

**INCORPORATING RESPONSIVENESS TO MARKETING EFFORTS WHEN
MODELING BRAND CHOICE**

DENNIS FOK, PHILIP HANS FRANSES, RICHARD PAAP

ERIM REPORT SERIES <i>RESEARCH IN MANAGEMENT</i>	
ERIM Report Series reference number	ERS-2001-47-MKT
Publication	August 2001
Number of pages	23
Email address corresponding author	dfok@few.eur.nl
Address	Erasmus Research Institute of Management (ERIM) Faculty of Economics Erasmus Universiteit Rotterdam P.O. Box 1738 3000 DR Rotterdam, The Netherlands Phone: +31 10 408 1182 Fax: +31 10 408 9640 Email: info@erim.eur.nl Internet: www.erim.eur.nl

Bibliographic data and classifications of all the ERIM reports are also available on the ERIM website:
www.erim.eur.nl

ERASMUS RESEARCH INSTITUTE OF MANAGEMENT

REPORT SERIES *RESEARCH IN MANAGEMENT*

BIBLIOGRAPHIC DATA AND CLASSIFICATIONS		
Abstract	<p>In this paper we put forward a brand choice model which incorporates responsiveness to marketing efforts as a form of structural heterogeneity. We introduce two latent segments of households. The households in the first segment are assumed to respond to marketing efforts while households in the second segment do not do so. Whether a specific household is a member of the first or the second segment at a specific purchase occasion is described by household-specific characteristics and characteristics concerning buying behavior. Households may switch between responsiveness states over time.</p> <p>We compare the in- and out-of-sample performance of our model with various versions of the MNL model. We conclude that, while using the smallest amount of parameters, our model outperforms all MNL variants on forecasting. This, together with the face validity of our parameter results, leads us to believe that incorporating responsiveness seems to be a worthwhile exercise.</p>	
Library of Congress Classification (LCC)	5001-6182	Business
	5410-5417.5	Marketing
	HF 6161.B4	Branding
Journal of Economic Literature (JEL)	M	Business Administration and Business Economics
	M 31	Marketing
	C 44	Statistical Decision Theory
	M 31 C 35	Marketing Discrete Regression and Qualitative Choice Models
European Business Schools Library Group (EBSLG)	85 A	Business General
	280 G	Managing the marketing function
	255 A	Decision theory (general)
	280 K	Marketing Science (Quantitative analysis)
Gemeenschappelijke Onderwerpsontsluiting (GOO)		
Classification GOO	85.00	Bedrijfskunde, Organiseatiekunde: algemeen
	85.40	Marketing
	85.03	Methoden en technieken, operations research
	85.40	Marketing
Keywords GOO	Bedrijfskunde / Bedrijfseconomie	
	Marketing / Besliskunde	
	Huishoudens, Merken, Koopgedrag, Econometrische modellen	
Free keywords	Marketing-instrument effectiveness, structural heterogeneity, state dependence, multinomial logit, mixtures	

Incorporating Responsiveness to Marketing Efforts when Modeling Brand Choice

Dennis Fok*

*Erasmus Research Institute of Management
Erasmus University Rotterdam*

Philip Hans Franses

*Econometric Institute and
Department of Marketing and Organization
Erasmus University Rotterdam*

Richard Paap

*Rotterdam Institute for Business Economic Studies
Erasmus University Rotterdam*

August 13, 2001

*We thank Michel Wedel for sharing useful ideas. The third author thanks the Netherlands Organization for Scientific Research (N.W.O.) for its financial support. The algorithms used in this paper are implemented in Ox 2.20 (Doornik, 1999). Address for correspondence: D. Fok, Erasmus University Rotterdam, Faculty of Economics, H16-19, P.O. Box 1738, NL-3000 DR Rotterdam, The Netherlands, e-mail: dfok@few.eur.nl.

Incorporating Responsiveness to Marketing Efforts when Modeling Brand Choice

Abstract

In this paper we put forward a brand choice model which incorporates responsiveness to marketing efforts as a form of structural heterogeneity. We introduce two latent segments of households. The households in the first segment are assumed to respond to marketing efforts while households in the second segment do not do so. Whether a specific household is a member of the first or the second segment at a specific purchase occasion is described by household-specific characteristics and characteristics concerning buying behavior. Households may switch between responsiveness states over time.

We compare the in- and out-of-sample performance of our model with various versions of the MNL model. We conclude that, while using the smallest amount of parameters, our model outperforms all MNL variants on forecasting. This, together with the face validity of our parameter results, leads us to believe that incorporating responsiveness seems to be a worthwhile exercise.

Key words: marketing-instrument effectiveness, structural heterogeneity, state dependence, multinomial logit, mixtures

1 Introduction

The application of brand choice models has become standard practice in marketing research (Guadagni and Little, 1983; Chintagunta *et al.*, 1991; Jain *et al.*, 1994; Keane, 1997). In many practical applications of these choice models, the random utility theory framework (McFadden, 1973, 1981) is used to represent the choice process. An often made assumption in these models concerns homogeneity of households¹. That is, it is often assumed that all households have similar tastes, and that they only differ in their (observed) characteristics. In the relevant literature there is however ample evidence that households do differ. They may differ in their preferences and/or in the way they make their decisions. Differences in base preferences are usually referred to as preference heterogeneity. Differences in decision processes are labeled as structural heterogeneity.

Preference heterogeneity can partly be explained by observable characteristics. This corresponds with so-called observed preference heterogeneity. Taste is in this case usually explicitly modeled, for example by including demographic variables (see e.g. Maddala, 1983). However, it may be that not all heterogeneity can be attributed to observed characteristics, and hence there might be so-called unobserved preference heterogeneity. There are two popular techniques to deal with unobserved preference heterogeneity, see Allenby and Rossi (1999) and Wedel *et al.* (1999) for a discussion. These are based on the notion that when there is unobserved heterogeneity in taste, there is a corresponding distribution in the population. One approach imposes a continuous distribution of a known form to capture the heterogeneity. The other approach tries to approximate a distribution of unknown form using a discrete distribution with a fixed number of probability masses. A choice model using the latter approach is an example of a mixture model, see for example Wedel and Kamakura (1999).

Differences across households may not be fully attributable to preference heterogeneity. Households may also differ in the actual decision process they use to make their choices, that is, there might be structural heterogeneity. For example, Kamakura *et al.* (1996) examine brand choice within a product category where the brands carry different product forms, like volume. A household might first choose a brand and then choose the specific form to purchase. Another household might first choose a specific product form and only then consider the different brands. A third household might completely ignore all this

¹As our illustration in Section 4 concerns households as the unit of analysis, we refer to the household as the decision maker throughout this paper.

and choose directly from all available brand and product form combinations. Kamakura *et al.* (1996) develop a model which combines preference and structural heterogeneity and they show that the inclusion of both types of heterogeneity leads to useful managerial insights for brand competition.

Structural heterogeneity is not only relevant to sequential choice processes as choice mechanisms of households can differ in many respects. Yang and Allenby (2000), for example, present a model in which households are allowed to differ in the reference point to which options are compared. These authors use a hierarchical Bayes model to model credit card adoption, where households are allowed to differ in their decision rule and where behavior can change over time. Yang and Allenby (2000) show that there is heterogeneity in decision rules and that it can be modeled using a mixture of sub-models.

In the present paper we extend their idea to a brand choice setting. Households, who choose amongst brands within a specific product category, may differ in their decision rules. For example, some households will spend more time and effort while making their choice than others do. If little time and effort is invested in this decision process, it is perhaps less likely that the household will respond to marketing instruments. For example, to be able to respond to price changes, one needs of course to recall the previous price of all brands. To be able to respond to advertising, one has to read for example the newspaper in which the advertisement is printed. It may be unrealistic to assume that all households are that much involved with the product category on all purchase occasions. Hence, it is likely that households will differ in the extent to which they are responsive to marketing efforts. These differences correspond to differences in the decision rules being used, and therefore they can be seen as an example of structural heterogeneity.

One reason why some households are unresponsive to marketing efforts could be just a lack of interest in marketing efforts issued by brand managers. On the other hand, economic motivations may also explain varying responsiveness across households. For example, search costs play an important role in the decision process of a household or an individual. As mentioned before, to be responsive to price changes one needs to remember the prices of every option on every purchase occasion. Additionally, people usually face time constraints. It takes time for a household to compare all prices of the options in a specific market at the time of purchase. Consider a household planning to buy many different items during the same shopping trip. There is obviously a limited amount of time available for this and therefore it may be unrealistic to assume that the household will allocate much time to each item. Following this reasoning, the more items a household

purchases at a shopping trip, the less responsive this household might be to marketing efforts. Hence, the monetary value of all products purchased at a shopping trip may be inversely related to the responsiveness to marketing efforts.

Taking the above arguments along, as the decision process differs across households and across purchase occasions, such structural heterogeneity implies that the observed choice of different households cannot be explained by the same variables. Choice behavior of responsive households can be explained by their base preferences, by marketing efforts, and by their purchase history. Brand choice by unresponsive households may only be described by base preferences and purchase history. Moreover, household characteristics are rarely seen to significantly contribute to explaining brand choice, but these might be informative for the type of decision process. As such, household characteristics might indirectly influence brand choice.

In this paper we put forward a brand choice model which incorporates responsiveness to marketing efforts as a form of structural heterogeneity. We assume that all brands are equal in appearance so that the source of structural heterogeneity studied by Kamakura *et al.* (1996) is not present, although our model can be extended to such a setting. We introduce two latent segments of households. The households in the first segment are assumed to respond to marketing efforts while households in the second segment are assumed not to do so. Whether a specific household is a member of the first or the second segment at a specific purchase occasion is described by household-specific characteristics and characteristics concerning buying behavior. Additionally, households are allowed to switch between the two segments over time.

The remainder of this paper contains the following. In Section 2, we discuss our model. We first introduce the basic model. Next, we extend the model to also capture preference heterogeneity. In Section 3, we discuss parameter estimation. Section 4 discusses the application of this model to panel data concerning purchases of liquid detergent, where we also compare the performance of our model to various related choice models. In Section 5, we conclude this paper with some remarks.

2 Model

In this section we first discuss the basics of our choice model, and then we discuss various extensions.

2.1 Preliminaries

To keep the presentation of the basic model simple, we will first ignore possible preference heterogeneity and only concentrate on structural heterogeneity. We assume that household $i = 1, \dots, I$ purchases brand $j = 1, \dots, J$, at occasion $t = 1, \dots, T_i$. Note that different households can have a different number of observed purchase occasions. Also note that purchase occasion t of household i not necessarily corresponds to the same period in time as purchase occasion t of household $l \neq i$. The variable $y_{i,j,t}$ denotes the chosen alternative, that is,

$$y_{i,j,t} = \begin{cases} 1 & \text{household } i \text{ purchases brand } j \text{ at occasion } t \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

Each household is, at any point in time, either responsive or unresponsive to marketing efforts. In case a household is unresponsive to marketing efforts, the choice can only be attributed to base preference, habit, state dependence or random influences. We introduce an indicator variable $Z_{i,t}$ to denote the state of a household at a specific point in time, that is,

$$Z_{i,t} = \begin{cases} 1 & \text{household } i \text{ is responsive to marketing efforts at purchase occasion } t \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

Note that we do not observe the values of $Z_{i,t}$, and hence that these have to be inferred from the data. To model the responsiveness, we consider a logit model (Maddala, 1983) which relates $Z_{i,t}$ to household characteristics collected in $W_{i,t}$. These characteristics may also include variables describing the general purchase behavior of the household. Examples of such variables are recency of the last purchase and the monetary amount spent on the shopping trip. The specification of the model for responsiveness becomes

$$\begin{aligned} Z_{i,t}^* &= \mu^{(z)} + W_{i,t}\gamma^{(z)} + \varepsilon_{i,t}^{(z)} \\ Z_{i,t} &= \begin{cases} 1 & \text{if } Z_{i,t}^* \geq 0 \\ 0 & \text{if } Z_{i,t}^* < 0. \end{cases} \end{aligned} \quad (3)$$

The disturbances $\varepsilon_{i,t}^{(z)}$ are assumed to follow a logistic distribution, such that,

$$P(Z_{i,t} = 1) = \frac{\exp(\mu^{(z)} + W_{i,t}\gamma^{(z)})}{1 + \exp(\mu^{(z)} + W_{i,t}\gamma^{(z)})}. \quad (4)$$

In case a household is responsive to marketing efforts, such as price and promotion, marketing instruments might have an effect on the choice made by this household. We denote the marketing instruments for brand $j = 1, \dots, J$, as experienced by household i at purchase occasion t , as $X_{i,j,t}$. To model the choice process of a marketing-responsive household we consider the Multinomial Logit Model [MNL] of McFadden (1973). Conditional on responsiveness, the utility of brand j for household i at purchase occasion t is denoted by $U_{i,j,t}^{(r)}$, and it is modeled as

$$U_{i,j,t}^{(r)} = \mu_j^{(r)} + X_{i,j,t}\beta^{(r)} + \alpha^{(r)}y_{i,j,t-1} + \varepsilon_{i,j,t}^{(r)}, \quad (5)$$

where $\varepsilon_{i,j,t}^{(r)}$ follows a type-I extreme-value distribution and where $y_{i,j,t-1} = 1$ if person i purchased brand j at purchase occasion $t - 1$. This last term is included to model state dependence. State dependence refers to a dynamic property of the choice process, as it incorporates if a household tends to buy the same brand as purchased at the previous occasion. The degree of state dependence is measured by $\alpha^{(r)}$.

Of course, in case a household is unresponsive to marketing activities, the marketing instruments will not have an effect on its choice behavior. The choice will be mainly determined by base preferences, recent behavior (state dependence), and random effects. This type of behavior can be modeled by a second MNL model, that is,

$$U_{i,j,t}^{(u)} = \mu_j^{(u)} + \alpha^{(u)}y_{i,j,t-1} + \varepsilon_{i,j,t}^{(u)}, \quad (6)$$

where, obviously, the $X_{i,j,t}$ are excluded. In sum, household i purchases brand j at purchase occasion t when, conditional on responsiveness, $U_{i,j,t}^{(r)}$ is the maximum utility among $U_{i,k,t}^{(r)}$, $k = 1, \dots, J$ or, when, conditional on unresponsiveness, $U_{i,j,t}^{(u)}$ is the maximum utility among $U_{i,k,t}^{(u)}$, $k = 1, \dots, J$. In short-hand, brand j is purchased when

$$\left(U_{i,j,t}^{(r)} = \max_{k=1,\dots,J} U_{i,k,t}^{(r)} \right) \Big| Z_{i,t} = 1 \text{ or} \\ \left(U_{i,j,t}^{(u)} = \max_{k=1,\dots,J} U_{i,k,t}^{(u)} \right) \Big| Z_{i,t} = 0. \quad (7)$$

As the random parts of the utilities are assumed to be independently extreme-value distributed, the probability of the purchasing brand j for household i , that is responsive at

purchase occasion t , is

$$P\left(U_{i,j,t}^{(r)} = \max_{k=1,\dots,J} U_{i,k,t}^{(r)} \mid Z_{i,t} = 1\right) = \frac{\exp(\mu_j^{(r)} + X_{i,j,t}\beta^{(r)} + \alpha^{(r)}y_{i,j,t-1})}{\sum_{k=1}^J \exp(\mu_k^{(r)} + X_{i,k,t}\beta^{(r)} + \alpha^{(r)}y_{i,k,t-1})}, \quad (8)$$

where $\mu_j^{(r)}$ is restricted to 0 for identification, see McFadden (1973). If the household is unresponsive at t , the probability of purchasing brand j is

$$P\left(U_{i,j,t}^{(u)} = \max_{k=1,\dots,J} U_{i,k,t}^{(u)} \mid Z_{i,t} = 0\right) = \frac{\exp(\mu_j^{(u)} + \alpha^{(u)}y_{i,j,t-1})}{\sum_{k=1}^J \exp(\mu_k^{(u)} + \alpha^{(u)}y_{i,k,t-1})}, \quad (9)$$

with $\mu_j^{(u)} = 0$ for identification. Finally, as we do not observe whether a household at purchase occasion t belongs to the responsive segment or not, the probability that it purchases brand j at purchase occasion t can only be obtained by summing the conditional probabilities over the segments, that is,

$$P(y_{i,j,t} = 1) = P(U_{i,j,t}^{(r)} = \max_{k=1,\dots,J} U_{i,k,t}^{(r)} \mid Z_{i,t} = 1)P(Z_{i,t} = 1) \\ + P(U_{i,j,t}^{(u)} = \max_{k=1,\dots,J} U_{i,k,t}^{(u)} \mid Z_{i,t} = 0)P(Z_{i,t} = 0), \quad (10)$$

where $P(Z_{i,t} = 1)$ is given in (4).

2.2 State dependence and preference heterogeneity

The model developed above does not include unobserved preference heterogeneity. Such preference heterogeneity can be assigned to differences in base preference. For example, some households prefer brand A over brand B, while others may have the opposite preference. In general, by allowing for different base preferences, one gets a more realistic description of consumer behavior.

There is also another reason why it is important to account for heterogeneity in preferences. It is known that when base preferences are not correctly taken into account in a model with state dependence, the state dependence of households can be overestimated, see Allenby and Lenk (1994) and Keane (1997), among others. State dependence and differences in base preferences both describe observed persistence in brand choice, but they refer to different patterns of behavior. State dependence refers to a causal link between brand choice at period t and brand choice at period $t + 1$. The fact that brand j is purchased at time t increases the probability that brand j will be purchased again at $t + 1$. Differences in base preferences also capture persistence in brand choice. In this

case however, there is no causal link between brand choice at time t and brand choice at $t + 1$. Stated differently, state dependence refers to a property of the dynamics of the choice process whereas differences in base preferences are related to exogenous factors.

If base preferences are ignored, the model component designed to capture state dependence will also capture the base-preference effect. Households with a large base preference for a certain brand do not switch often. The large base preference of these households might therefore easily be confused for state dependence. Therefore, it is important to not only model state dependence but also to model the base preferences of households for brands. To model possible differences in base preferences across households, we allow the intercepts in the brand choice models ($\mu^{(r)}$ and $\mu^{(u)}$) to be different across households. As mentioned in the introduction, there are two different approaches for modeling preference heterogeneity. In this paper we choose to approximate the population distribution of the utility constants using a few discrete probability masses. This corresponds with the mixture approach of modeling heterogeneity.

Implicitly, our model with responsive and unresponsive households may already capture part of the preference heterogeneity. In the proposed model the brand intercepts differ over the two responsiveness segments. For as far as the preference segments coincide with the responsiveness segments, there is no need to extend the model to incorporate differences in base preferences. However, it may be the case that two segments of base preferences may not be enough or that the preference segments do not correspond with the structural segments. Note that the base preferences of households are likely to be the same over the observation period, while households may switch between both responsiveness states. Concluding, extending the model to include additional preference heterogeneity does seem to be a good idea.

To separate the two types of heterogeneity and to facilitate interpretation, the base preferences for a given household should not depend on the responsiveness state of the household at a particular purchase occasion. Therefore, we should make sure that the base preferences for a specific preference segment are the same in case of responsiveness as in case of unresponsiveness. Utility intercepts in a MNL model have a dual purpose. First, the intercepts correct for the mean of the explanatory variables. Second, they capture base preferences. As the two MNL models contain different explanatory variables, we cannot simply restrict the utility intercepts to be equal for the responsive and the unresponsive households. Instead, we have to use another strategy. Assume that there are S segments of households in the market with the same base preferences. Each segment

$s = 1, \dots, S$ has its own vector of base preferences ω_s . We cannot observe to which preference segment a household belongs, and the membership of these segments also has to be inferred from the data. The probability for household i of belonging to segment s is denoted by π_s , with $0 < \pi_s < 1$ and $\sum_{s=1}^S \pi_s = 1$. Note that the segment probabilities π_s are independent of household characteristics and that we assume that households do not switch segments over time. To indicate that household i belongs to segment s , we use the notation $\delta_{i,s} = 1$. The utility intercepts, we will use for preference segment s , equal $\mu^{(r)} + \omega_s$ for the responsive households and $\mu^{(u)} + \omega_s$ for the unresponsive households, with $\omega_1 = 0$ and $\omega_{s,J} = 0$, $s = 1, \dots, S$, for identification. Hence, the base preferences are the same for the two responsiveness states, but the “mean correction” modeled by $\mu^{(r)}$ and $\mu^{(u)}$, differs.

We only observe the final choice of a household. The choice process, nor the segment the household belongs to, are observed. The probability that household i purchases brand j at purchase occasion t now has to be marginalized on both the preference segment as well as on the responsiveness segments, that is,

$$\begin{aligned}
P(y_{i,j,t} = 1) &= \sum_{s=1}^S \pi_s P(y_{i,j,t} = 1 | \delta_{i,s} = 1) \\
&= \sum_{s=1}^S \pi_s \left[P(y_{i,j,t} = 1 | \delta_{i,s} = 1, Z_{i,t} = 1) P(Z_{i,t} = 1) \right. \\
&\quad \left. + P(y_{i,j,t} = 1 | \delta_{i,s} = 1, Z_{i,t} = 0) P(Z_{i,t} = 0) \right],
\end{aligned} \tag{11}$$

where

$$\begin{aligned}
P(y_{i,j,t} = 1 | \delta_{i,s} = 1, Z_{i,t} = 1) &= \frac{\exp(\mu_j^{(r)} + \omega_{s,j} + X_{i,j,t} \beta^{(r)} + \alpha^{(r)} y_{i,j,t-1})}{\sum_{k=1}^J \exp(\mu_k^{(r)} + \omega_{s,k} + X_{i,k,t} \beta^{(r)} + \alpha^{(r)} y_{i,k,t-1})} \\
P(y_{i,j,t} = 1 | \delta_{i,s} = 1, Z_{i,t} = 0) &= \frac{\exp(\mu_j^{(u)} + \omega_{s,j} + \alpha^{(u)} y_{i,j,t-1})}{\sum_{k=1}^J \exp(\mu_k^{(u)} + \omega_{s,k} + \alpha^{(u)} y_{i,k,t-1})}
\end{aligned} \tag{12}$$

and

$$\begin{aligned}
P(Z_{i,t} = 0) &= \frac{\exp(\mu^{(z)} + W_{i,t} \gamma^{(z)})}{1 + \exp(\mu^{(z)} + W_{i,t} \gamma^{(z)})} \\
P(Z_{i,t} = 1) &= 1 - P(Z_{i,t} = 0).
\end{aligned} \tag{13}$$

An interesting by-product of our model concerns the possibility to calculate the pos-

terior probability of responsiveness at purchase occasion t , that is

$$P(Z_{i,t} = 1 | y_{i,j,t} = 1) = \frac{P(Z_{i,t} = 1, y_{i,j,t} = 1)}{P(y_{i,j,t} = 1)} \quad (14)$$

$$= \frac{P(y_{i,j,t} = 1 | Z_{i,t} = 1)P(Z_{i,t} = 1)}{P(y_{i,j,t} = 1 | Z_{i,t} = 1)P(Z_{i,t} = 1) + P(y_{i,j,t} = 1 | Z_{i,t} = 0)P(Z_{i,t} = 0)}.$$

This gives the probability that household i is responsive to marketing efforts at purchase occasion t , given the fact that brand j is purchased. In the application we will show a histogram of these posterior probabilities to give an impression of the average value and the dispersion of the responsiveness in the population.

3 Inference

To estimate the model parameters, Maximum Likelihood [ML] is used. The likelihood function of our model (4, 8 and 9) is given by

$$\mathcal{L} = \prod_{i=1}^I \mathcal{L}_i = \prod_{i=1}^I \sum_{s=1}^S \pi_s \mathcal{L}_{i|s}, \quad (15)$$

where \mathcal{L}_i equals the likelihood contribution of household i and $\mathcal{L}_{i|s}$ denotes this contribution conditional on $\delta_{i,s} = 1$, that is,

$$\mathcal{L}_{i|s} = \prod_{t=1}^{T_i} \prod_{j=1}^J P(y_{i,j,t} = 1 | \delta_{i,s} = 1)^{y_{i,j,t}}, \quad (16)$$

where $P(y_{i,j,t} = 1 | \delta_{i,s} = 1)$ follows from (11).

The likelihood function as well as the log likelihood function are non-linear functions of the model parameters. Direct maximization of the expression (15) turns out to be difficult. The complicating factor is that, due to the preference heterogeneity, the purchases are no longer independent. It seems to be a better idea to use the Expectation Maximization [EM] algorithm, see Dempster *et al.* (1977) or see Wedel and Kamakura (1999) for a presentation of this algorithm in the context of mixture models. In case the model does not include preference heterogeneity, the likelihood can be maximized directly.

The EM-algorithm is based on the idea of a complete data likelihood function \mathcal{L}^c . This function yields the likelihood of the model assuming that the preference segment allocation $(\delta_{i,s}, i = 1, \dots, I, s = 1, \dots, S)$ is observed. The complete data likelihood equals

$$\mathcal{L}^c = \prod_{i=1}^I \prod_{s=1}^S [P(\delta_{i,s} = 1) \mathcal{L}_{i|s}]^{\delta_{i,s}}. \quad (17)$$

In practice the segment allocation is not observed and \mathcal{L}^c therefore is a function of the stochastic variables $\delta_{i,s}$. Taking the expectation of the logarithm of the complete data likelihood with respect to the segment allocation, conditional on the observed purchases, yields

$$E(\log \mathcal{L}^c | y) = \sum_{i=1}^I \sum_{s=1}^S E(\delta_{i,s} | y_i) [\log \pi_s + \log \mathcal{L}_{i|s}], \quad (18)$$

where y and y_i denote all observed purchases and all observed purchases by household i , respectively. The posterior segment membership probability $E(\delta_{i,s} | y_i)$ will be abbreviated by $P_{i,s}$, and is defined as

$$P_{i,s} = E(\delta_{i,s} | y_i) = \frac{P(\delta_{i,s} = 1) \mathcal{L}_{i|s}}{\sum_{r=1}^S P(\delta_{i,r} = 1) \mathcal{L}_{i|r}}. \quad (19)$$

It can be shown that, under relatively general conditions, iteratively maximizing (18) and updating $P_{i,s}$ according to (19) yields at least a local optimum of the likelihood function in (15).

In essence, the EM-algorithm consists of the expectation step and the maximization step. Given provisional parameter estimates, the expectation of $\log \mathcal{L}^c$ is calculated using (18) and (19). The resulting function is then maximized over the model parameters in the second step. Maximizing $E(\log \mathcal{L}^c | y)$ over π_s yields

$$\hat{\pi}_s = \frac{1}{I} \sum_{i=1}^I P_{i,s}. \quad (20)$$

New estimates for the remaining parameters cannot be obtained analytically. To obtain these, we numerically maximize

$$\sum_{i=1}^I \sum_{s=1}^S P_{i,s} \log \mathcal{L}_{i|s} = \sum_{i=1}^I \sum_{s=1}^S P_{i,s} \sum_{t=1}^{T_i} \sum_{j=1}^J y_{i,j,t} \log P(y_{i,j,t} = 1 | \delta_{i,s} = 1). \quad (21)$$

The last expression is relatively easy to maximize, compared to the original likelihood in (15), because conditional on the preference segments, all purchase occasions are independent. Iterating over the two steps of the EM-algorithm provides the ML parameter estimates.

The ML estimator is asymptotically normal distributed, with as mean the true value of the parameters and as covariance matrix the inverse of the information matrix. To estimate this covariance matrix, we choose to use the outer product of gradients method, see

Berndt *et al.* (1974). This gives numerically more stable results compared to the method which computes the covariance matrix from the inverse of the negative of the Hessian evaluated at the estimates. The covariance matrix thus becomes

$$\widehat{\text{Var}}(\hat{\theta}) = \left[\sum_{i=1}^I \left(\frac{\partial \log \mathcal{L}_i}{\partial \theta} \Big|_{\theta=\hat{\theta}} \right)' \left(\frac{\partial \log \mathcal{L}_i}{\partial \theta} \Big|_{\theta=\hat{\theta}} \right) \right]^{-1}, \quad (22)$$

where θ denotes a vector containing all parameters, $\hat{\theta}$ denotes the ML estimates, and where the derivatives are obtained from the likelihood in (15).

4 Illustration

We apply our model to a data base containing liquid detergent purchases in Sioux Falls, South Dakota, during the period July 1986 – July 1988. The sample contains 400 households making 2,657 purchases. The last observed purchase of each household is used as a hold-out sample for model comparison and evaluation. All other recorded purchases are used for estimation. The same data are analyzed in Chintagunta and Prasad (1998) for other purposes. The households in the panel are selected to only purchase the top six national brands, Tide, Eraplus, Solo, Wisk, All and Surf. For each purchase occasion we know the time since the last liquid detergent purchase, the volume last purchased (in multiples of 32 oz.), the shelf prices of the six alternatives and which brands are featured or on display. Table 1 gives a brief overview of the number of purchases and the use of marketing instruments in this market. In our sample the most popular brands are Tide and Wisk, these two brands are also most often featured and on display. The two smallest brands in choice share are Solo and All, which are rarely featured. Next to these variables, we know the chosen brand and the expenditures on non-detergent products made on the same shopping trip. On average, households shop for detergent every 80 days, purchase 77 oz. of detergent and spend almost \$40 per shopping trip. The average household size in our sample is 2.8 persons. A preliminary analysis of the data using simple versions of our model, as well as using the MNL model, indicated that display does not have a relevant effect on explaining brand choice, and hence this variable will be discarded.

Before we turn to discussing the estimation results in Section 4.2, we first consider in Section 4.1 the selection of the optimal number of segments to use. Section 4.3 compares the performance of our model to various forms of the MNL model.

4.1 Selecting the number of segments

First of all we need to determine the optimal number of segments to be used to capture preference heterogeneity. Unfortunately, Likelihood-Ratio tests cannot be used to determine the optimal number of segments due to the Davies (1977) problem. Therefore, we rely on goodness-of-fit and forecasting performance measures to select the number of segments. Examples of such measures are the in-sample and the out-of-sample hit rate, which measure the percentage of correct forecasts. If multiple preference segments are used, out-of-sample forecasting can be done as if the purchases are made by unknown households or by using the posterior segment information that is available from the in-sample purchases. In the former case, the estimated prior mixing probabilities $\hat{\pi}_s$, $s = 1, \dots, S$, are used. In the latter case, the posterior mixing probabilities in (19) are used which give the probability that household i belongs to segment s , conditional on the observed (in-sample) purchases. Finally, we can use the likelihood of the model for the out-of-sample purchases, the so-called predicted (log) likelihood, as a measure to compare the models.

In Table 2 we give these performance measures for 1, 2 and 3 preference segments in our responsiveness model. All measures related to forecasting accuracy indicate the single segment model to perform best. Although households can switch between responsiveness segments but not between preference segments, it turns out that correcting for differences in the responsiveness to marketing efforts captures almost all of the heterogeneity in this market. The introduction of preference segments is not needed for our sample.

4.2 Estimation results

Table 3 shows the estimation results for the model equation concerning responsiveness in the model with one preference segment, see (4). The household characteristics and the marketing efforts are all normalized to have mean 0 and variance equal to 1. The variables concerning the previous shopping occasion have a significant effect on the responsiveness while household size does not have such an effect. If a household has not purchased liquid detergent for a long period relative to a more frequent shopper, it might better think about the next purchase and this increases the responsiveness to marketing efforts. The reverse holds for the volume of liquid detergent purchased previously. Households purchasing large quantities may have a strong preference for one of the brands, and therefore could be less responsive to marketing efforts of other brands.

Table 4 presents the parameter estimates of the brand intercepts in the MNL models

concerning the responsive and the unresponsive households as well as the estimated effect of marketing efforts and the effects of state dependence. From this table we can conclude that the choice behavior of unresponsive households is more difficult to predict compared to the responsive households. This is seen from the brand intercepts for unresponsive households which are closer to zero in absolute value. In fact, the brand intercepts for the unresponsive households are not significantly different from zero, with only one exception. All estimated effects of marketing efforts on purchase probabilities have the expected sign. The effect of price is clearly the largest. An interesting observation is that households who are unresponsive to marketing efforts usually act more state dependent than responsive households.

Finally we elaborate upon the posterior responsiveness. Figure 1 gives the distribution over the population of $P(Z_{i,t} = 1 | y_{i,d_{i,t},t} = 1)$, where $d_{i,t}$ denotes the index of the chosen brand. From this distribution we can conclude that, at a large proportion of the purchase occasions, the household must be responsive to marketing efforts. In almost 18 % of the purchase occasions, the household has a probability of almost 1 of being responsive to marketing efforts. Next to these, there is also a large proportion of purchases where the household was almost surely not responsive. This is reflected by the relatively large area under the distribution for probabilities close to 0. To evaluate whether households switch responsiveness segments, we use the posterior responsiveness to assign households to the two segments for every purchase occasion. If the posterior responsiveness probability exceeds 0.5 the household is, at that purchase occasion, assigned to the responsive segment. If the segment membership of a household is tracked over time, we see that of the 210 households that purchased detergent more than once, 131 households switched from unresponsive to responsive at least once.

Summarizing, we find for this data set that many households are responsive to marketing efforts, that households purchasing large amounts are less responsive to marketing efforts, that households who do not regularly purchase detergent are more responsive, that the behavior of responsive households is easier to predict due to a larger dispersion in $\mu^{(r)}$ compared to $\mu^{(u)}$, and that unresponsive households behave more state dependent than responsive households do.

4.3 Competing models

The results above show that the segmentation of purchase occasions in cases where households are responsive or unresponsive clearly separates the purchase occasions where the household acts state dependent from the cases where the household responds to marketing efforts. The other parameter estimates also seem to have a high face validity. All this together leads us to believe that the concept of responsiveness seems to be empirically useful. Of course, the ultimate test for a newly proposed model is whether it fits the data better and generates better forecasts compared to other models. In choice modeling, the MNL model seems to be the standard. In this section we will therefore compare our model to various forms of the MNL model. In the MNL models, all variables used in the responsiveness model are included, including the brand choice on the previous purchase occasion. We also consider various forms of heterogeneity. We use a MNL model without incorporating unobserved heterogeneity, and one where both the utility intercepts as well as the sensitivity to marketing efforts are allowed to be heterogeneous (*full heterogeneity*). Furthermore, we consider MNL models where either the brand intercepts (μ -heterogeneity) or the sensitivity (β -heterogeneity) are assumed to differ across households.

Table 5 shows the same model performance criteria as were used to determine the optimal number of preference segments. The in-sample forecasting results suggest that our model outperforms all MNL models. On out-of-sample forecasting, only the 2-segment full heterogeneity MNL model beats our model using posterior segment probabilities. This MNL model however does not provide the best forecasts when the forecasts are made as if the households were not previously observed. Overall our model seems to perform much better than the MNL model, notably with less parameters! We also compare the responsiveness model to a MNL where the state dependence parameter is allowed to differ in the population (results not included in this paper). These models also did not forecast better than the 1-segment responsiveness model. These results seem to indicate that the concept of responsiveness apparently captures most of the heterogeneity.

It is also interesting to compare the parameter estimates of some of the models. Especially the estimates of the marketing-effort effectiveness give information on the possible bias in a MNL logit model when potentially varying degrees of responsiveness are ignored. In Table 6, we compare parameter estimates for the responsiveness model, the simple MNL model, and the MNL model with two segments where the intercepts, the marketing sensitivity and the state dependence parameter are allowed to differ. The results confirm our conjecture that, if the responsiveness is ignored, marketing-effort effectiveness is un-

derestimated in a MNL model. An interesting observation for the 2-segment MNL model is that the second segment partly captures the unresponsive households, as sensitivity to marketing efforts is quite low compared to the first segment whereas there is a high degree of state dependence. The main difference between this model and our model is of course that in the responsiveness model the segment membership is correlated to household characteristics and households are allowed to switch between the segments over time.

5 Concluding remarks

Households might not respond to marketing-mix instruments at every purchase occasion. To be able to respond to these efforts, one needs to invest time and effort in, for example, remembering price changes and reading newspapers and leaflets to notice advertisements. Households differ in the amount of effort they wish to invest in a particular purchase, and therefore they will most likely also differ in their responsiveness to marketing efforts.

The choice model we developed in this paper incorporates the responsiveness of a household at a specific purchase occasion as a form of structural heterogeneity. Households differ in their decision-making process. In essence, we assume there are two decision processes. Households either take marketing efforts into account or they base their choice solely on base preference and their past experience. The specific decision process used can differ across households and across purchase occasions. To explain and forecast the decision process, used by a specific household at a specific purchase occasion, household characteristics can be used together with information on buying behavior. To take into account this form of structural heterogeneity, we extended the MNL model. Basically, we introduce two segments of households, one segment is unresponsive to marketing efforts whereas the other segment does respond to these efforts. The segment membership is separately modeled using a logit model. This “responsiveness model” can be further extended to also include preference heterogeneity, although our illustration did not suggest its necessity, at least for our data.

The main behavioral conclusions, we can draw from the application of our model to a six-brand market of liquid detergent, are that most households are responsive to marketing efforts, and that large basket shoppers and large volume shoppers tend to be less responsive. Infrequent shoppers however are more responsive. Further, an unresponsive household seems to act more state dependent.

Finally, we compared the in- and out-of-sample performance of our model to various

forms of the MNL model where we also correct for heterogeneity. From this comparison we conclude that, while using the smallest amount of parameters, our model outperforms all MNL variants on forecasting. This, together with the face validity of our parameter results, leads us to believe that incorporating responsiveness seems to be a worthwhile exercise.

Table 1: Data properties

Brand	# Purchases	Product characteristics		
		Avg. price per oz.	Feature ¹	Display ¹
Tide	701	0.0602	9.82%	9.17%
Wisk	703	0.0472	14.57%	10.11%
EraPlus	507	0.0610	3.34%	2.62%
Surf	406	0.0532	4.65%	3.73%
Solo	253	0.0594	1.90%	1.44%
All	87	0.0394	0.07%	0.07%

¹ Percentage of times over all observed purchase occasions that the product was featured or on display.

Table 2: Goodness of fit measures for the responsiveness model for 1 to 3 preference segments.

	Number of preference segments		
	1	2	3
Number of parameters	19	25	31
$\log \mathcal{L}$	-1403	-1338	-1308
pred $\log \mathcal{L}$	-192.6	-202.5	-207.3
Hit rate in-sample	77.61	76.68	75.34
Hit rate out-of-sample			
posterior segm. prob.	77.31	75.39	75.39
prior segm. prob.	77.31	76.15	74.62

Table 3: Parameter estimates in the responsiveness model (equation (4)), standard errors in parentheses

	Responsiveness	
Intercept ($\mu^{(z)}$)	0.60	(0.18)
Non-Detergent expenditures	-0.21	(0.09)
Household size	-0.07	(0.07)
Time since last purchase	0.65	(0.18)
Volume previously purchased	-0.47	(0.09)

Table 4: Parameter estimates for models (8) and (9) (standard errors in parentheses)

	Responsive		Unresponsive	
<i>Brand Intercepts</i>				
Tide	0.335	(0.131)	0.005	(0.523)
Wisk	0.367	(0.149)	-0.491	(0.761)
Surf	-0.035	(0.148)	-2.584	(0.726)
Solo	-0.688	(0.190)	-0.673	(0.745)
All	-2.274	(0.371)	-1.135	(0.874)
EraPlus	0*	-	0*	-
<i>Marketing efforts/state dependence</i>				
Price	-1.347	(0.09)	-	-
Feature	0.446	(0.048)	-	-
Lagged purchase	1.941	(0.132)	4.836	(0.507)

* Restricted for identification

Table 5: Comparison of new model versus various forms of the MNL model

Model	Responsiveness	MNL	MNL full heter.		MNL μ -heter.		MNL β -heter.	
No. segments	1	1	2	3	2	3	2	3
No. parameters	19	28	36	44	34	40	31	34
$\log \mathcal{L}$	-1403	-1429	-1343	-1285	-1349	-1306	-1373	-1364
pred $\log \mathcal{L}$	-192.6	-213.1	-206.5	-215.7	-215.8	-217.6	-207.9	-207.3
Hit rate								
in-sample	77.61	76.73	76.06	72.24	75.08	73.27	76.37	76.47
out-of-sample post.	77.31	74.23	75.39	77.69	74.62	74.23	74.23	75.00
out-of-sample prior	77.31	74.23	73.85	73.85	73.47	71.92	73.46	73.08

Table 6: Comparison of effectiveness of marketing instruments (standard errors in parentheses)

	Responsiveness model		MNL	2-segment MNL	
	Responsive	Unresponsive		Segment 1	Segment 2
Price	-1.35 (0.09)	- -	-1.02 (0.05)	-1.25 (0.07)	-0.67 (0.10)
Feature	0.45 (0.05)	- -	0.31 (0.03)	0.43 (0.05)	0.21 (0.06)
Lagged purchase	1.94 (0.13)	4.84 (0.51)	2.58 (0.07)	1.37 (0.12)	3.80 (0.15)

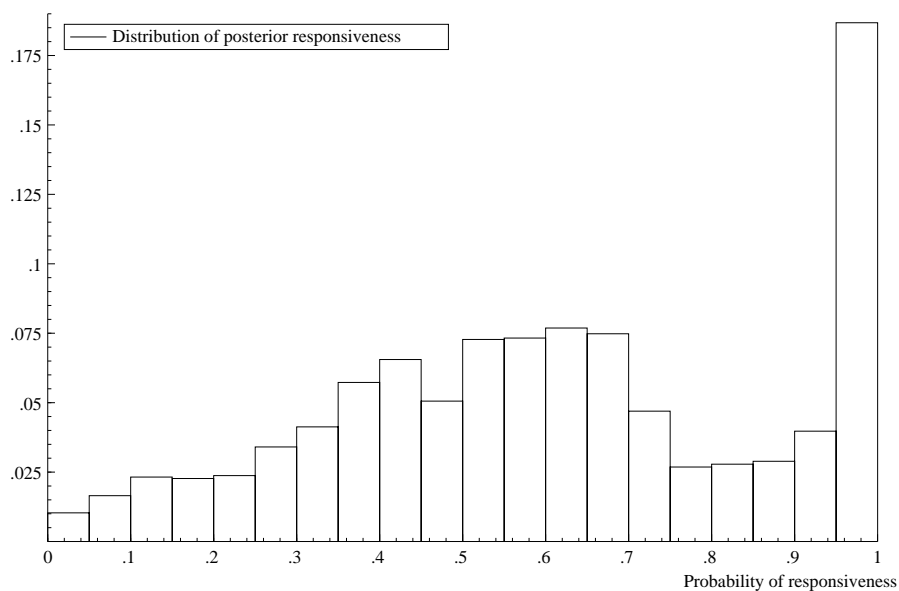


Figure 1: Distribution of posterior responsiveness

References

- Allenby, G. M. and P. J. Lenk (1994), Modeling Household Purchase Behavior with Logistic Normal Regression, *Journal of the American Statistical Association*, **89**, 1218–1231.
- Allenby, G. M. and P. E. Rossi (1999), Marketing Models of Consumer Heterogeneity, *Journal of Econometrics*, **89**, 57–78.
- Berndt, E., B. Hall, E. Hall, and J. Hausman (1974), Estimation and Inference in Non-linear Structural Models, *Annals of Economic and Social Measurement*, **3**, 653–665.
- Chintagunta, P. K., D. C. Jain, and N. J. Vilcassim (1991), Investigating Heterogeneity in Brand Preferences in Logit Models for Panel Data, *Journal of Marketing Research*, **28**, 417–428.
- Chintagunta, P. K. and A. R. Prasad (1998), An Empirical Investigation of the "Dynamic McFadden" Model of Purchase Timing and Brand Choice: Implications for Market Structure, *Journal of Business & Economic Statistics*, **16**, 2–12.
- Davies, R. B. (1977), Hypothesis Testing When a Nuisance Parameter Is Present Only Under the Alternative, *Biometrika*, **64**, 247–254.
- Dempster, A. P., N. M. Laird, and D. B. Rubin (1977), Maximum Likelihood from Incomplete Data Via the EM Algorithm (with Discussion), *Journal of the Royal Statistical Society B*, **39**, 1–38.
- Doornik, J. A. (1999), *Object-Oriented Matrix Programming Using Ox*, 3rd edn., London: Timberlake Consultants Press and Oxford: www.nuff.ox.ac.uk/Users/Doornik.
- Guadagni, P. M. and J. D. Little (1983), A Logit Model of Brand Choice Calibrated on Scanner Data, *Marketing Science*, **2**, 203–238.
- Jain, D. C., N. J. Vilcassim, and P. K. Chintagunta (1994), A Random-Coefficients Logit Brand-Choice Model Applied to Panel Data, *Journal of Business & Economic Statistics*, **12**, 317–328.
- Kamakura, W. A., B.-D. Kim, and J. Lee (1996), Modeling Preference and Structural Heterogeneity in Consumer Choice, *Marketing Science*, **15**, 152–172.

- Keane, M. P. (1997), Modeling Heterogeneity and State Dependence in Consumer Choice Behavior, *Journal of Business & Economic Statistics*, **15**, 310–327.
- Maddala, G. (1983), *Limited Dependent and Qualitative Variables in Econometrics*, vol. 3 of *Econometric Society Monographs*, Cambridge University Press, Cambridge.
- McFadden, D. (1973), Conditional Logit Analysis of Qualitative Choice Behavior, in P. Zarembka (ed.), *Frontiers in Econometrics*, chap. 4, Academic Press, New York, pp. 105–142.
- McFadden, D. (1981), Econometric Models of Probabilistic Choice, in C. Manski and D. McFadden (eds.), *Structural Analysis of Discrete Data: With Econometric Applications*, MIT Press, Cambridge, MA, pp. 197–272.
- Wedel, M. and W. A. Kamakura (1999), *Market Segmentation: Conceptual and Methodological Foundations*, Kluwer Academic Publishers, Dordrecht.
- Wedel, M., W. A. Kamakura, N. Arora, A. Bemmaor, J. Chiang, T. Elrod, R. Johnson, P. Lenk, S. Neslin, and C. S. Poulsen (1999), Discrete and Continuous Representations of Unobserved Heterogeneity in Choice Modelling, *Marketing Letters*, **10**, 219–232.
- Yang, S. and G. M. Allenby (2000), A Model for Observation, Structural, and Household Heterogeneity in Panel Data, *Marketing Letters*, **11**, 137–149.

Publications in the Report Series Research* in Management

ERIM Research Program: "Marketing"

2001

Predicting Customer Potential Value. An application in the insurance industry

Peter C. Verhoef & Bas Donkers

ERS-2001-01-MKT

Modeling Potentially Time-Varying Effects of Promotions on Sales

Philip Hans Franses, Richard Paap & Philip A. Sijthoff

ERS-2001-05-MKT

Modeling Consideration Sets and Brand Choice Using Artificial Neural Networks

Björn Vroomen, Philip Hans Franses & Erjen van Nierop

ERS-2001-10-MKT

Firm Size and Export Intensity: A Transaction Costs and Resource-Based Perspective

Ernst Verwaal & Bas Donkers

ERS-2001-12-MKT

Customs-Related Transaction Costs, Firm Size and International Trade Intensity

Ernst Verwaal & Bas Donkers

ERS-2001-13-MKT

The Effectiveness of Different Mechanisms for Integrating Marketing and R & D

Mark A.A.M. Leenders & Berend Wierenga

ERS-2001-20-MKT

Intra-Firm Adoption Decisions: Departmental Adoption of the Common European Currency

Yvonne M. van Everdingen & Berend Wierenga

ERS-2001-21-MKT

Econometric Analysis of the Market Share Attraction Model

Dennis Fok, Philip Hans Franses & Richard Paap

ERS-2001-25-MKT

Buying High Tech Products: An Embeddedness Perspective

Stefan Wuyts, Stefan Stremersch & Philip Hans Franses

ERS-2001-27-MKT

Changing Perceptions and Changing Behavior in Customer Relationships

Peter C. Verhoef, Philip Hans Franses & Bas Donkers

ERS-2001-31-MKT

How and Why Decision Models Influence Marketing Resource Allocations

Gary L. Lilien, Arvind Rangaswamy, Katrin Starke & Gerrit H. van Bruggen

ERS-2001-33-MKT

* A complete overview of the ERIM Report Series Research in Management:
<http://www.ers.erim.eur.nl>

ERIM Research Programs:

LIS Business Processes, Logistics and Information Systems

ORG Organizing for Performance

MKT Marketing

F&A Finance and Accounting

STR Strategy and Entrepreneurship

An Equilibrium-Correction Model for Dynamic Network Data
David Dekker, Philip Hans Franses & David Krackhardt
ERS-2001-39-MKT

Aggregation Methods in International Comparisons: What Have We Learned?
Bert M. Balk
ERS-2001-41-MKT

The Impact of Channel Function Performance on Buyer-Seller Relationships in Marketing Channels
Gerrit H. van Bruggen, Manish Kacker & Chantal Nieuwlaet
ERS-2001-44-MKT

Incorporating Responsiveness to Marketing Efforts when Modeling Brand Choice
Dennis Fok, Philip Hans Franses & Richard Paap
ERS-2001-47-MKT

2000

Impact of the Employee Communication and Perceived External Prestige on Organizational Identification
Ale Smidts, Cees B.M. van Riel & Ad Th.H. Pruyn
ERS-2000-01-MKT

Forecasting Market Shares from Models for Sales
Dennis Fok & Philip Hans Franses
ERS-2000-03-MKT

The Effect of Relational Constructs on Relationship Performance: Does Duration Matter?
Peter C. Verhoef, Philip Hans Franses & Janny C. Hoekstra
ERS-2000-08-MKT

Informants in Organizational Marketing Research: How Many, Who, and How to Aggregate Response?
Gerrit H. van Bruggen, Gary L. Lilien & Manish Kacker
ERS-2000-32-MKT

The Powerful Triangle of Marketing Data, Managerial Judgment, and Marketing Management Support Systems
Gerrit H. van Bruggen, Ale Smidts & Berend Wierenga
ERS-2000-33-MKT

Consumer Perception and Evaluation of Waiting Time: A Field Experiment
Gerrit Antonides, Peter C. Verhoef & Marcel van Aalst
ERS-2000-35-MKT

Broker Positions in Task-Specific Knowledge Networks: Effects on Perceived Performance and Role Stressors in an Account Management System
David Dekker, Frans Stokman & Philip Hans Franses
ERS-2000-37-MKT

Modeling Unobserved Consideration Sets for Household Panel Data
Erjen van Nierop, Richard Paap, Bart Bronnenberg, Philip Hans Franses & Michel Wedel
ERS-2000-42-MKT

A Managerial Perspective on the Logic of Increasing Returns
Erik den Hartigh, Fred Langerak & Harry Commandeur
ERS-2000-48-MKT

The Mediating Effect of NPD-Activities and NPD-Performance on the Relationship between Market Orientation and Organizational Performance
Fred Langerak, Erik Jan Hultink & Henry S.J. Robben
ERS-2000-50-MKT

Sensemaking from actions: Deriving organization members' means and ends from their day-to-day behavior
Johan van Rekom, Cees B.M. van Riel & Berend Wierenga
ERS-2000-52-MKT