Behavioral Price Discrimination in the Presence of Switching Costs

Koray Cosguner †

Assistant Professor of Marketing, Georgia State University

Tat Chan

Professor of Marketing, Washington University in St. Louis

P.B. (Seethu) Seetharaman

W. Patrick McGinnis Professor of Marketing, Washington University in St. Louis

Forthcoming at Marketing Science

Permalink: http://dx.doi.org/10.1287/mksc.2016.1024

† Corresponding Author: Koray Cosguner, email: kcosguner@gsu.edu; Ph: 404-413-7697. This article is based on the first author's doctoral dissertation, written at Olin Business School under the supervision of the second and third authors. The authors thank seminar participants at Olin, Georgia State, Yale, Rice, Alberta, Johns Hopkins, Emory, Carnegie Mellon, Koc, Sabanci, HEC, and Ozyegin for their comments and feedback on earlier versions of this article.

Behavioral Price Discrimination in the Presence of Switching Costs

Abstract

We study the strategic impacts of behavioral price discrimination (BPD) on manufacturers and retailers in a distribution channel when there are switching costs in consumer demand. Unlike previous empirical studies of behavioral price discrimination, which rely only on differences in price elasticity *across customers*, our pricing model allows the firm strategies to additionally account for differences in price elasticity *across time* (due to switching costs). We estimate a dynamic pricing model using empirical data from the cola category and, through a series of counterfactuals, we find that the retailer should simply outsource the data analytics and customization of coupons to manufacturers and improve its profit beyond what it can achieve by proactively couponing on its own. We further find that serving as an information broker to sell its customer database to manufacturers can be a vital source of profit to the retailer. In contrast, manufacturers end up worse off, illustrating that customer information is a potent source of channel power to the retailer. Finally, we show that simply using customers' most recent purchase information can significantly impact firms' profits. BPD based on this information is easy to implement and of low cost to manufacturers and retailers.

Keywords: Behavioral Price Discrimination, Dynamic Pricing, Targeted Coupons, Switching Costs, Inertia.

Introduction

Behavioral price discrimination (BPD henceforth) refers to the pricing strategy of firms charging different prices to customers with different purchase histories (Shaffer and Zhang 1995, Villas-Boas 1999, Fudenberg and Tirole 2000, Chen, Narasimhan and Zhang 2001, Villas-Boas 2004, Acquisiti and Varian 2006, Pazgal and Soberman 2008, Shin and Sudhir 2010; see Fudenberg and Villas-Boas 2006 for a detailed discussion and review of BPD). This strategy is enabled by the fact that customers' purchase histories are observed in retail transactional data collected by retailers using loyalty cards (such as the *Kroger Rewards* card). Using such customer-level transactional data, firms can estimate customer-specific price elasticities of demand and, therefore, tailor customer-specific retail prices using targeted coupons (see, for example, Elrod and Winer 1982, Rossi, McCulloch and Allenby 1996). Previous empirical research on BPD has assumed that while price elasticities can vary across customers, a given customer's price elasticity is time invariant. However, an extensive empirical literature in marketing has documented that on account of switching costs in demand, a customer's price elasticity for the most recently purchased brand decreases at her next purchase occasion (see, for example, Seetharaman 2004). This means that while determining the optimal BPD strategy, the retailer must recognize differences in brands' price elasticities not only across customers but also across time (due to switching costs). Taking switching costs into account, firms can better tune their BPD schemes by offering the right coupons not only to the "right customers" but also at the "right time."

One pricing implication of switching costs in demand is that manufacturers or retailers can target price-off coupons to consumers that are based on their most recently purchased brands. Suppose that the brand manager of Coke observes that John Doe's most recently purchased brand in the cola category is Pepsi. She will realize that, on account of switching costs, she should offer

John Doe a price discount on Coke in order to induce him to buy Coke instead of Pepsi. While such a price discount reduces the current profit margin of Coke, offering John Doe a coupon can still be profitable to Coke because, once John Doe buys Coke, he will be more likely to buy Coke again in the future. This future sales gain can dominate the cost of current price discount. The brand manager of Pepsi faces a similar incentive to lure previous buyers of Coke using price discounts. Whether such competitive pressures, which force firms to "poach" each other's customers (Chen 1997, Shaffer and Zhang 2000, Taylor 2003), could dominate the incentive to "milk" the installed customer base (i.e., most recent buyers) by charging higher prices, is an interesting empirical question. Without accounting for switching costs, previous empirical studies on BPD have concluded that competing firms benefit from BPD (Besanko, Dube and Gupta 2003, Pancras and Sudhir 2007). In this research, we study whether the dynamic competition discussed above yields different strategic implications for firms when they engage in BPD.

When studying BPD, the strategic role of the retailer cannot be ignored (Liu and Zhang 2006). This is because, as a category profit maximizer, the retailer's pricing incentive may not align with the manufacturers' pricing incentives. This means that the profit implications of BPD derived for competing manufacturers may unravel in the presence of the retailer. Furthermore, even if manufacturers can effectively by-pass the retailer by dropping direct-mail targeted coupons to end-consumers, the retailer can put paid to such efforts by not sharing its customer database with the manufacturers, which makes it impossible for the manufacturers to estimate customer-specific price elasticities (which are necessary for engaging in BPD). It is also logistically easier for the retailer to offer coupons to consumers at the store than for a manufacturer to mail coupons to consumers' homes. In this research, we study BPD within the distribution channel in a market with switching costs. In doing this, we are able to answer the question of whether it is profitable for the retailer to share its customer information with manufacturers and letting them engage in BPD instead of managing the BPD by itself.

We take an empirical approach. We estimate a structural model of consumer demand and firm pricing behavior, which is similar to Cosguner, Chan and Seetharaman (2016), and use the estimates to perform counterfactual experiments regarding BPD. In one counterfactual, we study a "partial" BPD scenario, where channel members offer targeted coupons based on customers' most recent purchases only. Another counterfactual is a "full" BPD scenario, where the offered coupons are based on customers' full purchase histories and channel members can correctly infer customers' price elasticities, switching costs, and brand preferences. Our key findings are as follows: First, significant switching costs exist among consumers; thus, optimal BPD should adjust over time even for the same set of consumers. Second, the retailer's profit increases when it engages in BPD. Interestingly, the profit increases even more when the retailer lets manufacturers employ BPD. In other words, the retailer can simply outsource the data analytics and customization of coupons to manufacturers and gain profit beyond what it can achieve by proactively engaging in BPD. Manufacturers' profits may increase or decrease when they engage in BPD, depending on how much of customers' purchase histories they can access. Third, the retailer's profit further improves if it can charge the manufacturers for access to its customer database. Serving as an information broker to sell transactional data to manufacturers can, therefore, be a vital source of business profit to the retailer (see Glazer 1991 for an early view on information as a business asset). However, the manufacturers end up being worse off than under the case of no price discrimination. Finally, we show that simply using customers' most recent purchase information can significantly impact firms' profits. We argue that BPD based on this information is easy to implement and of low cost to manufacturers and retailers.

Behavioral Price Discrimination in the Presence of Switching Costs

Our model is directly adopted from Cosguner, Chan and Seetharaman (2016), who study the harvesting and investing incentives of manufacturers and retailers. Our goal in this paper is to investigate the profit impact of BPD for manufacturers and retailers. To achieve this goal, we use the model estimates to conduct counterfactual simulations. Because model estimation is not the focus of this paper, we keep the discussion of the model setup brief, and skip the discussion on how the model is estimated. We refer interested readers to Cosguner, Chan and Seetharaman (2016).

We consider a monopolist retailer in a local market. There are *J* major manufacturers (Coca Cola and Pepsi in our application) selling their brands through the retailer. We assume a three-stage game: In the first stage (at the start of the game), the retailer, as the owner of the customer database that records customers' purchase histories, decides whether to directly engage in BPD, or to share the information with the manufacturers and help them employ BPD by distributing manufacturers' targeted coupons directly to consumers. In the second stage, manufacturers simultaneously choose wholesale prices for their brands and, if feasible, customized price-off coupons targeting different customers. In the third stage, taking manufacturers' wholesale prices and customized couponing decisions as given, the retailer chooses retail prices for the brands and, if feasible, customized price-off coupons targeting different customers.

Let S_{jt}^m denote a *state variable* that represents the segment *m*-specific ($m = 1, \ldots, M$) *installed base*, i.e., share of customers in latent segment *m* whose most recent purchase in the category was brand *j* ($j = 1, \ldots, J$), as of period *t* ($t=1, \ldots, T$). The set of state variables that determines manufacturers' and the retailer's BPD strategy is $\boldsymbol{S}_t = \{S_{1t}^1, S_{2t}^1, \ldots, S_{Jt}^1; \ldots; S_{1t}^M, S_{2t}^M, \ldots, S_{Jt}^M\}$, i.e., the collection of the installed bases of J brands within each of the M consumer segments. The objective of each firm in the channel is to maximize its long-term discounted profit. We assume that manufacturers compete in an infinitely repeated Bertrand pricing game and focus our attention on Pure-Strategy Markov-Perfect Equilibria (MPE). Let $\boldsymbol{a}_{jt}$ be the set of control variables for manufacturer j, which defines its BPD strategy. Suppose the manufacturer obtains customer information from the retailer in the first stage and, therefore, can identify that a customer bought brand k during his previous purchase and also that the customer belongs to latent segment m. The manufacturer can issue a coupon, $C\_mft_{jkt}^{m}$, to such a customer whose face value will be deducted from the retail price of brand j when the customer purchases brand j in period t. $C\_mft_{jkt}^{m}$ will directly impact the per-unit revenue for the manufacturer, w_{jkt}^{m}, and the "real" price (i.e., retail price minus coupon) that the customer has to pay for buying the brand, p_{jkt}^{m}. The manufacturer also has to decide wholesale price W_{jt}. Thus, the set of control variables is $\boldsymbol{a}_{jt} = \{W_{jt}, C\_mft_{jkt}^{m}, m = 1,..., M, k = 1,..., J\}$. If the retailer decides to directly engage in BPD, the retailer can issue targeted coupons, $C\_Ret_{jkt}^{m}$, $\forall j$, to customers who previously purchased brand k and belong to latent segment m. The coupon will impact the per-unit revenue for the retailer, r_{jkt}^{m}, and the "real" price that the customer has to pay for buying the brand, p_{jkt}^{m}. The set of control variables of the retailer is $\boldsymbol{a}_{Rt} = \{P_{jt}, C\_Ret_{jkt}^{m}, m = 1,..., M, j, k = 1,..., J\}$, where P_{jt} is the retail price.

Let $\boldsymbol{a}_t = \{\boldsymbol{a}_{Rt}, \boldsymbol{a}_{jt}, j=1,\ldots,J\}$ be the collection of control variables of the retailer and all manufacturers in the second and third stages of the pricing game. In the presence of switching costs, current decisions $\boldsymbol{a}_t$ will have a dynamic impact on the state variables in the next period,

$\boldsymbol{S}_{t+1}$. The following equation, called the *state equation*, captures the temporal evolution of the state variable, S_{jt}^{m}.

$$S_{j,t+1}^{m} = \sum_{k \neq j}^{J} S_{kt}^{m} * \Pr_t^m(k \to j \mid \boldsymbol{a}_t) + S_{jt}^{m} * \left(\Pr_t^m(j \to j \mid \boldsymbol{a}_t) + \Pr_t^m(j \to 0 \mid \boldsymbol{a}_t) \right), \tag{1}$$

where $\Pr_t^m(k \to j \mid \boldsymbol{a}_t)$ stands for the *switching probability*, for a customer in segment m, of switching from brand k to brand j, conditional on $\boldsymbol{a}_t$; $\Pr_t^m(j \to j \mid \boldsymbol{a}_t)$ stands for the customer's *repeat purchase probability* for brand j, conditional on $\boldsymbol{a}_t$; and $\Pr_t^m(j \to 0 \mid \boldsymbol{a}_t)$ stands for the probability of the customer choosing the outside option, conditional on $\boldsymbol{a}_t$, in which case the customer remains in the installed base of brand j. This equation suggests that brand j's installed base in a given period is the sum of those who bought brand j in the previous period and those who did not buy any brand in the previous period but whose most recently purchased brand was brand j. We use the inertial multinomial logit model (Seetharaman, Ainslie and Chintagunta 1999) to represent the brand choice probabilities, i.e., switching probability and repeat purchase probability, and include the outside option in the usual manner as a $(J+1)^{\text{th}}$ alternative. See Cosguner, Chan and Seetharaman (2016) for further details.

We assume that manufacturers and the retailer have a common discount factor $\rho < 1$, and for each manufacturer j, the discounted sum of profits can be written as the following Bellman equation.

$$V_j(\boldsymbol{S}) = \max_{\boldsymbol{a}_j} \left\{ \sum_{m=1}^{M} \pi_m \cdot \sum_{k=1}^{J} \left(w_{jk}^{m}\left(\boldsymbol{a}_j, \boldsymbol{a}_{-j}\right) - mc_j \right) \cdot S_k^m \cdot \Pr^m\left(k \to j \mid \boldsymbol{a}_j, \boldsymbol{a}_{-j}\right) + \rho \cdot EV_j(\boldsymbol{S}' \mid \boldsymbol{S}, \boldsymbol{a}_j, \boldsymbol{a}_{-j}) \right\}, \tag{2}$$

where π_m is the size of latent segment m, $V_j(\boldsymbol{S})$ the value function of the manufacturer under optimal policies, $\boldsymbol{a}_{-j}$ the policies of other players (i.e., other manufacturers and the retailer), S_k^m the installed base of customers in latent segment m for brand k, and $\boldsymbol{S}$ and $\boldsymbol{S'}$ the state variables in the current and next period, respectively. $w_{jk}^{m}\left(\boldsymbol{a}_{j},\boldsymbol{a}_{-j}\right)$ represents the actual revenue per unit sold for the manufacturer and is equal to the wholesale price minus the value of the coupon issued to customers in segment *m* whose previously purchased brand was *k*, and mc_j is the marginal production cost for the manufacturer.

Similarly, the retailer's discounted sum of profits can be written as

$$V_R(\boldsymbol{S}) = \max_{\boldsymbol{a}_R}\left\{\sum_{m=1}^{M}\pi_m \cdot \sum_{j=1}^{J}\sum_{k=1}^{J}\left(r_{jk}^{m}\left(\boldsymbol{a}_R,\boldsymbol{a}_{-R}\right)-W_j\right)\cdot S_k^m \cdot \Pr^m\left(k \to j \mid \boldsymbol{a}_R,\boldsymbol{a}_{-R}\right)+\rho\cdot EV_R(\boldsymbol{S}' \mid \boldsymbol{S},\boldsymbol{a}_R,\boldsymbol{a}_{-R})\right\}, \quad (3)$$

where $V_R(\boldsymbol{S})$ is the value function of the retailer under optimal policies, and $\boldsymbol{a}_{-R}$ the policies of other players (i.e., all manufacturers). $r_{jk}^{m}\left(\boldsymbol{a}_R,\boldsymbol{a}_{-R}\right)$ stands for the actual revenue per unit sold for the retailer and is equal to the retail price minus the value of the coupon issued to customers in segment *m* whose previously purchased brand was *k*, and W_j is the wholesale price charged by manufacturer *j*.

To investigate the implications of BPD on the pricing behavior and resulting profits of the manufacturers and the retailer, we study the different scenarios that correspond to different retailer decisions in the first stage regarding how to use the customer information. The different scenarios allow different members of the distribution channel to employ BPD. As each firm's pricing policy has strategic consequences for other channel members' pricing policies, we use the full equilibrium approach to examine the MPE pricing outcomes within each scenario.

Scenario 1: No BPD (Base Case)

This is the benchmark case. Manufacturers and the retailer do not offer coupons.

Scenario 2: Partial BPD - Couponing Based on Most Recent Purchases Only

We assume that the retailer or the manufacturers only use information on the most recently purchased brand of each customer. They use this information and the unconditional segment probabilities from the demand model to update, in a Bayesian manner, the posterior probability that the customer belongs to each segment. The BPD strategy is based on these posterior probabilities and to which brand's installed base the customer belongs. This scenario helps us understand the economic value to firms of the very limited information contained in the consumer's most recent purchase.

Case 1: Retailer Couponing: For each brand, the retailer issues different coupons to different customers (see Figure 1). In practice, the retailer can distribute these coupons to customers for future purchases, at the check-out counter inside the store. The retailer does not need to look at the long purchase history of the customer each time; it only has to know which brand the customer purchases in the current shopping trip.

Case 2: Manufacturer Couponing: The retailer shares customers' most recent purchase information with manufacturers and lets them issue coupons to different customer segments based on this information (see Figure 2). We assume that the retailer does not issue any coupons. For implementation, the retailer can help manufacturers distribute their coupons at check-out counters inside the store, like Case 1 above. We study 3 sub-cases, as explained below.

a. Both Coke and Pepsi Couponing: Retailer shares information with both manufacturers.

b. Exclusive Coke Couponing: Retailer shares information with Coke only.

c. Exclusive Pepsi Couponing: Retailer shares information with Pepsi only.

<u>Scenario 3: Full BPD - Couponing Based on the Full Purchase History</u>

In this scenario, the retailer or manufacturers can access the complete purchase history of each customer. Based on this, they can calculate each customer's posterior probabilities of segment membership. We assume that the firms classify each household into the segment for which the posterior probability is highest, and then average the posterior segment membership probabilities across all households within each segment in order to figure out the targeted coupon values for each segment.[1] We make this assumption for two reasons: (1) It significantly simplifies the computation for firms since otherwise they need to figure out a customized price for each household in the data (whose computation in our dynamic pricing set-up is not straightforward); (2) Our estimated distribution of posterior probabilities of segment membership for customers within each segment is close to bipolar (i.e., mostly takes values close to 0 or 1) and does not show much dispersion across households. The BPD strategy is, therefore, based on which brand's installed base the customer belongs to, as well as the probabilities of the customer's segment membership as inferred from the purchase history. Compared with Scenario 2, this scenario helps us understand the additional economic value to firms of knowing a customer's complete purchase history, above and beyond knowing the customer's most recent purchase.

Case 1: Retailer Couponing: The retailer issues coupons to customers and manufacturers do not have access to the information (see Figure 3).

Case 2: Manufacturer Couponing

The retailer does not issue any coupons. It shares the information with the manufacturers, and lets them issue coupons to different customer segments (see Figure 4). We study 3 sub-cases, as explained below.

a. Both Coke and Pepsi Couponing: Retailer shares information with both manufacturers.

[1] We acknowledge that BPD, as defined in the literature, would involve the retailer or manufacturer targeting each customer based on the customer's unique posterior probabilities, which is not what we do in this scenario.

b. Exclusive Coke Couponing: Retailer shares information with Coke only.

c. Exclusive Pepsi Couponing: Retailer shares information with Pepsi only.

Empirical Results

We use scanner panel data from Information Resources Incorporated's (IRI) scanner-panel database on cola purchases of 356 households making 32,942 shopping trips at a supermarket store (which is a local monopolist) in a suburban market of a large U.S. city from June 1991 to June 1993. Households choose among Pepsi, Coke, Royal Crown and a private label. Pepsi is the dominant cola brand (with an average 46% market share), while the Private Label is the smallest brand (with an average 7% market share). In our data, neither manufacturers nor the retailer uses BPD, which is the base case in Scenario 1.

We first estimate our demand model using the panel data on households' brand choices. In this application, we ignore the issue that consumers may purchase multiple units of product for inventory holding.[2] We also abstract away from the complications that consumers may be forward-looking and they may purchase multiple brands.[3] As previously noted, the model setup is the same

[2] To check whether ignoring quantity has a significant impact on the estimation results, we first calculate the correlation between the market share based on purchase incidences and the market share based on quantity outcomes over the entire span of weeks in the data, and find that the correlation is very high at 0.92. We then conduct an independent sample t-test in order to check whether the mean of the two market shares (based on quantity and incidence outcomes) differs, and find no statistically significant difference. We also conduct a paired sample t-test to see whether the two market shares differ in each week, and again find no statistically significant difference. These results suggest that ignoring multiple unit purchases may not lead to significant bias in our estimation results. Finally, to test whether inventory holding impacts our model estimates, we estimate a demand model with the variable of inventory that is calculated from the weekly purchases and the average consumption level over the whole sample period from each household. We find that the estimated price and inertia coefficients are very similar with and without the inventory. To test the pricing implications of ignoring inventory, we also simulate the steady-state equilibrium prices using the estimates from the demand model with inventory. We find no significant differences in the equilibrium retail and wholesale prices.

[3] As shown in Villas-Boas (2004), if customers are forward-looking, they might change their purchase behaviors in order to avoid revealing their true preferences. Our study ignores this issue. Furthermore, only on 93.1% of purchase observations are more than one cola brand purchased at the same time by the household. If one looks only at Coke and Pepsi purchases, this percentage goes up to 95.3%. Therefore, multiple brand buying is not a significant aspect of our data.

as in Cosguner, Chan and Seetharaman (2016).[4] Estimation results show that there are three segments among households: Segment 1 (35 % of households) has the highest price sensitivity (-7.53), followed by segment 3 (48% of households; price sensitivity: -7.30), and segment 2 (17% of households; price sensitivity: -4.31).[5] Segment 2 has the highest switching cost, as its estimated inertia is 2.39, followed by segment 1 (1.03) and segment 3 (0.73).[6]

For the supply side, we assume that Coke and Pepsi directly compete against each other and that their pricing strategies are independent of the pricing strategies of RC Cola and the Private Label.[7] The estimated marginal costs for Coke and Pepsi are \$0.433 and \$0.366, respectively.[8]

Given the estimated parameters, we run a series of counterfactual simulations to study the different price discrimination scenarios described in the model section. For simplicity, we assume that there are no coupons from the private label or Royal Crown, and focus on BPD by Coke and Pepsi only. We use the Pakes and McGuire (1994) algorithm to solve for the dynamic pricing equilibrium in the distribution channel. In the simulation, we assume that no customer is in any brand's installed base at the beginning, and forward simulate retail prices, wholesale prices and customers' brand choices for multiple future periods, until the prices and demands reach a steady

[4] Our model allows for 3 consumer segments, which is different from their model that only considers 2 segments.

[5] We also estimate a 4-segment model. The size of the additional segment is only 3.9%. As the profit gains from price discriminating this segment would be limited to manufacturers and retailers, we focus the counterfactuals based on the three-segment model.

[6] Our model assumes that consumer preferences are time-invariant. To test the validity of this assumption, we estimate our demand model using data from the second year only. Results show that most of the estimates are similar to those obtained using the full two years of data. Specifically, the price and inertia coefficients are similar under both cases. Still, there are some differences in the estimated brand preferences. We acknowledge that changing consumer preferences can be a possible explanation for the estimated inertia in brand choices.

[7] We test this duopoly assumption by regressing the price of one brand on the prices of other brands. For Coke (Pepsi), the main competitor Pepsi's (Coke's) price explains 40.54% (40.54 %) of its price variation. Adding the price of private label in the regression increases R-square by 5.56% for Pepsi and 1.67% for Coke, and adding the price of Royal Crown further increases R-Square by 6.07% for Pepsi and 1.83% for Coke. The results suggest that Coke only responds to Pepsi, and Pepsi mainly responds to Coke and to a lesser extent also to the private label and Royal Crown. The duopoly assumption captures the fact that competition appears to be mainly between Coke and Pepsi in this market.

[8] To save space we do not report all estimation results. The full results are available from the authors upon request.

state. Results reported below are based on the steady state equilibrium. In Scenario 2, the partial BPD scenario, there are two types of customers depending on whether they purchased Coke or Pepsi most recently. In Scenario 3, the full BPD scenario, for each brand there are six types of customers, depending on their segment membership (segment 1, 2 or 3), and previously purchased brand (own or competitor brand). The price elasticity of segment 2 for Pepsi decreases from -3.21 (when Coke was bought previously) to -2.25 (when Pepsi was bought previously), while the corresponding decreases for segments 1 and 3 are modest (-5.92 to - 5.71, and -5.84 to -5.80, respectively). The changes in price elasticities of the three segments for Coke are similar, highlighting that the optimal BPD should adjust over time even for the same set of customers.

We first simulate the steady state equilibrium outcomes, under the assumption that the same prices are offered to all customers in the market, i.e., Scenario 1. The second column in Table 1 reports the retail prices (P_{Coke} and P_{Pepsi}) and wholesale prices (W_{Coke} and W_{Pepsi}), as well as the profits of all channel members, averaged across all customers. Since Pepsi has lower average marginal cost and higher brand preference than Coke, it enjoys a higher profit than Coke. In addition, the retailer as a local monopolist enjoys a higher profit than the combined profits of Pepsi and Coke.

Next, we simulate the equilibrium outcomes in Scenario 2, under which firms offer customized coupons to customers based on their most recent purchases. We first study the case when the retailer offers coupons. Compared to the base case, customers will face a lower retail price if they switch brands. The reason that they receive the coupons is because the price elasticity for a brand among customers who did not purchase the brand previously is higher than those who purchased the brand previously. With such BPD, the retailer's profit per customer increases by 2.5 % (from \$0.933 to \$0.956). We then simulate the equilibrium outcomes under the assumption that the retailer shares the information (at no charge) with the manufacturers, and help manufacturers distribute their coupons (Case 2a). Comparing columns 2 and 4, both manufacturers' profits remain roughly the same, while the retailer increases its profit by 3.7 % from the manufacturers' couponing efforts. The profit increase is higher than the case when the retailer issues coupons by itself, suggesting that *the retailer is better off by simply outsourcing the customization of coupons to the manufacturers rather than proactively engaging in customization on its own.* Consequently, the retailer's optimal decision in the first stage (see our discussion in the model section) is to share the customer information with the manufacturers and help them employ BPD.

Since the information on customers' most recent purchases belongs to the retailer, we further explore whether the retailer can gain additional profit by charging manufacturers for access to the information, that is, if the retailer can serve as an information broker and sell the information to manufacturers (as in Pancras and Sudhir 2007). To answer this question, we first calculate the highest price that the retailer can charge each manufacturer, which is the manufacturer's willingness to pay. We employ the Nash equilibrium solution concept as follows: The highest price to charge Coke is the additional profit that Coke obtains, when both manufacturers engage in BPD (Case 2a, column 4) relative to when only Pepsi engages in BPD (Case 2c, column 6). Similarly, the highest price to charge Pepsi is the additional profit that Pepsi obtains in Case 2a relative to when only Coke engages in BPD (Case 2b, column 5). The simulation results suggest that the retailer can charge \$0.028 (= \$0.229 - \$0.201) to Coke and \$0.062 (= \$0.472 - \$0.410) to Pepsi, per customer, for access to the information. As a result, the retailer will be much better off outsourcing targeted couponing to manufacturers, with a net profit of \$1.057 (= \$0.967 + \$0.028 + \$0.062), which represents a 10.7 % increase over the profit obtained by managing BPD on its own, and 13.4 % higher than the profit obtained in the No BPD case. *In other words, serving as* *an information broker to sell transactional data to manufacturers can be a vital source of business profit to the retailer*. This finding is consistent with the findings in Pancras and Sudhir (2007), although they use a myopic pricing model in their study. Interestingly, this yields a net profit of \$0.201 and \$0.410 to Coke and Pepsi, respectively, both lower than the profits under Scenario 1. In other words, while the retailer benefits from inducing manufacturers to engage in BPD, the manufacturers are worse off than under the case of no price discrimination.[9] This is in contrast to the situation in Pancras and Sudhir (2007), where the authors find that manufacturers' profits also improve. This difference is probably because, when customers have switching costs in brand choices as in our model, the information becomes more valuable to manufacturers for engaging in BPD; consequently, the retailer can charge a higher price for the information.

The above results indicate that information on the customers' most recent purchases can have a significant profit impact for the retailer and the manufacturers. We now explore a further question: Is there any more profit gain from using the full customer purchase history? To answer that question, we assume that firms can use the full purchase history to infer the true segment membership of each customer in Scenario 3. Therefore, the retailer or manufacturers can offer targeted coupons based on six types of customers (segment 1, 2, or 3 customers who purchased own or the competitor brand previously). Similar to Scenario 2, we first study the case when the retailer offers customized coupons to customers (Case 1). The results are shown in column 7 of Table 1. Compared with column 2, allowing the retailer to target based on the full purchase history, in addition to the most recent purchases, bring a 12.3 % profit increase for the retailer. We then

[9] These results are based on the assumption that the retailer commits not to issue customized coupons to customers. To investigate whether this is a reasonable assumption, we conduct another simulation study, assuming that the retailer will also issue targeted coupons, after selling the information to the manufacturers. Results show that at the steady state equilibrium the retailer's profit will be lower (at \$1.009). This is because, once manufacturers know that the retailer will also offer customized coupons for their brands, they are not willing to pay such a high price for the customer information. Therefore, the retailer has an incentive to commit, when selling the information to the manufacturers, not to engage in BPD itself.

simulate the equilibrium outcomes under the assumption that the retailer shares customers' full purchase histories (at no charge) with the manufacturers, and lets them engage in couponing (Case 2). The profit increase for the retailer is 14.6 % in this case. The results again suggest that the retailer is better off by outsourcing BPD to the two manufacturers rather than proactively engaging in customization on its own. From the manufacturer perspective, BPD also benefits them since the profit increases by 2.6 % for Coke and 23.5 % for Pepsi. The gain for Coke is smaller since as a smaller brand in our data it has to cut its prices relatively more in order to be competitive.

Finally, we also explore whether the above findings change if the retailer can charge manufacturers for access to the information. Results show that the retailer can charge \$0.064 (= \$0.235 - \$0.171) to Coke and \$0.206 (= \$0.588 - \$0.382) to Pepsi per customer. Therefore, similar to Scenario 2, the retailer is much better off from outsourcing the customized couponing strategy to manufacturers, with a net profit of \$1.339 (= \$1.069 + \$0.064 + \$0.206), which represents a 27.8 % increase over the profit obtained by managing BPD on its own. From the manufacturers' perspective, compared to targeting based on most recent purchases, they become significantly worse off, as Coke (Pepsi) loses an additional profit of 18.3% (7.6%). Overall, the results indicate that having more information (a longer purchase history) brings significantly more benefit to the retailer but not to the manufacturers.

The comparison of the profit gains between Scenarios 2 and 3 suggests that, although the full purchase history of customers is more important to the retailer, the economic value of the information on which brands customers purchased previously is also significant. Furthermore, to implement BPD based on the most recent purchases is of low cost to the retailer and manufacturers, as coupons for future purchases can be distributed at the check-out counter simply based on which brand a customer has purchased during that shopping trip. To implement BPD that is based on the full purchase history, however, requires firms to access the customer database each time a customer makes a purchase. Implementing this strategy can be more costly to the retailer and manufacturers.

Conclusions

We study behavioral price discrimination (BPD), using targeted price-off coupons, within a distribution channel in a market with switching costs. Our study delineates the separate strategic roles of the manufacturers and the retailer. From the retailer's standpoint, we find that the retailer should simply outsource the data analytics and customization of coupons to manufacturers and improve its profit beyond what it can achieve by proactively engaging in customization on its own. We further find that serving as an information broker to sell its customer database to manufacturers can be a vital source of profit to the retailer. In contrast, cola manufacturers end up worse off, illustrating that customer information is a potent source of channel power to the retailer in markets with switching costs. Finally, we show that simply using the information contained in customers' most recent purchases can significantly impact firms' profits. BPD based on this information is easy to implement and of low cost to manufacturers and retailers.

Some caveats remain. We do not consider the case, usually considered in the literature on BPD and reflects many situations outside supermarket retailing (such as direct marketing), where each manufacturer only knows whether or not a customer has purchased their brand and not which of competing brands (if any) is purchased when the focal brand is not purchased. We also assume that coupon redemption rates are equal across customer segments. In reality, it is possible that more price sensitive customers are more likely to redeem coupons. If such were the case, our key results (such as, for example, that higher face value coupons must be dropped to more price sensitive customers) would be strengthened. To simplify the analysis, we assume that consumers are myopic in decision making. Future research may explore the implications for BPD of markets with switching costs where strategic, forward-looking consumers exist. Additionally, we ignore the effects of product inventory in the consumer's purchase decision. This could potentially raise an additional source of forward looking behavior for consumers. Another interesting area for future research is to study how consumer preferences may evolve over time and how ignoring such preference evolution may change our findings. Finally, it is also important to study the BPD strategies of retailers and manufacturers when consumers are variety seeking.

References

Acquisiti, A., Varian, H. (2006). Conditioning Prices on Purchase History, Marketing Science, 24, 3, 367-381.

Besanko, D., Dube, J. P., Gupta, S. (2003). Competitive Price Discrimination Strategies in a Vertical Channel Using Aggregate Retail Data, Management Science, 49, 9, 1121-1138.

Chen, Y. (1997). Paying Customers to Switch, Journal of Economics and Management Strategy, 6, 4, 877-897.

Chen, Y., Narasimhan, C., Zhang, Z. J. (2001). Individual Marketing with Imperfect Targetability, Marketing Science, 20, 1, 23-41.

Cosguner, K., Chan, T., Seetharaman, P. B. (2016). Dynamic Pricing in a Distribution Channel in the Presence of Switching Costs, forthcoming at Management Science.

Dube, J. P., Hitsch, G. J., Rossi, P. E., Vitorino, M. A. (2008). Category Pricing with State-Dependent Utility, Marketing Science, 27, 3, 417-429.

Elrod, T., Winer, R. S. (1982). An Empirical Evaluation of Aggregation Approaches for Developing Market Segments, Journal of Marketing, 46, 4, 65-74.

Fader, P., Hardie, B. G. S., Lee, K. L. (2005). RFM and CLV, Journal of Marketing Research, 42, 4, 415-430.

Fudenberg, D., Tirole, J. (2000). Customer Poaching and Brand Switching, The RAND Journal of Economics, 31, 4, 634-657.

Fudenberg, D., & Villas-Boas, J. M. (2006). Behavior-based price discrimination and customer recognition, Handbook on economics and information systems, 1, 377-436.

Glazer, R. (1991). Marketing in an Information-Intensive Environment, Journal of Marketing, 6, 4, 877-897.

Kamakura, W. A., Russell, G. J. (1989). A Probabilistic Choice Model for Market Segmentation and Elasticity Structure, Journal of Marketing Research, 26, 4, 379-390.

Liu, Y., Zhang, Z. J. (2006). The Benefits of Personalized Pricing in a Channel, Marketing Science, 25, 1, 97-105.

Pakes, A., McGuire, P. (1994). Computing Markov-Perfect Nash Equilibria, RAND Journal of Economics, 25, 4, 555-589.

Pancras, J., Sudhir, K. (2007). Optimal Marketing Strategies for a Customer Data Intermediary, Journal of Marketing Research, 44, 4, 560-578.

Pazgal, A., Soberman, D. (2008). Behavior-Based Discrimination, Marketing Science, 27, 6, 977-994.

Rossi, P. E., McCulloch, R. E., Allenby, G. M. (1996). The Value of Purchase History Data in Target Marketing, Marketing Science, 15, 4, 321-340.

Seetharaman, P. B. (2004). Modeling Multiple Sources of State Dependence in Random Utility Models, Marketing Science, 23, 2, 263-271.

Seetharaman, P. B., Ainslie, A., Chintagunta, P. K. (1999). Investigating Household State Dependence Effects Across Categories, Journal of Marketing Research, 36, 4, 488-500.

Shaffer, G., Zhang, Z. J. (1995). Competitive Coupon Targeting, Marketing Science, 14, 4, 395-416.

Shaffer, G., Zhang, Z. J. (2000). Pay To Switch or Pay To Stay: Preference-Based Price Discrimination in Markets With Switching Costs, Journal of Economics and Management Strategy, 9, 3, 397-424.

Shin, J., Sudhir, K. (2014). A Customer Management Dilemma, Marketing Science, 29, 4, 671-689.

Sudhir, K. (2001). Structural Analysis of Manufacturer Pricing in the Presence of a Strategic Retailer, Marketing Science, 20, 3, 244-264.

Taylor, C. R. (2003). Supplier Surfing: Competition and Consumer Behavior in Subscription Markets, RAND Journal of Economics, 34, 2, 223-246.

Villas-Boas, J. M. (1999). Dynamic Competition with Customer Recognition, RAND Journal of Economics, 30, 4, 604-631.

Villas-Boas, J. M. (2004). Price cycles in markets with customer recognition, RAND Journal of Economics, 486-501.

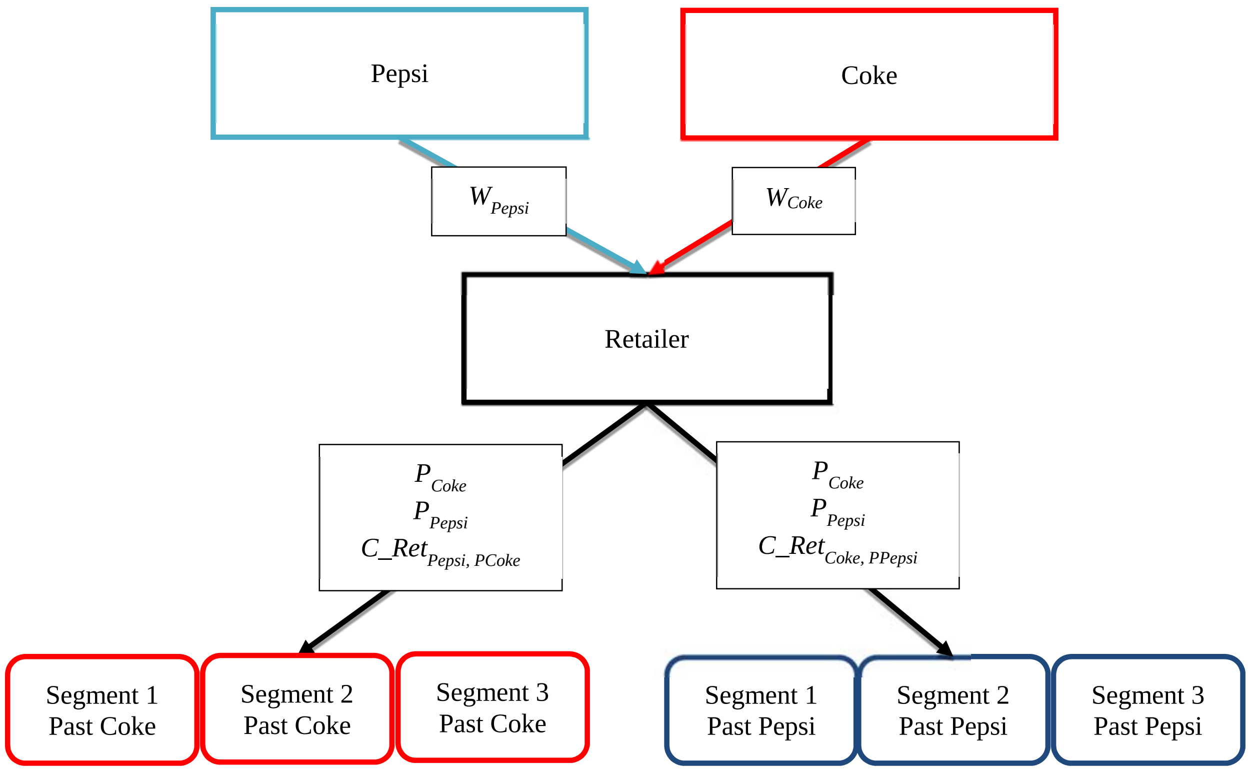

Figure 1: (Scenario 2; Case 1)
Retailer Couponing Based on the Most Recent Purchases

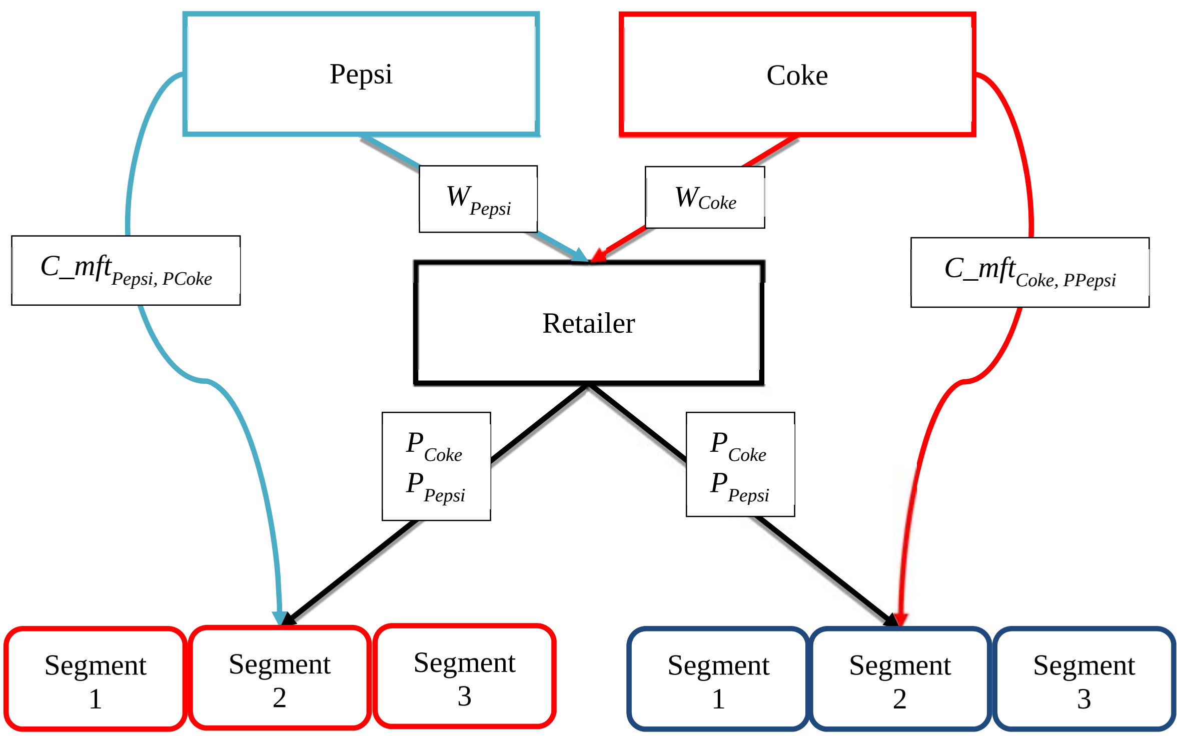

Figure 2: (Scenario 2; Cases 2a, b, c)
Manufacturer Couponing Based on Most Recent Purchases

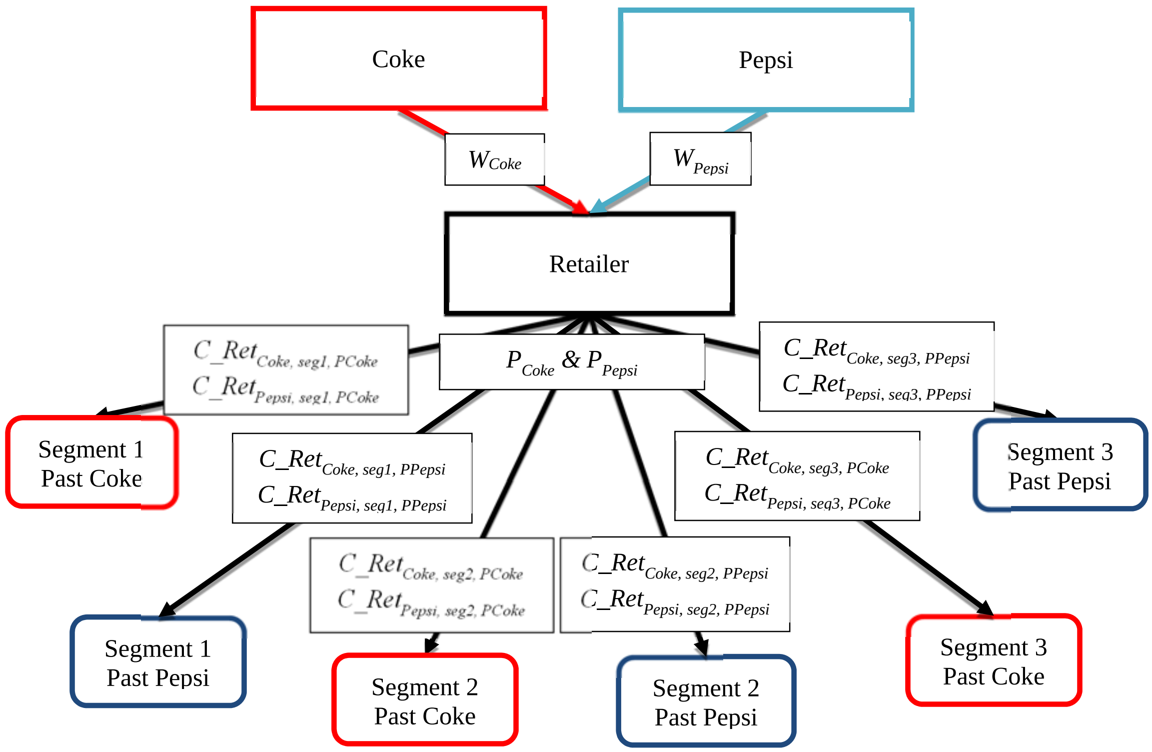

Figure 3: (Scenario 3; Case 1)
Retailer Couponing Based on Both Segments and Past Purchases

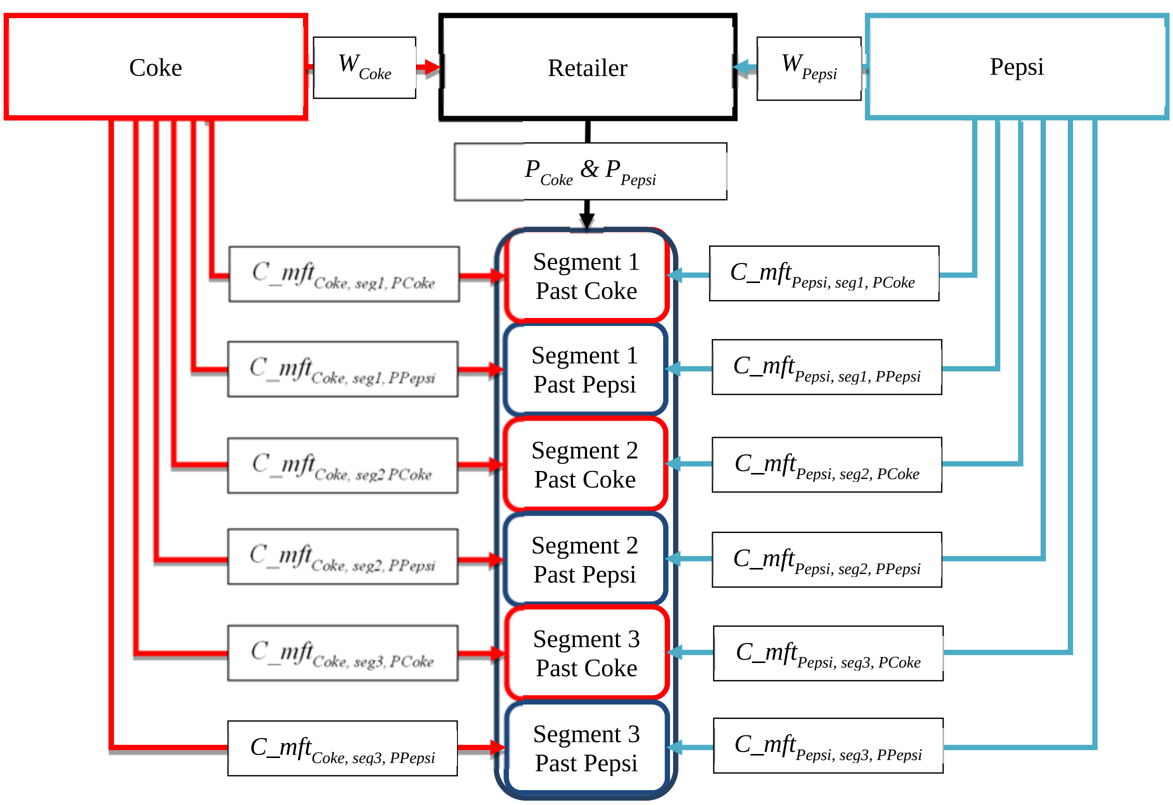

Figure 4: (Scenario 3; Case 2a, b, c)
Manufacturer Couponing Based on Both Segments and Past Purchases

TABLE 1: BPD COUNTERFACTUAL SIMULATION RESULTS

	Scenario 1: No targeting	Scenario 2: Targeting based on most recent purchase only				Scenario 3: Targeting based on segment membership and most recent purchase			
	(No Coupon)	Case 1 (Retailer Couponing)	Case 2 (Coke & Pepsi Couponing)	Case 2a (Coke Couponing)	Case 2b (Pepsi Couponing)	Case 1 (Retailer Couponing)	Case 2 (Coke & Pepsi Couponing)	Case 2a (Coke Couponing)	Case 2b (Pepsi Couponing)
P_{Coke}	0.847	0.851	0.889	0.879	0.837	0.965	0.855	0.911	0.837
P_{Pepsi}	0.798	0.800	0.811	0.781	0.820	0.877	0.861	0.784	0.884
W_{Coke}	0.602	0.597	0.657	0.644	0.594	0.584	0.648	0.709	0.584
W_{Pepsi}	0.558	0.552	0.575	0.538	0.579	0.558	0.649	0.534	0.687
$C_{Coke,seg1,PPepsi}$	Na	0.040	0.144	0.139	Na	0.218	0.141	0.251	0.188
$C_{Coke,seg2,PPepsi}$	Na	0.040	0.144	0.139	Na	Na	0.085	0.174	Na
$C_{Coke,seg3,PPepsi}$	Na	0.040	0.144	0.139	Na	0.220	0.114	0.188	0.185
$C_{Coke,seg1,PCoke}$	Na	Na	Na	Na	Na	0.234	0.094	0.162	0.324
$C_{Coke,seg2,PCoke}$	Na	Na	Na	Na	Na	0.066	Na	Na	0.247
$C_{Coke,seg3,PCoke}$	Na	Na	Na	Na	Na	0.239	0.085	0.148	0.245
$C_{Pepsi,seg1,PPepsi}$	Na	Na	Na	Na	Na	0.172	0.156	Na	Na
$C_{Pepsi,seg2,PPepsi}$	Na	Na	Na	Na	Na	Na	Na	Na	Na
$C_{Pepsi,seg3,PPepsi}$	Na	Na	Na	Na	Na	0.177	0.149	Na	Na
$C_{Pepsi,seg1,PCoke}$	Na	0.028	0.122	Na	0.116	0.188	0.242	Na	Na
$C_{Pepsi,seg2,PCoke}$	Na	0.028	0.122	Na	0.116	0.066	0.261	Na	Na
$C_{Pepsi,seg3,PCoke}$	Na	0.028	0.122	Na	0.116	0.196	0.198	Na	Na
Retailer Profit	0.933	0.956	0.967	0.962	0.943	1.048	1.069	0.988	1.001
Coke Profit	0.229	0.236	0.229	0.264	0.201	0.218	0.235	0.335	0.171
Pepsi Profit	0.476	0.459	0.472	0.410	0.508	0.587	0.588	0.382	0.695