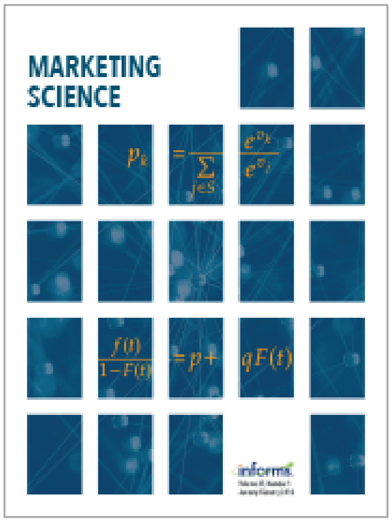

Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

An Integrated Choice Model Incorporating Alternative Mechanisms for Consumers' Reactions to In-Store Display and Feature Advertising

Jie Zhang,

MARKETING SCIENCE
Vol. 25, No. 3, May–June 2006, pp. 278–290
ISSN 0732-2399 | EISSN 1526-548X | 06 | 2503 | 0278

DOI 10.1287/mksc.1050.0170

Research Note

An Integrated Choice Model Incorporating Alternative Mechanisms for Consumers' Reactions to In-Store Display and Feature Advertising

Jie Zhang
Robert H. Smith School of Business, University of Maryland,
3446 Van Munching Hall, College Park, Maryland 20742, jiejie@rhsmith.umd.edu

The marketing literature has suggested two prominent decision mechanisms through which in-store display and feature advertising can affect brand choice, which I call *the price-cut proxy effect* and *the consideration set formation effect*. The primary objective of this study is to propose an integrated brand choice model that incorporates these two possible behavioral mechanisms, which have been shown to work by previous research. The model allows consumers to use various combinations of the decision mechanisms with different probabilities and thus enables one to *jointly* assess the extent to which each effect might occur in actual purchase data and to investigate how consumers might differ in their tendencies to engage in these decision processes. By incorporating these likely behavioral mechanisms, the proposed model alleviates the problems caused by multicollinearity and produces sensible parameter estimates for the joint effects of promotion vehicles (in-store display, feature ad, and price discount), which contributes to better managerial decision making.

Key words: brand choice model; behavioral mechanisms; joint promotion effects; consideration set formation; price-cut proxy; econometric models; promotion decisions
History: This paper was received July 19, 2005, and was with the author 5 months for 3 revisions; processed by Leigh McAlister.

1. Introduction

Retailers often offer temporary price reductions in combination with in-store displays and feature advertisements in their sales promotions programs. This strategy is driven by the belief that price discounts in combination with displays or feature ads will generate greater sales results than the sum of the effects when each promotion vehicle is offered separately. Due to this practice, marketing researchers might observe strong multicollinearity between these promotion variables in sales data. Unfortunately, standard brand choice models (such as multinomial logit or probit models) are likely to generate counterintuitive parameter estimates on the joint effects of promotion vehicles when there is strong multicollinearity in the data. Therefore, the very practice of combining price discounts with in-store displays and feature ads could hinder the ability to use a conventional brand choice model to evaluate the appropriateness of this strategy.

In a standard brand choice model, whether it makes sense to offer a price discount in combination with a display or feature ad depends on the signs of their interaction effects. A positive interaction effect would endorse the combination, while a negative interaction effect would suggest the opposite (see Lemon and Nowlis 2002 for such an application).[1] Despite the importance of a thorough understanding of these interaction effects, there has been little consensus in the literature regarding the direction of their effects. Most previous brand choice models that included display and feature ad effects did not explicitly measure their interactions with price cut. Those that did reported mixed patterns of the interaction effects. For example, Gupta (1988) found negative interactions of display and feature ad with price cut; Papatla and Krishnamurthi (1996) found a positive interaction of display and price cut but a negative interaction of feature ad and price cut, and Lemon and Nowlis (2002) reported a negative interaction of feature ad and price cut but small positive and negative interactions of display and price cut (which varied across brands). It is

[1] It should be pointed out that the multiplicative form of standard brand choice models inherently captures a certain degree of interaction among predictor variables without including explicit interaction terms, although such inherent interaction effects are restricted in the magnitude and sign.

not clear why displays and feature ads exhibit different patterns of interactions with price cut. Papatla and Krishnamurthi (1996) expected these interactions to be positive and considered the negative interaction of feature ad and price cut counterintuitive. Gupta postulated that the negative interactions in his study "suggest a possible overlap or substitutability among different promotional instruments" (1988, p. 348) but stopped short of exploring what may be causing the overlap.

To help shed light on these issues, we propose an integrated brand choice model that incorporates alternative decision processes for consumers' reactions to in-store displays and feature ads. Representations of these decision processes are built upon two behavioral mechanisms through which these promotional tools have been shown to affect brand choice by previous research. By incorporating these likely decision mechanisms, the proposed model alleviates the problems caused by multicollinearity and produces sensible parameter estimates for the joint effects of promotional vehicles, which makes it better suited for guiding managerial decisions. More important, this modeling approach helps reveal distinctive segments of consumers who are likely to engage in different decision processes, which is also beneficial for managerial decision making. In addition, our results provide an explanation for the mixed patterns of promotion interaction effects that have been reported in previous studies.

Many studies have documented that displays and feature ads can significantly increase brand choice even when the price discount effect is controlled for (e.g., Gupta 1988, Grover and Srinivasan 1992, Chintagunta 1992, Papatla 1996). Yet it is not obvious why merely being on display or a feature ad, net of price discount effect, would increase brand choice. The marketing literature has suggested two prominent behavioral explanations on the possible mechanisms through which their effects take place. One explanation for the observed effects of promotion signs was offered by Inman et al. (1990). They proposed that consumers on the peripheral route to persuasion do not engage in detailed information processing and simply interpret a promotion marker as a proxy for a price cut. Therefore, the mere presence of a promotion signal would lead these consumers to believe that the brand has been offered with a price discount. We refer to this mechanism as the *price-cut proxy effect.* Inman et al. (1990) found that this effect occurred only for consumers who exhibited low need for cognition, and a promotion sign without a price cut did not increase the choice probability for individuals with a high need for cognition.

At the same time, the growing literature on consideration sets provides an alternative explanation for why displays and feature ads affect brand choices. Behavioral research has observed that, for low-involvement product categories, consumers often rely on certain heuristics to form a consideration set first and then engage in more elaborate evaluation of the remaining alternatives in the consideration set (e.g., Payne 1976, Lussier and Olshavsky 1979, Hauser and Wernerfelt 1990, Roberts and Lattin 1991). Various studies in this stream of research have shown that in-store displays and feature ads can be utilized to form consideration sets (e.g., Fader and McAlister 1990, Allenby and Ginter 1995, Andrews and Srinivasan 1995, Bronnenberg and Vanhonacker 1996, Mehta et al. 2003). In other words, they increase a brand's probability of being chosen by making it more prominent to consumers. We refer to this mechanism as the *consideration set formation effect.*

The primary objective of this study is to construct a brand choice model that incorporates, in a single framework, these two behavioral mechanisms, which prior research has shown to work, and to use this added information to segment the market in terms of consumers' propensities to respond in different ways to the promotions. The proposed model allows consumers to use various combinations of the decision mechanisms with different probabilities, and thus enables us to *jointly* assess the extent to which each effect may occur in actual purchase data and investigate how consumers might differ in their tendencies to engage in these decision processes. Johnson et al. have called for models that are "better representations of decision processes" (1989, p. 268). This study attempts to make a step in that direction. By incorporating the two likely underlying behavioral mechanisms, we will show that the proposed model contributes to better managerial decisions on whether to combine or separate various promotion vehicles.

The proposed model also holds the promise of providing a possible explanation for the mixed patterns of promotion interaction effects reported in previous studies. If displays/feature ads mainly serve as price-cut proxies, combining them with price cuts would add little value in a consumer's mind. In other words, the joint effect of display/feature ad and price cut would be smaller than the sum of their stand-alone effects, and thus the interaction effect would be negative. On the other hand, if a display/feature ad affects brand choice mainly through the consideration set formation effect, which helps a brand to pass the first stage in the decision process, then combining it with a price cut would reinforce the effectiveness of both, because the price cut increases the attractiveness of a brand among those in the consideration set and enhances its chance of being chosen at the second stage of the decision process. This mechanism would suggest a positive interaction effect. We test these conjectures in the empirical analysis.

2. Model Formulation

We construct a brand choice model for the analysis of household scanner panel data. To assess the likelihoods of alternative decision mechanisms simultaneously, we allow consumers to use displays and/or feature ads as price-cut proxies and/or for consideration set formation in a probabilistic manner in the same model. In other words, with certain probabilities, consumers might use either or both promotion vehicles as price-cut proxies or consideration set formation devices, or they might ignore them and evaluate the entire set of alternatives based on quality and actual pricing information. The model also takes into account the possible correlation of using a display and feature ads for the price-cut proxy or consideration set formation effect.

$DP_i = 1$ if display is used as a price-cut proxy by household i; $DP_i = 0$ otherwise. $FP_i = 1$ if a feature ad is used as a price-cut proxy by household i, and $FP_i = 0$ otherwise. $DC_i = 1$ if a display is used for consideration set formation by household i, and $DC_i = 0$ otherwise. $FC_i = 1$ if a feature ad is used for consideration set formation by household i; $FC_i = 0$ otherwise. $\phi_{DP_i} =$ household i's probability of using a display as a price-cut proxy; $\phi_{FP_i} =$ household i's probability of using a feature ad as a price-cut proxy; $\phi_{DC_i} =$ household i's probability of using a display for consideration set formation; $\phi_{FC_i} =$ household i's probability of using a feature ad for consideration set formation.

A consumer who tends to use displays as price-cut proxies might also be likely to do this with feature ads. Similarly, the tendencies to use displays and feature ads for consideration set formation might be positively correlated as well. To account for the possible interdependence of the various decision mechanisms, we include the following covariance terms in the model:

$$\sigma_{DC,FC} = \mathrm{cov}(DC_i, FC_i), \quad \sigma_{DC,FP} = \mathrm{cov}(DC_i, FP_i),$$
$$\sigma_{DP,FC} = \mathrm{cov}(DP_i, FC_i), \quad \text{and} \quad \sigma_{DP,FP} = \mathrm{cov}(DP_i, FP_i).$$

A consumer can use display/feature ad as a price-cut proxy, a consideration set formation device, or neither at a given purchase occasion. The probability of the occurrence of each decision process is described in Chart 1 (the derivation is in the appendix).

For the conditional brand choice probabilities in the Chart 1, we use superscripts *DC* to denote display consideration set formation, *FC* to denote feature ad consideration set formation, *DP* to denote display price-cut proxy, and *FP* to denote feature ad price-cut proxy. The subscripts *ikt* refer to household i, alternative k, purchase occasion t. P^0_{ikt} represents the conditional brand choice probability when neither display nor feature ad is used in the decision, in which case the consumer chooses from the entire set of available alternatives and evaluates the actual price and price discount of each alternative.

The decision process assumed for the *price-cut proxy effect* is that a consumer infers a price cut from a display and/or feature ad regardless of whether there is an actual reduction in the price. It implies that, in a brand utility function with price cut, display, and feature ad dummy variables, the coefficient of display or feature ad should be the same as the coefficient of the price-cut variable. In addition, the model should avoid double counting by the analyst when price cut, display, and/or feature ad are offered in combination for an item. The decision process assumed in the *consideration set formation effect* is that a consumer relies on displays and/or feature ads to select alternatives and form a consideration set first, and then engages in more thorough evaluation of the alternatives in the consideration set. Like many previous studies

Chart 1 Decision Processes and Their Probabilities of Occurrence

Decision process			
Display	Feature ad	Probability of occurrence	Conditional brand choice probability
1. Consideration set formation	Consideration set formation	$\phi_{DC}\phi_{FC} + \sigma_{DC,FC}$	$P^{DC,FC}_{ikt}$
2. Consideration set formation	Price-cut proxy	$\phi_{DC}\phi_{FP} + \sigma_{DC,FP}$	$P^{DC,FP}_{ikt}$
3. Consideration set formation	No effect	$\phi_{DC}(1-\phi_{FC}-\phi_{FP}) - \sigma_{DC,FC} - \sigma_{DC,FP}$	P^{DC}_{ikt}
4. Price-cut proxy	Consideration set formation	$\phi_{DP}\phi_{FC} + \sigma_{DP,FC}$	$P^{DP,FC}_{ikt}$
5. Price-cut proxy	Price-cut proxy	$\phi_{DP}\phi_{FP} + \sigma_{DP,FP}$	$P^{DP,FP}_{ikt}$
6. Price-cut proxy	No effect	$\phi_{DP}(1-\phi_{FC}-\phi_{FP}) - \sigma_{DP,FC} - \sigma_{DP,FP}$	P^{DP}_{ikt}
7. No effect	Consideration set formation	$(1-\phi_{DC}-\phi_{DP})\phi_{FC} - \sigma_{DC,FC} - \sigma_{DP,FC}$	P^{FC}_{ikt}
8. No effect	Price-cut proxy	$(1-\phi_{DC}-\phi_{DP})\phi_{FP} - \sigma_{DC,FP} - \sigma_{DP,FP}$	P^{FP}_{ikt}
9. No effect	No effect	$(1-\phi_{DC}-\phi_{DP})(1-\phi_{FC}-\phi_{FP}) + \sigma_{DC,FC} + \sigma_{DP,FC} + \sigma_{DC,FP} + \sigma_{DP,FP}$	P^{0}_{ikt}

on consideration sets (e.g., Roberts and Lattin 1991, Andrews and Srinivasan 1995, Siddarth et al. 1995, Bronnenberg and Vanhonacker 1996, Gilbride and Allenby 2004), we assume that a consumer utilizes a compensatory strategy to choose the alternative that maximizes the perceived utility at the second stage of the decision process. We adopt an elimination-by-aspects approach to formulating the consideration set formation process. The reader is referred to Fader and McAlister (1990) as an example for details of the elimination-by-aspects model.[2]

Let $A_{DF,t}$, $A_{D,t}$, $A_{F,t}$, and A_0 be the set of alternatives defined as (1) those on display or feature ad at time t, (2) those on display only at time t, (3) those on feature ad only at time t, and (4) the entire set of available alternatives. The brand utility functions corresponding to the nine conditional brand choice probabilities can be summarized in a general expression as[3]

$$\begin{aligned} U_{ikt}^{x,y} &= V_{ikt}^{x,y} + \varepsilon_{ikt}^{x,y} \\ &= \alpha_{ik} + \gamma_i I_{ik,t-1} + \beta_{ap,i} AP_{kt} + \beta_{pc,i} B_{kt}(x,y) + \varepsilon_{ikt}^{x,y}, \\ &\qquad \forall k \in A_t(x,y), \quad (1) \end{aligned}$$

where

$$x = \begin{cases} 0 & \text{if display does not affect choice,} \\ C & \text{if display affects choice through consideration set formation,} \\ P & \text{if display affects choice through price-cut proxy;} \end{cases}$$

$$y = \begin{cases} 0 & \text{if feature ad does not affect choice,} \\ C & \text{if feature ad affects choice through consideration set formation,} \\ P & \text{if feature ad affects choice through price-cut proxy;} \end{cases}$$

$$B_{kt}(x,y) = \begin{cases} PC_{kt} & \text{if } x \neq P \ \&\ y \neq P, \\ PC_{kt} + D_{kt} - PC_{kt} \cdot D_{kt} & \text{if } x = P \ \&\ y \neq P, \\ PC_{kt} + F_{kt} - PC_{kt} \cdot F_{kt} & \text{if } x \neq P \ \&\ y = P, \\ PC_{kt} + D_{kt} + F_{kt} - PC_{kt} \cdot D_{kt} - PC_{kt} \cdot F_{kt} - D_{kt} \\ \quad \cdot F_{kt} + PC_{kt} \cdot D_{kt} \cdot F_{kt} & \text{if } x = P \ \&\ y = P; \end{cases}$$

$$A_t(x,y) = \begin{cases} A_{DF,t} & \text{if } x = C \ \&\ y = C, \\ A_{D,t} & \text{if } x = C \ \&\ y \neq C, \\ A_{F,t} & \text{if } x \neq C \ \&\ y = C, \\ A_0 & \text{if } x \neq C \ \&\ y \neq C. \end{cases}$$

In addition, α_{ki} = alternative-specific constant; $I_{ik,t-1} = 1$ if alternative k was chosen by household i at purchase occasion $(t-1)$, and 0 otherwise; AP_{kt} = alternative k's actual shelf price at t, that is, regular price minus the amount of price discount if there is any; $PC_{kt} = 1$ if alternative k has a price discount at t, and 0 otherwise; $D_{kt} = 1$ if alternative k is on display at t, and 0 otherwise; $F_{kt} = 1$ if alternative k is on feature ad at t, and 0 otherwise. The parameter γ_i captures a household's degree of state dependence. The parameter $\beta_{ap,i}$ represents a household's sensitivity to the *actual* shelf price, and the parameter $\beta_{pc,i}$ represents the additional effect on the household's brand choice when the shelf price is a *discounted* price as opposed to a regular price. Note that the depth of price discount, if any, is reflected in the actual shelf price variable AP_{kt}, because $AP_{kt} = RP_{kt} - APC_{kt}$, where RP_{kt} is the regular price and APC_{kt} is the amount of price discount for alternative k at time t. Therefore, when there is a price discount, its total effect is captured by $(-\beta_{ap,i} APC_{kt} + \beta_{pc,i})$, which means that the depth of price discount does affect brand choice in our model.[4]

$B_{kt}(x,y)$ and $A_t(x,y)$ in the above formulation are the key components of the model. $B_{kt}(x,y)$ captures the price-cut proxy effect. When neither a display nor a feature ad is used as a price-cut proxy, that is, $x \neq P$ and $y \neq P$, only the price-cut indicator PC_{kt} enters this part of the utility function. When a display is used as a price-cut proxy but a feature ad is not, that is, $x = P$ and $y \neq P$, the price cut-related effect is captured by $\beta_{pc,i}(PC_{kt} + D_{kt} - PC_{kt} \cdot D_{kt})$. In this case, in addition to the effect of the *actual* shelf price, the incremental effect of being on display is the same as that of a price discount in a consumer's mind, as measured by $\beta_{pc,i}$, and thus it reflects the notation of a price-cut proxy effect for display. The term $PC_{kt} \cdot D_{kt}$ is subtracted in the formulation to avoid double counting by the analyst when an item is both on display and having a price discount, in which case the incremental effect should still be $\beta_{pc,i}$ rather than $2\beta_{pc,i}$. The logic of the formulation is the same for the cases of ($x \neq P$ and $y = P$) and ($x = P$ and $y = P$).

[2] The consideration set formation process can also be captured in a structural model (see Mehta et al. 2003). Because a structural formulation of the price-cut proxy effect is not available, we choose to adopt a reduced form model, as in Fader and McAlister (1990), to incorporate the price-cut proxy effect in the same model.

[3] Our model can be extended to accommodate the possibility that price cut is used to form consideration sets, in which case there would be 18 decision process scenarios, and the same modeling approach applies. We choose to focus on formulating the alternative decision mechanisms of only displays and feature ads to keep the model relatively simple.

[4] A more conventional approach is to use regular price and the amount of price discount. Our formulation is driven by the need to accommodate the price-cut proxy effect in the same model. When a display or feature ad is used as a price-cut proxy, there is no actual amount of price discount, and thus the signaling effect should be captured by a dummy variable.

$A_t(x, y)$ captures the consideration set formation effect. When both display and feature ad are used to form consideration set, that is, $x = C$ and $y = C$, only alternatives in the set $A_{DF,t}$ are relevant for the brand choice decision at time t, and their utility functions are defined based on which condition in $B_{kt}(x, y)$ is satisfied. The same logic applies for the cases of ($x = C$ and $y \neq C$) and ($x \neq C$ and $y = C$). When neither display nor feature ad is used to form consideration sets, a consumer chooses from the entire set of available alternatives represented by A_0.

The utility function for each of the nine decision processes depicted in Chart 1 is thus a result of its combination of conditions in $A_t(x, y)$ and $B_{kt}(x, y)$. For example, $U_{ikt}^{DC,FC}$, the utility function when both display and feature ad are used as consideration set formation devices, is formulated according to ($x = C$ and $y = C$) in $A_t(x, y)$ and ($x \neq P$ and $y \neq P$) in $B_{kt}(x, y)$. Following is a complete list of the specific utility functions for the nine decision processes:

$$\begin{aligned} U_{ikt}^{DC,FC} &= V_{ikt}^{DC,FC} + \varepsilon_{ikt}^{DC,FC} \\ &= \alpha_{ik} + \gamma_i I_{ik,t-1} + \beta_{ap,i} AP_{kt} + \beta_{pc,i} PC_{kt} + \varepsilon_{ikt}^{DC,FC}, \\ &\qquad \forall k \in A_{DF,t}; \quad (1a) \end{aligned}$$

$$\begin{aligned} U_{ikt}^{DC,FP} &= V_{ikt}^{DC,FP} + \varepsilon_{ikt}^{DC,FP} \\ &= \alpha_{ik} + \gamma_i I_{ik,t-1} + \beta_{ap,i} AP_{kt} \\ &\quad + \beta_{pc,i}(PC_{kt} + F_{kt} - PC_{kt} \cdot F_{kt}) + \varepsilon_{ikt}^{DC,FP}, \\ &\qquad \forall k \in A_{D,t}; \quad (1b) \end{aligned}$$

$$\begin{aligned} U_{ikt}^{DC} &= V_{ikt}^{DC} + \varepsilon_{ikt}^{DC} \\ &= \alpha_{ik} + \gamma_i I_{ik,t-1} + \beta_{ap,i} AP_{kt} + \beta_{pc,i} PC_{kt} + \varepsilon_{ikt}^{DC}, \\ &\qquad \forall k \in A_{D,t}; \quad (1c) \end{aligned}$$

$$\begin{aligned} U_{ikt}^{DP,FC} &= V_{ikt}^{DP,FC} + \varepsilon_{ikt}^{DP,FC} \\ &= \alpha_{ik} + \gamma_i I_{ik,t-1} + \beta_{ap,i} AP_{kt} \\ &\quad + \beta_{pc,i}(PC_{kt} + D_{kt} - PC_{kt} \cdot D_{kt}) + \varepsilon_{ikt}^{DP,FC}, \\ &\qquad \forall k \in A_{F,t}; \quad (1d) \end{aligned}$$

$$\begin{aligned} U_{ikt}^{DP,FP} &= V_{ikt}^{DP,FP} + \varepsilon_{ikt}^{DP,FP} \\ &= \alpha_{ki} + \gamma_i I_{ik,t-1} + \beta_{ap,i} AP_{kt} \\ &\quad + \beta_{pc,i}(PC_{kt} + D_{kt} + F_{kt} - PC_{kt} D_{kt} - PC_{kt} F_{kt} \\ &\quad - D_{kt} F_{kt} + PC_{kt} D_{kt} F_{kt}) + \varepsilon_{ikt}^{DP,FP}, \\ &\qquad \forall k \in A_0; \quad (1e) \end{aligned}$$

$$\begin{aligned} U_{ikt}^{DP} &= V_{ikt}^{DP} + \varepsilon_{ikt}^{DP} \\ &= \alpha_{ki} + \gamma_i I_{ik,t-1} + \beta_{ap,i} AP_{kt} \\ &\quad + \beta_{pc,i}(PC_{kt} + D_{kt} - PC_{kt} \cdot D_{kt}) + \varepsilon_{ikt}^{DP}, \\ &\qquad \forall k \in A_0; \quad (1f) \end{aligned}$$

$$\begin{aligned} U_{ikt}^{FC} &= V_{ikt}^{FC} + \varepsilon_{ikt}^{FC} \\ &= \alpha_{ik} + \gamma_i I_{ik,t-1} + \beta_{ap,i} AP_{kt} + \beta_{pc,i} PC_{kt} + \varepsilon_{ikt}^{FC}, \\ &\qquad \forall k \in A_{F,t}; \quad (1g) \end{aligned}$$

$$\begin{aligned} U_{ikt}^{FP} &= V_{ikt}^{FP} + \varepsilon_{ikt}^{FP} \\ &= \alpha_{ki} + \gamma_i I_{ik,t-1} + \beta_{ap,i} AP_{kt} \\ &\quad + \beta_{pc,i}(PC_{kt} + F_{kt} - PC_{kt} \cdot F_{kt}) + \varepsilon_{ikt}^{FP}, \\ &\qquad \forall k \in A_0; \quad (1h) \end{aligned}$$

$$\begin{aligned} U_{ikt}^{0} &= V_{ikt}^{0} + \varepsilon_{ikt}^{0} \\ &= \alpha_{ki} + \gamma_i I_{ik,t-1} + \beta_{ap,i} AP_{kt} + \beta_{pc,i} PC_{kt} + \varepsilon_{ikt}^{0}, \\ &\qquad \forall k \in A_0; \quad (1i) \end{aligned}$$

where $V_{ikt}^{DC,FC}$, $V_{ikt}^{DC,FP}$, V_{ikt}^{DC}, $V_{ikt}^{DP,FC}$, $V_{ikt}^{DP,FP}$, V_{ikt}^{DP}, V_{ikt}^{FC}, V_{ikt}^{FP}, and V_{ikt}^{0} are the systematic components in each utility function, and $\varepsilon_{ikt}^{DC,FC}$, $\varepsilon_{ikt}^{DC,FP}$, ε_{ikt}^{DC}, $\varepsilon_{ikt}^{DP,FC}$, $\varepsilon_{ikt}^{DP,FP}$, ε_{ikt}^{DP}, ε_{ikt}^{FC}, ε_{ikt}^{FP}, and ε_{ikt}^{0} are the random terms in each utility function, respectively.

In the above utility functions, when consideration set formation takes place, only alternatives that satisfy the consideration set formation condition are evaluated, and the others have a zero chance of being chosen. For a price-cut proxy effect, the coefficient of a display/feature ad is the same as that of the price cut indicator variable, and the interaction term is subtracted to avoid double counting by the analyst. For example, Equation (1d) describes the utility function when a consumer uses feature ads to form consideration sets and treats displays as price-cut proxies. Under this decision process, only alternatives that are on a feature ad enter the consideration set and are compared against each other, and the consumer takes a display as a sign for a price cut when evaluating those alternatives. If alternative k, $\forall k \in A_{F,t}$, is having a price cut only or on display only, the term $(PC_{kt} + D_{kt} - PC_{kt} \cdot D_{kt}) = (1 + 0 - 0)$ or $(0 + 1 - 0) = 1$ and the effect on the brand utility is $\beta_{pc,i}$. If the alternative is both having a price cut and a display, the effect should still be $\beta_{pc,i}$, and this is reflected by the term $(PC_{kt} + D_{kt} - PC_{kt} \cdot D_{kt}) = (1 + 1 - 1) = 1$.

Assuming that each random term in the utility functions follows an identical and independent (IID) Type I extreme value distribution with a location parameter of 0 and a scale parameter of 1, we get the standard logit formulation of the conditional choice probabilities under each of the decision mechanisms described previously. We express the conditional probabilities in a general form:

$$P_{ikt}^{*} = \begin{cases} \exp(V_{ikt}^{*}) / \sum_{j \in A^{*}} \exp(V_{ijt}^{*}), & \text{if } k \in A^{*} \\ 0, & \text{otherwise} \end{cases}, \qquad (2)$$

where $P_{ikt}^{*} \in \{P_{ikt}^{DC,FC}, P_{ikt}^{DC,FP}, P_{ikt}^{DC}, P_{ikt}^{DP,FC}, P_{ikt}^{DP,FP}, P_{ikt}^{DP}, P_{ikt}^{FC}, P_{ikt}^{FP}, P_{ikt}^{0}\}$, $V_{ikt}^{*} \in \{V_{ikt}^{DC,FC}, V_{ikt}^{DC,FP}, V_{ikt}^{DC}, V_{ikt}^{DP,FC},$

$V_{ikt}^{DP,FP}, V_{ikt}^{DP}, V_{ikt}^{FC}, V_{ikt}^{FP}, V_{ikt}^{0}\}$, and $A^* \in \{A_{DF,t}, A_{D,t}, A_{F,t}, A_{0,t}\}$.

The unconditional probability of choosing alternative k by household i at purchase occasion t is obtained by taking the weighted average of the conditional choice probabilities under each decision process, weighted by the occurrence probabilities of the decision processes. Specifically,

$$\begin{aligned}P_{ikt} = {} & [\phi_{DC_i}\phi_{FC_i} + \sigma_{DC,FC}]P_{ikt}^{DC,FC} + [\phi_{DC_i}\phi_{FP_i} + \sigma_{DC,FP}] \\ & \cdot P_{ikt}^{DC,FP} + [\phi_{DC_i}(1-\phi_{FC_i}-\phi_{FP_i}) - \sigma_{DC,FC} - \sigma_{DC,FP}] \\ & \cdot P_{ikt}^{DC} + [\phi_{DP_i}\phi_{FC_i} + \sigma_{DP,FC}]P_{ikt}^{DP,FC} \\ & + [\phi_{DP_i}\phi_{FP_i} + \sigma_{DP,FP}]P_{ikt}^{DP,FP} \\ & + [\phi_{DP_i}(1-\phi_{FC_i}-\phi_{FP_i}) - \sigma_{DP,FC} - \sigma_{DP,FP}]P_{ikt}^{DP} \\ & + [(1-\phi_{DC_i}-\phi_{DP_i})\phi_{FC_i} - \sigma_{DC,FC} - \sigma_{DP,FC}]P_{ikt}^{FC} \\ & + [(1-\phi_{DC_i}-\phi_{DP_i})\phi_{FP_i} - \sigma_{DC,FP} - \sigma_{DP,FP}]P_{ikt}^{FP} \\ & + [(1-\phi_{DC_i}-\phi_{DP_i})(1-\phi_{FC_i}-\phi_{FP_i}) + \sigma_{DC,FC} \\ & \quad + \sigma_{DP,FC} + \sigma_{DC,FP} + \sigma_{DP,FP}]P_{ikt}^{0}. \end{aligned} \tag{3}$$

Note that all nine decision processes are assessed *simultaneously* in the model via Equation (3).

The unconditional choice probability function also depends on the category-level promotion situation at a given purchase occasion. Equation (3) applies to those purchase occasions on which there is at least one alternative on display and at least one alternative with a feature ad in the category. When there is no alternative on display in the category, the display price-cut proxy and consideration set formation effects and their covariance with a feature ad would not be relevant, in which case Equation (3) reduces to

$$P_{ikt} = \phi_{FC_i}P_{ikt}^{FC} + \phi_{FP_i}P_{ikt}^{FP} + (1-\phi_{FC_i}-\phi_{FP_i})P_{ikt}^{0}. \tag{4}$$

When there is no alternative on a feature ad in the category, the feature ad price-cut proxy and consideration set formation effects and their covariance with display would not be relevant, in which case Equation (3) reduces to

$$P_{ikt} = \phi_{DC_i}P_{ikt}^{DC} + \phi_{DP_i}P_{ikt}^{DP} + (1-\phi_{DC_i}-\phi_{DP_i})P_{ikt}^{0}. \tag{5}$$

Finally, when there is no alternative on display or feature ad in the category, all four decision mechanisms are irrelevant, and the household has to choose from the entire set of available alternatives. In this case, Equation (3) reduces to

$$P_{ikt} = P_{ikt}^{0}. \tag{6}$$

Directly estimating the price-cut proxy and consideration set formation probabilities ϕ_{DP}, ϕ_{FP}, ϕ_{DC}, ϕ_{FC}, and covariance terms $\sigma_{DC,FC}$, $\sigma_{DC,FP}$, $\sigma_{DP,FC}$, and $\sigma_{DP,FP}$ is possible but quite cumbersome because of the need to ensure that the resultant occurrence probabilities of the nine decision processes each fall between 0 and 1. Instead, we estimate the nine occurrence probabilities from the data first, using eight parameters, and then compute the parameters of interest (ϕ_{DP}, ϕ_{FP}, ϕ_{DC}, ϕ_{FC}, $\sigma_{DC,FC}$, $\sigma_{DC,FP}$, $\sigma_{DP,FC}$, and $\sigma_{DP,FP}$) from the eight equations that uniquely define the relationships between the parameters of interest and the occurrence probabilities. The details are provided in the appendix.[5]

So far, the model has been constructed at the individual household level. We adopt a finite mixture model approach to capturing unobserved consumer heterogeneity (see Kamakura and Russell 1989), in which parameters are segment specific, denoted by subscript $g = 1, \ldots, G$. The log-likelihood function is given by

$$LL = \sum_{i=1}^{N} \log\left(\sum_{g=1}^{G} q_g \prod_{t=1}^{T_i}\prod_{k=1}^{K}[P_{iktg}]^{I_{ikt}}\right), \tag{7}$$

where P_{iktg} is household i's probability of choosing k at time t according to Equations (3)–(6), given that it belongs to segment g, q_g is the probability of belonging to segment g, T_i is the number of total purchases by household i, and N is the number of households in the sample. The number of latent segments is determined empirically by comparing the Bayesian Information Criterion (BIC) of models with different G, and the one that yields the lowest BIC is selected. When estimating each model, we used 40 sets of different starting values to minimize the chance that the procedure ends at a local optimum.

We compare the proposed model to a multinomial logit model (MNL) with main effects and interactions of display and feature ad with price cut. The variables used in both models are exactly the same. For the MNL model, we also adopt the finite mixture model formulation to capture unobserved consumer heterogeneity. Household i's segment-specific brand utility function is

$$\begin{aligned}U_{iktg} = {} & \alpha_{kg} + \gamma_g I_{ik,t-1} + \beta_{ap,g}AP_{kt} + \beta_{pc,g}PC_{kt} \\ & + \beta_{d,g}D_{kt} + \beta_{f,g}F_{kt} + \beta_{pd,g}PC_{kt}\cdot D_{kt} \\ & + \beta_{pf,g}PC_{kt}\cdot F_{kt} + \varepsilon_{ikt}. \end{aligned} \tag{8}$$

The proposed model is constructed based on behavioral premises of the price-cut proxy and consideration set formation effects. It imposes certain constraints on

[5] Direct estimation of the parameters of interest involves a series of reparameterization. Our empirical results show that the two approaches produced almost identical parameter estimates, but those obtained from the indirect estimation approach are more stable.

Table 1 Descriptive Statistics

Alternative	Average price (cents/ounce)	% occasions with price cut[a] (%)	% occasions on display[a] (%)	% occasions on feature ad[a] (%)	% displays with a PC[a] (%)	% feature ads with a PC[a] (%)
Price and promotion information						
Kraft 12 oz.	18.8	46.5	11.5	21.8	64.4	98.9
Kraft 16 oz.	20.1	19.7	3.3	3.6	47.2	100
Velveeta 12 oz.	16.3	29.8	11.2	6.7	42.7	85.1
Borden 12 oz.	18.1	33.2	0.8	21.6	62.5	81.7
Private label 12 oz.	14.6	56.3	11.2	19.8	84.3	98.7
Private label 16 oz.	15.5	27.9	0.9	9.4	65.7	94.7

Alternative	Number of purchases	Share of purchases (%)	Purchases made on display[b]	Purchase made on feature ad[b]
Purchase information				
Kraft 12 oz.	678	30.3	143 (21.1%)	400 (59.0%)
Kraft 16 oz.	154	6.9	31 (20.1%)	39 (25.3%)
Velveeta 12 oz.	335	15.0	70 (20.9%)	25 (7.5%)
Borden 12 oz.	261	11.7	4 (1.5%)	91 (34.9%)
Private label 12 oz.	569	25.4	125 (22.0%)	246 (43.2%)
Private label 16 oz.	243	10.9	9 (3.7%)	74 (30.5%)

[a]Total number of purchase occasions in the sample = 2,240.
[b]Purchases that were made when the chosen alternative was on display/feature ad, with percentage of the total purchases of that alternative in parentheses.

the relationships among the effects of the price cut, display, and feature ad variables according to the underlying decision mechanisms. Because the parameter of each promotion variable in the MNL model is estimated freely, our theory-based model, which uses the same variables but with constraints on the relationships among them, is not likely to provide a better fit to the data.[6] These two types of models serve different research purposes. The issue at hand is whether the proposed model performs well relative to the MNL model in terms of fit to the data and predictive power, while providing better diagnostics of the underlying decision mechanisms and more reasonable parameter estimates. See Johnson and Meyer (1984), Johnson et al. (1989), and Fader and McAlister (1990) for more discussion on the value of decision-process-based models.

3. Empirical Analysis

3.1. Data Description

We calibrate the model using ACNielsen scanner panel data on single-wrap cheese slices. We chose this category because it had relatively frequent price discounts, displays, and feature ads, as well as fairly high variation in these promotion activities across brands. In addition, price discount, display, and feature ad activities were strongly correlated in the data, which makes it particularly suited for examining the performance of the proposed model. The data were collected in a midwest market during a 104-week period (January 1992 to December 1993). The top six brand-size combinations are included in the analysis: Kraft 12 oz., Kraft 16 oz., Velveeta 12 oz., Borden 12 oz., private label 12 oz., and private label 16 oz. These six items accounted for 77.1% of total category purchases. We use data of the first 52 weeks for model estimation and treat the second 52 weeks as the holdout period. There are 369 households in the data. Table 1 presents descriptive statistics of the estimation data. Note that both display and feature ad were often accompanied by a price cut in the data, and it occurred more frequently for feature ads (81.7%–100%) than for displays (42.7%–84.3%).

3.2. Model Estimation Results

A three-segment model appears to be the best one for the MNL and proposed models based on the method commonly used to determine the number of segments in a latent-class model according to the BIC statistics (Schwartz 1978). For models with one, two, three, and four segments, BIC is 2,611.9, 2,476.8, 2,455.8, and 2,483.7 for the MNL model and 2,652.6, 2,520.8, 2,498.3, and 2,545.9 for the proposed model, respectively. The estimation and holdout prediction results of the two models are reported in Tables 2 and 3.[7] As expected, the MNL model provides a better fit to the data, as indicated by the log-likelihood and BIC.

[6] Note that our model is not a restricted version of the standard MNL model.

[7] We estimated each mixture model with the full set of parameters first and found that in some segments all the estimated covariances turned out to be zero. Next, a model without the covariance terms for these segments was estimated, which generated exactly the same log-likelihood and estimates for the other parameters but with a smaller number of total parameters. Note that ϕ_{DC}, ϕ_{FC}, ϕ_{DP}, and ϕ_{FP} were estimated directly in segments with zero covariances.

Table 2 The Multinomial Logit Model

Variables/Parameters	Segment one	Segment two	Segment three
Constants (baseline: private label 16 oz.)			
Kraft 12 oz.	0.895***	3.833***	0.904***
Kraft 16 oz.	−0.263	3.013***	0.357
Velveeta 12 oz.	−0.951***	2.535***	0.599**
Borden 12 oz.	0.102	3.267***	−0.891**
Private label 12 oz.	0.553***	1.248***	−0.209
State dependence (γ)	0.324***	1.112***	2.799***
Actual price (β_{ap})	−0.293***	−0.329***	−0.113***
Price-cut indicator (β_{pc})	0.284**	−0.435**	−0.161
Display (β_D)	−0.088	1.218***	1.178**
Feature ad (β_F)	0.158	−0.116	1.725**
Price cut ∗ Display (β_{PD})	0.604*	−0.700**	−1.799***
Price cut ∗ Feature ad (β_{PF})	0.661*	1.305**	−1.267**
Segment size	43.1%	35.0%	21.9%
Number of parameters		38	
-Log-likelihood		2,309.2	
BIC		2,455.8	
Holdout -log-likelihood		3,096.2	
Holdout hit rate		69.6%	
RMSE (holdout market share prediction)		0.0243	

*p value < 0.10.
**p value < 0.05.
***p value < 0.01.

Table 3 The Proposed Model

Variables/Parameters	Segment one	Segment two	Segment three
Constants (baseline: private label 16 oz.)			
Kraft 12 oz.	1.401***	5.214***	0.721**
Kraft 16 oz.	0.088	4.487***	0.227
Velveeta 12 oz.	−0.160	3.487***	0.606***
Borden 12 oz.	0.389**	4.674***	−1.306***
Private label 12 oz.	0.658***	1.305**	−0.213
State dependence (γ)	0.477***	1.311***	2.743***
Actual price (β_{ap})	−0.329***	−0.395***	−0.095**
Price-cut indicator (β_{pc})	0.144**	0.125**	0.074
Pr{consideration set formation: display} (ϕ_{DC})	0.219	0.087	0.000
Pr{consideration set formation: feature} (ϕ_{FC})	0.356	0.041	0.000
Pr{price-cut proxy: display} (ϕ_{DP})	0.000	0.913	1.000
Pr{price-cut proxy: feature} (ϕ_{FP})	0.000	0.000	1.000
cov(DC, FC)	0.141	0.000	0.000
cov(DC, FP)	−0.000	0.000	−0.000
cov(DP, FC)	−0.000	0.000	−0.000
cov(DP, FP)	0.000	0.000	0.000
Segment size	56.5%	22.4%	21.1%
Number of parameters		42	
-Log-likelihood		2,336.3	
BIC		2,498.3	
Holdout -log-likelihood		3,096.8	
Holdout hit rate		70.1%	
RMSE (holdout market share prediction)		0.0228	

*p value < 0.10; **p value < 0.05; ***p value < 0.01, except for the probability and covariance terms, which were computed from other parameter estimates.

Nonetheless, the -log-likelihood of our model, which imposes constraints on the relationships among variables, shows a decent fit to the data, compared to the MNL model (2,336.3 vs. 2,309.2). To examine the models' predictive power for the holdout period data, we use three measures, as in Ailawadi et al. (1999): the holdout -log-likelihood, hit rate, and the root mean squared error (RMSE) in prediction of market share. The MNL model yields a holdout -log-likelihood of 3,096.2, a hit rate of 69.6%, and RMSE of 0.0243, while our model generates an almost identical holdout -log-likelihood of 3,096.8 and a slightly better hit rate of 70.1% and RMSE of 0.0228.[8]

The two models depict a similar segment structure in terms of segment sizes and estimates of the alternative-specific constants, state dependence parameter, and the actual shelf price coefficient within each segment. To test the stability of the segments, we have estimated several alternative models, including MNL models, using the regular price and depth of price discount with or without their multiplicative terms with display and feature ad, and the current MNL and proposed models estimated using the holdout period data. In addition, we have examined segment stability using the approach developed by Ailawadi et al. (1999). All results reveal very similar segment structures and indicate that the segments are stable.[9]

Despite the similar segment structure, the MNL model produces several counterintuitive inferences on promotion effects. For example, in segment two, the price cut coefficient is negative and significant. In addition, the effect of offering display and price cut together (0.083) is smaller than the effect of display alone (1.218). In segment three, the effect of offering feature ad and price cut together (0.297) is smaller than the effect of feature ad alone (1.725), and the effect of offering display and price cut together is even estimated to be negative (−0.782). Note that a negative interaction effect of display/feature ad and price cut *per se* is not problematic. The counterintuitive inferences discussed here result from a combination of a negative price cut coefficient and a negative

[8] We also did the model estimation and holdout prediction using data on split-half household samples. The model estimation results reveal similar patterns as in Tables 2 and 3. For holdout predictions, the proposed model performs better than the MNL model on all three measures. Detailed results are available upon request. The author thanks an anonymous reviewer for suggesting this alternative approach.

[9] The author thanks an anonymous reviewer for suggesting these tests.

interaction term. A similar problem would occur if a positive main effect of price cut or display/feature ad is smaller than the magnitude of a negative interaction effect. These problematic estimates are likely due to multicollinearity of the promotion variables.[10] Although the MNL model provides a somewhat better fit to the estimation data, the proposed model does not generate any counterintuitive parameter estimates in the empirical application, which attests to the advantage of our approach for the purpose of making promotion decision recommendations. In the proposed model, the constraints embedded in the underlying behavioral mechanisms prevent coefficients of the display and feature ad effects from taking on counterintuitive values.[11]

We now focus on the results of the proposed model. The parameter estimates reveal three distinct segments with regard to their tendencies to use displays and feature ads as price-cut proxies versus consideration set formation devices. Segment one is the largest in size (56.5%). It has the lowest degree of state dependence (0.477) and the largest effect of a real price cut (0.144). Consumers in this segment seem to use display and feature ad for consideration set formation sometimes, with probabilities of 21.9% and 35.6%, respectively, but never treat them as price-cut proxies. Segment two is the middle segment in terms of size (22.4%) and degree of state dependence (1.311) and has similar actual shelf price and price cut coefficients as segment one. This segment appears to be primarily influenced by in-store promotion activities. Consumers in this segment tend to see displays primarily as price-cut proxies (with a probability of 91.3%) and occasionally use them to form consideration sets (with a probability of 8.7%). They very occasionally use feature ads to form consideration sets (with a probability of 4.1%) but never seem to treat them as price-cut proxies. Segment three is the smallest in size (21.1%). It exhibits the highest degree of state dependence (2.743) and is least sensitive to the actual shelf price of an item. Unlike the first two segments, consumers in segment three appear to always treat display and feature ad as price-cut proxies (with estimated probabilities of 100%) and never use them to form consideration sets. Finally, the vast majority of the covariance terms between using display/feature ad as consideration set formation devices and price-cut proxies turned out to be zero (with rounding to the third decimal point). This is consistent with the corresponding probability estimates being zero or one in a segment. In other words, the covariance goes to zero when there is no variation (i.e., either yes or no) in the occurrence of one of the events involved. The only nonzero covariance term in our estimation indicates that there is a positive correlation of using displays and feature ads to form consideration sets.

The estimation results reveal some interesting patterns of the association between consumer characteristics and the tendencies to engage in different decision processes. It appears that consumers who are the least state dependent and are more sensitive to the actual shelf price (segment one) tend to use displays and feature ads to form consideration sets. Consumers who are most state dependent and least sensitive to the actual shelf price (segment three) tend to treat displays and feature ads as proxies for price cuts. Consumers who have an intermediate level of state dependence show a combination of the consideration set formation and price-cut proxy effects. A plausible cause for this pattern is that the state dependence, shelf price, and price discount coefficients could reflect how much effort consumers spend in processing information on the actual prices at a given purchase occasion. Consumers in segment one (the least state-dependent segment) are likely to be more involved in evaluating pricing and promotion information and thus know whether a display and feature ad is accompanied by a real price discount, and they are not likely to confuse them when they do not occur together. For these consumers, sometimes displays and feature ads can be used as heuristics to form consideration sets, which points to a conscious strategy to save cognitive effort and simplify purchase decisions, instead of unknowingly treating displays and feature ads as signs for price cuts. But most of the time (with a probability of 64.3%) consumers in segment one tend to evaluate the entire set of alternatives based on their actual price and discount information. Consumers in segment three (the most state-dependent segment) tend to rely on past purchase outcomes and are likely to be least involved in processing actual price and promotion information, and thus they seem to almost always take displays and feature ads as proxies for price cuts. This is consistent with the finding by Inman et al. (1990) that consumers with low need for cognition tend to see promotional signs as a cue for a price cut even when there is no actual reduction in the price. Consumers

[10] Using alternative variables, such as "display only (without feature ad and price cut)," "feature ad only (without display and price cut)," or "price cut only (without display and feature ad)," could reduce collinearity among promotion variables to some extent but does not seem to resolve the problems caused by multicollinearity completely. See, for example, Lemon and Nowlis (2002) and Gupta (1988). We estimated MNL models using promotion variables as defined in these two studies, respectively, as well as MNL models using regular price and depth of price cut with and without the multiplicative terms of price cut and display/feature ad, and found that the problem of counterintuitive parameter estimates exists in all four models.

[11] Note that the proposed model does not impose constraints on the coefficients of the actual shelf price and price-cut indicator variables.

in segment two appear to be in between the other two segments in terms of their involvement in the purchase decisions and thus might use displays and feature ads for both consideration set formation and price-cut proxies.

To summarize, we have found support for both the price-cut proxy and consideration set formation effects but show that they tend to occur for consumers with different characteristics. Even for those consumers who are aware of and sensitive to actual shelf prices, the mere fact that a brand is on display or has a feature ad can increase its choice probability by influencing their consideration set formation process. This phenomenon cannot be explained by the price-cut proxy effect.

To further examine the characteristics of the segments identified by our model, we classify households in the data into one of the three segments based on their posterior segment probabilities. Segment-specific descriptive information is presented in Table 4. Segment one has the highest percentage of purchases made on price discount and feature ad, the highest percentage of switching purchases, and the second highest percentage of purchase made on display. Segment two lies in the middle on these measures, except that it has the highest percentage of purchases made on display, which implies that this segment is the most responsive one to in-store display promotions. Segment three is the least responsive to price discounts, displays, and feature ads and also has the lowest percentage of switching purchases. These patterns are consistent with our model estimation results.

The different patterns across consumer segments also provide an explanation for the mixed patterns of promotion interaction effects in the literature. As speculated in §1, if display/feature ad affects brand choice mainly through the price-cut proxy effect, one would expect the interaction term of display/feature ad and price cut in a standard brand choice model to be negative, because the joint effect would be smaller than the sum of the two individual effects. On the other hand, if displays/feature ads are mainly used to form consideration sets, the interaction effect is likely to be positive, because a display/feature ad helps an item get into the consideration set, which is the first stage of the decision process, while a price cut helps it stand out among the remaining alternatives at the evaluation stage, and thus they reinforce each other's effect in the entire brand choice process. This conjecture is supported by a comparison of the estimation results from the MNL model and the proposed model, which depict a similar segment structure in terms of the brand constants, state dependence, and actual price coefficients.[12] For segment one, our model indicates that display and feature ad are used to form consideration sets, while their interactions with price cut are both positive in the MNL model. For segment two, feature ad is occasionally used to form consideration sets (but never treated as price-cut proxy) according to our model, and its interaction with price cut is positive in the MNL model, while display is more than 10 times more likely to be used as a price-cut proxy than a consideration set formation device, accordingly to our model (91.3% vs. 8.7%). Correspondingly, its interaction with price cut is negative in the MNL model. For segment three, the predominant effects of display and feature ad are price-cut proxies based on our model, and the interaction terms in the MNL model both turn out to be positive. The matching pattern between the MNL and our models suggests that *a positive interaction effect of display/feature ad and price cut in a compensatory brand choice model is likely due to the consideration set formation mechanism, while a negative interaction effect is likely attributable to the price-cut proxy mechanism*. It implies that the mixed patterns of the interaction effects reported in previous studies might simply be a result of which mechanism dominates at the aggregate level for a particular data sample. This explanation could help shed light on an unsolved puzzle in the literature.[13]

Table 4 **Segment-Specific Descriptive Information Based on the Proposed Model**

	Segment one	Segment two	Segment three
No. of households	222 (60.2%)	73 (19.8%)	74 (20.1%)
No. of purchases	1,242	528	470
Percent of purchases on price cut (%)	73.9	57.8	34.5
Percent of purchases on display (%)	19.0	19.7	8.9
Percent of purchases on feature ad (%)	50.2	37.5	11.3
Percent of switching purchases (%)	66.0	50.0	20.0

4. Discussion

The proposed model illustrates the value of incorporating likely behavioral mechanisms in model building for better decision making. Its managerial contribution can be demonstrated by a comparison of the

[12] To test how well the segments match up in the two models, we did a cross-classification of households' segment membership based on the two models and found that 89.2% of the households were classified to the same segment, which indicates that the segments in the two models do match up well. The author thanks an anonymous reviewer for suggesting this test.

[13] Note that accounting for heterogeneity does not solve the problem of counterintuitive estimates of promotion effects in standard MNL models, including the model without explicit interaction terms, based on our empirical analyses.

different implications for promotion decisions drawn from our model and the benchmark MNL model. It is a common practice by retailers to combine temporary price reductions with displays and feature ads, which is also the case in our data. Yet this practice would contradict with recommendations based on an MNL model that yields negative interaction effect(s) of price cut with display and/or feature ad. Previous studies as well as the current one indicate that standard brand choice models often generate negative interaction effects. For example, the MNL model estimated from our data predicts a negative overall interaction effect of display and price cut (averaged across segments), which implies that the retailers should not have offered price discounts with in-store displays for the category analyzed here. Yet it occurred in 43%–84% of the cases when an item was on display.

Results from the proposed model, however, reveal a very different picture. If a negative interaction effect in a standard brand choice model is due to the price-cut proxy mechanism, as corroborated by our empirical analysis, it would be unwise to completely eliminate price reductions from displays or feature ads. The price-cut proxy effect is caused by a lack of motivation or interest to engage in careful information processing. Its occurrence relies on consumers' lack of accurate or complete information contained in a display or feature ad. This effect would disappear once consumers realize that a display or feature ad is never accompanied by an actual price discount. It implies that, if a retailer completely eliminates price discounts from displays or feature ads, these two promotion vehicles will no longer have any effect (including the main effect) on brand choice in the long run for consumers who primarily use them as price-cut proxies, because their signaling effect would erode over time. Rather, some but not all displays and feature ads should be accompanied by price discounts to induce their usage as cues for price cuts, still taking advantage of the phenomenon that some consumers see them as signs for discounts even when there are no actual price reductions. It would be beneficial to combine price discounts with displays/feature ads if the consideration set formation mechanism dominates. (This recommendation is also supported by the MNL model in our application.)

Results from our model also indicate that it makes sense to bundle price cuts with feature ads more frequently than with displays for the category studied, which is exactly the case in our data, because the overall probability (averaged across segments) of using feature ads to form consideration sets is greater than that of using displays to form consideration sets (21.0% vs. 14.3%), while the overall probability of using displays as price-cut proxies is greater than that of using feature ads as price-cut proxies (41.6% vs. 21.1%). This finding suggests that, although retailers might not know the detailed decision mechanisms underlying consumers' brand choice processes, they seem to have the right intuition for frequently offering price discounts with displays and feature ads. The model we propose here offers an analytical tool to help retailers assess the tendencies to use various decision mechanisms by different consumers using actual purchase data, and our empirical results provide a rationale for a common practice employed by many retailers.

To summarize, the contributions of this study mainly lie in its substantive and managerial findings:

(1) It reveals distinctive segments of consumers who exhibit different tendencies to use display/feature ad as price-cut proxies and/or to form consideration sets.

(2) By generating sensible parameter estimates, it is better suited than standard brand choice models to assist managerial decisions.

(3) It provides a plausible explanation for the mixed patterns of promotion interactions reported in the literature.

This study suggests several directions for future research. For example, it would be interesting to look at whether displays/feature ads that make low-end brands more prominent would increase the sales of high-end brands due to the within-category ranking effect (see Leclerc et al. 2005). Another possible direction for future research is to apply the proposed model to study promotions in online stores, similar to the context examined by Zhang and Krishnamurthi (2004), where there might be a consideration set formation effect as well as a price-cut proxy effect of online coupons or other forms of promotions.

It should be pointed out that, although the two behavioral mechanisms we incorporate here are the prominent ones in the literature, there are other explanations for why displays and feature ads affect brand choice. For example, feature ads are typically seen before a consumer comes to the store. They might help the store choice decision and limit the consideration set to products available in the chosen store, in which case in-store displays enter the equation only after the first-stage effect of feature ads.[14] In addition, our model provides one approach that helps alleviate problems caused by multicollinearity in the context of brand choice models, but there have been other methods proposed to handle the multicollinearity issue. For instance, in their analysis of store-level sales data, van Heerde et al. (2001, 2004) either split the sample into four subsamples or equivalently use four separate price index variables (defined by "without display or feature ad support," "with display

[14] The author thanks an anonymous reviewer for this insight.

only," "with feature ad only," and "with both display and feature ad"), and separately estimate the unique demand function under each promotion scenario using semiparametric models. Their approach works very well in capturing complex effects.

A limitation of the proposed model is that the probabilities of using display/feature ad for price-cut proxy and consideration set formation do not vary across purchase occasions for the same consumer. These probabilities may depend on factors such as size of the feature ad, location or prominence of the display, major versus minor shopping trips, and depth of price discount. An extension of the current model is to formulate the probabilities as functions of these covariates. Of particular interest is whether the probabilities of consideration set formation vary by the depth of price discounts.[15]

Despite the limitations, we have shown that by incorporating two possible underlying behavioral mechanisms, the proposed model generates sensible parameter estimates on the joint effects of promotion vehicles (in-store display, feature ad, and price discount) and resolves a long-standing problem with standard brand choice models when it comes to evaluating the joint promotion effects. As a result, insights gained from the proposed model can help make better managerial decisions for promotion planning.

Acknowledgments
The author thanks ACNielsen for providing the data used in this study. She is also grateful to Aradhna Krishna, Lakshman Krishnamurthi, and Michel Wedel for their helpful input, and the editor, area editor, and three anonymous reviewers for their many valuable comments and suggestions.

Appendix

Derivation of the Occurrence Probabilities
For ease of exposition, the subscripts i and g are omitted in this section.

$$\because \quad E(DC\cdot FC) = 1\cdot \Pr(DC=1, FC=1) + 0\cdot[1-\Pr(DC=1, FC=1)] = \Pr(DC=1, FC=1),$$

and $\operatorname{cov}(DC, FC) = E(DC\cdot FC) - E(DC)\cdot E(FC)$

$$\therefore \quad \Pr(DC=1, FC=1) = E(DC)\cdot E(FC) + \operatorname{cov}(DC, FC) = \phi_{DC}\phi_{FC} + \sigma_{DC,FC}. \tag{A.1}$$

Similarly,

$$\Pr(DC=1, FP=1) = E(DC)\cdot E(FP) + \operatorname{cov}(DC, FP) = \phi_{DC}\phi_{FP} + \sigma_{DC,FP}, \tag{A.2}$$

[15] The author is grateful to two anonymous reviewers for their insights on this issue.

and,

$$\begin{aligned}\Pr(DC=1, FC=0, FP=0) &= \phi_{DC} - (\phi_{DC}\phi_{FC} + \sigma_{DC,FC}) - (\phi_{DC}\phi_{FP} + \sigma_{DC,FP})\\ &= \phi_{DC}(1-\phi_{FC}-\phi_{FP}) - \sigma_{DC,FC} - \sigma_{DC,FP}.\end{aligned} \tag{A.3}$$

Applying the same logic, we get the occurrence probabilities of the other six decision processes as described in Chart 1.

Indirect Estimation of ϕ_{DC}, ϕ_{FC}, ϕ_{DP}, ϕ_{FP}, $\sigma_{DC,FC}$, $\sigma_{DC,FP}$, $\sigma_{DP,FC}$, and $\sigma_{DP,FP}$
To ensure that the occurrence probabilities of the nine decision processes fall between 0 and 1 and sum to 1, we estimate them directly using a logit transformation. Let

$$\begin{aligned}
W &= 1 + \sum_{j=1}^{8}\exp(\theta_j)\\
p_1 &= \Pr(DC=1, FC=1) = \exp(\theta_1)/W\\
p_2 &= \Pr(DC=1, FP=1) = \exp(\theta_2)/W\\
p_3 &= \Pr(DC=1, FC=0, FP=0) = \exp(\theta_3)/W\\
p_4 &= \Pr(DP=1, FC=1) = \exp(\theta_4)/W\\
p_5 &= \Pr(DP=1, FP=1) = \exp(\theta_5)/W\\
p_6 &= \Pr(DP=1, FC=0, FP=0) = \exp(\theta_6)/W\\
p_7 &= \Pr(DC=0, DP=0, FC=1) = \exp(\theta_7)/W\\
p_8 &= \Pr(DC=0, DP=0, FP=1) = \exp(\theta_8)/W\\
p_9 &= \Pr(DC=0, DP=0, FC=0, FP=0) = 1/W.
\end{aligned}$$

$\theta_1, \ldots, \theta_8$ are parameters to be estimated from the data.

The relationships between the occurrence probabilities and the parameters of interest (ϕ_{DC}, ϕ_{FC}, ϕ_{DP}, ϕ_{FP}, $\sigma_{DC,FC}$, $\sigma_{DC,FP}$, $\sigma_{DP,FC}$, and $\sigma_{DP,FP}$) are defined by the following eight-equation system:

$$\begin{cases}
\phi_{DC}\phi_{FC} + \sigma_{DC,FC} = p_1\\
\phi_{DC}\phi_{FP} + \sigma_{DC,FP} = p_2\\
\phi_{DC}(1-\phi_{FC}-\phi_{FP}) - \sigma_{DC,FC} - \sigma_{DC,FP} = p_3\\
\phi_{DP}\phi_{FC} + \sigma_{DP,FC} = p_4\\
\phi_{DP}\phi_{FP} + \sigma_{DP,FP} = p_5\\
\phi_{DP}(1-\phi_{FC}-\phi_{FP}) - \sigma_{DP,FC} - \sigma_{DP,FP} = p_6\\
(1-\phi_{DC}-\phi_{DP})\phi_{FC} - \sigma_{DC,FC} - \sigma_{DP,FC} = p_7\\
(1-\phi_{DC}-\phi_{DP})\phi_{FP} - \sigma_{DC,FP} - \sigma_{DP,FP} = p_8.
\end{cases} \tag{A.4}$$

There are eight unknown parameters and eight equations in (A.4), and thus the solutions are unique. Solving (A.4) gives the following identities, which are used to estimate the parameters of interest:

$$\phi_{DC} = p_1 + p_2 + p_3, \tag{A.5}$$

$$\phi_{FC} = p_1 + p_4 + p_7, \tag{A.6}$$

$$\phi_{DP} = p_4 + p_5 + p_6, \tag{A.7}$$

$$\phi_{FP} = p_2 + p_5 + p_8, \tag{A.8}$$

$$\sigma_{DC,\,FC} = p_1 - (p_1 + p_2 + p_3)(p_1 + p_4 + p_7), \tag{A.9}$$

$$\sigma_{DC,\,FP} = p_2 - (p_1 + p_2 + p_3)(p_2 + p_5 + p_8), \tag{A.10}$$

$$\sigma_{DP,\,FC} = p_4 - (p_4 + p_5 + p_6)(p_1 + p_4 + p_7), \tag{A.11}$$

$$\sigma_{DP,\,FP} = p_5 - (p_4 + p_5 + p_6)(p_2 + p_5 + p_8). \tag{A.12}$$

It can be shown that estimates obtained based on the above equations conform to the regularity properties of probabilities and covariance terms.

References

Ailawadi, Kusum L., Karen Gedenk, Scott A. Neslin. 1999. Heterogeneity and purchase event feedback in choice models: An empirical analysis with implications for model building. *Internat. J. Res. Marketing* **16** 177–198.

Allenby, Greg M., James L. Ginter. 1995. The effects of in-store display and feature advertising on consideration sets. *Internat. J. Res. Marketing* **12** 67–80.

Andrews, Rick L., T. C. Srinivasan. 1995. Studying consideration effects in empirical choice models using scanner panel data. *J. Marketing Res.* **32**(1) 30–41.

Bronnenberg, Bart J., Wilfried R. Vanhonacker. 1996. Limited choice sets, local price response, and implied measures of price competition. *J. Marketing Res.* **33**(May) 163–173.

Chintagunta, Pradeep K. 1992. Estimating a multinomial probit model of brand choice using the method of simulated moments. *Marketing Sci.* **11**(4) 386–407.

Fader, Peter S., Leigh McAlister. 1990. An elimination by aspects model of consumer response to promotion calibrated on UPC scanner data. *J. Marketing Res.* **27**(August) 322–332.

Gilbride, Timothy, Greg M. Allenby. 2004. A choice model with conjunctive, disjunctive, and compensatory screening rules. *Marketing Sci.* **23**(3) 391–406.

Grover, Rajiv, V. Srinivasan. 1992. Evaluating the multiple effects of retail promotions on brand loyal and brand switching segment. *J. Marketing Res.* **29**(February) 76–89.

Gupta, Sunil. 1988. Impact of sales promotions on when, what, and how much to buy. *J. Marketing Res.* **25**(November) 342–355.

Hauser, John R., Birger Wernerfelt. 1990. An evaluation cost model of consideration sets. *J. Consumer Res.* **16**(March) 393–408.

Inman, Jeffrey J., Leigh McAlister, Wayne D. Hoyer. 1990. Promotion signal: Proxy for a price cut? *J. Consumer Res.* **17**(June) 74–81.

Johnson, Eric J., Robert J. Meyer. 1984. Compensatory choice models of noncompensatory processes: The effect of varying context. *J. Consumer Res.* **11**(June) 528–541.

Johnson, Eric J., Robert J. Meyer, Sanjoy Ghose. 1989. When choice models fails: Compensatory models in negatively correlated environments. *J. Marketing Res.* **24**(August) 255–270.

Kamakura, Wagner A., Gary J. Russell. 1989. A probabilistic choice model for market segmentation and elasticity structure. *J. Marketing Res.* **26**(November) 379–390.

Leclerc, France, Christopher K. Hsee, Joseph C. Nunes. 2005. Narrow focusing: Why the relative position of a good in its category matters more than it should. *Marketing Sci.* **24**(2) 194–205.

Lemon, Katherien N., Stephen M. Nowlis. 2002. Developing synergies between promotions and brands in different price-quality tiers. *J. Marketing Res.* **39**(May) 379–390.

Lussier, Denis A., Richard W. Olshavsky. 1979. Task complexity and contingent processing in brand choice. *J. Consumer Res.* **6**(September) 154–165.

Mehta, Nitin, Surendra Rajiv, Kannan Srinivasan. 2003. Price uncertainty and consumer search: A structural model of consideration set formation. *Marketing Sci.* **22**(1) 58–84.

Papatla, Purushottam. 1996. A multiplicative fixed-effects model of consumer choice. *Marketing Sci.* **15**(3) 243–261.

Papatla, Purushottam, Lakshman Krishnamurthi. 1996. Measuring the dynamic effects of promotions on brand choice. *J. Marketing Res.* **33**(February) 20–35.

Payne, John L. 1976. Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organ. Behav. Human Performance* **16**(August) 366–387.

Roberts, John H., James M. Lattin. 1991. Development and testing of a model of consideration set composition. *J. Marketing Res.* **28**(November) 429–441.

Schwartz, G. 1978. Estimating the dimension of a model. *Ann. Statist.* **6** 461–464.

Siddarth, S., Randolph E. Bucklin, Donald G. Morrison. 1995. Making the cut: Modeling and analyzing choice set restriction in scanner panel data. *J. Marketing Res.* **32**(August) 255–266.

van Heerde, Harald J., Peter S. H. Leeflang, Dick R. Wittink. 2001. Semiparametric analysis to estimate the deal effect curve. *J. Marketing Res.* **38**(May) 197–215.

van Heerde, Harald J., Peter S. H. Leeflang, Dick R. Wittink. 2004. Decomposing the sales promotion bump with store data. *Marketing Sci.* **23**(3) 317–334.

Zhang, Jie, Lakshman Krishnamurthi. 2004. Customizing promotions in online stores. *Marketing Sci.* **23**(4) 561–578.