European Journal of Operational Research 214 (2011) 340–347

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Stochastics and Statistics

The nested consideration model: Investigating dynamic store consideration sets and store competition ☆

Joseph Pancras *

University of Connecticut, School of Business, Marketing Department, 2100 Hillside Road, Unit 1041, Storrs, CT 06269-1041, United States

ARTICLE INFO

Article history:
Received 26 May 2010
Accepted 27 April 2011
Available online 30 April 2011

Keywords:
Marketing
Nested consideration
Store competition
Consideration sets
Nested logit

ABSTRACT

The nested logit model has been widely used to study nested choice. A typical example of such nested choice is store patronage and brand choice. An important limitation of the nested logit model is that it assumes that all alternatives at both levels of the nest are available in the choice set of the consumer. While there is a wide literature on the incorporation of restricted choice sets into the logit model, the logical extension of this analysis to nested restricted choice sets has not been pursued in the literature. In this study we develop a nested consideration model that integrates store choice and brand choice incorporating the formation of dynamic restricted choice sets of both stores and brands. We term the model the nested consideration model and derive the related probabilities in a manner analogous to the well-known nested logit model. In an empirical illustration, we find that the nested consideration model shows better prediction than nested logit models (with the same explanatory variables). We find that it is important to account for dynamic store consideration sets rather than static sets or store loyalty measures. We also conduct simulations to demonstrate the importance of the nested consideration set model for correct pricing and store location decisions of business managers.

1. Introduction

Nested choice has been widely studied in various areas of study such as marketing and transportation research. The widely used model used to study nested choice has been the nested logit model (Ben-Akiva and Lerman, 1985; Train, 2003). Applications in marketing of the nested logit model include Kannan and Wright (1991) and Bell and Lattin (1998). Recently there has also been increasing interest in the operations research literature on the theory and applications of the nested logit model and related choice modeling approaches, including Bekhor et al. (2006), Bierlaire (2004), Kalouptsidis et al. (2007), Baltas (2004), García-Ródenas and Ángel Marín (2009) and Schön (2010). While a robust literature has been developed in marketing and transportation literature on incorporating *restricted* choice sets into the logit model, the logical extension of incorporating restricted choice sets into a nested choice model has not, to our knowledge, been pursued in either literature.

Developing such a model is important for business managers for the following reasons. One, managers need to be able to predict store and brand choice accurately in making optimal decisions on marketing mix variables such as prices and promotions. Two, the manager of a focal store needs to estimate the effect of competition from nearby competing stores on the focal store. The recent travails of retail outlets such as Starbucks (which needed to close hundreds of outlets in the US) due to cannibalization between nearby stores (Kiviat, 2008) underscore the importance of better methods to predict store choice decisions of consumers. Three, while our empirical application is in the context of frequently purchased packaged goods, the problem of spatial competition between nearby outlets is relevant for any business with a widely dispersed distribution network. An important recent example is the auto industry in the US, where Chrysler dealerships were being shut down due to competition between dealers that were located closer together (Kiley, 2009). Understanding how store consideration sets and state dependence in store salience and/or choice affect spatial competition thus is critical for businesses.

This article makes the following methodological contributions to the study of store choice. One, we develop a model that incorporates restricted choice set formation at both the brand choice and store choice stages. Two, we demonstrate the importance of dynamic store consideration sets as compared to store loyalty or static store consideration sets, which have been traditionally utilized in explaining store choice. By incorporating store state dependence

☆ I thank David Bell (Wharton) for making the basket scanner data available for this research. I also thank K. Sudhir, Joel Steckel, Russell Winer, Yuxin Chen, Paris Cleanthous, Sha Yang and Skander Essegaier for their comments and suggestions. This article is based on an essay from my 2005 dissertation at New York University.

* Tel.: +1 860 486 0810; fax: +1 860 486 5246.
E-mail address: joseph.pancras@business.uconn.edu

0377-2217/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.ejor.2011.04.027

into the model of store salience and choice, we provide an integrated framework that store managers can use to set pricing strategies. The paper also makes the substantive contributions through two simulations that could be used by store managers, which could lead to better pricing and store location strategies.

The subsequent sections of this article are organized as follows. First, we discuss the relevant literature on store choice and consideration sets. Then we develop our model of nested consideration, building on assumptions and hypotheses about the choice process. Then we discuss our operationalization of the proposed model, and benchmark the performance vis-a-vis alternative, commonly used models of store choice. We also discuss two simulations that demonstrate the utility of this modeling approach to store managers. Finally we discuss the managerial implications of the proposed model and conclude.

2. Relevant literature

In Table 1 we lay out the contributions of this study with respect to the earlier literature.

These contributions can be classified into the following three broad streams:

(1) Incorporation of store salience into the store choice model, characterizing store salience by a threshold and store state dependence as opposed to earlier methods of using either only a store state dependence model (Popkowski Leszczyc et al., 2000) or using only store loyalty (Bell and Lattin, 1998; Bell et al., 1998) to explain store choice. Unlike earlier models of store state dependence, our approach also preserves the advantages of using the category inclusive value in explaining store choice.
(2) Integrating dynamic brand salience and store salience into the same choice model, unlike earlier papers that modeled either brand salience (Bronnenberg and Vanhonacker, 1996) or store salience (Fotheringham, 1988), but not both.
(3) Our approach, uniquely among the consideration set literature (Mehta et al., 2002), incorporates both the parsimonious sku (stock-keeping unit) characteristics approach and accounts for sku availability (Campo et al., 2003).

We next describe the relevant literature in these three streams.

2.1. State dependence in store choice and store consideration sets

The phenomenon of state dependence in brand purchase has been studied in the context of brand purchase (Seetharaman et al., 1999) as well as store choice (Popkowski Leszczyc et al., 2000, 2004). Positive state dependence leads to a higher probability of future purchase of the currently chosen brand/store and is termed ‘inertia’ while negative state dependence leads to a lower probability of future purchase of the currently chosen brand, a behavior termed ‘variety seeking’. Popkowski Leszczyc et al. (2000) use a hazard approach to model state dependence in store choice. This approach does not take into account store salience and store consideration sets. It also does not take into account category inclusive values. The other approach that has been used to study store choice is the nested logit approach of Bell and Lattin (1998). While this approach uses category inclusive values to explain store choice, it does not include dynamic store salience or store consideration sets. The approach of Bell and Lattin (1998) is rather to use ‘preprocessed’ (static) consideration sets, i.e., stores that have been visited in a calibration period, as variables to explain store choice. They also use store loyalty rather than store state dependence to explain store choice.

Our study incorporates dynamic store salience and consideration sets, with store state dependence being used to parameterize the dynamic store salience,[1] and we show that this approach dramatically improves the predictive power of the store choice model.

2.2. Two stage models of store and brand choice

Considerable evidence has been presented in the restricted choice set literature that households form consideration sets of brands, and then choose brands from the household’s consideration set, rather than choose brands in a single stage, from the set of all available brands (Mehta et al., 2002; Pancras, 2010). The literature on *store* consideration sets has however been somewhat limited. Fotheringham (1988) argues that consumers limit search not only with respect to brands but also with respect to stores. Consumers may not have the ability or time to evaluate all stores within a city, and may make an initial choice of a cluster of stores, a shopping district or perhaps a mall, then select a store from this reduced set of stores. An extreme example of such a set is where a consumer always buys from a single store or a single chain of stores. In general this phenomenon of limited search of stores is either ignored or accounted for in using simple indicator variables of which stores were chosen in an initialization period (Bell and Lattin, 1998). The latter approach may be suitable when the number of stores is smaller but may not capture the dynamics of store consideration sets which change over time, a lacuna that can become more pronounced when a larger number of stores/store formats are available to the consumer. Our study combines the advantages of the hierarchical choice approach (Bell and Lattin, 1998) with the advantages of accounting for dynamic store consideration sets.

2.3. Availability of SKUs across stores and the SKU characteristics approach

Recent research has pointed out that there is considerable variation in the availability of skus in stores (Bell et al., 2005). For the empirical researcher this variation in sku availability poses an issue since the competitive set in a store will vary over time. This has led researchers to use average prices across skus for a brand, a method used in several notable papers in the literature on store competition such as Bell and Lattin (1998). However, the varying availability of skus constitutes information on sku level competition between stores that should be utilized in modeling store choice and store competition. In our study, we adopt the sku characteristics approach of Fader and Hardie (1996) to account for sku level competition, and account for the varying availability of skus across stores over time, two aspects that distinguish our study from the earlier literature on consideration sets.

3. Model formulation

A hierarchical model of salience and choice for both the store and brand choice stages is shown in Fig. 1. A customer will first choose a store from a set of salient stores (the ‘store consideration set’), then choose a brand from the set of salient brands (the ‘brand consideration set’). We present below a short derivation of the nested consideration model (for the full derivation please see the Supplementary Materials on the journal website).

[1] In the restricted brand choice set literature the brand state dependence variable has been used to parameterize the ‘brand salience’ construct (Bronnenberg and Vanhonacker, 1996) which has been shown to directly impact the formation of restricted brand choice sets. We utilize this approach to parameterize the store salience construct with the store state dependence variable.

Table 1
Contribution of this study.

	Bell and Lattin (1998)	Bell et al. (1998)	Bronnenberg and Vanhonacker (1996)	Mehta et al. (2003)	Popkowski Leszczyc et al. (2000, 2004)	Fader and Hardie (1996)	Campo et al. (2003)	This Article
Store state dependence	No	No	No	No	Yes	No	No	Yes
Preprocessed store consideration sets	Yes	Yes	No	No	No	No	No	Yes
Dynamic store salience	No	No	No	No	No	No	No	Yes
Brand consideration sets	No	No	Yes	Yes	No	No	No	Yes
Category effect on store choice (inclusive value)	Yes	Yes	No	No	No	No	No	Yes
Preprocessed brand/size consideration sets	Yes	Yes	No	No	No		No	Yes
Sku characteristics approach	No	No	No	No	No	Yes	No	Yes
Accounting for sku non-purchase	No	No	No	No	No No	No	Yes	Yes

3.1. Derivation of the nested consideration model

Assume that every individual 'h' has a specific latent threshold of salience $\Theta^h_{\Delta t}$ to consider a store ($s = [1,\ldots,N_S]$). Denote salience for store 's' for consumer 'h' at occasion 't' by Δ^h_{st}.

From the econometrician's perspective, store salience Δ^h_{st} and the store cut-off value $\Theta^h_{\Delta t}$ are observed indirectly with some error $\xi^h_{.t}$, as follows:

$$\begin{aligned}\Delta^h_{st} &= \delta^h_{st} + \xi^h_{st}, \quad s = [1,\ldots,N_S],\\ \Theta^h_t &= \theta^h_t + \xi^h_{\Theta,t}.\end{aligned} \tag{1}$$

Assume that the $(N_S + 1)$ random components $[\xi^h_{s=1,t},\ldots,\xi^h_{s=N_S,t}, \xi^h_{\Theta,t}]$ are independent and identically distributed (i.i.d.) draws from a Type 1 extreme value distribution,[2] where $\xi^h_{\Theta,t}$ is common across stores so that Θ^h_t is a common threshold across stores for a given consumer on a given purchase occasion.

Store 's' is in the store choice set of consumer 'h' at purchase occasion 't' if $\Delta^h_{st} > \Theta^h_t$. Under the distributional assumptions, the probability that consumer 'h' includes store s, ($s = 1,\ldots,N_S$) in *store choice set* M_S at occasion 't' equals

$$\pi(s \in M_S)^{hS}_t = \left[1 + \exp\left(\theta^h_t - \delta^h_{st}\right)\right]^{-1}, \quad s = [1,\ldots,N_S]. \tag{2}$$

Let $\pi_s = \pi(i_s \in M_S)^{hS}_t$.

Consumers maximize utility among stores that are in the choice set. If the choice set were known to the econometrician and if the store utility is $V^h_{st} = v^h_{st} + \varepsilon^h_{st}$, and ε^h_{st} are draws from a Type 1 extreme value distribution, then the choice probabilities for a store can be shown under the above assumptions to be (Bronnenberg and Vanhonacker, 1996):

$$P(s) = \frac{\pi_s * e^{v_s}}{\sum_{i'_s} \pi_{s'} * e^{V_{s'}}} \tag{3}$$

In this equation we drop the household and time subscripts. While we derived the above equation in the context of store choice, consider a similar equation characterizing brand choice as derived by Bronnenberg and Vanhonacker (1996)[3]:

$$P(b) = \frac{\pi_b * e^{V_b}}{\sum_{b'} \pi_{b'} * e^{V_{b'}}} \tag{4}$$

π_b, the probability that brand b will be *considered* is a function of the salience S_b of brand b, so we can write the consideration probability as $\pi_b(S_b)$. Let $V_{S_b,V_b} = \pi_b * e^{V_b}$, the numerator of this expression. Then taking logarithms we get $\ln(V_{S_b,V_b}) = \ln(\pi_b) + V_b$.

Suppose we have a *multidimensional choice set* C_N *whose elements are brand-store combinations.* Let $U_{S_{bs},V_{bs}}$ be the total utility of an element of C_N consisting of store s and brand b. If some elements share common observed elements, we can write the total utility as a function of salience and choice utility as follows:

$$\begin{aligned}U_{S_{bs},V_{bs}} &= f(S_{bS}, V_{bS}) + \varepsilon_{S_{bs},V_{bs}} \quad \forall(s,b) \in C_N,\\ U_{S_{bs},V_{bs}} &= \ln(V_{S_{bs},V_{bs}}) + \varepsilon_{S_{bs},V_{bs}} \quad \forall(s,b) \in C_N.\end{aligned} \tag{5}$$

Using the form for $V_{S_{bs},V_{bs}} = f(S_{bS}, V_{bS}) = \exp[\ln(\pi_{b,s}) + V_{b,s}]$, we write the total utility as

$$U_{S_{bs},V_{bs}} = \ln(\pi_{b,s}) + V_{b,s} + \varepsilon_{S_{bs},V_{bs}} \quad \forall(s,b) \in C_N. \tag{6}$$

The salience probability for a store-brand combination (s,b) is $\pi_{b,s} = \pi_s * \pi_{b|s}$, where $\pi_{b|s}$ is the brand salience conditional on store salience.

Let $\tilde{\varepsilon}_{S_s,V_s}$ and $\tilde{\varepsilon}_{S_b,V_b}$ be the unobserved components of the total utility attributable to the store and the brand, respectively, which vary only across stores (for $\tilde{\varepsilon}_{S_s,V_s}$) or brands (for $\tilde{\varepsilon}_{S_b,V_b}$). The term $\tilde{\varepsilon}_{S_{bs},V_{bs}}$ is the remaining unobserved component of the total utility. The utility expression is:

$$\begin{aligned}U_{S_{bs},V_{bs}} = {} & \ln(\pi_s) + \ln(\pi_{b|s}) + \widetilde{V}_b + \widetilde{V}_s + \widetilde{V}_{bs} + \tilde{\varepsilon}_{S_s,V_s} + \tilde{\varepsilon}_{S_b,V_b}\\ & + \tilde{\varepsilon}_{S_{bs},V_{bs}} \quad \forall(s,b) \in C_N.\end{aligned} \tag{7}$$

If S_n and B_n are the marginal store and brand choice sets, assume $\tilde{\varepsilon}_{S_s,V_s}$ and $\tilde{\varepsilon}_{S_{bs},V_{bs}}$ are independent for all $s \in S_n$ and $b \in B_n$.

Let the terms $\tilde{\varepsilon}_{S_{bs},V_{bs}}$ be independent and identically Type 1 extreme value distributed with scale parameter μ^b.

Let B_{ns} be the conditional brand choice set, which is the subset of brands in B_n for the person 'n' who visits store 's'. $\tilde{\varepsilon}_{S_s,V_s}$ is distributed so that $max_{b \in B_{ns}}(U_{S_{bs},V_{bs}})$ is Type 1 extreme value distributed with scale parameter μ^s.

We can write the marginal probability as

$$P_n(s) = Pr\begin{bmatrix}\ln(\pi_s) + \widetilde{V}_s + \tilde{\varepsilon}_{S_s,V_s} + V'_s + \varepsilon'_{S_s,V_s}\\ >\\ \ln(\pi_{s'}) + \widetilde{V}_{s'} + \tilde{\varepsilon}_{S_{s'},V_{s'}} + V'_{s'} + \varepsilon'_{S_{s'},V_{s'}}, \quad \forall s' \in S_n,\ s' \neq s\end{bmatrix} \tag{8}$$

where $V'_s = \eta = \frac{1}{\mu^b} ln\left[\sum_{b \in B_{ns}} e^{(\ln(\pi_{b|s}) + \widetilde{V}_b + \widetilde{V}_{bs})^* \mu^b}\right]$ and where the disturbance term ε'_{S_s,V_s} is defined by $\varepsilon'_{S_s,V_s} = max_{b \in B_{ns}}(\ln(\pi_{b|s}) + \widetilde{V}_b + \widetilde{V}_{bs} + \tilde{\varepsilon}_{S_{bs},V_{bs}}) - V'_s$ and is also Type 1 extreme value distributed with scale parameter $\mu = \mu^b$. Since $\tilde{\varepsilon}_{S_s,V_s} + \varepsilon'_{S_s,V_s}$ is i.i.d.

[2] Note that the Extreme Value assumption is made to get convenient closed form expressions for the salience probabilities, though other parametric forms could also be used to obtain salience probabilities.

[3] We use the Bronnenberg and Vanhonacker (1996) approach as a starting point for our model since this approach is a simple implementation of fuzzy consideration sets, and the 'fuzzy approach' (as opposed to modeling 'crisp' consideration sets) does not suffer from the curse of dimensionality encountered when enumerating all possible consideration sets (the 'crisp approach') (Wu and Rangaswamy, 2003). Further, this model also has been shown to perform well on prediction.

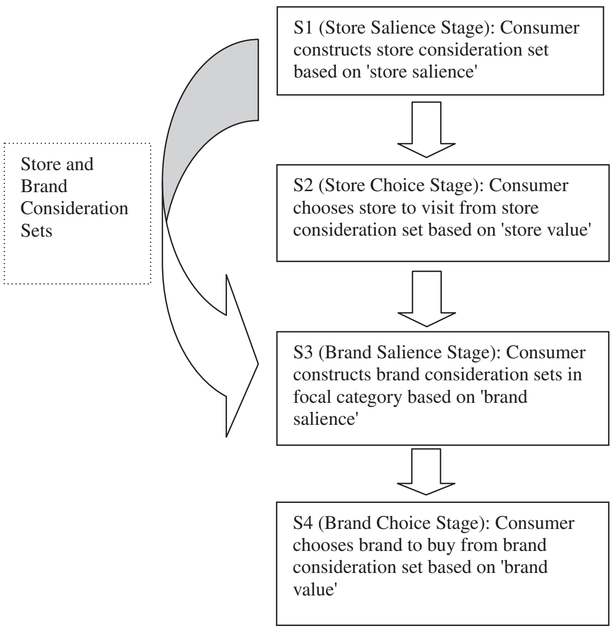

Fig. 1. Salience and choice stages for store and brand consumer choice.

Type 1 extreme value distributed with a scale parameter μ^s for all $s \in S_n$, the probability that store 's' is chosen is:

$$P_n(s) = \frac{e^{(\ln(\pi_s)+\widetilde{V}_s+V'_s)^* \mu^s}}{\sum_{s' \in S_n} e^{\left(\ln\left(\pi_{s'}\right)+\widetilde{V}_{s'}+V'_{s'}\right)^* \mu^s}}. \tag{9}$$

The probability that brand 's' is chosen conditional on store 's' being chosen can be derived as:

$$P_n(b|s) = \frac{e^{(\ln(\pi_{b|s})+\widetilde{V}_b+\widetilde{V}_{bs})^* \mu^b}}{\sum_{b' \in B_{ns}} e^{(\ln(\pi_{b'|s})+\widetilde{V}_{b'}+\widetilde{V}_{b's})^* \mu^b}}. \tag{10}$$

Combining the expressions for the marginal and conditional probabilities derived above, we get the expression for the joint probability of choice of the store brand combination (s,b) for the individual 'n' (see Supplementary Materials on the journal website for full derivation):

$$P_n(s,b) = \overbrace{\left[\frac{e^{(\ln(\pi_s)+\widetilde{V}_s+V'_s)^* \mu^s}}{\sum_{s' \in S_n} e^{(\ln(\pi_{s'})+\widetilde{V}_{s'}+V'_{s'})^* \mu^s}}\right]}^{P_n(s)} * \overbrace{\left[\frac{e^{(\ln(\pi_{b|s})+\widetilde{V}_b+\widetilde{V}_{bs})^* \mu^b}}{\sum_{b' \in B_{ns}} e^{\left(\ln\left(\pi_{b'|s}\right)+\widetilde{V}_{b'}+\widetilde{V}_{b's}\right)^* \mu^b}}\right]}^{P_n(b|s)}, \tag{11}$$

where

$$V'_s = \frac{1}{\mu^b} \ln\left[\sum_{b \in B_{ns}} e^{(\ln(\pi_{b|s})+\widetilde{V}_b+\widetilde{V}_{bs})^* \mu^b}\right]. \tag{12}$$

We term the above expression the *nested consideration model.* There are many similarities and some important differences between this model and the nested logit choice model. The similarities are in the model development and the way in which the second stage process depends on the 'inclusive value' V'_s provided by the choice alternatives in the lower nest. The differences are in the incorporation of dynamic salience effects into the choice processes in both the lower and the upper nests, through the salience probabilities π_s and $\pi_{b|s}$.

3.2. Two stage store and brand nested logit model with store loyalty

The baseline model that we will use is a nested logit model which is very similar to the model used in Bell and Lattin (1998). The equation that requires to be estimated for the nested logit model (Ben-Akiva and Lerman, 1985) is:

$$P_n(s,b) = \overbrace{\left[\frac{e^{(\widetilde{V}_s+V'_s)^* \mu^s}}{\sum_{s' \in S_n} e^{(\widetilde{V}_{s'}+V'_{s'})^* \mu^s}}\right]}^{P_n(s)} * \overbrace{\left[\frac{e^{(\widetilde{V}_b+\widetilde{V}_{bs})^* \mu^b}}{\sum_{b' \in B_{ns}} e^{(\widetilde{V}_{b'}+\widetilde{V}_{b's})^* \mu^b}}\right]}^{P_n(b|s)}, \tag{13}$$

where

$$V'_s = \frac{1}{\mu^b} \ln\left[\sum_{b \in B_{ns}} e^{(\widetilde{V}_b+\widetilde{V}_{bs})^* \mu^b}\right]. \tag{14}$$

Note that Eqs. 13 and 14 differ from Eqs. 11 and 12 in that the salience probabilities π_s and $\pi_{b|s}$ are not included in the two nested logit equations.

4. Data and empirical results

4.1. Data

We use data on store visits and on cola purchases from the Stanford Basket Market Database. We use the data from the suburb of a large US city, which covers a two year period from June 1991 to June 1993. There are 548 households in this database. We choose 140 households who bought cola three or more times in an initialization period. We use the first 24 weeks for initialization and the next 54 weeks for calibration. The calibration dataset for cola purchases for these 140 households consists of 2364 data points. The suburban market has 5 stores: stores 1 and 2 have an explicit EDLP (every day low price) positioning. Stores 3, 4, 5 are promotional pricing (also called HILO) stores. The top 16 skus which comprise more than 80% of cola sales were chosen for analysis. These 16 skus correspond to 5 discrete sizes and 5 different brands. Table 2 presents some descriptives on market share and availability for these 16 skus.

4.2. Comparison of variables used in the models

Table 3 lists the variables used in the four models being studied in this article. A detailed description of how each variable is constructed is provided in the Supplementary Materials on the journal website. The nested consideration model, termed NCDC ('Nested Consideration Dynamic Consideration') model is tested against three nested logit models, termed the NLSC ('Nested Logit Static Consideration'), NLSL ('Nested Logit Store Loyalty') and NLSR ('Nested Logit Store Recency') models.

As the table indicates one crucial condition for comparing predictive ability of different models is satisfied by the variables used, i.e., that the same number of independent variables be used in all the models. In addition the set of variables used in the NCDC model and NLSR model are *identical.* The NLSC model uses 'initially considered' stores variable representing the 'preprocessed consideration sets used by Bell and Lattin (1998) instead of 'store recency'. The NLSL model uses the store loyalty variable, which has also been used by Bell and Lattin (1998) instead of 'store recency'. The NCDC and NLSR models allow a 'head-to-head' comparison of prediction between the nested consideration and nested logit models since the same variables are used in estimating both models.[4]

[4] The two additional salience thresholds ('skusalience threhold' and 'store salience threshold') represent two additional parameters to be estimated in the NCDC model as compared to the nested logit models; the number of variables used in the NCDC model is however the same as in the nested logit models. All four models will also estimate the 'inclusive value coefficient'.

Table 2
Descriptives

sku	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
brand	Pepsi	Coke	RC	Wildwood	Canfield	Pepsi	Coke	RC	Pepsi	Coke	RC	Pepsi	RC	Pepsi	Coke	RC
size	67.6 oz	67.6 oz	67.6 oz	67.6 oz	67.6 oz	288 oz	288 oz	288 oz	144 oz	144 oz	144 oz	101.4 oz	101.4 oz	72 oz	72 oz	72 oz
share	12%	9%	13%	4%	2%	13%	11%	8%	5%	5%	3%	3%	6%	3%	2%	2%
Availability (across stores)	100%	100%	99%	39%	60%	60%	61%	58%	68%	82%	78%	44%	60%	65%	63%	61%

Table 3
Variables used in the NLSL and NCSM models.

	NCDC	NLSC	NLSL	NLSR
288 oz Intercept	Yes	Yes	Yes	Yes
144 oz Intercept	Yes	Yes	Yes	Yes
101.4 oz Intercept	Yes	Yes	Yes	Yes
72 oz Intercept	Yes	Yes	Yes	Yes
Initial size loyalty	Yes	Yes	Yes	Yes
Coke intercept	Yes	Yes	Yes	Yes
RC intercept	Yes	Yes	Yes	Yes
Wildwood intercept	Yes	Yes	Yes	Yes
Canfield intercept	Yes	Yes	Yes	Yes
Initial brand loyalty	Yes	Yes	Yes	Yes
Price	Yes	Yes	Yes	Yes
Feature	Yes	Yes	Yes	Yes
Lobby display	Yes	Yes	Yes	Yes
Aisle display	Yes	Yes	Yes	Yes
SKU recency	Yes	Yes	Yes	Yes
Store 1 intercept	Yes	Yes	Yes	Yes
Store 2 intercept	Yes	Yes	Yes	Yes
Store 3 intercept	Yes	Yes	Yes	Yes
Store 4 intercept	Yes	Yes	Yes	Yes
Distance	Yes	Yes	Yes	Yes
Initially considered store	No	Yes	No	No
Store loyalty	No	No	Yes	No
Store recency	Yes	No	No	Yes
Total number of independent variables	21	21	21	21

4.3. Estimation

We estimated the following model specifications: (1) Nested consideration model with dynamic store consideration sets (NCDC), in which choice set formation is modeled at both the store choice stage and the brand (sku) choice stage. (2) The benchmark nested logit model of store and brand (sku) choice with the three store-level parameterizations of 'static store consideration sets' (NLSC), store loyalty (NLSL) and 'store recency' (NLSR).

We obtained maximum likelihood estimates using a GAUSS-based steepest gradient search with the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm. Parameters of the NCSM model are identified up to a metric for the size constants and brand constants in the sku choice stage and identified up to a metric for the store constants in the store choice stage. We therefore set the corresponding parameters of one size, one brand and one store to zero. The threshold of sku salience does not need to be operationalized and is estimated latently, and the dynamic aspect of store salience being operationalized by the store recency variable.

4.3.1. Model fit

We present in Table 4 the results of the three nested logit models (NLSC, NLSL and NLSR), and in Table 5 the results of the nested consideration model (NCDC), using the cola category as an example. As the log likelihoods indicate, the nested consideration model outperforms all three nested logit model, in spite of the use of a similar set of explanatory variables. *In fact, in the case of NLSR, the nested logit model uses exactly the same set of variables as the nested consideration model, and the NCDC outperforms the NLSR model* in terms of both the log likelihood, and the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). The latter two information criteria penalize the NCDC model for the estimation of the two additional salience parameters ('store salience threshold' and 'sku salience threshold'), and thus the NCDC model is superior to the equivalent nested logit model (NLSR) on all predictive criteria. We also compare the NCDC model to two other commonly used nested logit models of store choice. First, we estimate the NLSL model with the store loyalty variable instead of store recency. Second, we estimate the NCSC model where we model 'static (or 'preprocessed') consideration' at the store choice stage using a variable which takes the value 1 for a store if the store were visited in the initialization period, 0 if it were not visited. The nested consideration model outperforms both these nested logit models. Next we examine some of the coefficients of these models and their implications.

Distance to store will also affect store choice, and we use this variable to explain store choice, as in Bell and Lattin (1998). Reflecting the higher travel cost of customers who live far away from a retail store, the distance coefficient is negative and significant for all the four models. The effect of price of skus also appears to be reduced (price coefficient of −1.206) in the nested consideration model when compared to the nested logit model (price coefficient of about −4.2). Brand and size loyalty positively affect 'brand value', as do feature and lobby and aisle displays and 'sku recency'. The three variables explaining store choice that are being studied in Table 4, i.e., store loyalty, store recency and 'preprocessed' store consideration sets, all positively affect probability of store choice.

4.3.2. Store and SKU consideration sets – post hoc analysis

The advantage of modeling store consideration sets using our approach as opposed to the 'preprocessed set' approach of Bell and Lattin (1998) is shown in Fig. 2. This figure plots the dynamic store consideration set probabilities for one household (coded '22') which has 19 weeks of purchase data in the dataset. In the calibration period this household had bought from both store 1 and store 2, and hence its 'preprocessed store consideration set would be [1 1 0 0 0]. Using this variable would only capture heterogeneity in the preprocessing across households. However this set will not vary over time for a specific household. In contrast to this, the dynamic consideration set approach that we use causes store consideration set probabilities to change over the 19 week period based on recency of the store visit. For example, store 1's consideration probability is initially high (0.94), dips, then recovers to 0.937 in week 3, and then drops steadily till week 10, reflecting how recent store visits have made the other store (store 2) more salient to household 22 in this period. The ability of our model to capture dynamic store consideration sets is a key reason why our model predicts better than the popular Bell and Lattin (1998) approach.

Fig. 3 demonstrates an important property of the model, that the brand-size consideration probability depends upon the store in which the sku is stocked. This figure shows that the sku level salience (or brand-size consideration probability) for the Pepsi 67.6 oz sku starts low (0.18) in both store 1 and store 2. In weeks 7 and 15

Table 4
Parameter estimates for static consideration and store loyalty nested logit models.

	Nested Logit Static Consideration (NLSC) model		Nested Logit Store Loyalty (NLSL) model		Nested Logit Store Recency (NLSR) model	
−log likelihood	4903.466		4773.867		4787.399	
BIC	9977.831		9718.632		9745.696	
AIC	9850.933		9591.734		9618.797	
	Coefficient	*t*-Value	Coefficient	*t*-Value	Coefficient	*t*-Value
288 oz Intercept	0.104	1.231	0.111	1.323	0.110	1.304
144 oz Intercept	0.350	3.415	0.333	3.258	0.340	3.317
101.4 oz Intercept	0.579	6.004	0.589	6.129	0.582	5.985
72 oz Intercept	1.065	7.611	1.024	7.332	1.081	7.681
Initial size loyalty	2.014	22.496	2.003	22.523	2.039	22.640
Coke intercept	0.162	2.378	0.165	2.431	0.166	2.435
RC intercept	−0.074	−1.160	−0.074	−1.159	−0.071	−1.123
Wildwood intercept	−0.043	−0.298	−0.019	−0.128	−0.024	−0.165
Canfield intercept	−0.607	−3.140	−0.599	−3.088	−0.595	−3.064
Initial brand loyalty	2.158	23.360	2.172	23.632	2.183	23.551
Price	−4.216	−7.014	−4.067	−6.780	−4.210	−6.967
Feature	0.564	9.032	0.606	9.702	0.579	9.177
Lobby display	1.362	16.541	1.309	15.619	1.338	15.996
Aisle display	0.735	10.195	0.717	9.996	0.713	9.744
SKU recency	1.459	22.197	1.452	22.004	1.436	21.791
Store 1 intercept	2.335	4.485	2.085	4.159	1.424	3.259
Store 2 intercept	3.110	5.100	2.549	4.645	2.095	4.092
Store 3 intercept	2.530	4.504	1.306	3.056	1.067	2.620
Store 4 intercept	−0.713	−1.566	−0.155	−0.375	−0.240	−0.573
Distance	−0.352	−5.120	−0.257	−4.066	−0.469	−5.236
Initially considered store	6.366	6.430	Not used	Not used	Not used	Not used
Store loyalty	Not used	Not used	4.968	6.772	Not used	Not used
Store recency	Not used	Not used	Not used	Not used	4.775	6.344
ln(inclusive value coefficient)	−0.413	−2.791	−0.303	−2.075	−0.335	−2.171

however, its probability in store 2 increases to a high value (0.88) due to the marketing activity for this sku in store 2 in those weeks. However the probability in store 1 never exceeds 0.38 throughout the period. This demonstrates the flexibility of the nested consideration model in accommodating varying sku consideration probabilities across stores.[5]

5. Simulations

In this section we implement two simulations: (1) pricing simulation to demonstrate the differences between the models in terms of pricing strategies and (2) store closing simulation to demonstrate how the results of the nested consideration model differ from the other models in studying overall store competition. For these simulations we compare the nested consideration (NCDC) model with the two nested logit models NLSC and NLSL.

5.1. Pricing simulation

Since there are 16 cola skus in the study there could be many possible price promotion strategies. However our objective is to study the effect of our model on store-level strategies, and we therefore study the effect of price cuts for all the skus stocked by a store. Fig. 4 shows the effect of price cuts by store 3 for the unit sales of all the skus stocked by that store. To implement this simulation, we compute the probability of choice of each sku stocked by store 3 at the actual prices observed in the market. Let this be Prob_NLSC, for the nested logit model with static consideration sets. Then we compute the probability of choice with a price cut of *x*% (2%, 4%, etc.) for all the skus stocked by store 3. Let this be Prob_NLSC_x, for the nested logit model with static consideration sets. Then the unit sales increase can be computed using the formula:

$$\Delta^{NLSC}_{unitsales} = \sum_{t=1}^{T}\sum_{j=1}^{16}(\text{Prob\_NLSC\_x} - \text{Prob\_NLSC}) * M, \quad (15)$$

where '*M*' is the marketsize, and '*T*' the total number of weeks in the dataset. In our application we calculate the market size as the product of the number of customers (140 in our dataset), and the average number of colas consumed in a week (5.3, as per Mintel Reports on US soda consumption, 2009), to yield a total marketsize of 742 consumption occasions per week.

Similarly the unit sales increase for the other two models is obtained as:

$$\Delta^{NCSR}_{unitsales} = \sum_{t=1}^{T}\sum_{j=1}^{16}(\text{Prob\_NCSR\_x} - \text{Prob\_NCSR}) * M, \quad (16)$$

$$\Delta^{NLSL}_{unitsales} = \sum_{t=1}^{T}\sum_{j=1}^{16}(\text{Prob\_NLSL\_x} - \text{Prob\_NLSL}) * M. \quad (17)$$

As expected, we see from Fig. 4 that unit sales increase with the price cut for all the models. However the magnitude of the increase is much lower in the nested consideration model than in either of the nested logit models (NLSC or NLSL) being studied. This figure highlights the erroneous inferences about incremental sales in the absence of a model of dynamic store consideration sets. This error can be substantial, e.g., for a 4% price cut at store 3, the sales increase from the nested logit models is more than five times that from the nested consideration model. Since in the hierarchical model the effect of the price cut on store choice will be through the 'inclusive value' (see Eqs. (5)–(8)), we present in Fig. 5 the ratio of inclusive values with and without the price cut for the three different models being studied. The trend is similar to the incremental unit sales that we observe in Fig. 4.[6]

[5] We thank an anonymous reviewer's comments about 'joint salience' between store and sku consideration probabilities that led to this analysis.

[6] See Supplementary Materials on the journal website for the Pricing Simulations with dollar sales.

Table 5
Parameter estimates for nested consideration model.

	Nested consideration dynamic consideration (NCDC) Model	
−log likelihood	4732.779	
BIC	9651.992	
AIC	9513.558	
	Coefficient	*t*-Value
Sku salience threshold	1.478	17.176
Sku recency	3.775	9.205
Lobby display	3.340	6.364
Aisle display	1.008	6.483
288 oz	0.206	2.791
144 oz	0.250	2.710
101.4 oz	0.850	8.935
72 oz	0.890	7.021
Initial size loyalty	2.266	28.447
Coke	0.295	4.590
RC	0.149	2.634
Wildwood	0.328	2.280
Canfield	−0.278	−1.437
Initial brand loyalty	2.420	28.986
Price	−1.206	−2.817
Feature	0.999	18.524
Store salience threshold	4.712	29.731
Store recency (store salience stage)	7.441	17.589
Store 1 intercept	1.587	7.661
Store 2 intercept	2.253	8.658
Store 3 intercept	1.422	6.297
Store 4 intercept	−0.998	−1.571
Distance	−0.293	−3.136
ln(inclusive value coefficient)	−0.538	−3.142

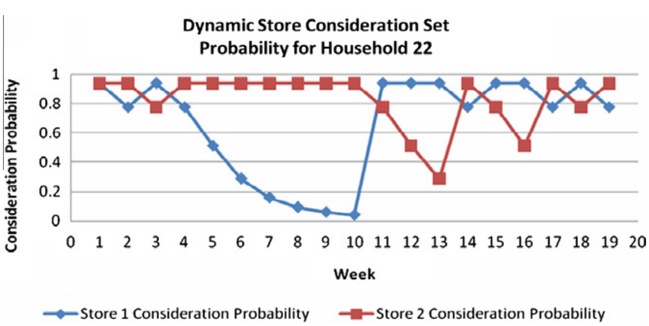

Fig. 2. Dynamic store consideration set probability.

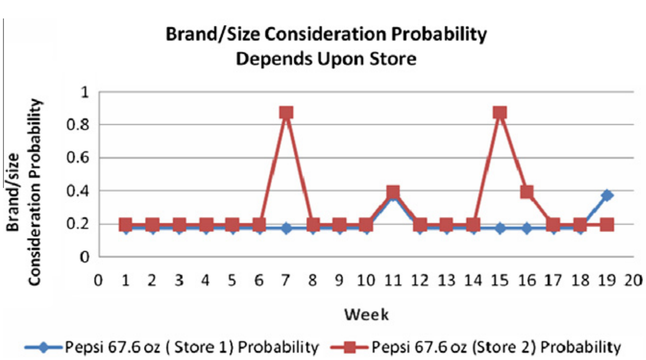

Fig. 3. Store-specific brand/size consideration probability.

5.2. Store closing simulation

We perform a second simulation to demonstrate the importance of the nested consideration model in studying the effect of

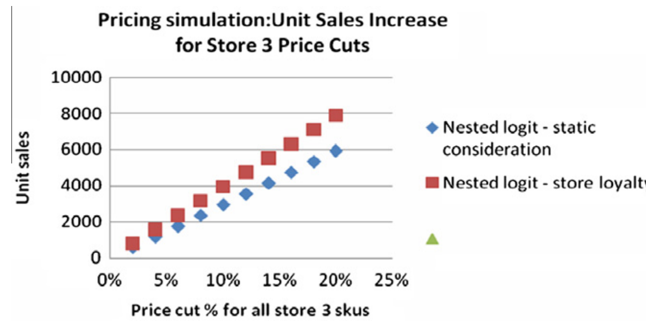

Fig. 4. Pricing simulation: unit sales increase.

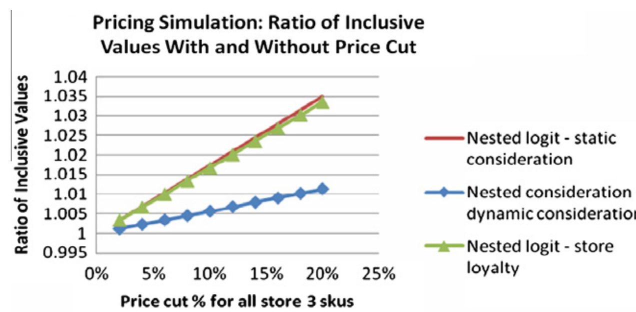

Fig. 5. Pricing simulation: ratio of inclusive values.

store competition. In this simulation we study the effect of the possible closure of one store (say, store 2) on the sales of the other stores. Unlike the earlier pricing simulation where we reported the effect of store 3's price cuts on its own sales, here we will report the effect of closing one of the stores (store 2) on the sales of all the other stores. We accomplish this simulation by removing all the skus for store 2 from the share equations for each of the models, and recomputing the market shares (and therefore sales) for the other four stores.

Table 6 presents the results of the store closing simulation for the nested consideration model and the two nested logit models. The two stores that are most affected by the closure of store 2 (which is an Every Day Low Price, or EDLP store) are store 1 (another EDLP store) and store 3 (a HILO store). Using the nested logit models will lead to higher prediction of sales for store 1 and a lower prediction of sales for store 3 than would be the case using the nested consideration model. Since the nested consideration model predicts the data better than all the nested logit models being studied, this points to the erroneous inferences about the nature of store competition that would result if the nested consideration model is not used in modeling market shares.

Table 6
Simulation results – incremental sales due to closure of store 2 (EDLP).

	NCDC	NLSL	NLSC
Incremental unit sales (oz)			
Store 1 (EDLP)	440,214	534,469	566,055
Store 3 (HILO)	214,054	188,542	176,807
Store 4 (HILO)	21,533	19,162	7,626
Store 5 (HILO)	54,341	35,017	25,921
Incremental dollar sales ($)			
Store 1 (EDLP)	79,732	95,417	99,540
Store 3 (HILO)	43,512	37,530	35,375
Store 4 (HILO)	4743	4219	1706
Store 5 (HILO)	12,960	8303	6094

6. Conclusion

In this article we develop a model that incorporates restricted store and brand choice set formation by consumers, as well as the relative effects of location, promotions and product assortment. We term the model the 'nested consideration' model, and derive the related probabilities in a manner analogous to the well-known nested logit model. This model also brings together two streams of marketing literature which study consideration effects, the more developed literature on restricted sku/brand choice sets, and the comparatively less studied literature on restricted store choice sets. Also uniquely among the literature, this model includes dynamic store salience (or store consideration sets), while controlling for state dependence, availability of skus in the store and brand and size loyalty. In empirical calibration, we find that the nested consideration model performs better than several commonly used nested logit models of store and brand choice.

We find that not taking the consideration stages into account overestimates the effect of price. This can lead to suboptimal pricing decisions due to erroneous inferences about incremental sales, as illustrated in a simulation. We also find in another simulation that not accounting for dynamic store consideration sets leads to erroneous conclusions about the magnitude of store competition. In our empirical illustration, we find that this error will be in the direction of over-predicting the effect on the same format (EDLP) store, and under-predicting the effect on the store with a different format (HILO).

It is important for retail managers take into account the multi-stage formation of store and brand consideration sets in decision-making, and the nested consideration model thus offers a quantitative tool that can be used for important retailing decisions such as pricing and retail store location.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ejor.2011.04.027.

References

Baltas, G., 2004. A model for multiple brand choice. European Journal of Operational Research 154 (1), 144–149.

Bell, D.R., Ho, T.H., Tang, C.S., 1998. Determining where to shop: Fixed and variable costs of shopping. Journal of Marketing Research 35 (3), 352–369.

Bell, D.R., Lattin, J., 1998. Shopping behavior and consumer preference for store price format: Why large basket shoppers prefer EDLP. Marketing Science 17 (1), 66–88.

Bell, D.R., Bonfrer, A., Chintagunta, P.K., 2005. Recovering stockkeeping-unit-level preferences and response sensitivities from market share models estimated on item aggregates. Journal of Marketing Research 42 (2), 169–182.

Bekhor, S., Ben-Akiva, M.E., Ramming, M.S., 2006. Evaluation of choice set generation algorithms for route choice models. Annals of Operations Research 144 (1), 235–247.

Ben-Akiva, M., Lerman, S., 1985. Discrete Choice Analysis: Theory and Application to Travel Demand. MIT Press, Cambridge, MA.

Bierlaire, M., 2004. A theoretical analysis of the cross-nested logit model. Annals of Operations Research 144 (1), 287–300.

Bronnenberg, B.J., Vanhonacker, W.R., 1996. Limited choice sets, local price response, and implied measures of price competition. Journal of Marketing Research 33 (2), 163–173.

Campo, K., Gijsbrechts, E., Nisol, P., 2003. The impact of retailer stockouts on whether, how much, and what to buy. International Journal of Research in Marketing 20, 273–286.

Fader, P.S., Hardie, B.G.S., 1996. Modeling consumer choice among SKUs. Journal of Marketing Research 33 (4), 442–452.

Fotheringham, A.S., 1988. Consumer store choice and choice set definition. Marketing Science 7 (3), 299–310.

García-Ródenas, R., Ángel Marín, 2009. Simultaneous estimation of the origin-destination matrices and the parameters of a nested logit model in a combined network equilibrium model. European Journal of Operational Research 197 (1), 320–331.

Kalouptsidis, N., Koutroumbas, K., Psarakic, V., 2007. Classification methods for random utility models with i.i.d. disturbances under the most probable alternative rule. European Journal of Operational Research 176 (3), 1778–1794.

Kannan, P.K., Wright, G.P., 1991. Modeling and testing structured markets: A nested logit approach. Marketing Science 10 (1), 58–82.

Kiley, D., 2009. Car dealers are fighting the Hammer as Chrysler and GM send out pink slips to retailers. Businessweek, May 13, 2009.

Kiviat, B., 2008. Starbucks Goes From Venti to Grande. Time, July 2, 2008.

Mehta, N., Rajiv, S., Srinivasan, K., 2002. Price uncertainty and consumer search: A structural model of consideration set formation. Marketing Science 22 (1), 58–85.

Popkowski Leszczyc, P.T.L., Sinha, A., Timmermans, H.J.P., 2000. Consumer store choice dynamics: An analysis of the competitive market structure for grocery stores. Journal of Retailing 76 (3), 323–345.

Popkowski Leszczyc, P.T.L., Sinha, A., Sahgal, A., 2004. The effect of multi-purpose shopping on pricing and location strategy for grocery stores. Journal of Retailing 80 (2), 85–99.

Pancras, J., 2010. A framework to determine the value of consumer consideration set information for firm pricing strategies. Computational Economics 35 (3), 269–300.

Seetharaman, P.B., Ainslie, A., Chintagunta, P.K., 1999. Investigating household state dependence effects across categories. Journal of Marketing Research 36 (4), 488–500.

Schön, C., 2010. Optimal dynamic price selection under attraction choice models. European Journal of Operational Research 205 (3), 650–660.

Train, K., 2003. Discrete Choice Methods with Simulation. Cambridge University Press, 2003.

Wu, J., Rangaswamy, A., 2003. A fuzzy set model of consideration set formation calibrated on data from an online supermarket. Marketing Science 22 (3), 411–434.