also to increases in R1's profits.

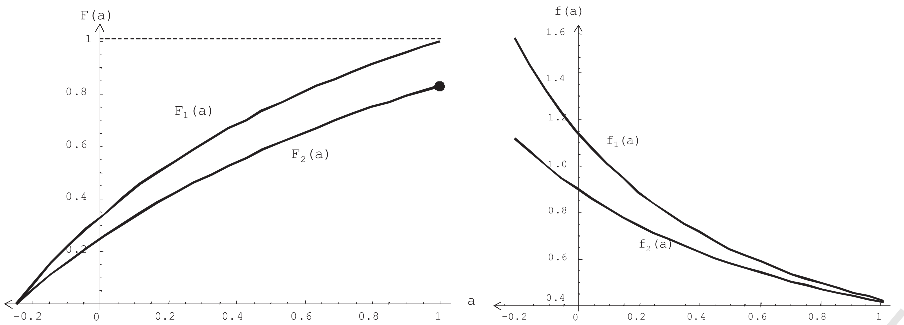
We note that four properties of the symmetric case remain valid for the asymmetric case: (1) the minimum and average prices decrease as $\frac{n}{N}$ increases; (2) the minimum and average prices for the best seller decrease as $\frac{b}{r}$ increases, though only b_2, the smaller average basket size, determines this ratio for both retailers; (3) a higher valued cross-selling parameter α increases the discount depths while leaving the frequency of discounts unchanged; and (4) loss leader prices are possible.

Next, we compare the retailer profits and range and frequency of prices across the symmetric and asymmetric duopolies with three propositions. Note that the default profit of a retailer is the profit it would make if it exclusively served the $\frac{N}{2}$ customers with the average basket and the $\frac{N}{2}\alpha$ customers with the best seller priced at r.

P_5. *The asymmetric equilibrium leads to higher profits for the larger retailer R1 than for R2. The profit of R2 is its default at* $\pi_2 = \frac{N(b_2+\alpha r)}{2}$, *and the profit of R1 is more than its default at* $E\pi_1 = \frac{N(b_1+\alpha r)}{2} + n\alpha(b_1-b_2)$.

In the asymmetric case, R1 improves its profit by $n\alpha(b_1-b_2)$. That is, commanding a larger basket size improves the profitability of the larger basket retailer. The traffic implications are also promising for this larger basket retailer. Formally:

P_6. *In the asymmetric equilibrium the larger retailer R1 enjoys higher traffic than R2.*



Retailer 1 has larger average basket size (b1>b2)

Fig. 3. The asymmetric mixed strategy equilibria: cumulative and probability distributions for best seller price a under parameter values $\frac{n}{N}=.25$, $\frac{b1}{r}=4$, $\frac{b2}{r}=3$, r=1, AND α=.5.

Prior research has demonstrated that loss leader pricing leads to neither a significant increase in store traffic nor an increase in profits (Lal and Matutes 1994; Walters and MacKenzie 1988). As P_1 shows, this argument holds in the symmetric duopoly case. However, P_5 and P_6 demonstrate that asymmetry between retailers leads to both increased profits and increased traffic for the retailer with the marginal advantage from cross-selling. The other retailer loses traffic, and its profit is unchanged.

The retailer asymmetry also has important implications on the pricing strategies. For comparison purposes, assume that relative to the average basket size under retailer symmetry (b_{sym}), the asymmetric case has $b_1>b_{sym}>b_2$. The asymmetry has the effect of lessening the overall severity of price competition between the two retailers. Formally,

P_7. *The severity of price competition is greater for symmetric retailers than under asymmetry in average basket size. Formally, assume $b_1>b_{sym}>b_2$. Then, $F_{sym}[a]>F_{1asym}[a]>F_{2asym}[a]$.*

In the asymmetric case, the dominance of the larger retailer R1 enables it to offer lower prices than R2. This asymmetry forces R2 to retreat to offering less frequent, less deep price discounts. Consequently, R1 follows suit and offers discounts only as deep as those offered by R2 at a higher frequency or, equivalently, on a greater number of products. Hence, discounts are shallower in the asymmetric case compared to the symmetric case (see Fig. 4). When retailers have similar basket sizes, severity of the competition leads to lower minimum and average prices. Recall that P_5 demonstrated the profitability of the larger retailer R1 being higher than that under symmetric competition, with R2 having the same profit regardless of retailer asymmetry. The larger retailer R1 takes full advantage of its ability to cross-sell by aggressively driving traffic through best seller discounts.

Empirical Support

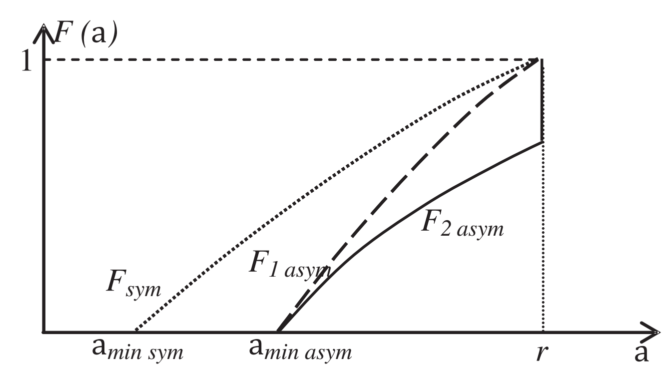
The theoretical propositions from our model make several predictions about retailer price discounting strategies we should observe in empirical price data. Although online pricing data are readily available, a lack of precise data on individual model parameters does not always allow direct tests of individual propositions. Nevertheless, our model findings do lead to several testable hypotheses, which if supported can further increase confidence in the model.

The dependent variable of interest is discounted price observations for best seller items sold by retailers. Price data for multiple products represent repeated observations of a mixed pricing strategy over time (e.g., Iyer and Pazgal 2003; Kocas and Bohlmann 2008; Raju, Srinivasan, and Lal 1990; Ratchford 2009). A price discount reflects an observed price for a specific product lower than the item's highest (list) price. We consider a retailer's average discounting behavior across a set of best seller items, in our case books.

Both the symmetric and asymmetric models predict that products with higher traffic generation potential, $\frac{n}{N}$, should be offered at deeper, more frequent discounts. The traffic generation potential of any product can be assessed by the sales rank, or popularity, of the item. Using sales rank as a proxy for traffic generation potential, we state our first hypothesis:

H1. Products with higher sales rank have a) deeper and b) more frequent discounts.

Moreover, our (symmetric and asymmetric) models predict that larger relative basket sizes $\frac{b}{r}$ lead to deeper, more frequent discounts. A larger relative basket size may be due to either a



Retailer 1 has larger average basket size (b1>b2)

Fig. 4. Comparison of the symmetric and asymmetric cases.

low reservation price (list price) on a best seller or a high relative average basket size. Thus:

H2. Products with lower list prices have a) deeper and b) more frequent discounts.

H3. Retailers with larger average basket sizes offer a) deeper and b) more frequent discounts.

To test our hypotheses, we gather three data sets: the first represents a time series of prices to test H1a (sales rank affects discount) and H2a (list price affects discount) in a model accounting for dual causality, the second represents a more comprehensive cross-sectional data set to test H1a and H2a for a wide range of best seller sales rank, and the third data set combines two online book price comparison sites to test H3a (retailer basket size affects discounts). We formally test our stated hypotheses only with respect to the depth of promotions, not frequency of promotions, because of the cross-sectional nature of our larger data set. The descriptive statistics for the first two datasets are given in Table 2; descriptive statistics for the third dataset are presented later in Table 5. In all our analyses, we standardize book prices with respect to their list (regular) prices by dividing the current price by the list price. An observed discount corresponds to any standardized price less than 1.

We present the details and corresponding analysis next.

Data Set 1: Amazon.com Time Series Data

The first data set runs from June 1, 2011 to Sept 3, 2011, a total of 3 months, on 7,332 books which were listed under *New releases > coming soon* at Amazon.com. The advantage of this data is that we can observe each book from the start of its availability. Data were collected on a rolling basis, and include price, Amazon Book Ranks (ABRank), the physical format of the book (hardcover or not), the average customer review, and number of sellers.

Analysis of Data Set 1

Our analysis proceeds in the steps of persistence modeling (Trusov, Bucklin, and Pauwels 2009) to explicitly analyze potential dual causality among price and sales rank (H1a). In the first step, we test for Granger Causality among price and sales rank. This test only reveals whether one variable drives another, not the direction (sign) nor magnitude of this relationship. To this end, we next estimate a vector autoregression (VAR) model with specification according to the unit root and cointegration tests. Based on this model, generalized impulse response functions (GIRF) track the over-time impact of a change in one variable to the other variables in the model. As in previous VAR applications (e.g. Trusov, Bucklin, and Pauwels 2009) we calculate the cumulative elasticity as the sum of all impulse response coefficients significantly different from zero at the 95% significance level.

The Granger Causality tests clearly show dual causality at the $p < 0.05$ significance level, considering up to 8 lags. Specifically, sales rank is both driven by and drives list price and discount at any lag ($p < 0.01$). Number of sellers also shows dual causality with both list price and discount at any lag ($p < 0.01$), as well as with sales rank ($p < 0.01$). List price drives discount at any lag ($p < 0.01$), although discount does not Granger cause list price ($p > 0.18$ for all lags). Number of sellers is also Granger caused by customer reviews at any lag ($p < 0.02$), but the reverse is not supported ($p > 0.05$). For customer reviews, dual causality with list price is supported only for 4 of the 8 tested lags ($p < 0.02$), while discount drives customer reviews at any lag ($p < 0.02$). Customer reviews Granger cause sales rank only starting lag 5 ($p < 0.03$), while sales rank causes customer reviews at any lag ($p < 0.03$). Because all variables are mean-stationary (as verified by unit

Table 2
Summary statistics for the Amazon.com data sets.

	Standardized price	Sales rank	Pub. year	# of sellers	Hardcover (1 = yes)	List price	Ave. customer review	Discount
Data Set 1: Amazon.com time series data with 7,332 books across 3 months								
Valid Obs.	847,405	613,439	847,405	360,369	847,405	847,405	366,050	847,405
Missing	0	233.966	0	487,036	0	0	481,355	0
Mean	.82	1,316,078	2011	20.68	.25	34.06	4.21	.18
Median	.78	529,057	2011	19.00	.00	19.99	4.30	.22
Std. Dev.	.16	1,929,478	0	11.88	.44	95.20	.58	.16
Minimum	.22	1	2011	1	.00	.00	1.57	.00
Maximum	1	10,517,303	2011	111	1.00	4,271.00	5.00	.78
Data Set 2: Comprehensive Amazon.com data with 819,377 books								
Valid Obs.	819,377	819,377	819,377	737,999	819,377	819,377	410,207	819,377
Missing	0	0	0	81,378	0	0	409,170 [a]	0
Mean	.93	3,010,871	1999	13.92	.37	37.04	4.23	.067
Median	1.00	2,163,100	2002	11.00	.00	20.00	4.40	.000
Std. Dev.	.12	2,712,880	9.32	13.61	.48	41.97	.81	.12
Minimum	.037	14	1913	1	.00	.39	1.00	.00
Maximum	1.00	9,999,948	2012	6,045	1.00	199.99	5.00	.96

[a] A specification omitting Avg. Customer Reviews vastly improves the number of valid cases; however, all coefficient signs and significances remain the same. We present the broader analysis here.

Please cite this article as: Cenk Kocas, et al., Pricing Best Sellers and Traffic Generators: The Role of Asymmetric Cross-selling, Journal of Interactive Marketing (2017), http://dx.doi.org/10.1016/j.intmar.2017.09.001

Table 3
Same-day and cumulative effects on discount depth (from var. model).

	Sales rank Same day	Sales rank 30 days	List price Same day	List price 30 days	Customer review Same day	Customer review 30 days
Response estimate	−0.044	−1.631	0.114	0.964	0.009	0.238
Standard error	*0.003*	*0.078*	*0.006*	*0.095*	*0.002*	*0.065*
Elasticity	−0.0005	**−0.017**	0.0093	**0.079**	0.0019	0.051

root tests), we specify the VAR model with Discount, Rank, Sellers, List price and Customer reviews as endogenous variables (explained by the model), and a constant and physical format (a dummy with 1 = hardcover) as exogenous variables, as shown in Eq. (4) below:

$$\begin{bmatrix} \text{Discount}_t \\ \text{Rank}_t \\ \text{Sellers}_t \\ \text{Listprice}_t \\ \textit{Cust Rev}_t \end{bmatrix} = \begin{bmatrix} \alpha_D \\ \alpha_R \\ \alpha_S \\ \alpha_L \\ \alpha_C \end{bmatrix} \times \text{Format} + \sum_{j=1}^{J} \begin{bmatrix} \phi^j_{11} & \phi^j_{12} & \phi^j_{13} & \phi^j_{14} & \phi^j_{15} \\ \phi^j_{21} & \phi^j_{22} & \phi^j_{23} & \phi^j_{24} & \phi^j_{25} \\ \phi^j_{31} & \phi^j_{32} & \phi^j_{33} & \phi^j_{34} & \phi^j_{35} \\ \phi^j_{41} & \phi^j_{42} & \phi^j_{43} & \phi^j_{44} & \phi^j_{45} \\ \phi^j_{51} & \phi^j_{52} & \phi^j_{53} & \phi^j_{54} & \phi^j_{55} \end{bmatrix} \times \begin{bmatrix} \text{Discount}_{t-j} \\ \text{Rank}_{t-j} \\ \text{Sellers}_{t-j} \\ \text{Listprice}_{t-j} \\ \textit{Cust Rev}_{t-j} \end{bmatrix} + \begin{bmatrix} \varepsilon_{D,t} \\ \varepsilon_{R,t} \\ \varepsilon_{S,t} \\ \varepsilon_{L,t} \\ \varepsilon_{C,t} \end{bmatrix} \tag{4}$$

Consistent with the Granger Causality tests, the Schwartz Bayesian Information Criterion (SBIC) selects 5 daily lags as the optimal balance between forecasting accuracy and parsimony. At this lag, the VAR-model passes the typical diagnostic tests (Franses 2005) and explains 97.3% of the variance in sales rank, 98.6% in customer reviews, 99.9% in list price and 96.6% in Discount. Table 3 shows the GIRF estimates of interest (both the same-day effects and the cumulative effects over 30 days) and their standard errors.

The GIRF of Discount shows that discounts are deeper for products with a better sales rank (cumulative elasticity = −.017) in support of H1a. Moreover, discounts are deeper for books with higher list price (.079) across all sales ranks, counter to H2a. We discuss this finding in detail in the analysis of the next data set. Finally, discounts are deeper for books with a better average customer review (.051). We further analyze these relations in the next data set.

Data Set 2: Comprehensive Amazon.com Data

The second dataset is cross sectional and has more books to test H1a and H2a for a wide range of best seller sales rank, including different bins of the data (i.e. books in the top 10^3 and the top 10^5). A web agent collected a random sample of 2,274,890 ISBN numbers in a 15-day period, ending on May 14, 2011. We collect the price and sales rank information, year of publication, number of sellers, the average customer review, and the physical format of the book. By removing formats other than paperbacks and hardcover books, items with missing prices, sales rank, publication year data, books with list prices higher than \$200, and books not sold by Amazon.com, we attain a sample of 819,377 books. Book prices are again standardized.

We run a linear regression on the whole data set to test H1a and H2a. We also run linear regressions based on logarithmic bins to demonstrate that bestseller status and effects on prices exist not only for the classical bestsellers (i.e. top 10^3), but also far down the sales ranks, even into one millionth sales ranks. Each bin represents a relatively homogeneous set of books according to sales ranks. The bins are the top 10^3, 10^3 to 10^4, 10^4 to 10^5, 10^5 to 10^6, and 10^6 to 10^7. We want to observe the signs and magnitudes of the coefficients in the regression equation:

$$\begin{aligned} \text{Discount} = {} & \alpha + \beta_1 \text{ Rank} + \beta_2 \text{ Year} + \beta_3 \text{ Sellers} \\ & + +\beta_3 \text{ Hardcover} + \beta_4 \text{ List Price} \\ & + \beta_5 \text{ Ave.Customer Review} + \varepsilon \end{aligned} \tag{5}$$

where Discount = 1 − standardized price, and Hardcover is a dummy variable (Hardcover = 1).

Analysis of Data Set 2

Results are shown in Table 4. For the control variables (i.e., year, sellers, hardcover, and average customer review), we find that newer books are offered at significantly deeper discounts than older books. Deeper discounts are observed for books carried by more Amazon sellers, probably because of heightened competition. Hardcover books are offered at significantly deeper discounts up to a sales rank of 100,000; however, this trend reverses between 100,000 and 1 million. Hardcover books with sales ranks higher than 1 million are sold at a significantly lower discount than paperbacks. We discuss this finding subsequently. Also, the higher the average customer review for a book, the higher is the discount.

We now test H1a and H2a on the basis of this data set. As the average discount column of the top panel of Table 4 shows, as well as the negative sign of the Sales Rank parameter in the overall regression of all books, better-selling books have significantly deeper discounts, as H1a predicts. The sales rank coefficients for all bins are significant and negative, in support of H1a. Best sellers with higher sales ranks have deeper discounts. The transition to best seller pricing is not discrete, as prior literature on loss leaders would suggest. Rather, we find that the prices of all books are affected by their inherent traffic potential, from the top 1,000 to the 10 millionth-ranked books in the long tail. The Frequency on Sale column of Table 4 also suggests that better-selling books are on sale more frequently, consistent with H1b.

Table 4
Linear regression, based on logarithmic bins.

		Model fit					
Bin	N	Average discount	Frequency on Sale	R^2	Adj. R^2	d.f.	F-value
1 to 10^3	351	38%	97%	0.13	0.114	350	8.53 ***
10^3 to 10^4	3,489	31%	88%	0.066	0.065	3,488	41.19 ***
10^4 to 10^5	29,646	24%	80%	0.058	0.058	29,645	302.99 ***
10^5 to 10^6	162,588	15%	58%	0.168	0.168	162,587	5,455.53 ***
10^6 to 10^7	195,215	4%	15%	0.084	0.084	195,214	2,970.72 ***
All	391,289	10%	29%	0.228	0.228	391,288	19,285.53 ***
Standardized beta coefficients							
Bin	Constant	**Sales rank**	Year of publication	Number of sellers	Hardcover	**List price**	Ave. cust. review
1 to 10^3	−3.92 **	**−.156** ***	.120 **	.225 ***	.171 ***	**−.092** *	0.077
10^3 to 10^4	−5.08 ***	**−.131** ***	.134 **	.035 **	.161 ***	**−.096** ***	0.015
10^4 to 10^5	−4.04 ***	**−.087** ***	.110 ***	.120 ***	.107 ***	**.023** ***	.019 ***
10^5 to 10^6	−5.38 ***	**−.168** ***	.138 ***	.256 ***	0.003	**.148** ***	.019 ***
10^6 to 10^7	−3.16 ***	**−.034** ***	.123 ***	.181 ***	−.027 ***	**.102** ***	.004 *
All	−5.44 ***	**−.236** ***	.150 ***	.251 ***	−.019 ***	**.102** ***	.020 ***
Hypothesis		**H1a**				**H2a**	
Predicted relationship		−				−	

Dependent variable is Discount.
* $p < .10$.
** $p < .05$.
*** $p < .01$.

To test H2a, we examine the coefficient of the list price considering all books, with additional analysis across the five bins (Table 4). The effect is significant and as expected for best-selling books with ranks up to 10^4. That is, for significant traffic generators, a lower list price leads to a significantly deeper discount on the book. However, for books with higher sales ranks, the effect is reversed. Thus, we find support for H2a, though only up to a point in the sales rankings. The hypothesis that products with lower list prices have deeper discounts is supported only if these products have relatively significant traffic generation potential. The interplay between list price and hardcover status depicts a more comprehensive picture, which we examine next.

In general, retailers discount a hardcover book less and a book with a higher list price more, as the overall regression parameters for the hardcover and list price in the last row of Table 4 confirm. Hardcover books target customers with lower price sensitivities, so it is not surprising that they are discounted less. A higher list price also provides more room for discounts (a given percent discount gives a higher discount value), given similar absolute cost structures for books; therefore, it is also not surprising that a book with a higher list price is discounted more.

The hardcover and list price columns at the bottom panel of Table 4 reveal a switch of the basis for discounting along the sales rank. In the long tail of the sales distribution, where sales ranks are in millions, a book is discounted less if it is a hardcover and is discounted more if its list price is high. However, as we show in the first and second rows, where sales ranks are up to 10,000, a book is discounted more if it is a hardcover or if its list price is low. Though contradictory to the general case, this finding is consistent with our model premises. Our model predicts that a book that acts as a traffic generator should be discounted heavily, which is true for books in the top 10,000. Moreover, when we control for list price, hardcover status is still an attractive attribute, so hardcover books with high sales ranks could still be offered at significantly deeper discounts. Although we do not model hardcover status explicitly in our model, the finding that hardcover books in the top 10,000 are discounted more is consistent with our model, given their relative attractiveness and traffic generation potential. Our previous finding that books with higher average customer reviews are discounted deeper also resonates with these results. Overall, these findings provide strong empirical support for H2a.

Data Set 3:Online Book Price Comparison Sites Data

To test for H3a, we collect data on Amazon.com's top-100 best-selling books on October 18, 2011, from multiple online retailers. We collect pricing data from 37 retailers in the three-day period ending with October 20, 2011, from two book price comparison sites, bookstores.com and addall.com. Dropping from the list marketplaces, auctions, and used-book sales, as well as retailers located outside the United States and those that carried fewer than 30 of the top-100 best-selling books, we obtained a final list of 19 retailers. Four retailers in this list are multicategory (MC) retailers (Walmart.com, Overstock.com, Amazon.com, and buy.com), and the remaining 15 are bookstores. Table 5 lists the 19 retailers and the average discounts they offered on the top-100 books sold. The four MC retailers fill the top spots with average discounts of 45%–48%. Bookstores occupy the remaining spots with average discounts of 7%–44%.

Table 5
Retailers' top-100 best-selling books statistics.

Rank	Retailer	Format	Average discount	Std. Dev.	Minimum discount	Maximum discount	Number of top-100 books on sale
1	Wal-Mart	Multicategory	48.2%	7.6%	28.0%	82.2%	98
2	Overstock.com	Multicategory	47.9%	7.8%	27.4%	67.9%	96
3	Amazon.com	Multicategory	47.8%	7.5%	33.3%	82.2%	99
4	Buy.com	Multicategory	44.9%	7.5%	30.9%	84.8%	99
5	Barnes & Noble	Bookstore	44.5%	9.3%	0.0%	82.2%	98
6	Alibris	Bookstore	44.1%	12.4%	2.9%	71.2%	92
7	AbeBooks	Bookstore	41.8%	13.8%	0.0%	66.5%	89
8	Books-A-Million	Bookstore	35.4%	14.2%	0.0%	80.6%	99
9	ValoreBooks.com	Bookstore	35.2%	14.7%	0.0%	64.2%	85
10	TextbookX	Bookstore	31.5%	9.5%	10.0%	58.6%	83
11	Book Byte	Bookstore	28.0%	7.6%	11.7%	55.6%	77
12	Better World Books	Bookstore	26.2%	13.9%	0.0%	60.6%	88
13	Strand Bookstore	Bookstore	25.0%	20.6%	0.0%	71.0%	59
14	Bookstores.com	Bookstore	24.4%	13.7%	0.0%	59.8%	65
15	TextbooksRus	Bookstore	22.4%	8.2%	6.5%	45.4%	89
16	Borders	Bookstore	22.2%	18.0%	0.0%	77.9%	96
17	BiggerBooks	Bookstore	21.5%	10.6%	2.0%	74.8%	99
18	eCampus	Bookstore	19.9%	10.8%	0.0%	74.2%	99
19	Powell's Books	Bookstore	7.1%	12.5%	0.0%	69.7%	88

Analysis of Data Set 3

We run paired samples t-tests to determine whether the average prices of MC retailers are lower than those of bookstores, as our model would predict. The t-values and corresponding significance levels appear in Table 6. With 4 MC retailers (columns in Table 6) and 15 bookstores (rows in Table 6), there are 60 comparison pairs; as the t-values show, the MC retailer prices are significantly lower for 58 of these 60 pairs. Thus, we find significant support for H3a ($\chi^2 = 52.26$, $p < .01$); retailers with larger average basket sizes offer significantly deeper discounts. Table 5 also presents the number of books available on sale for each retailer that are among the top-100 books sold by Amazon.com. If we consider the percentage of books available for sale in the top 100 as a proxy for frequency of discounts, we find that of the 60 pairs, 8 have an equal number of books, 7 have more books sold by the bookseller than the MC retailer, and 45 pairs have more books sold by the MC retailer than the bookseller. A chi-square test for frequencies (grouping 8 pairs with an equal number of books with 7 pairs against H3b versus 45 pairs for H3b) shows that MC retailers carry significantly more books in the top 100 than booksellers ($\chi^2 = 15$, $p < .01$), consistent with H3b. Given their larger basket sizes, the MC retailers also carry more best seller products to increase their cross-selling efforts.

The data sets provide empirical support for the findings from our theoretical model, supporting all of our hypothesized relationships for discount depth. Our empirical work shows that books with higher sales ranks have deeper discounts, and this relationship holds farther down the best seller list. Books with lower list prices also have deeper discounts, though this relationship does not hold farther down the best seller list. We also show that larger basket size (multicategory) retailers offer deeper discounts on the top best sellers, as our opening example suggests.

Table 6
Paired t-test results of comparisons of multicategory retailer prices with bookstore prices.

	Walmart.com	Overstock.com	Amazon.com	Buy.com
Barnes & Noble	6.44***	5.15***	5.39***	0.40
Alibris	2.26**	2.30**	2.12**	1.53
AbeBooks	2.95***	3.88***	2.85***	2.07**
Books-A-Million	7.29***	7.17***	6.76***	5.75***
ValoreBooks.com	4.35***	8.11***	4.30***	3.50***
TextbookX	19.93***	15.02***	19.14***	14.28***
Book Byte	21.65***	15.81***	20.57***	15.79***
Better World Books	6.47***	6.42***	6.13***	6.06***
Strand Bookstore	7.27***	7.27***	7.55***	6.46***
Bookstores.com	6.61***	13.19***	6.58***	5.82***
TextbooksRus	42.45***	27.65***	40.03***	31.11***
Borders	13.75***	12.12***	13.52***	12.52***
BiggerBooks	26.21***	21.07***	22.87***	22.52***
eCampus	26.38***	21.43***	23.15***	22.99***
Powell's Books	23.36***	20.16***	21.26***	21.53***

Notes: Multicategory retailer prices are lower than bookstores' prices at *p < .10, **p < 0.05, ***p < .01.

Discussion

In this research we set out to examine how profit-maximizing online retailers should price traffic generators in a competitive market. Our analytical model treats traffic generation potential as a continuous variable and is unique in differentiating and modeling attraction (traffic generation potential), cross-selling (conversion incidence), and the effects of promotions when the best seller is included in a larger shopping basket (inclusion incidence). Uncovering the tensions of this linkage between the motivation to lower prices of traffic generators and the motivation to increase their prices in anticipation of higher-margin basket incidences is a unique contribution of our model. We show that the frequency and the depth of discounts are higher for products with higher conversion-to-inclusion ratios, such as seasonal items or best-selling books. Our empirical

Please cite this article as: Cenk Kocas, et al., Pricing Best Sellers and Traffic Generators: The Role of Asymmetric Cross-selling, Journal of Interactive Marketing (2017), http://dx.doi.org/10.1016/j.intmar.2017.09.001