B. Wierenga AN INVESTIGATION OF BRAND CHOICE PROCESSES

AN INVESTIGATION OF BRAND CHOICE PROCESSES

B. WIERENGA

UNIVERSITAIRE PERS ROTTERDAM

An investigation of brand choice processes

Dit proefschrift met stellingen van Berend Wierenga, landbouwkundig ingenieur, geboren te Zoutkamp op 28 juni 1945, is goedgekeurd door de promotor dr. ir. M.T.G. Meulenberg, hoogleraar in de Marktkunde en het Marktonderzoek.

Wageningen, 27 maart 1974

De rector-magnificus
van de Landbouwhogeschool,
H.A. Leniger

B. WIERENGA

An investigation of brand choice processes

Proefschrift ter verkrijging van de graad
van doctor in de landbouwwetenschappen
op gezag van de rector magnificus,
prof. dr. ir. H. A. Leniger, hoogleraar in de technologie,
in het openbaar te verdedigen op vrijdag 10 mei
1974 des namiddags te vier uur in de aula van de
Landbouwhogeschool te Wageningen

1974
Universitaire Pers Rotterdam

To Tjitske, Anieke
and Corette

Acknowledgements

I would like to thank Prof. M.T.G. Meulenberg for the stimulating interest with which he followed my research and for the very useful discussions we had while this book was in preparation.

I am also indebted to Prof. L.C.A. Corsten for a number of helpful comments, and to Mr. J.A. Bijkerk for his much-appreciated assistance in preparing the data for the computer and in writing a number of the computer programs.

The Computer Centre of the Agricultural University provided great help in processing the data.

I would also like to thank Attwood Nederland N.V., Dongen, and the Bureau voor Budgetten Contrôle, Amsterdam, for making available the data used in this study.

Finally, I wish to express my gratitude to the Typing Department of the Agricultural University, in particular to Mrs. P.W. van den Anker-Gerbrands, and to Mrs. A.J. Nijsse of the Department of Marketing for the excellent job they did in typing the manuscript.

Contents

I. Brand loyalty and brand choice processes: An introduction to the study

1.1 THE PHENOMENON OF BRAND LOYALTY

An important aspect of the buying behavior of a consumer making succesive purchases of a certain product is the brand, which is chosen at the consecutive purchase occasions. Generally, more than one brand is offered of the same product within a given product field, and a buyer has to choose one of these brands at the moment of purchase. It may be expected that this is not a random choice from the alternatives available, because the various brands of the same product will differ more or less from each other in the eyes of the consumer. One indication of this is the widely-spread use of brands in marketing policy nowadays. If the brands were all the same to a consumer, all efforts of manufacturers and retailers to give their product its own identity via a brand and accompanying marketing activities would be in vain. The fact that a great many products today, consumer goods in particular, are sold as branded products, shows the great confidence placed in the effectivity of branding as a tool for marketing policy.

Amongst the first to publish research results, showing that consumers do not choose a brand at random but that – for a certain consumer – different brands have different probabilities to be bought, were Brown (1952) and Cunningham (1956). Cunningham, using data from a consumer panel, calculated for every family in the panel the percentage represented by the most favored (= most bought) brand in the purchases of various products during 3 years. For the seven products studied, he found that individual households bought the favorite brand much more frequently than would be expected under the assumption of a random choice from the brands available.

For every family there is usually one brand (sometimes more than one) which has a rather high probability of being chosen on a purchase occasion, while for other brands this chance is low, perhaps almost zero.

In general, the brand chosen at many previous purchase moments has a high probability of being bought again on subsequent occasions.
So one often stays with the same brand for quite a number of purchases, and to cover this phenomenon the term brand loyalty has come into use.

There is not one exclusive measure of *brand loyalty*. The quantity, calculated by Cunningham i.e., the percentage in the purchases represented by the most favored brand, seems to be a quite natural measure, but there are other possibilities, e.g., the percentage in the purchases represented by the first and second favored brand together, or: the number of different brands bought during a certain number (e.g. 100) of purchase occasions, etc.

Following the work of Brown and Cunningham a number of studies on brand loyalty appeared. In these studies attention was especially paid to the question of whether or not there are general factors which determine the degree to which a consumer is brand loyal. We will return to this subject in chapter 8.

1.2 THE BRAND CHOICE PROCESS

The notion of brand loyalty, dealt with in the previous section, has the character of a summary measure. After a family has completed a sequence of purchases, a quantity can be computed that measures the family's brand loyalty. As noted above, such a measure is not unique.

Such kinds of brand choice measures are derived from the *brand choice process*. Here we define a brand choice process as: *the consecutive buying of certain brands of a product by a consumer*. It is this brand choice process which forms the subject of study in this book.

Generally consumer behavior is not deterministic. At a purchase moment each brand has a certain probability of being chosen, and we can speak about the probability distribution over the various brands. Brand choice processes are essentially stochastic processes. Let $\{X_t\}$ represent the brand choice process of a certain consumer for a certain product, where X_t denotes the brand chosen on purchase occasion t. (A purchase occasion occurs every time the consumer makes a purchase of the product under study, irrespective of the brand chosen; more generally, such an occasion can be called a response occasion.) When there are m different brands in a particular market and the brands are represented by the numbers 1 to m, the state space, the values X_t can take, is the set of numbers 1 to m; t, the indexing number of the purchase occasions, can take the values 1, 2, 3, ... So the brand to be chosen at a purchase occasion is a random variable and the brand choice process $\{X_t\}$ is a stochastic process with a discrete state space in discrete time. The probability that brand i will be chosen at purchase occasion t, given the brands chosen at the k previous occasions, can be noted as follows:

$$\text{Prob}\,(X_t = i \mid x_{t-1}, x_{t-2}, \ldots, x_{t-k}) \qquad (1.1)$$

In real market situations m can be a rather big number (e.g. greater than 100). It is clear that in empirical research one must condense the state space of the process by combining the brands into a smaller number of classes.

For practical applications (1.1) is too general and one has to specify the way in which the probability of buying a certain brand depends on the purchase history and which part of the purchase history is relevant to this probability. The various brand choice models developed in marketing literature represent such specifications. As an example here we will briefly discuss the Markov Model, which has received much attention over the years.

Markov processes constitute an important class of stochastic processes. Like other social phenomena such as voting behavior and social mobility (see Coleman (1965)), brand choice behavior seems to be suitable for description by a Markov model. The characteristic feature of a Markov process is, that the probability distribution over the state space at a certain response occasion, depends only on a limited number of previous realisations of the process. In brand choice terms: the probability that a certain brand will be chosen at a given purchase occasion depends only on the brands chosen at the recent purchase occasions. For instance for a so-called first order Markov process the probability distribution of X_t is conditional only on the realisation of the process at $(t-1)$, i.e.

$$\text{Prob}\,(X_t = i \mid x_{t-1}, x_{t-2}, \ldots) = \text{Prob}\,(X_t = i \mid x_{t-1}) \qquad (i = 1, \ldots m)$$

In the sixties quite a number of papers appeared which treated the brand choice process within a Markovian framework. We will discuss this approach more extensively in chapter 3, where the Markov Model will be treated as one of a number of different brand choice models, amongst which the so-called Linear Learning Model will also be included.

1.3 SCOPE OF THE STUDY

In this study, the subject of research is the brand choice process. For this investigation the structure of the purchase histories of consumers is studied. These purchase histories are realisations of the brand choice process. It can be said that the purchase histories reflect the brand choice process by which they were brought about. Here the concept *structure* points to the configuration of the different brands in a purchase history. The characteristics of a brand choice process, e.g., the relative frequency of a certain brand in the purchase history, the number of transitions to another brand, the number of times the same brand is consecutively bought, etc. refer to the structure of the purchase history. These characteristics give information about the brand choice processes, which produce the purchase history.

One way of studying brand choice processes is to formulate mathematical models – stochastic in nature of course – which represent certain assumptions about the underlying mechanism of the brand choice process, and then to examine in how far these models give a good description of empirical brand choice processes.

Generally, different brand choice models have different implications for the structure of the resulting purchase histories and the fit of a specific brand choice model to an empirical brand choice process gives information about the structure of the purchase histories studied. A part of this book, namely chapters 3 to 5, is dedicated to this kind of model approach.

A drawback of this approach is, however, that with the present state of the art of brand choice models one is mostly forced to condense the markets under study to two-brand markets so that only a summarized form of the original brand choice process can be studied.

Another possibility is to omit the formulation of exact models and just observe what happens in a brand choice process. In this way the attempt can be made to answer questions like: how many brands are involved in a brand choice decision? do consumers rather straight-forwardly switch from one brand to another? do they exhibit search behavior, etc. This method of studying the brand choice process has the advantage that it requires less condensation of the original processes. It can act as a supplement to the model approach and this is the method followed in chapter 6, where the poolsize approach is treated.

In the general formulation (1.1) it is assumed that the probability of choosing a particular brand depends only on the purchase history of a consumer. Obviously, however, the brand choice process does not take place in isolation but within a certain environment. A consumer makes his purchases in certain shops, has certain time intervals between successive purchases and is exposed to the marketing policy of the various sellers, which implies certain prices and certain levels of advertising and other promotional activity for the different brands. In chapter 7 the influence of these environmental variables – as we call them – on brand choice will be examined.

Finally, it is interesting to know if different households have different types of brand choice processes. To this end the relationships between characteristics of the brand choice process and household variables (socio-economic and purchase variables) will be examined in chapter 8.

For the investigation the purchase histories of 600–1100 families, all members of a consumer panel, were used. The purchases of 3 different products – frequently purchased consumer goods – over a period of 2 years were studied.

The following limits were set to the study.

The prime interest was in brand choice; inter-purchase times (times between subsequent purchases) were only considered to see if they had a bearing on brand choice.

The behavioral manifestation of brand choice was studied, not the underlying state of mind of the consumers making the subsequent brand choices. Note that the definitions of brand loyalty and brand choice process in the preceding sections are behavioral: they look at the behavior which is observed, not at the mental state behind it.

Of course the psychological workings of the consumer, which lie behind his observable behavior, is important. For example Nicosia (1966), Engel, Kollat & Blackwell (1968) and Howard & Sheth (1968) have built models in which the psychological processes assumed to underly purchasing behavior are described in detail. These models are of a high degree of complexity however, and require very detailed observations of the buying process for their validation. In the current state of brand choice research, such models have the important merit that they offer a framework for thinking about buying behavior.

With respect to the marketing-relevance of this study the following can be remarked. It is clear that knowledge about brand choice processes is important from the marketing point of view. To implement a marketing policy for a product, a company should be aware of the position of its brand vis-à-vis other brands in the market. For example, it is important to know if there is a process working which might eventually lead to an unfavorable position of the own brand in the market. When we can adequately describe the brand choice process by a certain model, an unfavorable development can be detected quite early by calculating the equilibrium state from the parameters, or, if necessary, by a simulation of the process. When one only considers summary statistics, e.g., market shares, the underlying process may temporarily remain concealed. With regard to the environmental variables, it can be said that when the relationship of the brand choice process to such marketing variables as price and advertising is known, this information can be used for an attempt to turn the movement in the market in a more favorable direction. Knowledge of the relationships between the brand choice process and family characteristics can be useful as a basis for market segmentation.

1.4 PLAN OF THE BOOK

We will briefly outline here the contents of the various chapters of this book. In the next chapter – chapter 2 – a description of the data used and the main characteristics of the markets for which the brand choice processes were studied are given.

In chapter 3 a number of brand choice models are presented, with the corresponding estimation and testing procedures. The theory in this chapter is based for the greater part on existing literature, but at a number of points original contributions are made.

In chapter 4 the brand choice models are applied to the empirical brand choice processes and it is determined if – with these models – satisfactory descriptions of brand choice processes can be obtained. This is done in the first place by means of appropriate testing procedures, corresponding to the models. Moreover a simulation study is carried out in order to evaluate reproduction of the brand choice processes by the various models.

In chapter 5 the application of learning models to brand choice processes is discussed at some length. This is done because of the good results for the so-called Linear Learning Model in chapter 4. A number of different learning models taken from mathematical psychology are treated and the possibilities of applying them to brand choice processes discussed. Some additional properties of the Linear Learning Model are given, especially with respect to equilibrium behavior and some possible generalisations of this model shown.

In chapter 6 an approach to the brand choice process is proposed, whereby every consumer is assumed to have a set of alternatives (a pool) from which he chooses a brand. This set is generally smaller than all brands in the market. Looking at the brand choice process from this angle provides additional insights into the way a consumer makes his choices.

In chapter 7 the relationships between the brand choice process and environmental variables are considered: the relationships of brand choice with store choice, prices, advertising, deal-offers and inter-purchase times are studied here.

In chapter 8 the relations between household variables and characteristics of the brand choice process are examined.

Chapter 9 contains the major conclusions of the research.

2. The data used and some characteristics of the markets

2.1 THE CONSUMER PANEL

To study the brand choice process for a certain product, one must be able to trace the purchases of individual consumers in time. Therefore so-called purchase histories of individual consumers are needed. Members of a consumer panel provide such continuous recordings of their purchases.

Thanks to the cooperation of 'Attwood Statistics Nederland', it was possible for us to obtain available purchase histories for 3 frequently purchased food-products by members of the Attwood panel in the Netherlands. The purchases were made during the years 1967 and 1968. These data are used throughout this study.

The Dutch Attwood panel consists of 2000 households which make weekly reports of their purchases of a number of products. The 2000 households are representative for the population of Dutch consumers in so far as they live in households. Because they are spread all over the country, it can be assumed that the individual households in the panel are mutually independent with regard to their buying behavior.

For each purchase a certain amount of information is recorded (see section 2.2); moreover for each household in the panel, information is available with respect to socio-economic variables and buying attitudes.

A general treatment of the possibilities and difficulties regarding consumer panels can be found in Boyd & Westfall (1960). Two critical points, mentioned by these authors and relevant to our research, are accuracy of registration and the possible conditioning effect of panel-membership on the purchase behavior of members. In this respect the following research results can be mentioned.

Sudman (1964) tested the influence of a number of variables on the recording accuracy of members of the Market Research Corporation of America (MRCA) Panel by comparing sales, computed on the basis of panel recordings, with shipments from individual companies. He admits that this comparison is rather crude, because a manufacturer usually sells to a universe broader than that which a consumer panel measures and moreover other factors may cause differences. With respect to one of the variables considered, viz., type of product, he concluded that frequently bought food products and non-food grocery products were recorded relatively accurately compared with other products. This conclusion favours the data used in our study because they refer to frequently bought food products.

Ehrenberg (1960) and McGloughlin (1971), examining data from the Attwood panels in England, Germany and the Netherlands, concluded that no systematic differences could be found between new and old panel members with respect to such variables as: number of purchases, brand shares, number of different brands bought, prices paid, etc.

Morrison et al. (1966), found some differences in purchasing behavior between new and old members of the Chicago Tribune Panel, but they caution against a too hasty conclusion because the 2 groups they compared differed considerably in socio-economic characteristics.

2.2 THE PRODUCTS CHOSEN AND THE HOUSEHOLDS SELECTED

The products for which we studied brand choice processes are:

1. a frequently purchased food product, for reasons of confidentiality indicated by the pseudonym: fopro;
2. beer;
3. margarine (often referred to as: marg).

Considerations of sufficient purchase frequency and a reasonable number of brands influenced the choice of these products, but of course other choices would have been possible.

The number of products chosen is a compromise between the requirements of studying different situations to get as broad a picture as possible and of keeping the research project manageable.

As said before, the data at our disposal were the purchase histories for the 3 products mentioned above of the members of the Attwood panel during the years 1967 and 1968.

The 2000 households in the panel during the 2 years did not remain constant. Naturally for our analysis purchase histories should be as long as possible. Furthermore, it is necessary that recorded purchases of different households refer to the same period, otherwise households are not mutually comparable. These considerations led us to the decision to use in the analysis only those households which recorded their purchases over the whole 2 years. Households for which no complete registration was present were discarded.

A further condition for a household to be accepted in the analysis for a product is, of course, that purchases of the product in question were actually made. For fopro and beer the requirement was that a household should have made at least 10 purchases during the 2 years to be included in the analysis, for margarine (because of the greater intensity with which this product is generally bought) the threshold was 20 purchases.

The consequence of this selection procedure, in which only households recording over the whole 2 years and making a certain minimum number of purchases, is that it can no longer be assumed that the purchase data are representative for Dutch households, as was the original Attwood panel. Thus figures about purchasing levels, brand shares, etc. cannot directly be compared with corresponding data from other sources.

In Table 2.1 the resulting number of households and their purchases are given.

Table 2.1 General information about households and purchases

Product	*Number of households in the analysis*	*Total number of purchases recorded*	*Mean number of purchases per household*
Fopro	672	59297	88
Beer	627	27265	43
Marg	1059	130572	123

Evidently the households present in the analysis for the respective products are to a certain extent the same. 378 households were present in the analyses for all 3 products.

2.3 INFORMATION ABOUT PURCHASES

For every purchase of a product by a household the following information is recorded:
— date
— brand
— volume of the unit bought

— number of units bought
— price paid per unit
— special offer (deal) or not
— shop-type

All different combinations of these 7 purchase variables are treated as separate purchases. For example, when on the same day different brands or different volume units were bought of the same product, or the product was bought in different shop-types, such purchases are recorded as different purchases. So it is possible that more than one purchase of a product is recorded for the same day. Purchases, made together with other purchases representing different combinations of the 7 variables mentioned above, of the same product on the same day are here called: multiple purchases.
Only the shop-types (like supermarket, dairyshop, department store, etc.), not the individual shops visited, were recorded by the Attwood panel. In this study, the word shop is used, when shop-type is meant, and it should therefore be remembered that in this way the real number of different shops for a household is under-estimated, in so far as different shops of the same type are visited.

2.4 GENERAL CHARACTERISTICS OF THE MARKETS

It is perhaps pertinent to give some general features of the markets for which the brand choice processes were studied.

Some figures are given in Table 2.2. They all refer to the 2-year period, and are based on the purchases made by families selected for the study.

From (1) and (2) it is clear that the concentration with respect to brands is strongest in the beer market. In the margarine market a great number of brands each have a small share of the market.

The figures for (7) and (8) indicate that with respect to brand switching, fopro is the least dynamic market. Only in 8.9% of all purchases is the brand chosen different from the brand at the preceding purchase. The fact that the mean number of different brands (3) is somewhat higher for fopro than for beer does not imply greater brand switching activity, because the number of purchases per household (= number of possibilities to switch brands) is twice as high for fopro. (7) and (8) clearly indicate that the margarine market shows the most activity on the point of brand mobility. One point that makes the figure for (8) especially high, should be mentioned here. Margarine is a product with two different ways of

Table 2.2 Characteristics of the markets

	Fopro	*Beer*	*Marg*
1. Number of biggest brands, representing 85% of the market by volume	14	8	29
2. Number of biggest brands, representing 85% of the market by number of purchases	12	6	24
3. Mean number of different brands per household	2.88	2.57	4.26
4. Standard deviation of the number of different brands per household	1.87	1.46	3.13
5. Mean number of different shops per household	2.78	2.92	3.36
6. Standard deviation of the number of different shops per household	1.61	1.51	1.90
7. Mean percentage in the purchases represented by a household's favorite brand	84.2	80.0	72.9
8. Percentage of the purchases with brand different from the brand of the previous purchase	8.9	11.4	26.3
9. Deal-purchases as percentage of total purchases	2.5	3.0	6.4
10. Multiple purchases as percentage of total purchases	1.7	14.8	9.5

household use, viz.: on sandwiches and as an ingredient in the preparation of other food. Some households have different brands for different uses and show a kind of alternating behavior between 2 or more brands. These households can be loyal to more than one brand at a time and are called here: multi-loyal. Households not exhibiting this kind of behavior are called: mono-loyal. In chapter 4 we will meet this phenomenon of multi-loyalty again. In that chapter it is shown that to describe the brand choice process for margarine by stochastic models it is necessary to separate households into two groups: mono-loyal and multi-loyal; this separation is effectuated by means of a criterion, developed there. For mono-loyal households, selected according to this criterion, the figure for variable (8) is 14.1.

Multiple purchases can constitute a problem in the study of brand choice processes, because for these purchases the order in which they were made is not known. For fopro this problem is not relevant, because of the very low percentage (10).

At first glance it seems to be a problem for beer, but a closer scruting, reveals that only one third of the multiple beer purchases were made on days in which more than one different brand was bought. In the case of other multiple purchases there were different shops or different volume units, for example, but no different brands. So the percentage of multiple purchases, which constitute a problem in the sense that the order in which different brands were bought is not known, is only 4.9 for beer.

For margarine the percentage of multiple purchases is rather high. As might be expected this percentage is much lower for the sub-class of monoloyal households, i.e.: 4.5.

In the brand choice analyses reported in this study the multiple purchases with unknown brand order are left out, or the order is established at random.

2.5 THE BRANDS CONSIDERED

In most of the brand choice analyses reported in the following chapters special brands of the 3 products are studied. Sometimes the brand under study is the favorite brand for each individual household, and the problem is then that different households have different favorite brands, which makes interpretation of the results and especially the marketing consequences somewhat difficult.

Therefore we have worked mainly with real brands, as they appear in the markets. For reasons of confidentiality they are referred to by symbols, not by their real names. For fopro they are: F1, F2, F3 and F4; for beer: B1, B2, B3 and B4; and for margarine: M1, M2, M3 and M4. The first 3 brands from these sets are real brands; these are national brands, distributed all over the country and promoted by newspapers, television and radio on a national scale. F4, B4 and M4 stand for all other brands. There is one further exception: brand F3 stands for all private brands in the fopro market.

In Table 2.3 information about the market shares over the two years is given. To obtain some idea of the magnitude of change in the markets, the market shares were also computed separately for the 24 4-weekly reporting periods over the 2 years. This resulted in the variable: range, defined as the highest brand share minus the lowest share in the 24 periods.

With brand choice models we are usually forced (to keep the models workable) to use purchases (numbers) as the unit and cannot be concerned with the volume of purchases. Fortunately there is a rather close relationship between the 2 concepts. In this study the market share by volume and the market share by number of purchases per 4-weekly period were calculated. Thus for every brand there were 24 observations of both variables, for which the correlation coefficient was computed. This was done for each of the 4 brands for every product. The 4 computed correlation coefficients for fopro range from .72 to .89, for beer from .50 to .78 and for margarine from .52 to .96.

Table 2.3 Information about brands

Product	*Brand*	*Market share during 2-year period (based on number of purchases)*	*Range of market share over 24 periods*
Fopro	F1	.4127	.0322
	F2	.1362	.0379
	F3	.2046	.0328
	F4	.2466	.0308
Beer	B1	.4318	.0455
	B2	.1155	.0478
	B3	.1526	.0273
	B4	.3001	.0654
Marg	M1	.2621	.0438
	M2	.0794	.0418
	M3	.0437	.0131
	M4	.6149	.0437

This means that – volumes being ultimately the interesting quantities to a company – the approach by number of purchases can give useful results.

2.6 DATA FOR MARKETING VARIABLES

One of the purposes of this study is to examine the influence of the marketing variables, price and advertising, on the brand choice process. Prices could be derived from the panel data, but for advertising figures additional information was needed, and the 'Bureau voor Budgetten-Controle' in Amsterdam obligingly provided estimated advertising expenditures for the different brands in the fopro, beer and margarine markets. These were monthly data over 1967 and 1968, calculated from observed advertising in newspapers, magazines, radio, television, etc.

It is admitted that total expenditure constitutes a crude measure for advertising efforts, because everything is reduced to the money spent, while the quality of advertising with respect to theme, copy, etc. does not play a role. But as there are no objective weights for these elements in advertisements, using money spent as a basis appears to be the only possibility.

3 Models for brand choice processes

3.1 GENERAL OBSERVATIONS

In this chapter a number of brand choice models which have appeared in the marketing literature during the last 15 years are presented. We do not, of course, pretend to give all the different models that have been published, but we believe that with the 5 types to be discussed landmarks in the field of brand choice models are presented. Moreover these are the models applied to the empirical data in chapter 4.

A common feature of the models is that they are stochastic in nature: it is assumed that a consumer does not deterministically take one of the available brands of a product, but that he has a probability distribution over the choice possibilities represented by the various brands.

For each model the formulation, estimation and testing procedures will be given, some applications, if present, mentioned and possible advantages and shortcomings discussed. In section 7 of this chapter some special aspects are dealt with, viz., the measure for goodness of fit of a brand choice model and the numerical minimization procedure to be used.

A rather comprehensive treatment of brand choice models is Massy, Montgomery and Morrison (1970). When reference is made to this work we indicate it as MM & M for convenience. A treatment of brand choice models of earlier date is the survey given by Herniter & Howard (1964). This chapter has mainly the character of a review of literature, although at a number of points original contributions are put forward. As such the demonstration of heterogeneity effects for first order Markov chains (in 3.3.5.2) can be mentioned, the EOS-test on the existence of purchase feedback (3.4.2.3), the estimation procedure for the parameters of the distribution of p in the Heterogeneous Bernoulli Model based on the ROS-concept (a concept earlier used by MM & M for parameter estimation of the Linear Learning Model) (3.4.3.1), the use of the beta-density for the distribution of initial p-values for the Linear Learning Model (in 3.6.2), and the use of measures for goodness of fit of brand choice models, which incorporate the effect of sample size (3.7.1).

Below some definitions and principles, which hold for all models, are given.

A *purchase occasion* or *purchase moment* takes place every time a purchase of the product in question is made. At every purchase occasion a brand choice is made. Choosing a brand is only possible when a purchase occasion takes place. When we speak of 'the probability of choosing a brand' we mean this in the *conditional sense:* i.e., *given that a purchase is actually made.*

A *purchase sequence* or *purchase history* is a series of brands chosen at consecutive purchase occasions. For example, in a two brand market with brands A and B the purchase sequence of a family making 10 purchases may take the form: AABBAAABBB.

When a brand choice process (b.c.p.) or purchase history is spoken of, this is always the b.c.p. or purchase history of *one product.* When it is said, for example, that only one brand was bought during a certain time period, this always refers to the purchases of the product for which the b.c.p. is studied.

In this study we use the words *family*, *consumer* and *household* interchangeably for the performer of the brand choice process.

The models usually require that the number of different brands in a market be reduced to 2. These brands are then indicated by 1 and 0, where 1 stands for the brand we are especially interested in, while 0 indicates all other brands.

By the *parameter p* of a consumer we indicate the probability that he (or she) will choose brand 1. (So with probability $(1-p)$ brand 0 is chosen).

The equi-distant points on the time scale are the indexing numbers of the purchases. The first purchase of a consumer gets the number 1, the second gets the number 2, etc. So a time interval of 2 days between 2 subsequent purchases cannot be differentiated from a time interval of 2 weeks. When a consumer does not make his purchases regularly (for instance one purchase a week), the time scale of the models has no direct relation to calendar time. Also, the time scales of different households are not identical. This makes it difficult to directly relate aggregate figures e.g. market shares, produced by a model, to corresponding quantities in the market place. This is a drawback, but for the purpose of this study, analysing the structure of the brand choice process in which the sequences of brands bought is primarily studied, this is not a great problem.

No attention is paid to the volumes bought on the various purchase occasions. Every purchase counts for one, irrespective of size. In this context the market share is the portion in number of purchase moments, not in volume.

The models used mainly differ in two respects.

Firstly, some models assume homogeneity among consumers with regard to the probability distribution over brands or with regard to the transition probabilities. Other models assume that different consumers may have different probability distributions or different transition probabilities, i.e., these models assume heterogeneity.

The second source of difference is the assumptions made about purchase feedback with respect to brand choice probabilities. Some models assume that a brand chosen at a certain purchase occasion has a definite influence on brand choice probabilities at subsequent purchase occasions. Other models assume that there is no such influence at all: i.e., that subsequent brand choices are independent of each other. These differences refer to the order of the brand choice process assumed by the model.

3.2 THE HOMOGENEOUS BERNOULLI MODEL (HOBM)[1]

A random variable is said to have a Bernoulli distribution when it can take only 2 values: 1 and 0, with probabilities of p and $(1-p)$ respectively.

Now the brand choice process is said to be a homogeneous Bernoulli process when every consumer chooses brand 1 with probability p and brand 0 with probability (1-p), regardless of the purchase history. Then:

$$\text{Prob}\,(X_t = 1 \mid x_{t-1}, x_{t-2}, \ldots) = \text{Prob}\,(X_t = 1) = p,$$

for all values of t.

In this model there is no purchase feedback and every consumer is assumed to have the same probability distribution over the possible brands. In this respect there is homogeneity. It is not necessary to restrict the number of brands in the market to two. In an m-brand market the assumption of the HOBM is that all consumers have the same vector of probabilities $(p_1 \ldots p_m)$, indicating the chances that the respective brands will be chosen.

Because there is no feedback of the brand choices made, p, or the vector $(p_1 \ldots p_m)$, is assumed to remain constant over time.

1. A list of abbreviations and variable names can be found at the back of the book.

Estimation and testing procedures for the HOBM are simple applications of the properties of the binomial distribution.

The stringent requirement that all consumers buy specific brands in the same proportions – an assumption that is contradicted by almost all empirical evidence – means that this model must be qualified as being too simple to offer a realistic description of the brand choice process.

We give the model only for completeness and do not use it further in this study, although the HOBM is a special case of all the models that follow.

3.3 THE HOMOGENEOUS MARKOV MODEL (HOMM)

3.3.1 *Formulation*

In an HOMM a consumer's brand choice depends only on the recent purchase history of that consumer.

The case most commonly treated is the first order Markov Model, where only the last purchase influences current brand choice:

$$\text{Prob}\,(X_t = i \mid x_{t-1}, x_{t-2}, x_{t-3}, \ldots) =$$
$$\text{Prob}\,(X_t = i \mid x_{t-1}).$$

Purchases made before purchase occasion $(t-1)$ are assumed to be irrelevant for the choice made on occasion t.

In a two-brand market there are 2 different purchase histories of length 1 possible. Also 2 different purchases can be made, so that we have probabilities for 4 events:

$$p_{11} = \text{Prob}\,(X_t = 1 \mid x_{t-1} = 1)$$
$$p_{10} = \text{Prob}\,(X_t = 0 \mid x_{t-1} = 1)$$
$$p_{01} = \text{Prob}\,(X_t = 1 \mid x_{t-1} = 0)$$
$$p_{00} = \text{Prob}\,(X_t = 0 \mid x_{t-1} = 0)$$

These so-called transition probabilities can be written in the transition matrix P:

$$\begin{array}{cc|cc} & \text{from} & \text{to } 1 & 0 \\ P = & \begin{array}{c} 1 \\ 0 \end{array} & \multicolumn{2}{c}{\begin{bmatrix} p_{11} & p_{10} \\ p_{01} & p_{00} \end{bmatrix}} \end{array}$$

With the states 1 and 0 put in the margins, the meaning of the probabilities is clear. For example, the element (2, 1) is the probability of going from state 0 to state 1. The rows of the matrix P must sum to unity of course.

The feature of homogeneity in this model lies in the fact that every consumer is assumed to have the same transition matrix.

The application of the HOMM is not restricted to a two-brand market. When there are m different brands in the market, the transition matrix P is an $(m \times m)$ matrix:

$$P = \begin{bmatrix} p_{11} & \cdots & p_{ij} & \cdots & p_{1m} \\ \vdots & & \vdots & & \vdots \\ p_{i1} & \cdots & p_{ij} & \cdots & p_{im} \\ \vdots & & \vdots & & \vdots \\ p_{m1} & \cdots & p_{mj} & \cdots & p_{mm} \end{bmatrix}$$

The element p_{ij} is the probability that brand j will be chosen, given that brand i was chosen at the last purchase.

These transition probabilities have an appealing marketing interpretation. If, for example, p_{ii} is high, this means that a consumer buying brand i has a high probability of remaining a customer for brand i. And if p_{ki} is low, this indicates that it is difficult to get a consumer to switch from brand k to brand i.

For higher order Markov models a greater number of previous purchases influences the current probability distribution over the brands.

A second order Markov model, for example, assumes that the last two purchases have an impact on current brand choice:

$$\text{Prob}\,(X_t = i \mid x_{t-1}, x_{t-2}, x_{t-3}, \ldots) =$$
$$\text{Prob}\,(X_t = i \mid x_{t-1}, x_{t-2}).$$

For this model, there are 4 different purchase histories in a two-brand market: 11, 10, 01 and 00. By defining these histories as states, the transition probabilities can still be written in a one step transition matrix of the following form:

$$\begin{array}{c|cccc} \text{from} \backslash \text{to} & 11 & 10 & 01 & 00 \\ \hline 11 & p_{111} & p_{110} & 0 & 0 \\ 10 & 0 & 0 & p_{101} & p_{100} \\ 01 & p_{011} & p_{010} & 0 & 0 \\ 00 & 0 & 0 & p_{001} & p_{000} \end{array}$$

Here p_{ijk} is the probability that brand k will be chosen, given that brand j was chosen at the last and brand i at the last but one purchase occasion.

This transition matrix contains a number of zero's because some transitions are impossible. For example, when the current purchase history of length two is 11, the purchase history one purchase later can never begin with a 0.

When the brand choice process follows the HOMM, the brand choice process is a Markov process. Because it is a process in discrete time with discrete state space, the brand choice process can be called a Markov chain.

3.3.2 *Some properties of the model*

The theory about Markov processes is well-known – see for example Feller (1957), Kemeny & Snell (1960), Cox & Miller (1965).

We will briefly mention some properties of first order Markov processes here. Proofs can be found in the literature mentioned above.

1. The k-step transition probabilities

The element p_{ij} of the transition matrix P is the probability of making a transition from state i to state j, i.e., of going in one step from i to j. It is also possible to speak of the probability that a consumer who is now in state i (last brand bought being i) will be in state j after k transitions (in the Markov framework every purchase constitutes a transition). This probability is called the k-step transition probability from i to j and is denoted as $p_{ij}^{(k)}$. The matrix containing the k-step transition probabilities is then indicated as $P^{(k)}$.

Now we have the property:

$$P^{(k)} = P^k,$$

which means that the k-step transition probability matrix is the kth power of the one-step transition probability matrix.

2. The steady state distribution

When the Markov chain has m states and Π is a probability (row) vector (with elements $\geqslant 0$ and summing to unity), then we have:

$$\Pi P^k_{k \to \sim} = \alpha, \text{ for every probability vector } \Pi,$$

subject to 2 conditions, which are – in technical terms – that the chain is ergodic and aperiodic. Because in brand choice applications each state is – in one step – attainable from all other states, these conditions are clearly met there. Like Π, α is a probability (row) vector. α gives the distribution over the states after a great number of transitions. Therefore α is called the steady state vector. α is exclusively determined by the elements of P and can be computed from the relation:

$$\alpha P = \alpha,$$

which is in the m state case a question of solving a system of m equations with m unknowns.

The latter property has the following marketing interpretation. Suppose that the m-dimensional vector Π contains the brand shares of the m different brands in the market at a certain point in time. When the brand choice process is a first order Markov process with transition matrix P, then α can be interpreted as being the vector of brand shares after a great number of purchases. α is no more dependent on Π and can be called the vector of equilibrium brand shares.

It is especially this property of a steady-state distribution, which offers the possibility of long term market share forecasting, that makes the Markov process so attractive for brand choice applications. This explains the emphasis put on Markov-processes in brand choice literature.

Of course it only makes sense to compute the equilibrium market shares – for a market where the brand choice process follows the HOMM – when the transition matrix P does not change over time. In that case the Markov chain is said to be stationary.

3.3.3 *Estimation and testing*

In their classic article, Anderson and Goodman (1957) give a rather complete treatment of estimation and testing procedures with respect to Markov chains. An extensive paper by Billingsley (1961) on this subject has also appeared. In this section we will present some of the results which are important for brand choice processes.

For the application of these procedures in empirical brand choice situations, the type of data must be such that individual consumers' purchasing processes can be followed. It must be possible to relate the current brand choice and preceding purchase behavior of a particular consumer. This type of data can be provided by a consumer panel.

3.3.3.1 *Estimation of transition probabilities*

We assume that each individual behaves as a first order Markov chain with m states. Let $N_{ij}(t)$ be the number of individuals who are in state i at time $(t-1)$ and in state j at time t. $N_i(t)$ is the number of individuals in state i at time t.

Now the distribution of the $N_i(t-1)$ individuals, who are in state i at $(t-1)$ over the m possible states: 1, 2, 3, ... j ... m at time t, is a multinomial distribution with probabilities: $p_{i1}, p_{i2}, p_{i3}, ..., p_{ij}, ... p_{im}$. A maximum likelihood estimator for p_{ij} is then the fraction:

$$\hat{p}_{ij} = \frac{N_{ij}(t)}{N_i(t-1)}.$$

When the chain is stationary, i.e., p_{ij} is constant over time, we can use all transitions made at $t = 1, 2, ...$ for the estimation of p_{ij}.

When T transitions of every individual are observed, the maximum likelihood estimator for p_{ij} is then:

$$\hat{p}_{ij} = \frac{\sum_{t=1}^{T} N_{ij}(t)}{\sum_{t=0}^{T-1} N_i(t)} \tag{3.1}$$

In brand choice terms: we find the transition probability from brand i to brand j by dividing the number of times that consumers go from brand i to brand j by the number of times consumers have brand i as the last purchase and make a subsequent purchase.

Generally we work with the empirical purchase histories of different consumers, and these have different lengths. The number of observed transitions T is thus different for different consumers. When for T is read: 'the number of transitions made by the household with the longest purchase history', equation (3.1) is still valuable. In the terms of this expression, with t near to T, only the purchases of a small number of households (with the longest purchase histories) are then used. This constitutes no problem for in the HOMM all consumers are assumed to have the same transition probabilities.

Here we treated the estimation of the transition probabilities for a first order Markov chain. The case of higher order is a quite natural extension. Equation (3.1) can still be used, but now i is not conceived of as a single state but as a combination of states. In brand choice terms, this combination is a certain purchase history and p_{ij} is the probability of a brand j purchase after history i.

3.3.3.2 *Testing the order of the process*

The length of the preceding purchase history, which influences subsequent brand choices, determines the order of the brand choice process.

When the process has the order 2, for example, the probability distribution over the brands on a purchase occasion is determined by the brands chosen at the last and last but one purchase.

Suppose we want to test the null-hypothesis that a brand choice process has order zero (no purchase feedback at all) against the alternative that the order is one (an influence of only the last purchase).

Then it is useful to put the brand switch data into the following scheme (using a two brand market here for simplicity):

from \ to	1	0	
1	N_{11}	N_{10}	$N_{1.}$
0	N_{01}	N_{00}	$N_{0.}$
	$N_{.1}$	$N_{.0}$	N

Here N_{ij} denotes the number of transitions from brand i to brand j $N_{i.}$ is the number of times brand i was purchased and a subsequent purchase was made, $N_{.i}$ is the number of times transitions to brand i were made, irrespective of the brand previously chosen.

When the process is of order zero, i.e., when the last purchase has no influence, the fraction $N_{11}/N_{1.}$ should be equal to $N_{01}/N_{0.}$ (Apart from random deviations, of course).

Anderson and Goodman have shown that to test this the table can be treated as a conventional contingency table. When $\hat{p} = N_{.i}/N$ under the null-hypothesis then:

$$\chi^2 = \sum_{i=0}^{1} \left\{ \frac{(N_{i1} - \hat{p}N_{i.})^2}{\hat{p}N_{i.}} + \frac{(N_{i0} - (1-\hat{p})\, N_{i.})^2}{(1-\hat{p})\, N_{i.}} \right\}$$

has a χ^2 distribution with 1 degree of freedom.

With m brands in the market instead of 2, we get an $(m \times m)$ contingency table with $(m-1)^2$ degrees of freedom.

When the conclusion of the zero order test is that the process is not of order zero, the next question to be answered is: is the process of the order one or higher? i.e.: does only the last purchase have an influence or are former purchases also of importance? Now we test the null-hypothesis that the process is first order against the alternative of order 2.

We must therefore have a test scheme in which it is possible to isolate the influence of the last but one purchase from the last purchase. In the two-brand case therefore we form the following tables:

from | to	1	0
01	N_{011}	N_{010}
11	N_{111}	N_{110}

from | to	1	0
00	N_{001}	N_{000}
10	N_{101}	N_{100}

Here N_{ijk} is the number of times a consumer with purchase history ij made a transition to brand k.

When the last but one purchase has no influence, then the distribution over the columns is not dependent on the row (quite analogous with the table for the test on zero order, just discussed).

We can now treat these two tables as contingency tables again. For each table separately we can compute as before the usual test statistic for independence, which is each distributed according to a χ_1^2 variable, under the hypothesis of order one. As Anderson and Goodman have proved, the sum of these two statistics has the χ_2^2 distribution, under the null-hypothesis mentioned.

When there are m different brands, we get m different tables; the test statistic under the first order hypothesis then has $m(m-1)^2$ degrees of freedom.

Generalisations to a higher order will now be clear. When we want to test the null-hypothesis that an m-state chain is of order k against the alternative that the order is $(k+1)$, we get m^k different tables (one for each purchase history of length k), which each produce – under the null-hypothesis – a χ^2 statistic with $(m-1)^2$ degrees of freedom. The summary statistic then has $m^k(m-1)^2$ degrees of freedom.

3.3.3.3 *Testing the stationarity of the process*

A Markov chain is said to be stationary when the transition probabilities do not change over time. In marketing applications it can be useful to test for stationarity; it is to be expected that the brand choice process will not always remain the same. Transition probabilities may change because of changing marketing conditions.

As outlined above a chi-square test can be developed for this test. Suppose we want to test whether or not a first order Markov chain has the same transition matrix during r subsequent periods. In the two-brand case the following tables can be formed.

from 1 | to	1	0
period		
1	$N_{11}(1)$	$N_{10}(1)$
2	$N_{11}(2)$	$N_{10}(2)$
⋮		
r	$N_{11}(r)$	$N_{10}(r)$

from 0 | to	1	0
period		
1	$N_{01}(1)$	$N_{00}(1)$
2	$N_{01}(2)$	$N_{00}(2)$
⋮		
r	$N_{01}(r)$	$N_{00}(r)$

Here N_{ij} (1) denotes the number of transitions from i to j in period 1.

When the period has no influence, the distribution over the columns is independent of the row in both tables. Again these tables may be treated as the usual contingency tables. For each table the conventional test statistic can be computed, which is under the null-hypothesis of no period influence a χ^2 variable with $(r-1)$ degrees of freedom. The complete stationarity hypothesis is tested by computing the sum of the χ^2_{r-1}-statistics for the separate tables. Under stationarity this sum has a χ^2 distribution with $2(r-1)$ degrees of freedom.

When testing the stationarity of a first order chain in the general m-brand case, we get m different tables, each with $(m-1)$ $(r-1)$ degrees of freedom. This results in a summary χ^2-statistic with $m(m-1)$ $(r-1)$ degrees of freedom.

For extensions and proofs with respect to the tests given in these sections, the reader is referred to Anderson and Goodman (1957) or Billingsley (1961).

3.3.4 *Application of Markov chains to brand choice processes*

In the early sixties the idea of studying brand choice behavior within a Markov framework was put forward by a number of authors.

Amongst the first of these were Herniter and Magee (1961). They gave an exposition of the theory of Markov chains and showed some possible applications to the brand choice process. They also mentioned the possibility for optimization with the aid of Markov programming. The Markov programming procedures outlined by R.A. Howard (1960) might, for example, be applicable in promotional decisions. It can be assumed that different types of promotional strategies give rise to different transition matrices. When these matrices are known, which consumer (dependent on his state) should receive which type of promotion in order to maximize a certain quantity, for example sales, can be determined.

Papers introducing the use of the Markov chain to brand switching in a similar way as Herniter & Magee are those of Maffei (1960) and Harrary & Lipstein (1962).

Examples of applications of Markov chains to empirical consumer buying situations are the following: Styan & Smith (1964) have applied Markov chains to purchases of washing powder by housewives. In their paper the states were not brands but the different types of use of washing powder by housewives. The process thus defined was found to be higher than zero order and stationary; in a paper by Draper and Nolin (1964) the states were different brands of cake mix. They concluded that a first order Markov chain gives an acceptable description of the brand choice process. Again the chain was found to be stationary.

Massy (1966) studied the coffee purchases of members of the Chicago Tribune panel. He computed Markov transition matrices for individual households and for households together. His most important conclusion was that population heterogeneity can cause spurious higher order test results. This subject will be further discussed in the next section.

3.3.5 *Problems of homogeneous Markov models in brand choice processes*

3.3.5.1 *Problems of time and quantity*

In the applications mentioned in the preceding section the Markov chain which described the brand choice process was mostly defined by equidistant epochs, the time period between two subsequent points of time being taken as real calendar periods, e.g., one week. This causes some problems. Firstly, in which state should a consumer be classified when he makes more than one purchase in a certain period and these purchases consist of different brands? A possible solution is to define the state as the brand most purchased during a period, but then the interpretation of the results is of course difficult. Secondly, when a consumer does not buy at all in a certain period, what is then his state? Some authors have solved this question by introducing an artificial state of 'did not buy' (see e.g. Harrary and Lipstein (1962)). This is not so satisfactory, because the state 'did not buy' is quite different in character from the state: 'the brand bought'. When this approach is followed, the transition matrix may change merely because of a change in frequency of purchases, i.e., without a change in switching patterns between real brands, which is confusing. Thirdly, only the brands bought count, not the number of units bought.

R.A. Howard has suggested solving these problems by using a Semi-Markov process instead of a Markov process to describe brand choice behavior (see Howard (1963) and Herniter & Howard (1964)). In the Markov process thus far treated, transitions were assumed to take place at equi-distant points in time, i.e., the time space was discrete. In Semi-Markov processes, also called Markov Renewal processes, on the contrary, the time between transitions is not fixed but is a random variable with a probability density function (see Cox & Miller (1965) or Cox & Lewis (1966)). In the most general case there is a p.d.f. for each type of transition. So the probability of making a transition from i to j, which we denoted by p_{ij}, is supplemented by the p.d.f. of the time (holding time) to move from i to j: $h_{ij}(t)$. A special class is formed by the so-called 'Markov processes in continuous time'. In this case $h_{ij}(t)$ is the p.d.f. of an exponential distribution.

It is clear that with this extension the problems of fitting the brand choice process into a Markov model are considerably diminished, especially when it is assumed that inter-purchase times may take all values $\geqslant 0$. On the other hand, estimation problems regarding the distributions of the holding times should not be disregarded, a large number of observations on all types of inter-purchase times being necessary.

No applications to brand choice processes of this Semi-Markov process in its general form seem to be present in the literature. Herniter's (1971) approach has some features of a Semi-Markov model when he describes a brand choice process in which the inter-purchase times follow an Erlang distribution while brand choice occurs according to a first-order Markov process. However, in this application brand choice and inter-purchase times are assumed to be independent of each other, which constitutes a major simplification of the general Semi-Markov assumptions. In section 7.4.4. we return to the application possibilities of Semi-Markov processes to brand choice processes.

As stated in section 3.1, in this study we work with a transformed time scale in which the equi-distant points are the purchase moments irrespective of the number of days that constitute the inter-purchase time, while only the brand bought and not the quantity is of interest. This means, that we are hardly troubled by the difficulties mentioned. But – as pointed out – the direct relation to real market figures is sacrificed.

In an extensive paper Ehrenberg (1965) outlines a pessimistic view on the possibilities of Markov chain applications in brand choice processes. He mentions the problems referred to above, but a new argument he puts forward is the following:
In the Markov chain concept of brand choice the crucial quantities are the transition probabilities. For example p_{ij} is the probability of purchasing brand j, given that i was bought on the last occasion. Now Ehrenberg's question is: why do we look at those 'forward' probabilities? Another approach, in Ehrenberg's view not more unnatural, is to study the so-called backward probabilities: i.e., q_{ij}, where: $q_{ij} = \text{Prob}$ (j being chosen at a certain purchase occasion | i is chosen at the next purchase occasion).

He then demonstrates that for a Markov chain with stationary forward transition probabilities, the backward transition probabilities (called Vokram probabilities by Ehrenberg) are not stationary at all.

This non-symmetric behavior of the Vokram probabilities makes Ehrenberg doubt the possibilities of Markov chain applications in brand choice processes.

In this connection the following observations can be made. The phenomenon of backward transition probabilities not being stationary is not surprising. Kemeny & Snell (1960) proved that the backward transition probabilities of a chain with constant forward transition probabilities are only constant when the Markov chain is in equilibrium.

Of course it can not be taken on trust that a Markov chain with stationary forward transition probabilities should be used to describe a brand choice process. As holds for the use of all models in general: an important condition for application must be that the model gives a satisfactory description of empirical reality.

3.3.5.2 *The aggregation problem*

In this section a major difficulty with respect to the HOMM, which can arise with estimation and testing in empirical situations, is discussed. This problem essentially refers to the assumption of homogeneity, and it becomes relevant when this assumption is not justified. Unfortunately, it can not be directly observed when the population is not homogeneous in relation to transition probabilities, and in this case aggregation can lead to misleading conclusions.

The problem was first detected by Frank (1962); he discovered the phenomenon and called it: spurious contagion.

The problem can best be illustrated by a numerical example.

Suppose that a consumer population consists of two different types of consumers: type I and type II. There is a two brand market with brands 1 and 0. Both groups of consumers buy the two brands according to a Bernoulli process with parameter p, and the only difference between type I and type II is that they have different values for p.

Let:

$$\text{Prob (brand 1 is purchased} \mid \text{consumer is type I)} = \text{Prob}(1 \mid \text{I}) = .9$$
$$\text{Prob (brand 1 is purchased} \mid \text{consumer is type II)} = \text{Prob}(1 \mid \text{II}) = .2$$

Let the population contain 50% type I and 50% type II consumers.

Suppose that for this population of consumers the brand choice process is studied empirically and the test on a zero versus a first-order Markov process (section 3.3.3.2) is performed, where the population is considered as homogeneous. The purchase histories are collected and mixed, irrespective of the type of consumer. In the test the transition probabilities, like $p_{11} = \text{Prob}(1 \mid 1)$ and $p_{01} = \text{Prob}(1 \mid 0)$, are estimated with the fractions empirically found. Under the null-hypothesis of zero order $p_{11} = p_{01}$, while in the case of a positive influence from the preceding purchase we can expect: $p_{11} > p_{01}$.

For the population just described we can predict the values of $\hat{p}_{11}$ and $\hat{p}_{01}$ (apart from random disturbances of course). Let us take the computation of $\hat{p}_{11}$. For this computation we consider all purchase sequences of length two. The computation then consists of 2 steps. First, all sequences starting with a 1 are collected. Let there be $N_{1.}$ of such sequences. Then how many of this latter set of sequences also have a 1 on the second place is counted. Let this number be N_{11}.

Now $\hat{p}_{11} = \dfrac{N_{11}}{N_{1.}}$, according to equation (3.1)

How does this procedure work out for the population described above? First we examine the origin (with respect to the type of consumer) of a purchase sequence starting with a 1. Therefore we consider:
$\text{Prob}(\text{I} \mid 1) = \text{Prob (consumer is of type I} \mid \text{first purchase} = 1)$
Following Bayes' theorem we get:

$$\text{Prob}(\text{I} \mid 1) = \frac{\text{Prob}(1 \mid \text{I})\ \text{Prob}(\text{I})}{\text{Prob}(1)} =$$

$$= \frac{.9 \times .5}{.9 \times .5 + .2 \times .5} = .82$$

and $\text{Prob}(\text{II} \mid 1) = 1 - .82 = .18$

So we conclude, that a sequence starting with a 1 stems with probability .82 from a Type I consumer and with a probability of .18 from a Type II consumer.

A Type I consumer purchases on the next purchase occasion brand 1 with probability .9 and a Type II consumer with probability .2. In general the probability of a brand 1 purchase on the second occasion is:

$$\begin{aligned} &\text{Prob}\,(1\,|\,\text{I})\,\text{Prob}\,(\text{I}) + \text{Prob}\,(1\,|\,\text{II})\,\text{Prob}\,(\text{II}) \\ &= .9 \times \text{Prob}\,(\text{I}) + .2 \times \text{Prob}\,(\text{II}) \end{aligned}$$

So the probability of a brand 1 purchase after a brand 1 purchase =

$$P(1\,|\,1) = .9 \times .82 + .2 \times .18 = .77.$$

In the same way we get:

$$\text{Prob}\,(\text{I}\,|\,0) = \frac{\text{Prob}\,(0\,|\,\text{I})\,\text{Prob}\,(\text{I})}{\text{Prob}\,(0)} = .11$$

and $\text{Prob}\,(\text{II}\,|\,0) = .89$

Therefore:

$$\begin{aligned} &\text{Prob}\,(1\,|\,0) \\ &= .9 \times .11 + .2 \times .89 = .28 \end{aligned}$$

So when, for the population specified, the Markov transition probabilities are computed in the usual way, we get:

$$\begin{aligned} \hat{p}_{11} &= \text{Prob}\,(1\,|\,1) = .77 \\ \hat{p}_{01} &= \text{Prob}\,(1\,|\,0) = .28 \end{aligned}$$

This might lead to the conclusion that the most recent purchase has a definite influence and the process is of an order greater than zero. But this conclusion would be clearly wrong, because we have a consumer population which follows a Bernoulli brand choice process.

In fact it is true that a brand 1 purchase has a greater chance of being followed by a 1 than by a brand 0 purchase. This is not because the probabilities of the consumers change as a consequence of the brand 1 resp. 0 purchase, but because there is a high probability for a brand 1 purchase that it originates from a consumer who already has a high probability of a brand 1 purchase.

So when the homogeneous Markov model is assumed for the brand choice process the process may appear to be first order, while it is in fact zero order, as a consequence of the aggregation from a heterogeneous population. Massy (1966) has empirically demonstrated this possibility.

The argument extends to higher order situations, as the following example demonstrates. Suppose a consumer population which consists of 50% Type I and 50% Type II consumers.

Both types follow a first order Markov brand choice process in a two-brand market, with, however, different transition matrices.

The transition matrices are:

Type I:

from \ to	1	0
1	.6	.4
0	.1	.9

with steady state distribution (.2, .8)

Type II:

from \ to	1	0
1	.8	.2
0	.4	.6

with steady state distribution (2/3, 1/3)

Let the process be in equilibrium for both groups.

Suppose that for an empirically observed brand choice process of this population the test (discussed in 3.3.3.2) with a null-hypothesis of first order versus the alternative of order 2 is performed. We assume that the brand choice processes are in the steady state. As an example we take the comparison of Prob $(1 \mid 00)$ and Prob $(1 \mid 10)$, which is a part of the test. Under the null-hypothesis we have:

$$\text{Prob}\,(1 \mid 00) = \text{Prob}\,(1 \mid 10)$$

Under the alternative of a positive influence of the last but one purchase we expect:

$$\text{Prob}\,(1 \mid 00) < \text{Prob}\,(1 \mid 10)$$

For the population specified we have:
Prob $(00 \mid \text{I})$ = Prob (first purchase $= 0$ and then a transition to $0 \mid \text{I}) = .8$ (steady state proportion) $\times .9 = .72$.
In the same way:

$$\text{Prob}\,(00 \mid \text{II}) = .20$$
$$\text{Prob}\,(10 \mid \text{I}) = .08$$
$$\text{Prob}\,(10 \mid \text{II}) = .13$$

For the marginal probabilities we have:
$\text{Prob}\,(00) = \text{Prob}\,(00 \mid \text{I})\,\text{Prob}\,(\text{I}) + \text{Prob}\,(00 \mid \text{II})\,\text{Prob}\,(\text{II}) = .46$

Analogously Prob (10) = .11

$$\text{Now:}\quad \text{Prob}\,(\text{I}\,|\,00) = \frac{\text{Prob}\,(00\,|\,\text{I}) \times \text{Prob}\,(\text{I})}{\text{Prob}\,(00)} = .78$$

and Prob (I | 10) = .36

So the probability that the sequence 00 will be followed by a 1 is: $.78 \times .1 + .22 \times .4 = .17$

The probability that 10 will be followed by 1 is: $.36 \times .1 + .64 \times .4 = .29$

Thus we get:

Prob (1 | 00) = .17
Prob (1 | 10) = .29

The values estimated for these probabilities in an empirical situation will be approximately equal to these figures. Therefore the brand choice process of the population given will appear to be at least second-order, while all consumers in the population have a first-order brand choice process. Again this is a consequence of aggregation from a heterogeneous population.

Because the order is an important characteristic of the brand choice process, a major difficulty with the use of homogeneous Markov models is that heterogeneity in the population may make the process seem to be of a higher order than it actually is.

This difficulty can be overcome in two ways. One can divide the population into homogeneous sub-populations. In the extreme case this means taking each household separately. The other possibility is to build the heterogeneity into the model, as in the models of section 3.5.

3.4 THE HETEROGENEOUS BERNOULLI MODEL (HEBM)

3.4.1 *Formulation*

In the HEBM a consumer is assumed to buy brand 1 with probability p and brand 0 with probability $(1-p)$, irrespective of the purchase history:

$$\text{Prob}\,(X_t = 1\,|\,x_{t-1}, x_{t-2}, x_{t-3}, \ldots) = \text{Prob}\,(X_t = 1) = p$$

There is no purchase feedback.

Thus far the model is identical with the HOBM discussed in section 3.2. However, an essential difference is that in the HEBM p may differ among consumers. It can be said that p is in a certain way distributed over the consumers, let us say with p.d.f. $f(p)$.

One possible standard distribution which can be used for the distribution of p is the beta distribution. The beta distribution has two parameters, indicated here as α and β. The probability density function (p.d.f.) of the beta distribution is:

$$b(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\,\Gamma(\beta)}\, x^{\alpha-1}(1-x)^{\beta-1} \; (0 \leqslant x \leqslant 1) \tag{3.2}$$

with raw moments:

$$\mu_1 = \frac{\alpha}{\alpha+\beta} \tag{3.3}$$

and

$$\mu_{i+1} = \mu_i \left(\frac{\alpha+i}{\alpha+\beta+i}\right) \quad (i = 1, 2, \ldots) \tag{3.4}$$

The values a beta variable can take lie between 0 and 1, which is an essential condition for the probability p. Further, the beta distribution is very flexible which means that a great number of different forms can be taken, depending on the values of α and β. The p.d.f. can take the U-form $(0<\alpha,\ \beta \leqslant 1)$, the inverse U-form $(\alpha, \beta > 1)$, the J-form $(0<\beta \leqslant 1<\alpha)$, the inverse J-form $(0<\alpha \leqslant 1<\beta)$ and the rectangular form $(\alpha = \beta = 1)$. See Kendall & Stuart I (1969, p. 150, 151).

Because of this flexibility the beta distribution is suitable for describing the distribution of p in the population.

MM&M use the name Compound Bernoulli Model for the HEBM. They speak of beta-heterogeneity, when $f(p)$ is the beta-density $b(p)$.

3.4.2 *Testing procedures*

The characteristic of the brand choice process checked by the following tests, is the influence of previous purchases on current brand choice. The null-hypothesis is the HEBM-assumption, viz., that there is no influence i.e., that the process is zero order. Three different tests which can be used are discussed below.

3.4.2.1 *The run test*

Run tests are well-known in non-parametric statistics (see for example Bradley, 1968, chapter 11). They can be used to analyse sequences of events of a binary type, such as brand choice processes in a two-brand market. Suppose a consumer has the purchase history: 011001011100010. In run test terminology every unbroken sequence of the same brand constitutes a run. In this example there are 9 runs.

The test statistic studied is U, which stands for the number of runs in a sequence. The expected value and the variance of U, $E(U)$ and Var (U) respectively, under the null-hypothesis of constant probability of the occurrence of a 1, can be expressed by the variables n_1 and n_0, denoting the numbers of ones, resp. zeros in the sequence.

The expressions are:

$$E(U) = \frac{2n_1 n_0}{n_1+n_0} + 1 \tag{3.5}$$

$$\mathrm{Var}(U) = \frac{2n_1 n_0(2n_1 n_0 - n_1 - n_0)}{(n_1+n_0)^2\,(n_1+n_0-1)} \tag{3.6}$$

Now we can compare the actual number of runs of a sequence with the number expected under the hypothesis that the probability of the occurrence of a one is constant during the series of purchases which constitute the sequence. Should the latter probability not be constant but increase after a brand 1 purchase and decrease after a zero – as might be expected with purchase feedback – then there will be a tendency for the ones and zeros to stick together, which will result in a relatively small number of runs.

Whether or not the actual value of U differs more from the expected values than would occur by chance can be tested. For small values of n_1 and n_0 there are tables available, for higher values the approximation of

$$\frac{U-E(U)}{\sqrt{\mathrm{Var}(U)}} \tag{3.7}$$

by a standard normal variable can be used. By applying a correction for continuity this approximation can be slightly improved (see Bradley, ibid, p. 262).

As mentioned previously, in brand choice processes the question is usually whether or not there is a tendency for U to be lower than expected, which means that the test is against a left-sided alternative.

Because every family can have its own p, the test must be performed for every family separately. For an overall indication of the importance of purchase feedback in the population of buyers, the portion of families for whom the null-hypothesis is rejected can be considered.

The use of run tests in brand choice processes was first demonstrated by Frank (1962).

One difficulty attached to their use is their low power, at least when sample sizes (lengths of sequences) are modest. From the power functions given by Bateman (1948) an impression can be obtained. Suppose that a consumer chooses the brands 1 and 0 according to a first-order Markov chain with: $p_{11} = .80$ and $p_{01} = .40$. When the purchase process of this consumer is observed over 20 purchases, containing 10 brand 1 and 10 brand 0 purchases, and the run test is performed, the probability that the hypothesis of sequential independence will be rejected is about .60. This is rather low for a process which is so obviously non-Bernoulli. Of course the power of the runtest does increase with sample size; Bateman, however, does not give results for the effect of sample size.

3.4.2.2 *Conditional model probabilities* (CMP) *test*

The CMP test, discussed here was developed by MM&M. In this test the essential quantity considered is – given a certain purchase history – the probability that the next purchase (immediately following that history) is brand 1. In general, of course, this probability is dependent on the preceding purchase history but it is also dependent on the mechanism which produces the purchase sequence in question.

When it is assumed that this mechanism works according to a certain brand choice model M, the probability just mentioned is conditional on M. Therefore MM&M call it: the conditional model probability.

We denote a purchase history by x. When sequences of length 4 are considered x can take such forms as:

0100, 1001, 1100, etc.

The CMP, i.e., the probability of a brand 1 purchase given purchase history x, and the model M can be indicated as $P_M(1 \mid x)$.

In what follows, we assume that $M =$ HEBM and drop this symbol for convenience. The assumption is then that the brand choice process follows the HEBM.

3.4.2.2.1 *Arbitrary distribution*

Here p is assumed to have the arbitrary p.d.f.: $f(p)$. Suppose we observe the purchase history x of a certain family. Now we have 2 pieces of information about this family: the observed purchase history x and the fact its p-value is a sample from a distribution with p.d.f. $f(p)$.

This information can be combined by means of a Bayesian approach (see Raiffa & Schlaifer (1961)). By using Bayes' Rule the knowledge of the family's distribution of p is updated by the information contained in x.

We call $f(p)$ the prior distribution. The posteriori distribution, which arises after the information from x has been incorporated, is denoted as $f_1(p)$. So $f_1(p)$ is the p.d.f. of p, given that x is observed. Let $l(x|p)$ denote the likelihood that purchase history x is observed, given that the parameter of the Bernoulli brand choice process is p.

By Bayes' Theorem:

$$f_1(p) = \frac{l(x \mid p)\, f(p)}{\text{Prob}\,(x)} = \frac{l(x \mid p)\, f(p)}{\int_0^1 l(x \mid p)\, f(p)\, \mathrm{d}p} \tag{3.8}$$

The conditional model probability is:

$$\text{Prob}\,(1 \mid x) = \int_0^1 p f_1(p)\, \mathrm{d}p = \frac{\int_0^1 p l(x \mid p)\, f(p)\, \mathrm{d}p}{\int_0^1 l(x \mid p)\, f(p)\, \mathrm{d}p}$$

From this equation it can be seen that the CMP's for different x's, say x_1 and x_2, are equal when their likelihood functions are equal: i.e.,
When

$$l(x_1|p) = l(x_2|p)$$

then

$$P(1|x_1) = P(1|x_2)$$

Now consider for example: $x_1 = 1001$ and $x_2 = 0110$.
Under the hypothesis of the HEBM, we have:

$$l(x_1 \mid p) = l(x_2 \mid p) = \binom{4}{2} p^2(1-p)^2,$$

a conventional binomial probability.

Thus, under the HEBM hypothesis, the purchase sequences 1001 and 0110 have the same likelihood and the CMP's corresponding with these 2 purchase histories are equal.

The more general conclusion is now clear, i.e., that the CMP's – under the HEBM hypothesis – for purchase sequences of the same length – are equal when the numbers of ones in the sequences are equal. The exact configuration of zeros and ones is then no longer important.

This constitutes the basis of the test. The test examines if the equality of CMP's as expected under the HEBM hypothesis is in agreement with the data. If this is not the case, the HEBM is rejected.

When sequences of length 4 are considered, the 3 following groups of purchase histories have – under the HEBM hypothesis – the same CMP's.

Group I	*Group II*	*Group III*
0111	0011	0001
1011	0101	0010
1101	1001	0100
1110	0110	1000
	1010	
	1100	

With empirical data it can be tested, for each group separately, if the CMP is equal for all sequences in a group. This test is performed with a contingency table. Under the HEBM hypothesis the resulting statistic has a χ^2 distribution with resp. 3, 5 and 3 degrees of freedom. The overall statistic (the sum of the 3 from the respective tables) is then a χ^2_{11} variate.

Sequences of length 4 were used in the above presentation because the purchase histories on which the test is applied in chapter 4 have that length. But this length is, of course, optional. Generally speaking, shorter sequences result in less degrees of freedom for the χ^2 test, while the use of longer sequences diminishes the number of observations.

3.4.2.2.2 *Beta distribution*

Thus far no assumptions have been made about the distribution of p. The p.d.f. $f(p)$ was arbitrary. Now we assume that $f(p)$ is the beta density $b(p)$ with parameters α and β. Let $x = (n, r)$ indicate a purchase sequence of length n, containing r ones $(r \leqslant n)$. As seen above, the number of ones is the only information about a purchase sequence needed.

For the expressions in equation (3.8) we have now:

$$l(x \mid p) = l((n, r) \mid p) = \binom{n}{r} p^r (1-p)^{n-r},$$

and

$$f(p) = b(p) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\,\Gamma(\beta)}\, p^{\alpha-1}(1-p)^{\beta-1}$$

It can be shown that in this case $f_1(p)$, as derived from equation (3.8), is again a beta density with parameters $(\alpha+r)$ and $(\beta+n-r)$.

So when we know of a household a priori that its p-value is a sample from the beta distribution with parameters α and β, and then a purchase sequence of length n is observed which contains r ones, we can update our knowledge about that family's p by specifying its distribution as being a beta distribution with parameters $(\alpha+r)$ and $(\beta+n-r)$.

The CMP in this case is simply:

$$\text{Prob}\,(1\,|\,(n, r)) = \int_0^1 p f_1(p)\,\mathrm{d}p = \frac{\alpha+r}{\alpha+\beta+n},$$

which is the expected value of the beta variable with parameters $(\alpha+r)$ and $(\beta+n-r)$.

For example, when $x = 1010$, we have

$$\text{Prob}\,(1\,|\,(4, 2)) = \frac{\alpha+2}{\alpha+\beta+4}$$

It will be clear that under this beta distribution assumption the power of the test on order zero is increased, compared with that of an arbitrary distribution. Now it can be tested for each possible sequence of a certain length, if the observed fraction of ones which follow that sequence is in agreement with the CMP under the Bernoulli-beta hypothesis. This is an extension, compared with the test outlined in the previous chapter in which only the equality of the CMP for particular sequences was tested.

When α and β are given they can be directly used; if not they have to be estimated from the data, which results in the loss of 2 degrees of freedom.

3.4.2.3 *A test on equal occurrence of sequences* (EOS)

Here we give a test which can be conceived of as a variant of the previous one. In this test not the CMP of the occurrence of a one after a certain purchase sequence is studied, but the relative occurrence of the sequence itself.

Suppose we study purchase sequences of length 4. For a family which purchases brand 1 with constant probability p, the probability that an observed sequence will have the configuration 1000 is $p(1-p)^3$.

When all purchase sequences of length 4 of a buying population with an HEBM brand choice process and p.d.f. $f(p)$ are considered, the expected relative frequency of the sequence 1000 is:

$$E(1000) = \int_0^1 p(1-p)^3 f(p)\,\mathrm{d}p$$

But in the same way, for example

$$E(0010) = \int_0^1 p(1-p)^3 f(p)\,\mathrm{d}p,$$

which is exactly the same.
This can directly be extended to:

$E(1000) = E(0010) = E(0100) = E(0001)$, under the hypothesis of the HEBM. This holds irrespective of the p.d.f. $f(p)$.

An analogous reasoning can be applied to sequences of length 4 with 2, resp. 3, ones.

Thus, under the null-hypothesis of the HEBM, every sequence within each of the 3 groups of sequences, I, II, and III (given in 3.4.2.2.1) is expected to occur with the same frequency.

For group I, for example, the distribution of the numbers of sequences 0111 to 1110 is then multinomial with $p_{0111} = \dots = p_{1110} = \frac{1}{4}$, where the sample size is the total number of sequences in group I. The EOS test uses these implications of the HEBM to test if the HEBM is compatible with the data. For each of the 3 groups mentioned it can be tested whether or not the expected equal frequencies are in agreement with the data.

Again this leads to a chi-square test with 3, 5 and 3 degrees of freedom for groups I, II and III respectively. The summary statistic has – under the HEBM hypothesis – a χ^2_{11} distribution.

3.4.3 *Estimation*

In the HEBM the estimation problems refer to the parameters of the distribution of p. When the type of distribution (e.g. normal, beta or gamma) is not known in advance, a choice regarding this type has to be made and the parameters must be estimated. It then remains to be tested whether or not the distribution in question gives a satisfactory fit.

3.4.3.1 *Estimation based on the relative occurrence of sequences* (ROS)

The procedure presented here is based on the relative frequency with which certain purchase sequences occur. The concept of using the relative occurrence of sequences, which is briefly designated here as the ROS-concept, was applied earlier by MM&M to develop a parameter estimation procedure for the Linear Learning Model.

The estimation method to be presented in the following is the so-called Minimum Chi-square Method.

Consider Prob $(r\,|\,(n, p))$, i.e., the probability that a purchase sequence of length n for a family following a Bernoulli brand choice process with parameter p contains r ones.

We have:

$$\text{Prob}\,(r\,|\,(n, p)) = \binom{n}{r} p^r (1-p)^{n-r}$$

Over the whole population of buyers, the corresponding quantity is:

$$\text{Prob}\,(r\,|\,(n, f(p)) = \int_0^1 \binom{n}{r} p^r (1-p)^{n-r} f(p)\,\mathrm{d}p \tag{3.9}$$

Because r can take the values 0, 1, ... n, there are $(n+1)$ such probabilities for the sequence length n, for sequences with resp. 0, 1, ... and n ones.

These types of sequences, viz., those with 0, 1, ... n ones – are the types which are considered in the following. Let us call these types: $s_0, s_1, \ldots, s_r, \ldots, s_n$ – so that s_r is a sequence type with r ones in a sequence of length n.

When $f(p)$ is specified with given parameters, the expected relative frequencies of the sequence types $s_0 \ldots s_n$ can be computed from equation (3.9). They can be called theoretical relative frequencies and can then be compared with the actual relative frequencies of $s_0 \ldots s_n$, so that by means of a χ^2 test it can be tested whether or not agreement exists.

Usually, however, the parameters of $f(p)$ are not known. Then the principle discussed above can be used for their estimation.

Let the distribution of p (with p.d.f. $f(p)$) have s parameters, contained in the v-dimensional vector θ.

Now the probability Prob $(r\,|\,(n, f(p))$ $(r = 0, \ldots, n)$ is a function of θ. We denote this as:

$$g_r(\theta) = \text{Prob}\,(r\,|\,(n, f(p)))\;(r = 0, \ldots n)$$

Let the observed (actual) frequencies of $s_0, \ldots s_n$ be indicated as: $a_0, a_1, \ldots, a_n$.

The test statistic G for the goodness of fit between theoretical and observed relative frequencies of $s_0, \ldots s_n$, for a given parameter vector θ, is:

$$G(\theta) = \sum_{r=0}^{n} \frac{(g_r(\theta) - a_r)^2}{g_r(\theta)} \cdot N, \tag{3.10}$$

where N is the total number of sequences of length n that is observed.

Now the parameters contained in θ can be estimated by minimization of $G(\theta)$ with respect to the elements of θ, i.e., the estimated parameter values are such that they make the discrepancy between theoretical and observed relative frequencies as small as possible. This so-called Minimum Chi-square Method of estimation produces estimates which are consistent and asymptotically efficient (see Cramer, 1957, p. 424–434).

The functions $g_r(\theta)$ will often be such that a numerical minimization procedure must be applied. One possible algorithm that can be used is given in section 3.7.2.

Min $G(\theta)$ is distributed according to a χ^2 variable with $(N-v-1)$ degrees of freedom, under the distribution given. This provides us with a test for the fit of the distribution specified. So here estimation and testing are performed by the same procedure.

In the case that $f(p)$ is the beta density $b(p)$ and θ is the two-dimensional vector containing the parameters α and β, the estimation procedure goes as follows:

We consider purchase sequences of length 4.
So we have: $v=2$, $n=4$.

For $g_r(\theta)$ we get:

$$g_r(\alpha, \beta) = \text{Prob}\,(r\,|\,(4, b(p))) = \int_0^1 \binom{4}{r} p^r(1-p)^{n-r} b(p)\,\mathrm{d}p,$$

which can be reduced to the following expressions:

r	$g_r(\alpha, \beta)$
0	$1-4\mu_1+6\mu_2-4\mu_3+\mu_4$
1	$4(\mu_1-3\mu_2+3\mu_3-\mu_4)$
2	$6(\mu_2-2\mu_3+\mu_4)$
3	$4(\mu_3-\mu_4)$
4	μ_4

Now the moments μ_1, μ_2, μ_3 and μ_4 are simple functions of α and β (see equations (3.3) and (3.4)). Hence, given α and β, $g_r(\alpha, \beta)$ $(r=0, \ldots 4)$ can be computed.

When the actual relative frequencies $a_0, \ldots a_4$ are given, $G(\alpha, \beta)$ can also be directly evaluated for a given α and β.

Then by means of a minimization procedure the values of α and β which minimize $G(\alpha, \beta)$ can be found.

Those values are the estimates for the parameters of the beta distribution, while the lowest value of $G(\alpha, \beta)$ itself is the test statistic, which – under the beta distribution hypothesis – has a χ_2^2 distribution.

3.4.3.2 *Estimation of* p *for individual consumers*

When the individual purchase histories are long enough, the fraction of purchases represented by brand 1 can be computed for every consumer and taken as an estimate for that consumer's p-value.

Such individual p values can then be considered as observations from the distribution with p.d.f. $f(p)$. With these observations the mean and variance of the distribution can be estimated and, if necessary, a test for goodness of fit performed.

3.4.4 *Applications of the* HEBM

The idea that a consumer may have a constant probability to choose a certain brand, while different consumers are allowed to have different probabilities, was put forward by Frank (1962). He placed this concept against the concept of a higher order process and suggested that the conclusion that a brand choice process is higher order may be a consequence of what he calls: 'spurious contagion'. We discussed this in section 3.3.5. On the basis of run test results Frank concluded that for the population of buyers he examined the brand choice processes of many families were consistent with the constant probability hypothesis.

Massy (1966) also suggested the possibility that for the purchase data he considered the b.c.p. was Bernoulli. The latter data were investigated again by MM&M (chapter 4) and it was found, with the test given in 3.4.2.2, that there were considerable departures from the Bernoulli hypothesis.

The assumption of the HEBM, i.e., that a consumer always chooses a brand with the same probability, is rather stringent. If it holds, this can only be so for a limited time interval, for it is hardly imaginable that a consumer would buy a certain brand with the same probability during his whole lifetime.

A great advantage of the HEBM is the possibility for each consumer to have his own p-value.

Therefore, as a starting point for more complex models which have the same heterogeneity with respect to p, the HEBM is very useful.

3.5 HETEROGENEOUS MARKOV MODELS (HEMM)

Since the HEBM has the appealing feature that consumers are assumed to be heterogeneous with respect to p, while the HOMM has the attractive property that purchase feedback is built into it, it seems sensible to combine these two models: it is reasonable to expect that in many brand choice processes the properties of both population heterogeneity and purchase feedback are present.

The HEMM's proposed by Morrison (1966) are an attempt to accommodate this 'hybridization'.

3.5.1 *Formulation*

In the HEBM, every consumer has a parameter p which may be different for different consumers; analogously in the HEMM every consumer has his own first order Markov transition matrix and different consumers may have different transition matrices.

So every individual consumer follows a first order Markov brand choice process with his own specific transition matrix, which has the following general form:

$$\begin{array}{c|cc} \text{from} \backslash \text{to} & 1 & 0 \\ \hline 1 & p & (1-p) \\ 0 & (1-q) & q \end{array}$$

This transition matrix is characterized by the two parameters p and q.

But the heterogeneity with respect to the transition matrices cannot go too far. When p and q are both allowed to vary, without an *exact* relation between them, a bivariate distribution is needed to describe the distribution of p and q in the population. This makes handling of the process very complicated. Therefore in the models which follow restricting assumptions are introduced.

When a fixed relation between p and q is assumed one parameter disappears and the distribution of the different transition matrices in the population is determined only by the distribution of p.

The most simple relation is: $q = p$.
Then the transition matrix becomes:

$$\begin{array}{c|cc} \text{from} \backslash \text{to} & 1 & 0 \\ \hline 1 & p & (1-p) \\ 0 & (1-p) & p \end{array}$$

This model can be called: the Symmetric Heterogeneous Markov Model.

It is possible however to give the model a somewhat more general character. Below follows a presentation of the types given by Morrison (1966a).

1. The Brand-Loyal Model (BLM)
Here the relation assumed between p and q is:

$q = 1 - kp \ (0 \leqslant k \leqslant 1)$.

Then the transition matrix becomes:

$$\begin{array}{c|cc} \text{from} \backslash \text{to} & 1 & 0 \\ \hline 1 & p & 1-p \\ 0 & kp & 1-kp \end{array}$$

Every consumer may have his own value for p, but the coefficient k is the same for all consumers.

According to this model, consumers who have a high probability of going from brand 1 to brand 1 also have a (relatively) high probability of going from brand 0 to brand 1.

Note that when $k = 1$, we have the HEBM-situation again.

2. The Last-Purchase Loyal Model (LPLM)
Now the relation between q and p is:

$q = kp \qquad (0 \leqslant k \leqslant 1)$

The transition matrix becomes:

$$\begin{array}{c|cc} \text{from} \backslash \text{to} & 1 & 0 \\ \hline 1 & p & (1-p) \\ 0 & (1-kp) & kp \end{array}$$

As in the BLM, k is equal for all consumers. In this model a consumer with a high repeat purchase probability for brand 1 also has a high repeat probability for brand 0.

For both models it holds that the value of p determines a family's transition matrix. p is distributed in the population according to a certain distribution, let us say with p.d.f. $f(p)$.

In this study of HEMM's only the BLM and the LPLM are used.

3.5.2 *Testing and estimation*

The basic idea in testing and estimation procedures with respect to the BLM and the LPLM is that of conditional model probabilities (CMP's).

This concept was encountered in section 3.4.2.2. There, the formation of such probabilities for specific purchase sequences was extensively treated.

A CMP was denoted as:

$$P_M(1\,|\,x),$$

which is the probability of a brand 1 purchase, given that the observed purchase history is x and the model for the brand choice process is M.

In section 3.4.2.2 the CMP's were conditional on the HEBM. There the M used was the HEBM. In this section M=BLM and LPLM respectively; now the CMP's are conditional on these 2 models.

In a way analogous to that followed in the HEBM case it can be derived that – regardless of the p.d.f. $f(p)$ and regardless of the parameter k – different purchase sequences with the same CMP's can be found. Of course these sequences with equal CMP's are different for the BLM and the LPLM. Details of this derivation can be found in MM&M, chapter 4; the following results are obtained there.

When purchase sequences of length 4 are considered, the sequences within the 4 following groups have the same CMP's under the BLM.

Group I	*Group II*	*Group III*	*Group IV*
0111	1010	0011	0010
1011	0110	1001	0100
1101			

Under the LPLM we have the following groups of sequences with equal CMP's:

Group I	*Group II*	*Group III*	*Group IV*
0111	1110	0110	1011
0011	1100	0100	1101
0001	1000	0010	1001

It can be tested for empirical data by means of χ^2 tests whether the equal CMP's within each group are in agreement with the data. The overall χ^2 statistic, which is the summation of the χ^2-statistics per group, has 5 degrees of freedom in the BLM-case and 8 in the LPLM case.

Using these tests we can check the goodness of fit of the BLM resp. the LPLM, irrespective of k and $f(p)$. When k is known and $f(p)$ is a known p.d.f. with given parameters, the exact CMP for each sequence can be computed and compared with the fraction actually found. This makes the test more powerful than when nothing is known about k and $f(p)$.

k and the parameters of $f(p)$ can also be estimated from the data. A minimum chi-square procedure, as described in section 3.4.3.1 can be used.

3.5.3 *Applications*

MM&M give an application of the BLM and the LPLM to a regular coffee brand choice process. For these data the HEBM (as tested with the CMP test) did not describe the b.c.p. very well. The BLM gave the best fit here, at least for some specific segments of the market, while the LPLM was no better than the HEBM.

The fact that the fit for both models was not very good, may have been a consequence of the stringent relations assumed between the elements of the individual transition matrices.

The attempts made by MM&M to treat a general first order Markov chain – without the restrictions of the BLM or the LPLM – in the HEMM-framework and to use conditional model probabilities for testing it, however failed.

3.6 THE LINEAR LEARNING MODEL (LLM)

3.6.1 *Formulation*

Like the Markov Models, the LLM assumes that the brand chosen at a certain purchase occasion has an influence on the brand choice in subsequent purchases.

This influence is expressed by the fact that, generally, the probability of a brand 1 purchase is different for different purchase histories.

In a first-order Markov model this probability can take only two values, dependent on the two possible brands chosen on the previous purchase occasion.

In the LLM the situation is somewhat more complicated. There, the probability of a brand 1 purchase at purchase occasion $(t+1)$, p_{t+1}, is a linear function of the corresponding quantity at occasion t of the type:

$$p_{t+1} = \varepsilon + \gamma p_t.$$

The LLM used in brand choice processes is a special type of the so-called linear operator models for learning, well-known in mathematical psychology (see Bush & Mosteller, 1955; Luce et al. 1963, chapter 9; or Coombs et al. 1970, chapter 9). In psychology these models are mainly used to describe learning processes in experimental situations. The word learning has a very broad interpretation in these models. Bush and Mosteller (ibid, p. 3) define learning as being 'Any systematic change in behavior, whether or not this is not adaptive, desirable, etc.'. Coombs (ibid, p. 231) gives the definition: 'Learning is any change in the probability of a response'.

The general situation is that an individual respondent makes a sequence of choices, with time intervals, from a number of alternative responses. Here only the case of two responses, response 1 and response 0, will be treated. p_t is the probability that the respondent chooses response 1 at response occasion t.

The choice made at response occasion t works out as a so-called event. In the two-response case there are two types of events: a 1-event and a 0-event. For every event there is a corresponding operator, which determines the influence of the event on p.

Following Bush and Mosteller, these operators are indicated by the symbol Q. Q_1 and Q_0 are the operators corresponding to a 1-event and a 0-event, respectively.

As stated earlier, the operators bring about a linear transformation of p_t, which results in p_{t+1}:

$$p_{t+1} = Q_1 p_1 = a_1 + \alpha_1 p_t \text{ (when at occasion } t \text{ a type 1-event occurred)}$$

$$p_{t+1} = Q_0 p_1 = a_0 + \alpha_0 p_t \text{ (when at occasion } t \text{ a type 0-event occurred)}$$

This general framework can be transferred to the brand choice process. The respondent is then the consumer, who makes subsequent responses (brand choices) from the 2 response possibilities: brand 1 and brand 0. p_t is now the probability of a brand 1 purchase at occasion t. In brand choice applications Q_1 is called the purchase operator, Q_0 is the rejection operator.

The LLM, used here for the b.c.p., has the following simplification:

$$\alpha_1 = \alpha_0 = \alpha,$$

i.e., the slope parameters of the purchase and the rejection operator are equal.

It is reasonable to assume that after a brand 1 purchase the probability that the next purchase will again be brand 1 is greater than after a brand 0 purchase. Therefore, we expect:

$$a_1 > a_0 .$$

We define:

$$a = a_0$$
$$b = a_1 - a_0$$
$$c = \alpha$$

and get then for the operators:

Purchase operator: $p_{t+1} = a + b + cp_t$

Rejection operator: $p_{t+1} = a + cp_t$.

Because the p's are probabilities – which lie between 0 and 1 – and positive learning is assumed with $c \geq 0$, a, b and c must satisfy the following conditions:

$$\begin{aligned} &0 \leq a \leq 1 \\ &0 \leq (a+b) \leq 1 \\ &0 \leq c \leq (1-a-b) \end{aligned} \tag{3.11}$$

The brand choice models treated in the previous sections were put into the stochastic process framework. This can also be done with the LLM. Here we have:

$$\begin{aligned} p_t &= \text{Prob}\,(X_t = 1 \mid x_{t-1}, p_{t-1}, x_{t-2}, x_{t-3}, \ldots) \\ &= \text{Prob}\,(X_t = 1 \mid x_{t-1}, p_{t-1}) \\ &= a + bx_{t-1} + cp_{t-1} . \end{aligned}$$

That this last equation holds will be clear when it is remembered that x_{t-1} can only take the values 1 and 0, which lead to the corresponding operators. An alternative way for expressing p_t is:

$$\begin{aligned} p_t &= \text{Prob}\,(X_t = 1 \mid x_{t-1}, x_{t-2}, p_{t-2}) \\ &= a + bx_{t-1} + ac + bcx_{t-2} + c^2 p_{t-2} . \end{aligned}$$

Generally:

$$p_t = \text{Prob}\,(X_t = 1 \mid x_{t-1}, x_{t-2}, \ldots, x_{t-k}, p_{t-k})$$
$$= a + bx_{t-1} + ac + bcx_{t-2} + ac^2 + c^2 bx_{t-3} + \ldots$$
$$+ ac^{k-1} + bc^{k-1} x_{t-k} + c^k p_{t-k}$$

$$= a \sum_{i=0}^{k-1} c^i + b \sum_{i=0}^{k-1} c^i x_{t-i-1} + c^k p_{t-k} \tag{3.12}$$

Note that for p_t we need a purchase sequence of a certain length k and the probability of a brand 1 purchase before that sequence was made, i.e., the probability at the $(t-k)$th purchase.

The probability p_{t-k} incorporates the effect of the purchases before the $(t-k)$th purchase. For example, when $k = 1$, besides x_{t-1} only p_{t-1} is needed. When p_{t-1} is given, the purchase history before $(t-1)$ is irrelevant; it does not matter by which purchase history the given value for p_{t-1} was brought about. This phenomenon is called: independence of path.

Because when shorter purchase histories are used the effect of the preceding purchases is contained in the given p-value, all the expressions for p_t given above are equivalent.

In principle the whole purchase history of a consumer is relevant for his current p-value, but it can be seen from (3.12) that the influence of a purchase is less the longer ago it is that the purchase was made.

When the probability p_t itself is considered as the state variable, it can be seen that the series: $p_t, p_{t+1}, p_{t+2}, \ldots$ constitute a stochastic process, in casu a first-order Markov process. This can be useful for the computation of the equilibrium distribution for p.

Some implications of the model should be mentioned. When a large number of consecutive brand 1 purchases takes place, p approaches an upper value $p_U = (a+b)/(1-c)$.

The p-value of a consumer who makes a large number of consecutive brand 0 purchases, approaches a minimum value: $p_L = a/(1-c)$. In practice one can expect p_U to be near 1 and p_L to be near zero. p can only take values between p_L and p_U.

The effect of the brand chosen at a purchase occasion on subsequent brand choices declines according to the number of purchases made since that purchase occasion. As can be inferred from (3.12) the decrease is exponential, according to the powers of c. This refers to the order of the brand choice process. A high c-value indicates an influence from previous purchases which is long-lasting, for example.

As seen already, in principle, the whole purchase history of a consumer influences his current p-value, thus essentially the LLM assumes higher order brand choice processes.

One property of the LLM, viz., that the effect of former purchases gradually diminishes, is particularly appealing. This appears to be a more reasonable description of the way in which the purchase experience works than is given, for example, by a first order Markov model where the effect is assumed to work out at once.

At a given moment an arbitrary consumer may have any one of all possible p-values lying between p_L and p_U. Different consumers may have different p-values, so in this respect the LLM assumes heterogeneity. However, the parameters a, b and c are the same for all consumers.

The LLM has been rather briefly presented here. A more extensive discussion of learning models is given in chapter 5. In this and the next chapter, the LLM is used as formulated above.

3.6.2 *Testing and estimation*

The procedure for testing and estimation given here is based on the work of MM&M (chapter 5). Essentially it is a procedure based on the concept of relative occurrence of different purchase sequences, the ROS-concept (see section 3.4.3.1).

For a fixed sequence length and given parameters of the LLM, the theoretical relative frequency for each sequence can be computed.

This theoretical relative frequency can then be compared with the relative frequency actually found. This can be done for all different sequences of a given length.

Now the parameters of the LLM can be estimated by finding those values for them which make the differences between actual and theoretical relative frequencies as small as possible, in a minimum chi-square sense. So again we use a minimum chi-square method of estimation here.

Let us consider sequences of length 2, as an example. In this case there are 4 possible different sequences. The initial probability of a brand 1 purchase by a consumer (i.e., before the sequence of length 2 is made) is denoted as p_0. Now the probability that a consumer, starting with p_0, will produce each of the 4 different sequences is given below.

Purchase Sequence	Probability
00	$(1-p_0)(1-(a+cp_0)) = 1-a-cp_0-p_0+ap_0+cp_0^2$
01	$(1-p_0)(a+cp_0) \quad = a+cp_0-ap_0-cp_0^2$
10	$p_0(1-(a+b+cp_0)) \quad = -bp_0+p_0-ap_0-cp_0^2$
11	$p_0(a+b+cp_0) \quad = bp_0+ap_0+cp_0^2$
	Total $= 1$

Different consumers will generally have different p_0's. Let p_0 be distributed in the population with p.d.f. $f(p_0)$. Then the theoretical relative frequency of the sequence 00, for example, (relative to the total number of sequences of length 2) is:

$$\int_0^1 (1-a-cp_0-p_0+ap_0+cp_0^2)\, f(p_0)\, \mathrm{d}p_0 = (1-a)+(a-c-1)\mu_1+c\mu_2,$$

where μ_1 and μ_2 are the first two moments of $f(p_0)$. We see that the moments of the distribution of initial p-values besides the parameters of the purchase and rejection operator, also determine the relative frequency of a sequence. Thus these moments should also be estimated.

We might try to estimate, via the procedure discussed above, the parameters a, b, c, μ_1 and μ_2 from the observed relative frequencies of sequences of length two. But it is immediately clear that this would not be successful. The 4 relative frequencies of the sequences leave 3 degrees of freedom, while 5 unknown parameters have to be estimated. This is impossible.

The situation improves when we consider purchase sequences of length 3. There are $2^3 = 8$ such sequences, which means that the 8 relative frequencies deliver 7 degrees of freedom.

It is possible to write out expressions for the theoretical 8 relative frequencies in terms of a, b, c and the moments of $f(p_0)$, as we did above for the sequences of length 2. In these expressions μ_3 is now also present. So in this case we have to estimate 6 parameters with 7 degrees of freedom. This is possible, although only 1 degree of freedom is left.

When we take purchase sequences of length 4, we have 16 different sequences (15 d.f.) and 7 parameters to be estimated, which means that 8 degrees of freedom are left. Because this is a reasonable proposition, we will work for the rest with sequences of length 4 here. It is hardly possible to write out the expressions (in a, b, c, μ_1, μ_2, μ_3 and μ_4) for the theoretical relative frequencies of the 16 different purchase sequences of length 4. The expressions given by MM&M for the relative frequencies for sequence length 3 are already discouraging. When a computer is used, it is not necessary to write out these cumbersome expressions by hand.

The discrepancy measure analogous to (3.10), is formed as follows. Let $u_1, u_2, \ldots, u_{16}$ be the observed relative frequencies of the sequences 0000, 0001, ..., 1111.

Let the parameters a, b, c, μ_1, μ_2, μ_3 and μ_4 be contained in the 7-dimensional vector θ, and let $t_1, t_2, \ldots t_{16}$ indicate the theoretical relative frequencies of the sequences mentioned above. The t-values are functions of θ of course.

The measure of discrepancy G which must be minimized is:

$$G(\theta) = N \cdot \sum_{i=1}^{16} \frac{(u_i - t_i(\theta))^2}{t_i(\theta)} \tag{3.13}$$

where N is the total number of sequences of length 4. Now the θ which minimizes $G(\theta)$ contains the estimates for a, b, c, μ_1, μ_2, μ_3 and μ_4.

Min $G(\theta)$ itself is a χ^2-statistic with 8 degrees of freedom, which can be used to test the goodness of fit of the LLM.

The estimated parameter values should satisfy the restrictions in (3.11) and the condition: $\mu_i > \mu_{i+1}$ ($i = 1, 2, 3$).

So estimating the parameters and testing the goodness of fit of the LLM can be performed in one procedure, which needs as input the observed relative frequencies of the 16 different purchase sequences of length 4.

When it can be assumed that the p.d.f. of p_0, the distribution of the initial probability p_0, is the beta density, then μ_3 and μ_4 can be expressed in μ_1 and μ_2 with equations (3.3) and (3.4). In this case the number of unknown parameters is reduced to 5, which simplifies the minimization procedure.

3.6.3 *Applications of the* LLM

The author who put forward the idea of using linear learning models derived from mathematical psychology for brand choice processes was Kuehn (see Kuehn, 1962; and Kuehn and Day, 1964). Kuehn examined the purchase data for frozen orange juice. He calculated the probability that the brand 'Snow Crop' would be chosen at a certain purchase occasion in dependence of the purchase history preceeding that purchase. These purchase histories were binary coded: 1 = Snow Crop, 0 = Other brand. He concluded that the influence of a Snow Crop purchase on brand choice thereafter diminished exponentially with the number of purchases made between the two brand choice occasions concerned. This led him to the use of the Linear Learning Model to describe this brand choice process. As seen in section 3.6.1, this LLM has the property of exponentially diminishing influence of former purchases. Kuehn calculated the influence of one brand choice on subsequent brand choices by a kind of a factorial analysis.

Carman (1966) used the LLM to describe the brand choice process for tooth paste, with special reference to the brand Crest. He used a regression technique for the parameter estimation. Carman's procedure assumes that every consumer starts with the same probability of a brand 1 purchase. As we saw in section 3.6.1, however, this is ordinarily not the case; in the LLM different consumers may have different probabilities. Carman concludes that the probability of purchasing Crest does vary, depending on past purchase experience, in a manner consistent with the LLM.

McConnel (1968) tried to fit the LLM to brand choice data obtained in an experiment during which persons were asked to make beer purchases from 3 imaginary brands. He used the estimation method, described in section 3.6.2. For these experimental purchases, in only 1 out of 3 cases was the fit satisfying.

MM&M applied the LLM to a number of data sets, among which were included the frozen orange juice purchases from Kuehn and the tooth paste data used by Carman. Their overall conclusion is that the LLM does a rather good job in the description of brand choice processes and 'can be used with confidence as a tool for analyzing brand switching data'.

Haines (1969) gives an application of the LLM on brand choice data for a recently introduced product. He uses a simplified version of the rejection operator and makes the further assumption that all consumers start with the same probability of buying the new product. The different estimation methods used by him give rather divergent results.

3.6.4 *The* HOMM *and the* HEBM *as special cases of the* LLM

Two of the models treated earlier in this chapter are special cases of the LLM. This will be shown in this section. The LLM is defined by the operators:

$p_{t+1} = Q_1 p_t = a + b + cp_t$, if the brand chosen at purchase occasion t is 1.

$p_{t+1} = Q_0 p_t = a + cp_t$, if the brand chosen at purchase occasion t is 0.

The initial value of $p : p_0$ is distributed in the population according to the p.d.f. $f(p_0)$.

We consider two special cases here.

(1) $c = 0$

When the common slope has the value 0, we get:

$$Q_1 p_t = a + b$$
$$Q_0 p_t = a$$

Now we are back in the first order HOMM-situation with transition matrix:

$$\begin{array}{c|cc} \text{from} \backslash \text{to} & 1 & 0 \\ \hline 1 & a+b & 1-a-b \\ 0 & a & 1-a \end{array}$$

So the first-order HOMM is a special case of the LLM, viz., when $c = 0$.

(2) $c = 1,\ a = b = 0$

With these parameter values we get:

$$p_{t+1} = p_t,$$

regardless of the purchase made at purchase occasion t.

The distribution of p in the population remains the distribution of the initial value p_0. So we see that the HEBM is a special case of the LLM, viz., when $c = 1$ and $a = b = 0$.

The fact that the LLM incorportates the HOMM and the HEBM is an attractive feature. When the parameter values for a, b and c, in applications of the LLM, show a tendency to approximate one of the special value sets mentioned above, it is useful to investigate if one of these simpler models can describe the brand choice process in question.

3.7 SPECIAL POINTS REGARDING TESTING AND ESTIMATION

3.7.1 *Goodness of fit measures*

In the preceding sections we discussed procedures which tested whether or not a specific brand choice model was compatible with empirical data. Of course the models always remain more or less successful approximations to the real process. It is impossible to find models which exactly describe the brand choice behavior of consumers. The question then arises: what do we call a reasonable fit? When is the description of reality by the model so bad that the model should be rejected?

Many of the tests treated above led to a goodness of fit test which used a χ^2-statistic. In these tests the null-hypothesis always is that the brand choice process under consideration can be described by the model that is specified. The probability of finding – under the null-hypothesis – a χ^2-value as high or higher than the value actually found serves as a measure of fit for the model. This probability is called: *p*-level. A minimum value, for example .05 (critical level), can be specified and if in an application the *p*-level found is lower than this minimum then the conclusion might be that the model is rejected. There is a problem involved when such a procedure to test the fit of a model is followed.

It is a general feature of the χ^2-value that its power increases with the sample size N. This sample size should therefore be considered when judging a calculated test statistic and its *p*-level.

For example: the fact that – when testing the fit of a model – the χ^2-statistic is only just in the critical region in the case of a high N, may be more in favour of the model under the null-hypothesis than a somewhat lower χ^2-value (not leading to rejection) with a small sample size.

Following the procedure that every model which produces a χ^2-value above a certain specified level be rejected means that when $N \rightarrow \infty$, only models exactly covering reality will be accepted. For a large N very small differences between the model and reality lead to rejection.

It will be clear that N should be incorporated into the results of the model tests. One way in which this can be done is to use the so-called coefficient of contingency, which can be applied to contingency tables (see Kendall & Stuart, II, 1967, p. 557).

Pearson's coefficient of contingency is:

$$P = \left(\frac{\chi^2}{N+\chi^2}\right)^{\frac{1}{2}},$$

which is in some sense similar to the usual correlation coefficient.

As is easily seen, the lowest possible value for P in the case of no association at all is 0. But, generally, the upper value in the case of total association is lower than 1 and depends on the number of rows (r) and columns (c) in the contingency table.

Cramer (1957) gives a modified contingency coefficient which, as formulated by Kendall and Stuart, is

$$C = \left(\frac{\chi^2}{N \times \min(r-1, c-1)}\right)^{\frac{1}{2}},$$

which can be proved always to lie between 0 (no association) and 1 (complete association).

To judge the χ^2-values resulting from the contingency tables, we use in the next chapter the quantity F_c defined as:

$$F_c = 1 - C.$$

We call F_c: the coefficient of fit. F_c lies between 0 (an extremely bad fit of the model) and 1 (an extremely good fit). We use this complement of the contingency coefficient, because high χ^2-values mean bad fit and vice versa. F_c-values for contingency tables with different numbers of rows and/or columns can be mutually compared.

In cases like the procedures based on the ROS-concept for the HEBM (3.4.3.1) and the LLM (3.6.2), the χ^2-value results not from a contingency table but from a test on the parameters of a multinomial distribution.

As is immediately clear from, for example, equation (3.13), the χ^2-value resulting from such a multinomial test is linearly dependent on the sample size N, when the observed and theoretical relative frequencies (and therefore the discrepancies between them) are fixed.

So χ^2/N is an interesting quantity, in relation to judgment of the goodness of fit. Instead of this quantity itself, we use a monotonic function of it:

$$F_p = 1 - \left(\frac{\chi^2}{N+\chi^2}\right)^{\frac{1}{2}} = 1 - \left(\frac{1}{1+(\chi^2/N)^{-1}}\right)^{\frac{1}{2}}$$

This F_p happens to be analogous to the coefficient of fit, derived from the Pearson instead of the Cramer contingency coefficient, which makes it in some way similar to F_c.

F_p lies between 0 and 1. The latter value means an extremely good fit. The value 0 can never be reached however, the lower limit of F_p depends on the number of classes in the multinomial distribution in question.

Although the absolute value of F_p may be somewhat difficult to interpret, it can be useful for mutual comparisons of χ^2-values resulting from multinomial distributions with the same number of categories.

For the test results in the next chapter, the χ^2-values, the corresponding p-levels and the F_p, resp. F_c coefficients will be given. It is on these figures together that judgments about the performance of a model in an empirical situation should be based.

3.7.2 *The minimization procedure*

In sections 3.4.3 and 3.6.2 we encountered so-called minimum chi-square estimators. In this procedure the estimates are found by minimizing a function of a type following equations (3.10) and (3.13) with respect to a number of parameters.

In those cases the function is not a simple expression of the parameters: the relation between the parameters and the value of the function is complicated and clearly not linear. Moreover, not all values are acceptable for the parameters.

We are concerned here with a non-linear minimization problem with constraints. There are a great number of different algorithms which can be used for this type of problem (see, for example, Wilde, 1964; Wagner 1969, chapter 15; and Powell, 1971). Many of these procedures require first or higher order derivatives, for which the numerical computation in this case is rather complicated.

Following MM&M we used a so-called Pattern Search procedure in our minimization problems. A description of this can be found in Wilde (1964, pp. 145–150). As stated there, this procedure 'appears admirably adapted to non-linear curve fitting problems involving minimization of a sum of squares', which is exactly our situation. The pattern search procedure has the advantages that no derivatives are needed, the constraints can be easily built-in and the computation procedure is simple. As is often the case in non-linear programming problems, there can be no certainty that the minimum found is the absolute minimum. So when the minimization procedure produces a low χ^2-value it can be concluded that the model in question fits well, but when the lowest χ^2-value is high it is not completely certain that the model does not fit, because there might be a lower χ^2-value that has been overlooked in the minimization procedure.

3.8 CONCLUDING REMARKS

Hereunder will be given a brief recapitulation of this chapter. The following brand choice models, given here in the context of a 2-brand market with the brands 1 and 0, have been discussed:

a. *The Homogeneous Bernoulli Model* (HOBM)
According to this model every consumer purchases brand 1 with the same constant probability p.
b. *The Homogeneous Markov Model* (HOMM)
In this model the probability that a consumer purchases brand 1 at a certain purchase occasion depends on the brand(s) he bought at the preceding purchase occasion(s). The probabilities of moving from one brand to another are called transition probabilities, which together form the transition matrix. The buyers population is assumed to be homogeneous in the sense that all consumers have the same transition matrix.
c. *The Heterogeneous Bernoulli Model* (HEBM)
This is the same model as the HOBM, except that in the HEBM different consumers are allowed to have different values for p.
d. *The Heterogeneous Markov Model* (HEMM)
This is the same model as the HOMM, except that in the HEMM different consumers are allowed to have different transition matrices.
e. *The Linear Learning Model* (LLM)
In this model, at each purchase moment a consumer has a certain probability p of buying brand 1. After a purchase this probability is transformed in a way which is dependent on the brand bought at that purchase.

After outlining the different models, some goodness of fit measures were discussed which can be used to verify the fit of the models in empirical situations; then a minimization procedure which can be applied to obtain the minimum chi-square estimates of parameters needed for certain models was mentioned.

The models outlined in this chapter will be applied to the empirical brand choice processes in chapter 4.

4. Application of the brand choice models to the empirical data

4.1 GENERAL OBSERVATIONS

In this chapter the models of chapter 3 are confronted with the brand choice data discussed in chapter 2. The purpose is to find out which of the models give a good description of the empirical brand choice processes[1] for fopro, beer and margarine, for which observations over a 2-year period are possessed. In the following sections the results of the various model tests outlined in chapter 3 are given. After this a simulation study, done to examine the behavior of the brand choice process under different underlying models, is reported. In the last section of this chapter the various results are discussed.

First some general observations which hold for all models will be made. Essentially, it is the influence of a brand chosen at a certain purchase occasion on subsequent brand choices which is studied. Therefore, from the empirical brand choice processes, only those purchases can be used as observations for which a preceding purchase history of sufficient length is known. For example, the first purchase of a household cannot be used as an observation to calculate a one step transition probability, because the brand chosen before is not known. Similarly, when there are breaks in the observed purchase histories of consumers some purchases must be discarded for the analysis. This occurs when a household does not report for one or more weeks, because of illness, etc. There are at least 2 breaks in each individual purchase history because for a 4-weekly holiday period in the summer purchases were not recorded.

The purchase histories of the different households are not of the same length. For example, in the case of fopro the shortest history has length 10, the longest counts 492 purchases.

1. In this chapter, the word 'brand choice process' is used for the stochastic process, which produces a purchase history, as well as for the purchase history itself, which is the realisation of the stochastic process.

When all purchases of all households are included in the analysis, households with long purchase histories (for convenience called long households) have a relatively great influence. If the purpose is to draw conclusions about the brand choice process in general, where all households are equally important, the procedure just mentioned may bias the results, namely when long and short households have different brand choice processes.

On the other hand long households, which are generally the heavy users, are more important from the marketing point of view than short households, which is an argument which might justify their greater contribution to the results. In the model tests to be reported, two types of data sets were used:

1. The purchase histories of individual households for the first 10 (fopro and beer), respectively first 20 (margarine) purchases. The limits were chosen such that all households could be used in this data set. (Compare the selection criteria for households in 2.2).
2. The entire purchase histories of all households.

In data set (1) the influence of household length has been removed, at the expense of discarding a considerable part of the data.
For the results reported, which of the two data sets was used will always be indicated: 'first 10(20) purchases' or 'all purchases'.

As stated earlier, binary coded brand choice processes are mainly used. Because 4 brands in each market are distinguished (section 2.5), 4 different brand choice processes for each product can be formed.

For example, in the fopro market we have the following possibilities:

brand 1 = F1	brand 0 = all others
brand 1 = F2	brand 0 = all others
brand 1 = F3	brand 0 = all others
brand 1 = F4	brand 0 = all others

Even a fifth b.c.p. (brand choice process) can be formed, viz., when brand 1 is coded as 'the favorite brand of a household'. Sometimes this latter process will also be considered.

It will be clear that it is impossible to give in this chapter all the detailed results for each b.c.p. Therefore, in most cases the detailed results for one or a small number of brands is given to demonstrate the procedure. For the other brands summary results are then given.

Unless stated otherwise, in the following analyses for margarine only the mono-loyal households are included. As mentioned in 2.4, the margarine households are divided into mono- and multi-loyal households. The criterion used for this is given in section 4.3.1.

4.2 THE HOMOGENEOUS MARKOV MODEL

Of the models outlined in chapter 3, the model with the most appealing marketing interpretation is the HOMM. The assumption that each household has the same matrix of transition probabilities is essential. Whether or not there is a Markov chain of a certain order that describes the brand choice processes in a satisfactory way will be investigated. Thereafter the possible effect of the homogeneity assumption, which – as seen in section 3.3.5.2 – might lead to 'spurious contagion', will be examined.

4.2.1. *Test on order zero*

4.2.1.1 *Two brand market*

In Table 4.1 the first order transition matrices, the χ^2-value for the test on order zero and the coefficient of fit F_c can be found for the b.c.p.'s formed from F1, B1 and M1, over the first 10(20) purchases. The N refers to the total number of observations in the contingency table for the product under consideration. In Table 4.2 summary results are given for all brands of the 3 products.

Table 4.1. First-order transition matrices and zero order test results, first 10(20) purchases

		Transition Matrix					
Product	*Brand* 1 =	*from* \ *to*	*0*	*1*	*N*	χ_1^2	F_c
Fopro	F1	0	.9622	.0378			
		1	.0648	.9352	6048	4873.4	.10
Beer	B1	0	.9047	.0953			
		1	.1178	.8822	5643	3481.6	.21
Marg	M1	0	.9645	.0355			
		1	.0898	.9102	16093	12302.7	.13

It can be directly inferred from Table 4.1 that the preceding purchase has a very definite influence. The probability of a brand 1 purchase is about .9 higher when the last purchase is brand 1 than when it is brand 0. We found that this also holds for the case where other brands were coded as brand 1. These transition matrices are not given here for practical reasons. From the summary results in Table 4.2 it can be inferred from the very low F_c values that the zero order hypothesis is strongly rejected for all cases.

Table 4.2 Test on zero order, summary results

		First 10(20) purchases			*All purchases*		
Product	*Brand 1 =*	*N*	χ_1^2	F_c	*N*	χ_1^2	F_c
Fopro	F1	6048	4873.4	.10	55465	47202.7	.08
	F2	6048	4815.8	.11	55465	47730.1	.07
	F3	6048	4215.0	.17	55465	41265.7	.14
	F4	6048	4308.4	.16	55465	43182.3	.12
Beer	B1	5643	3481.6	.21	20296	15378.7	.13
	B2	5643	3054.1	.26	20296	13932.5	.17
	B3	5643	4314.3	.13	20296	14427.9	.16
	B4	5643	3724.0	.19	20296	15229.1	.13
Marg	M1	16093	12302.8	.13	85929	67877.8	.11
	M2	16093	10061.0	.21	85929	58195.6	.18
	M3	16093	11073.5	.17	85929	62481.2	.15
	M4	16093	11840.9	.14	85929	64872.9	.13

4.2.1.2 *Four brand market*

Unlike most other models, for the HOMM the number of brands distinguished in the market can be greater than 2. In fact there is no restriction with respect to the number of brands distinguished, except that the number of purchases made of a brand should be sufficient to draw conclusions. As an illustration here, the first order transition matrices and the results of the zero order test for each market with 4 brands distinguished are given. These can be found in Table 4.3.

The transition matrices for the four brand case give some additional information about the directions of market movements. For example, it can be inferred from the transition matrix for beer that a relatively large number of B2-customers go over to B1 (9%), while this is not the case in the opposite direction.

In Table 4.3 the transition matrix of the multi-loyal margarine households is also given. It is clear that the structure of this matrix is quite different from that in the mono-loyal case. Sometimes the directions of transitions are opposite for both cases. For example, after an M2 pur-

Table 4.3 First order transition matrices in a four-brand market and zero order test results, all purchases

Product	from \ to					*N*	χ^2	F_c
Fopro		F1	F2	F3	F4			
	F1	.9545	.0065	.0194	.0196			
	F2	.0186	.9377	.0240	.0198			
	F3	.0394	.0131	.8905	.0550			
	F4	.0336	.0115	.0437	.9112	55465	134354.4	.10
Beer		B1	B2	B3	B4			
	B1	.9267	.0210	.0221	.0303			
	B2	.0907	.8464	.0208	.0421			
	B3	.0568	.0109	.8686	.0637			
	B4	.0439	.0171	.0324	.9066	20296	43961.4	.15
Marg (mono-loyal)		M1	M2	M3	M4			
	M1	.9187	.0106	.0029	.0678			
	M2	.0396	.8356	.0087	.1161			
	M3	.0157	.0141	.8591	.1110			
	M4	.0304	.0121	.0078	.9498	85929	18848.4	.15
Marg (multi-loyal)		M1	M2	M3	M4			
	M1	.4543	.1629	.0151	.3678			
	M2	.4067	.3094	.0045	.2795			
	M3	.1119	.0097	.2648	.6136			
	M4	.1587	.0510	.0379	.7525	25567	6776.5	.70

Transition Matrix — rows: from; columns: to.

chase, 84% of mono-loyal households buy brand M2 again while the majority of the multi-loyal households (69%) switches to another brand. An overall transition matrix for margarine, i.e., for mono- and multi-loyal households together, would constitute a kind of weighted average, the interpretation of which would be quite difficult. This different kind of purchase behavior makes it necessary to treat mono-and multi-loyal margarine households as different groups when one wishes to study the structure of the brand choice process.

This is done throughout this chapter, where most attention is paid to the biggest group: the mono-loyal households (847 households, as compared to 212 multi-loyal).

As expected, the four brand analysis leads to the same strong rejection of the zero order hypothesis as occurred in the two brand case. In the case of Marg-multi-loyal, the departure from zero order is less.

Table 4.4 Second order transition matrices and first order test results, first 10(20) purchases

Product	*Brand 1=*	*Transition Matrices* *from* | *to*	*0*	*1*	*N*	χ_1^2	F_c	χ_2^2	$\bar{F}_c$
Fopro	F1	00	.9762	.0238					
		10	.6308	.3692	3361	416.4	.65		
		01	.4688	.5313					
		11	.0376	.9624	2015	366.5	.57		
					5376			728.9	.61
Beer	B1	00	.9414	.0586					
		10	.5462	.4538	2893	422.8	.62		
		01	.4403	.5597					
		11	.0712	.9288	2123	307.1	.62	729.9	
					5016				.62
Marg	M1	00	.9824	.0176					
		10	.4780	.5220	10972	2784.5	.50		
		01	.5532	.4468					
		11	.0455	.9545	4274	1089.1	.50		
					15246			3873.6	.50

4.2.2 *Test on order one*

In the previous section it was ascertained that the probability of choosing a certain brand depends on the brand bought at the last purchase occasion. The next question is: does only the last purchase have an influence or is the last but one also important? This is tested here. Second order transition matrices and first order test results for F1, B1 and M1 – first 10(20) purchases – are given in Table 4.4. Summary results for all cases can be found in Table 4.5.

The $\bar{F}_c$ value for each brand is the unweighted mean of the two F_c values, for the two contingency tables of the test on first order.

Table 4.5 Test on first order, summary results

		First 10(20) purchases			*All purchases*		
Product	*Brand 1 =*	*N*	χ^2_2	$\bar{F}_c$	*N*	χ^2_2	$\bar{F}_c$
Fopro	F1	5376	782.9	.61	52871	9988.9	.56
	F2	5376	806.6	.59	52871	11729.1	.51
	F3	5376	1116.6	.55	52871	11737.4	.55
	F4	5376	966.6	.57	52871	9044.3	.50
Beer	B1	5016	729.9	.62	18035	3124.7	.58
	B2	5016	652.0	.60	18035	2194.3	.59
	B3	5016	836.8	.60	18035	3193.2	.60
	B4	5016	581.8	.66	18035	3019.5	.59
Marg	M1	15246	3873.6	.50	81973	19372.3	.51
	M2	15246	3004.2	.53	81973	15457.8	.50
	M3	15246	4567.9	.46	81973	24417.5	.48
	M4	15246	3688.3	.51	81973	19940.5	.50

The conclusion to be drawn from Tables 4.4 and 4.5 is the following. Although the fit of the first order Markov model is better (as reflected by the higher F_c-values) than the fit of the zero order model, the last but one purchase has a definite influence. Keeping the last brand constant, we see that the probability of choosing brand 1 is .4 to .5 higher when the last but one brand is 1 than when it is 0.

4.2.3 *Test on order two*

In the previous sections it was found that the last and the last but one purchase influence the brand choice at a certain purchase occasion. The next question is: is the last but two purchase still important as well? In this section the hypothesis that the process is of order two against the alternative of order 3 is tested. According to the rule given in section 3.3.3.2, in a two brand market 4 different transition matrices would be obtained for which the influence of the last but two purchase should be tested separately.

However, some sequences did not occur with a frequency sufficient to justify a chi-square test. Here only those sequences which occurred most frequently will be considered.

This results in two instead of four transition matrices, for which the separate tests are performed. This means, of course, that the test on order two is only partial.

Also because of the requirement of a sufficient number of observations per purchase sequence, the tests on order two were only performed for

Table 4.6 Third order transition matrices and second order test results, all purchases

Product	*Brand 1=*	*Transition Matrices* to / from	*0*	*1*	*N*	χ_1^2	F_c	χ_2^2	$\bar{F}_c$
Fopro	F1	000	.9871	.0129					
		100	.7687	.2313	26853	1275.0	.78		
		011	.2813	.7187					
		111	.0187	.9813	20231	1344.3	.74		
					47084			2619.3	.76
Beer	B1	000	.9776	.0224					
		100	.7923	.2077	8171	314.7	.80		
		011	.2620	.7380					
		111	.0252	.9748	6686	384.5	.76		
					14857			699.2	.78
Marg	M1	000	.9895	.0105					
		100	.7071	.2929	53261	4309.1	.72		
		011	.3272	.6728					
		111	.0274	.9726	19852	1757.2	.70		
					73113			6066.	3.71

the case of all purchases. Results for F1, B1 and M1 can be found in Table 4.6; Table 4.7 contains summary figures for all brands.

As can be inferred from these tables, there is still a definite influence of the last but two purchase. Keeping the last and the last but one brand constant, we see that the probability of a brand 1 purchase is .2 to .3 higher when the last but two purchase is brand 1 than when it is brand 0.

The next question that comes naturally to mind is if the last but 3 purchase also has an influence on the brand choice. To test this, we would need purchase histories of length 4 with the subsequent purchases belonging to them.

Table 4.7 Test on second order, summary results, all purchases

Product	*Brand* 1 =	*N*	χ^2_2	F_c
Fopro	F1	47084	2619.3	.76
	F2	47084	1140.8	.81
	F3	47084	3341.6	.74
	F4	47084	2814.0	.74
Beer	B1	14857	699.2	.78
	B2	14857	524.0	.78
	B3	14857	707.8	.78
	B4	14857	653.8	.79
Marg	M1	73113	6066.3	.71
	M2	73113	7766.6	.66
	M3	73113	6106.2	.67
	M4	73113	6167.2	.70

However, the number of different sequences increases with increasing length and the number of observations per sequence decreases. In the case of purchase histories of length 4, most transition matrices would be based on a number of observations that is too small to perform the chi-square test. For this reason we did not test on an order higher than 3.

4.2.4 *The aggregation problem*

The conclusions drawn in the previous sections are of the type: a purchase sequence with brand 1 as last purchase has a greater probability of being followed by a brand 1 purchase than a purchase history with brand 0 as the last purchase. Analogous statements can be made for a brand 1 as the last but one, respectively as the last but two purchase.

The influence of former purchases is smaller when those purchases are further back in time. The probability p is about .9 higher when brand 1 is the last purchase compared with brand 0 as the last purchase. For a brand 1 compared with a brand 0 on the last but one purchase, the difference in p is .4 to .5 and for the last but two purchase this difference is .2 to .3.

Of course these effects of brand 1 at different preceding purchases can not be added, e.g. to obtain the probability of a brand 1 purchase after 3 successive brand 1 purchases, because in the analysis above we did not take into account the interaction of brand 1 purchases at different places in the purchase history.

The figures above suggest that purchases further removed than 3 also have an influence, albeit a gradually diminishing one. In Markov chain terminology our conclusion would be that the Markov chain which describes the brand choice process is of a higher order; the order is at least 3.

But now we come back to the aggregation problem that exists for the HOMM, as discussed in section 3.3.5.2. The conclusion above, that the process is at least of the order 3, is reached under the assumption of homogeneity, i.e., every household is assumed to have the same transition matrix.

As we saw in section 3.3.5.2, if there are different sub-populations with different p-values (in a Bernoulli model) or different transition matrices (in a Markov model) and the brand choice processes of these sub-populations are mixed together, this may result in aggregate transition matrices which are misleading because they suggest an order of the process higher than that which any of the homogeneous sub-populations separately has.

So the question with respect to our conclusion of a higher order process is: do individual households have such high order brand choice processes, or is this high order character only due to spurious results as a consequence of aggregation?

We approach this question, within the Markov framework here, in two ways, viz., by considering specific sub-populations and considering individual households.

4.2.4.1 *Transition matrices for specific sub-populations*

To examine the effect of heterogeneity, we divide the households into two sub-groups which can be assumed to be more homogeneous than the whole population. This is done for each product separately.

By means of the first order transition matrices for each sub-group, it is then tested if the brand choice processes for these sub-groups have an order greater than zero. The division is performed two times, according to different criteria which are given hereunder.

Because the division is carried out by applying these criteria to the observed brand choice processes of individual households, it is tautological to state that the brand choice processes of the sub-groups distinguished will be different. Even if all households had exactly the same mechanism underlying the brand choice process, the realisations of the b.c.p. for different households would be different as a consequence of its stochastic character. Therefore, it is not the simple fact that the transition matrices of the sub-groups are different which proves the heterogeneity; the important questions are: what is the magnitude of the differences and what is the order of the b.c.p.'s of the more homogeneous sub-populations?

4.2.4.1.1 *Partition according to favorite brand*

In 3.3.5.2 we saw that when a population consists of households with a high (constant) probability and of households with a low (constant) probability of choosing brand 1, aggregation of the brand choice processes of these two types of households may result in transition matrices which misleadingly point to a high order brand choice process.

To examine if such a phenomenon might have influenced our conclusions, transition matrices for the following sub-groups of households were computed:

I. Brand 1 favoring households: households for which brand 1 is the brand most purchased over the 2 years.

II. Brand 1 not-favoring households: households for which the favorite brand is not brand 1.

When the 2 groups with constant probability, indicated above, exist, the type I households can be roughly conceived of as the households with high p and the type II households as the households with low p.

This analysis is only performed for the first (being the biggest) brand for each product coded as 1. For the other brands the group of households with such a brand as its favorite brand is rather small. The results are given in Table 4.8.

It can be seen that the brand choice processes for these more homogeneous sub-groups are far from zero order, although comparison with Table 4.2 shows that the F_c values are somewhat higher than for the aggregate population. From the differences in transition matrices we conclude that the brand choice processes in the two sub-groups are not the same.

Table 4.8 Transition matrices of households favoring brand 1 versus those not favoring brand 1, all purchases.

		Transition Matrices								
		I = Brand 1 Fav. Households				*II = Brand 1 Not-Fav. Households*				
Product	*Brand 1 =*	*from* | *to*	*0*	*1*	F_c	*from* | *to*	*0*	*1*	F_c
Fopro	F1	0	.6851	.3149		0	.9812	.0188	
		1	.0288	.9712	.34	1	.4188	.5812	.44
Beer	B1	0	.6499	.3501		0	.9712	.0288	
		1	.0401	.9599	.39	1	.3792	.6208	.40
Marg	M1	0	.6907	.3093		0	.9852	.0148	
		1	.0485	.9515	.36	1	.3680	.6320	.38

Table 4.9 Transition matrices of households with more than 2 brands versus all households, all purchases

		Transition Matrices							
		Households with more than 2 brands				*All Households*			
Product	*Brand 1 =*	*from* | *to*	*0*	*1*	F_c	*from* | *to*	*0*	*1*	F_c
Fopro	F1	0	.9502	.0498		0	.9679	.0321	
		1	.1181	.8819	.17	1	.0455	.9545	.08
Beer	B1	0	.9240	.0760		0	.9437	.0563	
		1	.1825	.8175	.26	1	.0733	.9267	.13
Marg	M1	0	.9574	.0426		0	.9697	.0303	
		1	.1547	.8453	.20	1	.0813	.9187	.11

4.2.4.1.2 *Partition according to number of different brands*

Another criterion which can be used to divide the population into possibly more homogeneous sub-populations is the number of different brands bought by a household in the 2-year period.

The transition matrix for the sub-group households with more than 2 brands were computed separately and compared with the transition matrix for all households. The resulting figures can be found in Table 4.9. From these results it can be inferred that the transition matrices for sub-populations of households with more than two brands do not differ much from the aggregate matrices. For these sub-populations the order of the brand choice process is obviously also greater than zero.

It can be concluded that the finding of the processes being of an order greater than zero is not a spurious result of the types of heterogeneity involved in the two partitions above. The more homogeneous sub-groups have brand choice processes which are far from zero order.

4.2.4.2 *Order tests for individual households*

The simplest solution for the problem involved in aggregating different households is to apply the order test only to the purchases of individual households. But then we have the requirement, that the purchase histories of individual households be sufficiently long to enable inferences to be made about their brand choice processes. Only the margarine purchase histories are of a length which makes it possible to perform the Markov order tests for individual households in a number of cases.

The margarine households for which the tests were individually performed, were selected as follows. To be able to carry out the tests two requirements are important. Firstly, the number of purchases should not be too small. Secondly, those purchases should be sufficiently dispersed over the various cells of the contingency tables used for the order tests.

Therefore, for the individual tests we selected from all households only those making at least 175 purchases and for which the portion in the purchases represented by the most favored brand was not higher than .85. The coding of the brands was such that the favorite brand of a household was coded as 1, all other brands as 0. Of course, the limits mentioned above are somewhat arbitrary. Proceeding in this way, we were left with 44 mono-loyal margarine households, on which Markov order tests were individually performed. For each of these 44 households the test on zero order and, as far as was possible, the test on second order were executed.

As can be seen from 4.2.1 and 4.2.2, the zero order test produces one χ_1^2 value, while the test on first order produces two χ_1^2 values, which separately form partial tests on first order. The sum of the latter χ_1^2 values is a χ_2^2 value, the overall test statistic for the first order test. (Always under the assumption that the corresponding null-hypotheses are true). In our case we would have gotten 44 zero-order test statistics, 88 partial and 44 overall first-order test statistics. However, in spite of the selection criteria used, the expected number of purchases in each cell of the respective contingency tables was not always sufficient to permit execution of the test on first order. The zero order tests could be performed for all households.

The following results were obtained:
— In 34 out of 44 cases (77%) the hypothesis of zero order was rejected and the higher order alternative accepted.
— For 24 households the complete test on first order could be carried out. In all 24 cases (100%) the hypothesis of first order was rejected and it was decided that the process was of higher order.
— 54 partial first order tests were performed. 42 of them (78%) led to rejection of the first order hypothesis.

A somewhat striking point is that of the 24 households for which the first order hypothesis was rejected, 9 households did not show a rejection of the zero order hypothesis. So it is possible that there is a feedback of earlier purchases, while the last purchase shows no influence. Although we have – in the sense of our criterion – only mono-loyal households here, the finding, just mentioned may be due to a remaining multi-loyalty, which results in the influence of former purchases being farther away.

The conclusion is then that for these individual households, the brand choice process is generally higher order (at least second order). This confirms the conclusions of the aggregate brand choice processes. It should be kept in mind of course, that the individual tests were carried out only for the b.c.p. of one product, and only for a limited number of households which were selected in a specific way.

4.2.5 *Conclusions regarding the* HOMM

Within the homogeneity assumption, it can be concluded that the brand choice processes are higher order. The zero-order model fits very badly, the first and second order models give an increasingly better fit. For a completely satisfying description of the brand choice processes by a Markov model, a Markov chain of an order higher than 2 is needed. This holds for all 3 products and all brands.

Of course not all consumers have the same brand choice process, as can be seen from section 4.2.4.1. This being true, the results of section 4.2.4 justify the presumption that not all higher order evidence is a spurious result due to aggregation, but that the brand choice processes of individual households are also generally of an order higher than zero or one.

In the next section the zero order hypothesis is extensively tested with procedures which are not sensitive to household-heterogeneity.

4.3 THE HETEROGENEOUS BERNOULLI MODEL

In this section we examine if the HEBM is a model by which the brand choice processes under study can be satisfactorily described. As discussed in 3.4.1, according to the HEBM each household is assumed to choose brand 1 with a constant probability p, where p is distributed in the population according to the p.d.f. $f(p)$. The model assumes that the brand choice processes of individual households are zero order, i.e., that there is no purchase feedback. Heterogeneity of families is explicitly built-in.

It is essentially the zero order hypothesis which is examined in the following tests; the p.d.f. $f(p)$ may take all possible forms. Three different tests, as discussed in section 3.4.2, will be applied. Finally the parameter estimation for the distribution of p will be given.

4.3.1 *The run test; development of the mono-loyalty criterion for margarine*

A description of the use of the run test in brand choice processes was given in section 3.4.2.1. The symbols defined there will be used here. The run test is performed for each family separately. When n_1 or n_0 is too small for a family, the test cannot be carried out. The extreme case here, of course, is a purchase history of only ones or zeros ($n_0 = 0$ or $n_1 = 0$). Similarly, when n_1 or n_0 equals 1 and for some combinations of small values of n_1 and n_0, the test also cannot be carried out.

To keep the power of the test as high as possible, all purchases of a family are used. It would not make sense to consider only the first 10 or 20 purchases here. Moreover, the danger that long households would 'over-rule' short households is not present here, because the test is carried out for each family separately.

When one specific brand is coded as 1, the same for each family, there will be many families with $n_1 = 0$, unless the brand under consideration has a large market share. To be able to do the test for as many households as possible, the coding was done in such a way that brand 1 was the favorite brand of a household and brand 0 denoted all others.

A second analysis was made with the first brand in a market, F1, B1 and M1 respectively, being coded as brand 1. It appears that the number of households for which the test can be carried out then becomes considerably smaller.

To test the null-hypothesis of random alternation, i.e., that the probability of a brand 1 purchase is the same after a brand 1 as after a brand 0 purchase tables were used for small n_1 and n_0. For n_1 or $n_0 > 20$, the normal approximation, given by (3.7) was applied.

The null-hypothesis was tested against a left-side alternative; purchase feedback was assumed to be such that purchases of the same brand stick together.

The mono-loyalty criterion was developed as follows.
It was observed that in the case of margarine there were many households with a high U-value, i.e., for these households the number of runs was much greater than expected under the null-hypothesis. These households exhibited a kind of alternating brand choice process; obviously they bought different brands for different uses of margarine. This point was discussed in section 2.3.

This run test statistic was used for discriminating between mono-loyal and multi-loyal households. Those households for which U was in the upper 25% region of the distribution of this test statistic under the null-hypothesis, were classified as multi-loyal because they apparently had too many runs. All other households were called mono-loyal. Of course, the boundary of 25% is somewhat arbitrary.

In Table 4.10 the results of the run test in the case where the favorite brand was coded as 1 are given. The percentage points mentioned there relate to the distribution of U under the hypothesis of random alternation. $U \leqslant$ lower 5% point, means that such a small U value would occur under the null-hypothesis by chance in only less than 5% of the cases. With $\alpha = .05$ we reject the null-hypothesis then and conclude that purchase feedback exists.

From Table 4.10 it can be concluded that the percentage of households for which purchase feedback was ascertained is considerable. The somewhat lower figure for beer than for fopro must be seen in the light of the shorter purchase histories for beer, which result in a diminished power of the run test.

For an honest comparison, the number of households in the margarine case for which the null-hypothesis was leftsidely rejected should be related to the total number of mono-loyal households for which the run test was performed, instead of to all households. The percentage is then 71%, which is rather high again. The general conclusion for all products is that the number of households for which the hypothesis of random alternation is rejected is much too high to be compatible with the zero-order hypothesis.

Table 4.10 Results of the run test; brand 1 = favorite brand

	Fopro		*Beer*		*Margarine (all households)*	
	Abs	*% of (2)*	*Abs*	*% of (2)*	*Abs*	*% of (2)*
(1) Number of Households	672	—	627	—	1059	—
(2) Number of Households for which the test could be executed	429	100	399	100	836	100
(3) Number of Households with $U \leq$ lower 5% point	315	73	231	58	445	53
(4) Number of Households with $U \geq$ upper 5% point	—	—	—	—	157	19
(5) Number of Households with $U \geq$ upper 25% point (Multi-loyal Households)	—	—	—	—	212	25

It can be concluded from Table 4.10 that the number of multi-loyal households is 212, which leaves 847 mono-loyal families. As additional information, the number of margarine households for which the U-values lie above the upper 5% point is also given. Of course, these are all multi-loyal households. The evident concentration of their U-values, right from the 5% point, means that – with respect to these U-values – the multi-loyal households form a rather clearly distinguishable group. To give the multi-loyal households – as distinguished by the criterion just given – somewhat more profile vis-à-vis the mono-loyal ones, the following can be said about the relative importance of their various brands. It was found for the multi-loyal households that the average share in their purchases of the first favored brand was .57, of the second favored brand .31 and of the third favored brand .09. For the mono-loyal households, these figures were respectively: .79, .17 and .08. So for the multi-loyal households the importance of other brands, besides the one most bought, was much greater than for the mono-loyal households, which was to be expected.

For a demonstration of the differences in brand choice processes between both types of households, see Table 4.3.

In Table 4.11 the results of the run test application with brand 1 = F1, B1 and M1 respectively are given.

Table 4.11 Results of the run test; brand 1 = first brand in the market

	Fopro	*Beer*	*Marg*
(1) Brand 1 =	F1	B1	M1
(2) Number of Households	672	627	847
(3) Number of Households for which the test could be executed	254	288	297
(4) Number of Households with $U \leq$ lower 5% point	165	140	177
(5) Idem, as a percentage of (3)	65	49	60

Now the number of tests that could be carried out was much smaller than for the favorite brand. The rejection rate was somewhat less, but still it is clear that generally the zero order hypothesis cannot be maintained for the brand choice processes of these products.

4.3.2 CMP *test*

As discussed in section 3.4.2.2, under the null-hypothesis of a zero order process and an arbitrary distribution of p, there are 3 groups of purchase sequences of length four, for which every sequence in the group has the same probability of being followed by a one. With a chi-square test, it can be verified if this is compatible with empirical data.

This test was performed for each product, with the 4 brands distinguished throughout this study and the favorite brand being respectively coded as brand 1. Each observation consists of a series of 5 consecutive purchases. The first 4 purchases make up the purchase history (they determine the group of sequences to which the observation belongs), while the 5th purchase from the series indicates the brand chosen after that purchase history.

For each household the following purchases were taken: 1–4 with the 5th, 2–5 with the 6th, etc. Of course, within a family there is no absolute independence of consecutive observations, but it is not likely that – with the large number of different households (600–900) – this would harmfully bias the results. In the case of all purchases, the results for the analysis using only non-overlapping sequences are also given. In that analysis, purchases 1–4 with the 5th, purchases 6–9 with the 10th, etc., were taken as observations of a household.

To give an idea of how the test statistic is built up, the detailed results for brand F1 of fopro are given in Table 4.12.

Table 4.12 Test on equal CMP'*s for separate groups of purchase histories; product: fopro, brand 1 = F1, first 10 purchases*

Group	*Sequence*	*Probability of next purchase being* 0	1	*N*	χ^2	*Df*	F_c
I	0111	.1667	.8333				
	1011	.3077	.6923				
	1101	.1740	.8260				
	1110	.5405	.4595	128	15.3	3	.65
II	0011	.1852	.8148				
	0101	.6667	.3333				
	1001	.4444	.5556				
	0110	.5000	.5000				
	1010	.6429	.3771				
	1100	.7036	.2964	110	17.9	5	.60
III	0001	.5000	.5000				
	0010	.8214	.1786				
	0100	.7000	.3000				
	1000	.8636	.1364	149	15.9	3	.67

The summary χ^2-statistic (with 11 degrees of freedom) has the value 49.1. The upper 5% of the χ^2_{11} distribution is 19.7, so the hypothesis of zero order was rejected. The mean F_c value: $\bar{F}_c = .64$. It can be inferred from Table 4.12 that, generally speaking, for a fixed number of ones in the purchase history, the probability of a brand 1 purchase is higher, as the previous brand 1 purchases are more recent. Mutatis mutandis, the same holds for a brand 0 purchase. This means that the ones and zeros tend to stick together, which agrees with the results of the previous section.

Summary results for all products and brands are given in Table 4.13.

Table 4.13 Test on equal CMP'*s, summary results*

Product	Brand 1 =	First 10 (20) purchases			All purchases all sequences			All purchases, non-overlapping sequences		
		N_{tot}	χ^2_{11}	$\bar{F}_c$	N_{tot}	χ^2_{11}	$\bar{F}_c$	N_{tot}	χ^2_{11}	$\bar{F}_c$
Fopro	F1	387	49.1	.64	4692	222.1	.78	969	27.5	.83
	F2	220	33.5	.66	1966	213.9	.67	409	30.5	.73
	F3	314	42.0	.67	5086	180.1	.81	1039	41.2	.80
	F4	414	46.5	.68	5338	186.6	.81	1102	45.8	.79
	Favorite	812	79.4	.72	8523	355.2	.78	1743	66.3	.79
Beer	B1	756	63.2	.72	4119	204.5	.78	872	60.0	.74
	B2	474	28.5	.68	2408	125.4	.78	505	36.2	.74
	B3	145	19.6	.74	1518	44.6	.83	321	15.2	.79
	B4	609	76.2	.71	3666	193.6	.77	752	39.9	.77
	Favorite	1141	80.5	.71	5900	286.2	.77	1237	66.9	.76
Marg	M1	1208	41.3	.82	8879	297.7	.83	1804	73.2	.81
	M2	557	14.8	.84	4414	208.7	.82	888	48.9	.80
	M3	282	19.8	.75	2447	144.8	.80	496	37.9	.76
	M4	1608	47.8	.83	12203	416.5	.84	2470	102.9	.81
	Favorite	1321	45.5	.82	19053	777.7	.82	3854	187.3	.80

In this table N_{tot} stands for: total number of observations in the 3 contingency tables for each test.

It can be seen that in the case of the first 10(20) purchases, the zero order hypothesis was rejected in 13 out of 15 tests. In the case of all purchases and all sequences, the zero order hypothesis was always rejected.

The slightly higher coefficients of fit suggest that the departure from zero order is smaller for margarine than for the other products. In the case of all purchases, it does not make much difference if all purchase sequences are used or only the non-overlapping ones. This means that there is no systematic influence of the dependence between successive observations present when all sequences are used.

4.3.3 EOS *test*

As was pointed out in section 3.4.2.3, under the hypothesis of a zero order process the relative frequency of occurrence of sequences 0001, 0010, 0100 and 1000 should be equal. The same holds for 4-purchase sequences within the groups with 2 ones and 3 ones respectively. A chi-square test, to check if these equalities are in agreement with the empirical data, can be considered as a test on zero order.

The test is performed only for the case of all purchases; the absolute frequencies in the case of the first 10(20) purchases would be rather small. In Table 4.14 the execution of the test is shown in detail for brand F1 of fopro.

Table 4.14 EOS *test. Product: fopro; brand 1 = F1; all purchases*

Group	*Sequence*	*Actual Frequency*	*Expected Frequency*	*N*	χ^2	*Df*	F_p
1 one	0001	366	323.6				
	0010	277	323.6				
	0100	276	323.6				
	1000	376	323.6	1295	27.8	3	.86
2 ones	0011	217	159.5				
	0101	133	159.5				
	0110	137	159.5				
	1001	117	159.5				
	1010	135	159.5				
	1100	218	159.5	957	64.8	5	.75
3 ones	0111	350	315.8				
	1011	272	315.8				
	1101	272	315.8				
	1110	369	315.8	1263	24.8	3	.84
				3515	117.4	11	

We see that the summary χ^2_{11}-statistic in this table has the value 117.4, which means that the zero order hypothesis was rejected. The observations show that there is a general tendency for sequences in which ones and zeros stick together to be relatively over-represented, which again points to the presence of purchase feedback. The mean value for the coefficient of fit, defined in 3.7.1: $\bar{F}_p = .82$.

In Table 4.15 summary results for all products and brands are given.

Table 4.15 EOS *test, summary results, all purchases*

Product	*Brand 1 =*	*N*	χ^2_{11}	$\bar{F}_p$
Fopro	F1	3515	117.4	.82
	F2	1547	89.1	.78
	F3	3889	65.7	.86
	F4	3924	85.1	.85
Beer	B1	1746	53.7	.82
	B2	862	38.3	.78
	B3	687	31.4	.79
	B4	1651	60.1	.80
Marg	M1	6031	42.8	.92
	M2	2811	38.6	.91
	M3	1555	40.0	.88
	M4	8226	76.8	.92

From Table 4.15 it can be seen that the hypothesis of zero order was rejected in all cases.

As in the case of the CMP test of the previous section, the departure from zero order was somewhat less for margarine than for fopro and beer. Note that the coefficients of fit follow those of the corresponding CMP tests (all purchases and all sequences) shown in Table 4.13. The general level of these coefficients of fit is somewhat higher for the EOS test.

4.3.4 *Estimation of the parameters of $f(p)$*

Although the test results of the previous sections showed that the zero order assumption was not justified for the brand choice processes under study, in this section the parameters of $f(p)$ are estimated according to the procedure given in 3.4.3, where the purchase feedback is disregarded. These parameters are needed for the simulation study given further in this chapter.

Because the simulation is only performed for the first 10(20) purchases, the parameters are estimated only for that case.

In section 3.4.3 we discussed 2 different procedures for estimating the parameters of $f(p)$. One method is based on the relative occurrence of certain sequences; the other method uses the portions of brand 1 purchases of individual households as 'direct' observations of p. Because the estimation of a p-value per household on the basis of a small number of purchases is rather unreliable, only the first method is appropriate in the case of the first 10(20) purchases. As discussed in section 3.4.3.1 the basic idea of this method is to find values for the parameters of the distribution of p such that the theoretical relative frequencies of certain purchase sequences agree as good as possible with the relative frequencies actually found (equation (3.10)).

We assume here that $f(p)$ is a beta density, so we have to estimate the parameters α and β. Purchase sequences of length four are considered.

In Table 4.16 the actual relative frequencies and the theoretical relative frequencies, which agree as good as possible with the actual values, are given for fopro, brand 1 = F1. The α and β values, also given, were found by the pattern search minimization procedure described in 3.7.2.

Table 4.16 Actual and theoretical relative frequencies of different purchase sequences. Product: fopro; brand 1 = F1, first 10 purchases

Number of ones in sequence	*Actual relative frequency*	*Theoretical relative frequency*	
0	.5757	.5758	$N = 4704$
1	.0376	.0353	$\chi_2^2 = 1.9$
2	.0270	.0267	$F_p = .98$
3	.0308	.0337	$\alpha = .04713$
4	.3289	.3285	$\beta = .07822$

From this table it can be inferred that the fit of the beta distribution is quite satisfactory for these relative frequencies. It is important to note here that no inference can be made about the zero order character of the brand choice process from the χ^2-quantity computed. The result of Table 4.16 should be interpreted thus: when purchase feedback is disregarded, the beta distribution, with the parameters given, satisfactorily describes the distribution of p in the population.

The results for all products and brands are given in summary form in Table 4.17.

As can be seen from this table, the fit of the beta distribution is on the whole very good. The χ^2 values are very low, relative to the sample sizes.

One striking point is that the estimated values of parameters α and β always lie between zero and one. As seen in section 3.4.1, this means that

Table 4.17 Goodness of fit statistics and parameter estimates for f(p) as a beta density, first 10 (20) purchases

Product	*Brand 1 =*	χ_2^2	*N*	F_p	α	β
Fopro	F1	1.9	4704	.98	.04713	.07822
	F2	5.4	4704	.97	.02039	.11161
	F3	.9	4704	.99	.05038	.15163
	F4	1.5	4704	.98	.04450	.15459
Beer	B1	1.1	4389	.98	.12313	.16789
	B2	14.4	4389	.94	.05727	.30743
	B3	.1	4389	.99	.01821	.12582
	B4	.3	4389	.99	.07844	.18618
Marg	M1	9.6	14399	.97	.03907	.09990
	M2	3.9	14399	.98	.01560	.25392
	M3	10.6	14399	.97	.00758	.19065
	M4	10.2	14399	.97	.10028	.06061

the p.d.f. of the beta distribution takes the *U*-form, with much of the probability mass being concentrated near 0 and 1. In our context this means that there are many households with $p \approx 0$ and many households with $p \approx 1$ in the population. As an illustration, in Figure 4.1 the p.d.f. of the beta distribution with $\alpha = .04713$ and $\beta = .07822$ (from Table 4.16) has been drawn.

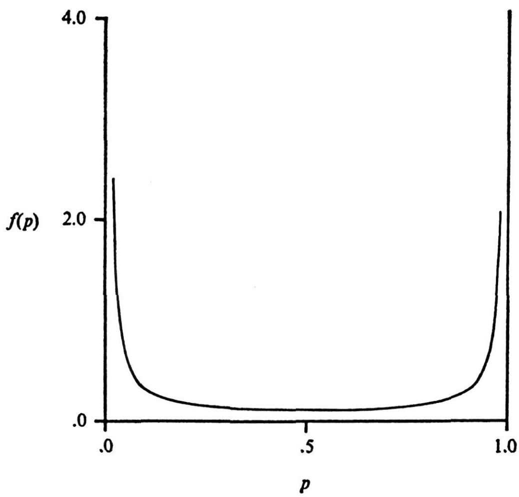

Fig. 4.1 P.d.f. of the beta distribution with parameters .04713 and .07822.

4.3.5 *Conclusions regarding the* HEBM

The tests used here all allow population heterogeneity, i.e., different households may have different p-values. The 3 tests used have their own characteristics and limitations. It is likely that they differ with respect to the probability, to detect certain departures from the zero order situation. Therefore, it is important to note that all the test results unanimously rejected the hypothesis of no purchase feedback. The conclusion is then that the probability p of a household is not constant, but changes in response to the brands chosen.

This is in agreement with the findings in the Markov section, where it was found that the probability p is dependent on the brands bought on at least the last 3 purchase occasions. Those results were obtained under the hypothesis of homogeneity with respect to transition matrices.

Because of this purchase feedback, it can be concluded that the HEBM is not an adequate model to describe the brand choice processes under study.

We will now proceed with models that combine purchase feedback and heterogeneity.

4.4 HETEROGENEOUS MARKOV MODELS

The HEMM discussed in section 3.5 assumes that every consumer's brand choice process is a first order Markov chain in which each individual has his own transition matrix. However, to keep the model tractable several limitations must be imposed, which together ensure that an individual transition matrix is determined by only one parameter.

This section examines if the two variants, the BLM and LPLM can describe the brand choice processes under study. As is pointed out in section 3.5.2, this can be accomplished by a CMP test, which is based on the fact that if a model (for example the LPLM) fits the brand choice process there are certain groups of purchase sequences which have the same probability of being followed by a 1. Using chi-square tests it can be verified if these equal probabilities are in agreement with the data.

Again purchase sequences of length four were used. The way in which these sequences, with the corresponding subsequent purchases, were collected is the same as that described in section 4.3.2 for the CMP test on the HEBM. As was done there, in the case of all purchases the tests were performed with all purchase sequences, as well as with only the non-overlapping sequences. It was found that these results do not differ much, and only the results for non-overlapping sequences will be reported here.

In Table 4.18 detailed results are given for the test on the BLM in the case of fopro; brand 1 = F1; first 10 purchases.

Table 4.18 Test on BLM *product: fopro; brand 1 = F1; first 10 purchases*

Group	*Sequence*	*Probability of next purchase being* 0	1	*N*	χ^2	*Df*	F_c
I	0111	.1667	.8333				
	1011	.3077	.6923				
	1101	.1740	.8260	91	2.2	2	.84
II	1010	.6429	.3571				
	0110	.5000	.5000	26	.5	1	.86
III	0011	.1852	.8148				
	1001	.4444	.5556	45	3.5	1	.72
IV	0010	.8214	.1786				
	0100	.7000	.3000	67	.3	1	.96
				229	6.5	5	

Table 4.19 Test on BLM, *summary results*

		First 10(20) purchases			*All purchases, non-overlapping sequences*		
Product	*Brand 1 =*	N_{tot}	χ^2_5	$\bar{F}_c$	N_{tot}	χ^2_5	$\bar{F}_c$
Fopro	F1	229	6.5	.85	545	10.3	.88
	F2	111	4.2	.80	215	5.4	.86
	F3	289	17.0	.74	630	5.5	.91
	F4	229	13.8	.74	635	9.0	.90
	Favorite	523	18.8	.81	1095	17.2	.90
Beer	B1	442	5.6	.91	519	7.5	.94
	B2	262	15.2	.77	277	13.6	.78
	B3	113	5.4	.79	180	5.2	.87
	B4	335	6.6	.87	429	8.7	.87
	Favorite	724	17.7	.82	781	23.1	.88
Marg	M1	761	21.6	.85	1137	24.1	.87
	M2	343	9.2	.84	531	18.9	.81
	M3	165	6.4	.80	297	9.2	.86
	M4	978	6.4	.93	1547	25.8	.89
	Favorite	838	22.5	.87	2496	28.2	.91

For this relatively small number of observations, the BLM was not rejected, $\bar{F}_c = .85$ here.

Summary results for the tests on the BLM and the LPLM are given in Tables 4.19 and 4.20.

The critical values for a χ_5^2 and a χ_8^2 variable ($\alpha = .05$) are 11.1 and 15.5 respectively. It can be seen that in the case of the first 10(20) purchases, the BLM is rejected 7 out of 15 times and the LPLM 14 out of 15 times. In the case of all purchases, non-overlapping sequences, the BLM is rejected 7 times while the LPLM is always rejected.

It can be concluded that, generally, these HEMM's do not give a good description of the brand choice processes observed, although the BLM obviously makes a better impression than the LPLM.

Table 4.20 Test on LPLM, *summary results*

Product	Brand 1 =	First 10(20) purchases			All purchases, non-overlapping sequences		
		N_{tot}	χ_8^2	$\bar{F}_c$	N_{tot}	χ_8^2	$\bar{F}_c$
Fopro	F1	361	31.3	.72	899	99.2	.73
	F2	200	18.9	.74	366	58.9	.67
	F3	409	47.4	.67	937	105.7	.68
	F4	374	50.9	.65	1011	120.2	.68
	Favorite	731	13.3	.80	1600	86.7	.80
Beer	B1	689	71.3	.72	766	105.4	.67
	B2	423	42.5	.73	428	43.0	.71
	B3	180	19.4	.74	283	27.1	.74
	B4	571	49.2	.75	675	89.0	.66
	Favorite	1032	23.7	.84	1083	79.2	.71
Marg	M1	1054	142.0	.64	1528	228.1	.64
	M2	491	58.8	.68	738	60.5	.72
	M3	226	29.5	.66	383	48.0	.66
	M4	1404	160.0	.67	2069	301.3	.64
	Favorite	1150	142.0	.66	3158	229.4	.74

4.5 THE LINEAR LEARNING MODEL

In this section we examine if the LLM described in section 3.5 can be used fruitfully as a model for the brand choice processes being investigated. The estimation and testing procedure used has been described in section

3.6.2. The starting-point for this procedure is the vector containing the observed relative frequencies of the 16 different purchase sequences of length 4. Via an iterative procedure, those values for the parameters of the LLM are found which make the discrepancies between actual and theoretical relative frequencies as small as possible.

First the various estimation- and testing results will be reported; after that the parameter values found will be discussed.

4.5.1 *Check on Kuehn's data*

To verify our subroutine for the estimation procedure described in section 3.6.2, which generates the theoretical relative frequencies, and to test the minimization procedure, these two procedures were first applied to the brand choice data used by Kuehn in his analysis.

The observed relative frequencies required as the data-input can be derived from the figures given in Kuehn (1962, Table 1). As can be inferred from this source, the data refer to more than 15,000 purchases of the product frozen orange juice by about 600 families.

MM&M computed parameter estimates and a test statistic for these data. Their results and those obtained by our procedure are given in Table 4.21.

Table 4.21 Parameter estimates and test statistic for the LLM. *Data: Kuehn (1962)*

Parameter	MM&M's *results*	*Our results*
a	.015	.0150
b	.305	.3050
c	.612	.6125
μ_1	.214	.2138
μ_2	.108	.1081
μ_3	.075	.0745
μ_4	.048	.0477
χ^2_8	1.16	1.28
(N)	13519	13519

It can be seen that these parameter estimates are practically equal to those of MM&M and that the test statistic differs only slightly. This gives confidence in the numerical procedures used.

Kuehn himself does not give the exact procedures via which he estimated the parameters. As far as Kuehn's estimates could be reconstructured, MM&M found that the values agreed pretty well with the figures they obtained.

4.5.2 *First 10(20) purchases*

In this section the results of the estimation– and testing–procedures applied to vectors of relative frequencies, based on the first 10(20) purchases, are given.

Here all sequences of length 4 of a household are used as observations. The first observation is the sequence formed by purchases 1–4, the next observation is the sequence constituted by purchases 2–5 etc. When all sequences of length 4 from a purchase history are taken, these sequences are not independent of each other. When the sequence 1111 is observed, for example, the next sequence can only be 1111 or 1110, but not 0000. With enough different families for which purchase histories have been observed, it is not likely that this dependence will cause a systematic bias, i.e., it is not likely that the relative frequency of certain purchase sequences would be systematically over or under-represented. In section 4.5.4 we examine the possible effect of this dependence in observed purchase sequences.

The procedure for brand F1 of the product fopro will be treated in some detail. In this case the number of observations is $(10-3) \times 672$ (number of households) $= 4704$.

The vector of observed relative frequencies (u in the notation of section 3.6.2) and the vector of theoretical relative frequencies (t), which gives the best fit possible, are presented in Table 4.22.

The measure of discrepancy (equation (3.13)) has the value:

$$G(\theta) = 4704 \sum_{i=1}^{16} \frac{(u_i - t_i(\theta))^2}{t_i(\theta)} = 2.67$$

The values of the parameters in θ : a, b, c, μ_1, μ_2, μ_3 and μ_4 respectively, which produce the lowest value for G are: (.0052, .4393, .5448, .3762, .3379, .3259, .3241).

The χ^2_8-statistic ($= G(\theta)$) is 2.67. From a χ^2-table it can be concluded that the probability of finding – under the LLM-hypothesis – a higher value by chance is about .95. This is expressed by saying: the p-level is .95.

So the LLM is evidently not rejected as a model here. On the contrary: the model describes the observed brand choice processes very well.

Table 4.22 Observed and theoretical relative frequencies for different purchase sequences. Product: fopro; brand 1 = F1; first 10 purchases

i	*Sequence*	u_i = *observed relative frequency*	$t_i(\theta)$ = *theoretical relative frequency*
1	0000	.5757	.5755
2	0001	.0100	.0101
3	0010	.0070	.0072
4	0011	.0072	.0069
5	0100	.0089	.0080
6	0101	.0028	.0034
7	0110	.0031	.0034
8	0111	.0092	.0093
9	1000	.0117	.0125
10	1001	.0040	.0032
11	1010	.0030	.0032
12	1011	.0062	.0059
13	1100	.0068	.0068
14	1101	.0057	.0061
15	1110	.0098	.0096
16	1111	.3289	.3288

The results for all products and brands are given in a summary form in Table 4.23.

Table 4.23 Parameter estimates and test statistics for the LLM: *first 10 (20) purchases*

Product	*Brand 1*	a	b	c	μ_1	μ_2	μ_3	μ_4	χ^2_8	p-level	N	F_p
Fopro	F1	.0052	.4393	.5448	.3762	.3379	.3259	.3241	2.67	.95	4704	.98
	F2	.0022	.4138	.5652	.1547	.1350	.1318	.1317	5.67	.68	4704	.97
	F3	.0033	.2625	.7302	.2476	.2079	.1875	.1755	5.20	.74	4704	.97
	F4	.0058	.3844	.5925	.2212	.1841	.1683	.1596	10.57	.23	4704	.95
Beer	B1	.0168	.4034	.5619	.4200	.3448	.3071	.3057	6.20	.63	4389	.96
	B2	.0030	.2852	.6903	.1580	.1135	.1013	.1012	6.20	.63	4389	.96
	B3	.0012	.4244	.5547	.1338	.1137	.1062	.1058	4.98	.76	4389	.97
	B4	.0083	.3984	.5715	.2984	.2381	.2098	.1989	5.97	.65	4389	.96
Marg	M1	.0002	.1157	.8829	.2809	.2533	.2405	.2330	15.41	.05	14399	.97
	M2	.0015	.2788	.7174	.0554	.0413	.0350	.0320	15.33	.05	14399	.97
	M3	.0003	.1712	.8238	.0380	.0310	.0277	.0267	26.10	.001	14399	.96
	M4	.0000	.1266	.8714	.6253	.5887	.5709	.5604	17.51	.03	14399	.96

From these figures it can be inferred that the LLM offers a good description of the brand choice processes studied here.

Only 2 out of 12 cases produce a χ^2-value that has a p-level lower than 5%. Moreover, these cases both occur for margarine, where the sample size is biggest. At first glance, looking at only the χ^2-values the fit of the LLM appears to be worse for margarine than for the other products. But, as seen in section 3.7.1, because the sample size for margarine is about 3 times the sample sizes for fopro and beer, the same deviations between observed and theoretical relative frequencies would result in a chi-square value for margarine which is three times the value for the other products. When this sample size is taken into account, the fit for margarine is not worse than those for the other products. This can be seen by inspecting the F_p-values in Table 4.23.

On the whole it can be concluded that the fit of the LLM is very good for the first 10(20) purchases for all products and brands. Of course, one should realize that this fit is measured here and – throughout this section about the LLM – by reproduction of the vector of relative frequencies.

The constraints built into the minimization procedure were: all parameter values should lie between zero and one and $\mu_1 > \mu_2 > \mu_3 > \mu_4$. As can be inferred from the parameter estimates, these constraints were only occasionally active.

4.5.3 *All purchases*

Because of the good results of the LLM for the first 10(20) purchases, it was also examined if the LLM gives a good reproduction of the vector of relative frequencies in the case when this vector is computed on the basis of all purchases. The results are given in Table 4.24.

With these large sample sizes it does not make sense to give the p-level; except in the case of beer (with the smallest sample size of the 3 products), this level is well below 5%. Notwithstanding this, the χ^2-values are small relative to the sample sizes, which is reflected by the high F_p-values. It can be concluded that in the case of all purchases also, the LLM gives a good description of the brand choice processes considered.

Although the corresponding parameter values of Tables 4.23 and 4.24 show some differences, the figures are roughly of the same order of magnitude. Some difference between both cases is possible because in the case of all purchases long households get a heavier weight.

Table 4.24 Parameter estimates and test statistics for the LLM; *all purchases*

Product	*Brand 1*	*a*	*b*	*c*	μ_1	μ_2	μ_3	μ_4	χ^2	*N*	F_p
Fopro	F1	.0044	.4317	.5574	.4202	.3935	.3808	.3737	22.43	50437	.98
	F2	.0009	.3691	.6244	.1391	.1271	.1222	.1215	61.93	50437	.96
	F3	.0021	.3187	.6709	.1957	.1679	.1529	.1446	27.04	50437	.98
	F4	.0028	.3313	.6588	.2487	.2213	.2118	.2068	14.04	50437	.98
Beer	B1	.0058	.4197	.5668	.4429	.3986	.3826	.3658	12.37	16117	.97
	B2	.0024	.4172	.5604	.1004	.0800	.0765	.0733	6.68	16117	.98
	B3	.0012	.3061	.6880	.1674	.1522	.1436	.1401	19.66	16117	.97
	B4	.0050	.3867	.5983	.2895	.2507	.2289	.2137	10.90	16117	.97
Marg	M1	.0012	.2565	.7389	.2747	.2477	.2352	.2278	64.17	78311	.97
	M2	.0004	.2053	.7907	.0651	.0530	.0488	.0474	54.79	78311	.97
	M3	.0000	.0790	.9200	.0416	.0355	.0327	.0310	50.27	78311	.97
	M4	.0018	.1822	.8149	.6169	.5826	.5647	.5533	88.55	78311	.97

4.5.4 *The influence of the dependence between purchases sequences*

As was pointed out in section 4.5.2, the way in which the different sequences of length four are collected causes dependence between subsequent observations. The margarine case offers the possibility of verifying if this dependence can be expected to bias the results.

For the parameter estimates given in Table 4.23, the first 20 purchases of every margarine household were used. Each household delivered 17 observations, viz., the sequences formed by the purchases 1–4, 2–5, ..., 17–20. The vector of observed relative frequencies, which produced the parameter values for margarine in Table 4.23, was based on these observations.

To remove the dependence here we drop the overlapping sequences and use only the sequences formed by the purchases 1–4, 5–8, 9–12, 13–16 and 17–20 as a basis for computing the vector of relative frequencies. Now every household delivers 5 observations. It will be clear that this procedure is not appropriate for fopro and beer, because the ten purchases we have there only produce 2 non-overlapping sequences per household, which is very few.

To get an impression of the effect of the dependence, the results obtained in this way were compared with those obtained when all sequences were used.

In Table 4.25 the vector of relative frequencies when brand M1 is coded as one, computed with all 17 sequences per household and with the 5 non-overlapping sequences, is given. The LLM parameters estimated for the two sets of data are also given.

Table 4.25 Relative frequency vectors and estimates for LLM-parameters for overlapping and non-overlapping sequences. Product: margarine; brand 1 = M1; first 20 purchases

	All 17 sequences per household used	*Only 5 non-overlapping sequences used*
Sequence	Relative Frequency	Relative Frequency
0000	.6738	.6741
0001	.0083	.0083
0010	.0079	.0071
0011	.0042	.0047
0100	.0085	.0078
0101	.0055	.0059
0110	.0044	.0043
0111	.0067	.0071
1000	.0081	.0071
1001	.0040	.0043
1010	.0060	.0068
1011	.0069	.0078
1100	.0040	.0052
1101	.0078	.0080
1110	.0067	.0059
1111	.2370	.2357
LLM-parameter	Estimate	Estimate
a	.0002	.0007
b	.1157	.1148
c	.8829	.8854
(χ^2_8)	15.41	6.54
(N)	14399	4235

It can be seen that both vectors are very similar; the LLM-parameters, computed with these vectors as starting-points, are also very identical. Apparently, using all sequences, instead of only the non-overlapping ones, does not cause a bias here. The deviations between the vectors based on overlapping and non-overlapping sequences for the margarine b.c.p. with M_2, M_3 and M_4 respectively coded as brand 1 were also inspected. In all these cases the differences were in the same order of magnitude as those in Table 4.25.

These findings do not show a biasing influence from the dependence in sequences, which is an argument in favour of using all sequences, included the overlapping ones, for the estimation of the LLM-parameters.

4.5.5 *Discussion of the parameter estimates*

The estimated parameters can be divided into two groups. a, b and c are the proper LLM-parameters, they are the coefficients of the purchase and rejection operator, the essential relations in the learning model. μ_1, μ_2, μ_3 and μ_4 are parameters of the distribution of p_0, the probability of a brand 1 purchase at the beginning of a purchase sequence of length four.

When the learning process starts, the distribution of p_0 is independent of a, b and c, of course. After the process has been going on for some time, the probability p of a brand 1 purchase is influenced by the learning process. The distribution of p then depends on the starting distribution of p_0 and the learning parameters a, b and c.

In chapter 5 it will be shown that, after a great number of purchase occasions, the influence of the initial value p_0 diminishes to zero and the distribution of p is completely determined by a, b and c.

4.5.5.1 *The proper* LLM-*parameters*

4.5.5.1.1 *Margarine versus the other products*

Looking at parameters a, b and c in Tables 4.23 and 4.24, we note that the values of the corresponding parameters for fopro and beer are of the same order of magnitude, while margarine gives a somewhat different picture. For this product the c-values are higher and the b-values lower than in the other cases. As seen in section 3.6.1, a higher c-value means that a certain value for the probability p has a longer lasting influence with respect to subsequent purchase occasions. We saw that the latter influence decreases according to the powers of c. In this sense it can be said that a higher c-value means a higher order process.

On the other hand, the higher c-values for margarine are accompanied by lower values for b. This is inevitable because one of the constraints for the LLM-parameters is (equation 3.11): $(a+b) \leqslant (1-c)$. (Although we did not impose this constraint in our minimization procedure, it is fullfilled for all sets of the parameter values found).

From section 3.6.1 we know that the difference between the effects of the purchase and the rejection operator on p is exactly b. Therefore, a small b-value means that the difference between the effects of a brand 1 and a brand 0 purchase is smaller.

So it can be said that for margarine the 'direct' impact of a chosen brand is smaller, but that the influence of a given p-value works out over a longer period than for the other products. The consequence is that the probability behaves in a more stable way; fluctuations are much smaller. This can be illustrated as follows.

Let a consumer start with $p_0 = .5$ and then exhibit the following purchase sequence: 00101100. When the parameters of the LLM, a, b and c, are respectively .0052, .4393 and .5448 (the case of fopro, brand 1 = F1), the values of p during the sequence are: .5000, .2776, .1564, .5297, .2938, 6046, .7739, .4268, .2377. With parameter values. 0002, .1157 and .8829 (the case of margarine, brand 1 = M1) the p-values are: .5000, .4417, .3902, .4604, .4067, .4750, .5352, .4728, .4176. It can be immediately seen that in the latter case the range of values for p is much smaller.

These smaller variations in the probability p mean that brand choice processes with the parameter values of margarine resemble zero order processes (in which p does not change at all) more than do processes with the parameter values of fopro and beer. This fact was already encountered in sections 4.3.2 and 4.3.3, where it was observed that in the tests on zero order the margarine brand choice processes showed smaller departures from zero order than those of the other products.

4.5.5.1.2 *Limits of p*

In section 3.6.1 it was seen that the probability p always lies between two limits which are determined by the LLM-parameters.

The lower limit is: $p_L = a/(1-c)$

and the upper limit: $p_U = (a+b)/(1-c)$

As can be computed from the parameter values in Table 4.23, these limits for the case of fopro (brand 1 = F1) are: $p_L = .0114$ and $p_U = .9765$. These limiting or boundary values were calculated for all parameter sets presented in Tables 4.23 and 4.24. The highest value found for p_L was .0383; the lowest value for p_U was .9306, while the mean values were: $\bar{p}_L = .0084$ and $\bar{p}_U = .9743$.

Therefore it can be concluded that for the processes studied p can take almost all values between 0 and 1, except those falling near to these extremes.

This is in agreement with what we would expect. It is hardly imaginable, for example, that a consumer would always choose a certain brand with probability 1. According to the LLM, however many times a consumer buys the same brand, there is always a (perhaps very small) probability that he will switch to another one.

4.5.5.1.3 *Imbedded other models*

In section 3.6.4 it was shown that for special parameter values the LLM reduces to the first order HOMM or the HEBM. When $c = 0$ the LLM is the first order HOMM.

As can be seen from an inspection of Tables 4.23 and 4.24, none of the cases show a c-value in the neighbourhood of zero. On the contrary, there is no c-value below .5. Here we have a confirmation of the conclusion drawn in section 4.2, viz., the first order HOMM does not give a good description of the brand choice processes studied.

For $c = 1$ and $a = b = 0$, the LLM reduces to the HEBM. As already discussed, the figures are closer to these values for margarine than for the other products. Although there remains a definite purchase feedback ($b > 0$), it can be said that the margarine brand choice process is more zero order in character than the others.

4.5.5.1.4 *Comparison with* MM&M*'s estimates*

The parameter values found by MM&M for the brand choice processes they considered (given in the Tables 5.4 to 5.7 of their study) are roughly as follows:

For the greater part, their estimates for a lie between .0100 and .1000, for b between .2000 and .4000, while the values for c range from .4000 to .7000. Comparing this with our estimates, it can be concluded that for the brand choice processes studied here, the values for a are generally lower and the figures for b are of the same order of magnitude as MM&M's results. The c-values found for fopro and beer agree with MM&M's figures, while the c-values for margarine are high compared to their findings. As demonstrated in 4.5.5.1.1 a higher c-value means a smoother development of p. The magnitude of changes in p from purchase to purchase is then generally smaller.

4.5.5.2 *The moments of* $f(p_0)$

We now come to the second set of parameters found in the LLM-estimation procedure: the raw moments of the distribution of p_0: μ_1, μ_2, μ_3 and μ_4. As is immediately seen by observing of Tables 4.23 and 4.24, the values for these parameters differ considerably over the various brands and products. This is not surprising, because these parameters reflect the position of a brand in the market. Thus μ_1 should be roughly equal to the market share of the brand concerned. A comparison of the μ_1 values from Table 4.24 with the market shares given in Table 2.3 leads to the conclusion that the corresponding values are indeed rather similar.

4.5.5.3 *Use of the beta-density for* $f(p_0)$

So far no assumptions have been made about the distribution of p_0. Here the case of $f(p_0)$ having a particular density will be discussed.

When a special distribution can be assumed for p_0, there will exist known relations between μ_1, μ_2, μ_3 and μ_4. This makes it possible to estimate the values of only a subset of these parameters, from which the others can be derived. This limits the number of parameters in the minimization procedure, which can be a time-saving advantage. Moreover, when one wishes to perform a simulation with the LLM, as in the next section, one has to assume a particular distribution for p_0 to be able to draw starting values for p.

Here we investigate if the beta distribution – some properties of which were discussed in section 3.4.1 – can be used for the distribution of p_0. In section 4.3.4, the beta-form was used for the distribution of p, the constant probability for individual households to buy brand 1. Here we use the beta distribution for p_0: the initial probability of a brand 1 purchase, at the beginning of a purchase sequence of length four. Note that the way in which μ_1 to μ_4 are estimated implies that they are the moments of the distribution of the probability of a brand 1 purchase at the beginning of any observed purchase sequence of length 4, and not only of the first of these sequences for each household. In so far as the distribution of p_0 changes during the brand choice process – which change does not amount to much for the empirical processes, as can be inferred from the stable market shares – the estimated moments have the character of average values over the brand choice processes.

By means of equations (3.3) and (3.4), parameters α and β of the beta distribution can be computed from μ_1 and μ_2. The appropriateness of the beta distribution for the distribution of p_0 can then be judged as follows.

α and β, computed from the estimated μ_1 and μ_2, can be used to compute 'theoretical' values for μ_3 and μ_4 (equation (3.4)). These theoretical values can then be compared with the values resulting from the estimation procedure. When the beta distribution offers a reasonable fit, the differences between corresponding theoretical and estimated figures should not be great.

In Table 4.26 this comparison is presented for the case of the first 10(20) purchases.

Table 4.26 Theoretical and estimated moments for verifying the fit of the beta distribution for $f(p_0)$; first 10(20) purchases

Product	*Brand* 1=	μ_1	μ_2	α	β	μ_3 *(theor)*	μ_3 *(estim)*	μ_4 *(theor)*	μ_4 *(estim)*
Fopro	F1	.3762	.3379	.0734	.1217	.3192	.3259	.3070	.3241
	F2	.1547	.1350	.0274	.1499	.1257	.1318	.1198	.1317
	F3	.2476	.2079	.0671	.2038	.1892	.1875	.1774	.1755
	F4	.2212	.1841	.0607	.2138	.1668	.1683	.1559	.1546
Beer	B1	.4200	.3448	.1876	.2590	.3083	.3071	.2851	.3057
	B2	.1580	.1135	.0794	.4232	.0943	.1013	.0829	.1012
	B3	.1338	.1137	.0281	.1817	.1043	.1062	.0984	.1058
	B4	.2984	.2381	.1207	.2838	.2099	.2098	.1925	.1989
Marg	M1	.2809	.2533	.0445	.1138	.2399	.2405	.2313	.2330
	M2	.0554	.0413	.0204	.3484	.0352	.0350	.0316	.0320
	M3	.0380	.0310	.0090	.2278	.0278	.0277	.0259	.0267
	M4	.6253	.5887	.1158	.0694	.5700	.5709	.5576	.5604

The overall conclusion to be drawn from this table is that the agreement between the estimated and theoretical values of μ_3 and μ_4 is rather good. Therefore, these results suggest that when the beta density is used for $f(p_0)$, a reasonable approximation is obtained.

Of course the description of $f(p_0)$ by the beta density can never be perfect, because according to the LLM p_0 can never attain the values 1 or 0. In the beta distribution, however, the latter points are attainable. This discrepancy will be not harmful when the lower and upper limits of p approach 0 and 1, as is the case for the processes under study here (4.5.5.1.2).

It is seen that $0<\alpha$, $\beta<1$ always, which means that the fitted beta distributions are U-shaped, with the probability mass being concentrated near to 0 and 1. So most consumers are either very much or not at all inclined to buy brand 1. A similar result was encountered in section 4.3.4.

In chapter 5 it is demonstrated that for the equilibrium situation the fit of the beta density for $f(p_0)$ is still better. Because every LLM process moves towards equilibrium, this is an additional argument for using the beta distribution.

One computational advantage of the use of the beta distribution is that when a beta distribution can be assumed, only 5 instead of 7 parameters have to be estimated. As shown above, μ_3 and μ_4 can then be derived from μ_1 and μ_2.

4.5.6 *Conclusions regarding the* LLM

From the results of the sections discussing the LLM, it can be concluded that the LLM offers a good description of the brand choice processes studied. Both in the case of the first 10(20) purchases and in the case of all purchases, the agreement between observed and theoretical relative frequencies of purchase sequences, which is the criterion used, is very satisfactory.

Having completed the tests for the various models, we now turn to a simulation study in order to arrive at a better idea of how different models of the brand choice process work out in practice.

4.6 A SIMULATION STUDY

4.6.1 *Introduction*

In the preceding sections a number of brand choice models were applied to the fopro, beer and margarine data and the fit of these models examined with respect to the empirical processes. The test statistics used for each of the models are different. For example, the tests for the HOMM's verify the equality of certain transition probabilities; in the HEBM, whether or not certain conditional model probabilities are equal is tested, while in the LLM-case the agreement between observed and theoretical relative frequencies of purchase sequences is checked. So different tests look at different features of the brand choice process and the results for the different model tests are not entirely comparable. For example, it is questionable if the non-rejection of the first-order Markov hypothesis would imply the same about the fit of the first-order Markov model as acceptance of the HEBM in the corresponding relevant test would say about the fit of the HEBM.

Therefore, the need was felt for additional information to check the conclusions drawn regarding the fit of the various models obtained from the model tests in the previous section. What we would like to know is how the brand choice process behaves when each of the various models underlies it, and how much this behavior resembles the brand choice processes actually observed. To this aim, each of the empirical brand choice processes was simulated with the earlier used models, and an examination was made of to what extent the processes thus generated were similar to the corresponding processes, as originally observed. The model parameters used in this simulation study were those estimated in the previous sections, so they are directly derived from the original processes.

The simulation study is performed with the following models: HEBM, first order (f.o.) HOMM, second order (s.o.) HOMM and LLM. The HEMM's were not used in the simulation study; their parameters were not estimated and it was decided – so that the simulation would not become too complicated – to leave these models out.

4.6.2 *Way of comparing original and simulated processes*

When two brand choice processes are to be judged on their mutual similarity, it is necessary to have a yardstick for comparing this similarity. The brand choice processes considered here are all 0–1 processes. It will be clear that practically no two processes are exactly identical in the sense that they have zeros and ones on all corresponding places. The fact that brand choice processes are stochastic processes makes such a coincidence very unlikely. So it does not make sense to compare two processes in this way.

We want to compare the structure of the various processes. Roughly speaking, this refers to the picture, the configuration, of ones and zeros. This configuration cannot be uniquely expressed in one figure. Different measures can be imagined, each of which measures special aspects of the configuration. A very crude measure is the portion of ones in the observed process. For a brand choice process this means: the market share of brand 1. This quantity only looks at the number of ones and zeros and does not pay attention to the placement of the ones between the zeros; e.g., it is not sensitive to those situations where all ones and zeros stick together or where they completely alternate.

A statistic giving more information about the structure of the process would be: the mean run-length, i.e., the mean number of consecutive purchases of the same brand.

In the tests described in sections 4.2 to 4.5, different test statistics were used for the respective tests. Each of these statistics constituted a measure for the configuration of ones and zeros in the brand choice process. Now we can also use these statistics for comparing the structure of the original and simulated brand choice processes. Thus we have the statistics for

— the test on equal CMP's for the HEBM (4.3.2)

— the test on the zero order Markov Model (4.2.1)
— the test on the first order Markov Model (4.2.2)
— the test for the LLM (4.5.2)

The numerical values of those statistics for the original b.c.p.'s are known, and have been given in the sections mentioned. After a brand choice process has been simulated according to a specific model, the statistics mentioned above can be computed for the simulated process and compared with the values for the corresponding original process.

In fact, the similarity of the simulated and original processes is judged with the four statistics mentioned above and two additional statistics. The latter ones are: the mean run-length and the standard deviation of the run-length, both for runs of ones. These were added because it was felt that they especially point to the configuration of ones and zeros in the process.

As seen already, the market share does not say much about the structure of the process in which we are interested here, therefore it is not used in the comparison. Moreover, the simulation procedure, as described in section 4.6.3, was such that the original market shares were generated rather accurately.

To be able to compare corresponding statistics of the simulated and the original process, the processes should be equal in terms of the number of households and number of purchases per household. Therefore, we simulated the purchases of a number of households, which amounted to exactly the same as the original number: 672 for fopro brand choice processes, 627 for beer and 847 for margarine. For fopro and beer 10 purchases per household were simulated, for margarine this number was 20.

The statistics of the brand choice processes thus generated were directly comparable with the corresponding ones for the original processes in the case of the first 10(20) purchases.

Because of the varying number of purchases per household, it served no purpose to perform the simulation for the case of all purchases. To obtain statistics directly comparable with those for the original process, it would have been necessary to build into the simulation the same configuration of numbers of purchases per household, which would have required a much more detailed analysis of the brand choice process.

After having reported here the way in which the simulated processes were judged once they were produced, in the next section we describe how the simulation was executed for each of the four models.

4.6.3 *Generation of the simulated processes*

4.6.3.1 *Random numbers*

The numbers, randomly drawn from the interval [0, 1], required for the simulation were produced by a pseudo-random number generator working according to the so-called congruence method (see Abramowitz & Stegun, 1968, p. 949).

In this procedure, a starting-value must be given the first time and then the subroutine produces a pseudo-random number each time it is called on. Although this procedure is deterministic in the sense that the whole sequence of numbers depends on the starting-value given, the next number cannot be predicted by the preceding one (i.e., by a person who does not know the generation parameters). After a drawing of .999, for example a following outcome of .999 could be as likely as a drawing of .007. Because it is not possible to completely exclude dependence in the series, the numbers thus produced are called pseudo-random numbers. For practical purposes they can be used as if they were random numbers.

4.6.3.2 *Starting-values for p*

For every (imaginary) household a starting-value for p should first be determined. This was done as follows.

For the HEBM, where $f(p)$ is assumed to be the beta density, the starting-value is simply a drawing from the beta distribution, the respective parameter sets of which are given in Table 4.17.

Such a drawing from a special distribution starts with a drawing from the standard uniform distribution. Let this drawing be d. When $Bet(x)$ stands for the distribution function of the beta variate with the relevant parameters, then the drawing from the beta distribution belonging to d is z, which is such that: $Bet(z) = d$.

Because no direct procedure for the inverse beta distribution function was available, an indirect method, which is not completely exact, was used. In this procedure, for the relevant parameters α and β an array was first made of values of $Bet(x)$ for $x = .000\ (.001)\ 1.000$. Then for a certain d value x_1 and x_2 were determined from the array, such that:

x_1 was the highest value of x, for which $Bet(x) \leq d$
x_2 was the lowest value of x, for which $Bet(x) \geq d$.

The drawing from the beta distribution corresponding to d was then found by interpolating between x_1 and x_2.

The fact that we had U-shaped density functions made it impossible to use more straightforward procedures, such as the acceptance-rejection method (see Abramowitz & Stegun, 1968, p. 952), to produce beta-drawings.

In the case of the first order HOMM, a consumer starts in one of the two following states: last purchase = 0 or last purchase = 1. For example: for fopro, brand 1 = F1, this means: $p = .0378$ or $p = .9352$ (Table 4.1). It was decided to make the portion of imaginary households starting in the state last purchase = 1, equal to the observed market share of brand 1 in the original brand choice process. The remaining households started in the state: last purchase = 0. It appears that the observed market shares are very near to the equilibrium market shares, as can be computed for Markov chains (3.3.2).

For the second order HOMM there are 4 different states in which a consumer can start. They are formed by the purchase histories: 00, 10, 01 and 11. The corresponding p-values in the case of fopro, brand 1 = F1, are for example: .0238, .3692, .5313 and .9624 (Table 4.4). Here again the portions of households starting in each of the 4 possible states were made equal to the relative frequencies of these states in the original brand choice processes.

For the LLM the starting value p for an individual household is a drawing from the distribution of p_0, which was assumed (4.5.5.3) to be a beta distribution. The moments of this distribution were produced by the estimation procedure for the LLM. The parameters α and β, which can be derived from the moments, are given in Table 4.26. The drawings from the beta distribution were produced in the same way as for the HEBM.

One additional point should be made for the LLM. It was seen that in the LLM p never can go below a lower limit p_L and above an upper limit $p_U \cdot p_L$ and p_U are determined by the LLM-parameters a, b and c. But the domain of the beta distribution, derived from μ_1 and μ_2, is the whole interval [0, 1]. To remove this discrepancy it is assummed that the probability mass of the fitted beta distribution, left from p_L and right from p_U is concentrated in p_L and p_U, respectively. So in fact a truncated beta distribution is used. For example, in the case of fopro, brand 1 = F1: $p_L = .0114$ and $p_U = .9765$. Whenever the drawing procedure delivered a $p_0 < .0114$ or $> .9765$, the values .0114, respectively .9765 were taken for p_0. As long as p_L is near to zero and p_U is near to one, which in section 4.5.5.1.2 was seen to be the case for all processes studied, the deviation caused by this discrepancy will not be great.

4.6.3.3 *Simulation of the purchases*

Let us suppose that the nth purchase of a household must be simulated.

The state of the household is characterized by the probability that the nth purchase is brand 1, denoted as p_n.

Then a drawing d is made from the standard uniform distribution and the brand chosen is decided as follows.

When, $d \leq p_n$: the brand chosen at occasion n is brand 1;
if $d > p_n$: the brand chosen at occasion n is brand 0.

Afterwards p_{n+1} must be determined. p_{n+1} is computed from p_n and the outcome of the nth purchase, according to the underlying model. This is done as follows: For the HEBM: $p_{n+1} = p_n$ irrespective of the nth purchase.

In the case of the first order HOMM p_n and p_{n+1} can have only two different values. Let us take the example of fopro, brand 1 = F1.

When $p_n = .0378$ and the nth purchase is brand 1: $p_{n+1} = .9352$
When $p_n = .9352$ and the nth purchase is brand 0: $p_{n+1} = .0378$.
In the other cases: $p_{n+1} = p_n$.

For the second order HOMM the procedure is analogous, except that the number of possible combinations is greater.

In the LLM, the purchase or rejection operator works on p_n. In the fopro case just mentioned this means:

$p_{n+1} = .0052 + .4393 + .5448\, p_n$, when the nth purchase is brand 1
and
$p_{n+1} = .0052 + .5448\, p_n$, otherwise.

After the determination of p_{n+1}, the whole procedure can start again for the $(n+1)$th purchase. When the starting-value for p is put equal to p_1, the procedure above can be said to be executed for $n = 1, 2, \ldots, T$, where T is the number of simulated purchases per household. The number of households for which the purchases are simulated can be denoted by M.

The simulation procedure is schematically given in Figure 4.2. This scheme is not very detailed, of course; e.g., input and output are not indicated, while generally one statement in the scheme stands for quite a number of statements in the corresponding computer program.

4.6.4 *Results of the simulation*

4.6.4.1 *The fopro case as an example*

4.6.4.1.1 *Background of the results*

As discussed in section 4.6.2, for each simulated process 6 statistics are computed. Four of them are test statistics for individual model tests; the other two are: mean and standard deviation of the run length for runs of ones. These 6 statistics are called here *comparison points*, designated as

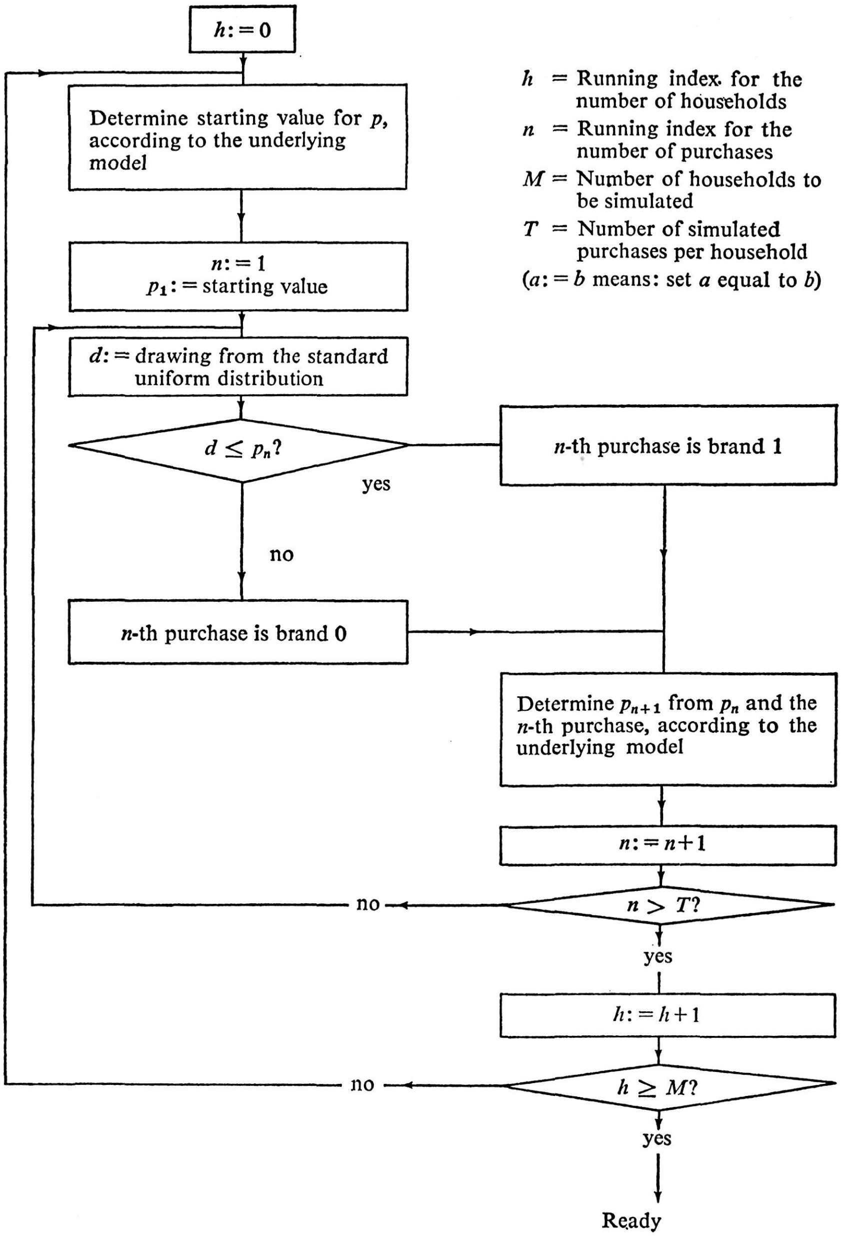

Fig. 4.2. Flow chart for the simulation of purchases

$c_1, c_2, \ldots, c_6$. We argue, that the agreement between the simulated and original process is better according as the differences between their corresponding comparison points are smaller.
In Table 4.27 the results can be found for fopro, brand 1 = F1.

Table 4.27 Comparison of simulated processes with the original process for fopro, brand 1 = F1

Comparison Point \ Generating model	HEBM	HOMM (*first order*)	HOMM (*sec. order*)	LLM	*Original process*
$c_1 = \chi^2_{11}$-statistic for the test on equal CMP's (HEBM)	14.08	440.89	197.63	84.96	49.2
$c_2 = \chi^2_1$-statistic for the test on zero order Markov Model	4885.83	4911.96	4992.43	4927.52	4873.37
$c_3 = \chi^2_2$-statistic for the test on first order Markov Model	1160.33	2.77	495.71	513.71	782.89
c_4 = Mean absolute value of the difference between simulated and original rel. frequencies $\times 10000$ (LLM)	17	90	22	21	0
c_5 = Mean length of runs of ones	6.33	6.55	6.61	6.54	6.37
c_6 = Standard deviation of runs of ones	4.08	3.33	3.71	3.83	3.96
Market share of brand 1	.379	.377	.374	.381	.374

Two comments regarding this table are in order here.

The comparison point with respect to the LLM-character is not the χ^2_8-statistic or one of the LLM-parameters, obtained as a result of the minimum chi-square estimation procedure for the original brand choice processes. It would have been a much too time-consuming affair to perform this minimization procedure for each simulated process. It should be noted that equal vectors of relative frequencies of purchase sequences produce the same parameter and χ^2-values in the LLM estimation procedure. So instead of using the results of the estimation procedure for the comparison of simulated and actual processes, the quantities from which they are derived, viz., the vector of relative frequencies of purchase sequences of length 4, was used. Apart from the point that the parameters of the LLM can be derived from these vectors, comparison of vectors of relative frequencies for the simulated and the original process makes sense in itself, of course. In fact, the quantity computed is: the mean absolute value of the difference between the 16 relative frequencies in the original and simulated process. In the case of complete agreement the value of this comparison point is zero.

The market share of brand 1 is not used as a comparison point. From the way in which the starting-values for p were determined (section 4.6.3.2), it becomes clear that the simulated market share will generally not be far from the original one.

In fact what was studied here was the (dis-)agreement between the *structure* of the brand choice processes for the simulated and the original case, conditional on the circumstance that the shares of brand 1 in both processes are approximately equal. This was done to prevent a difference in the market share itself causing differences in comparison points. Therefore, when for a special starting value in the random number generation process (this starting value itself being produced by a table of random numbers), the market share of brand 1 was more than 2 points removed from the original value, the simulation of the process was repeated using another starting value. Because a model can sometimes describe the process so badly that even the same market share cannot be produced, the repetition was done maximally 2 times. Also when the resulting market share was still more than 2 points removed, the resulting process was used for the comparison then.

Anticipating the general results, it can be remarked here that in only a small number of cases was this repetition necessary. There were 4 out of 48 cases, where the ultimate difference between original and simulated market share was more than 2%. Three of these cases showed a difference of less than 3%, in one case the difference was 3.2%. It can be stated, therefore, that, on the whole, the comparison of the processes was made under the condition of approximately equal market shares.

4.6.4.1.2 *Analysis of the figures for fopro, brand 1 = F1*

From Table 4.27, the following can be inferred with respect to the various comparison points.

c_1 = test on constant p. As expected, in the HEBM case the hypothesis of constant p is accepted; for the other models it is obviously rejected. Of those cases the LLM produces the figure which is nearest to the figure for the original process.

c_2 = test on zero order Markov Model. For all cases the hypothesis of zero order is rejected, the c_2-values do not differ much.

One striking point here is that the value for the HEBM is as high as for the other models. The strong rejection of the zero order hypothesis for this essentially zero order model must here be due to population heterogeneity.

c_3 = test on first order Markov Model. As expected, this hypothesis is accepted for the first order HOMM process.

c_4 = measure of agreement between simulated and original relative frequencies. Although we would expect the LLM to produce the lowest value here, the HEBM gives an even better figure.

c_5 and c_6: comparison points for the runs of ones. We see that the differences are not very big here.

When the values of the different comparison points are considered, the same process is looked at from different viewpoints. Of course, what one would like to have is an overall figure expressing the agreement between simulated and original process. For that purpose we proceed in the following way. For each comparison point c_i, it can be ascertained which model produces the figure nearest to the figure for the original process. We give this model rank 1 for the comparison point in question, the model which is second nearest gets rank 2, etc. This is done for the results in Table 4.27, the ranks are presented in Table 4.28.

Table 4.28 Rankings for the models according to performance on different comparison points

Comparison point \ *Model*	HEBM	HOMM (*f.o.*)	HOMM (*s.o.*)	LLM
c_1	1	4	3	2
c_2	1	2	4	3
c_3	3	4	2	1
c_4	1	4	3	2
c_5	1	3	4	2
c_6	1	4	3	2
Total score T_j	8	21	19	12
General ranking	1	4	3	2

When the rankings of each model on the respective comparison points are added, a total score T is obtained for each model, which can then be used as the basis for a general ranking.

To ascertain the agreement of the different comparison points with respect to the general ranking of the models, the Kendall Coefficient of Concordance, W (see Siegel, 1956, p. 229-235), can be used. The comparison points are conceived as judges ranking the four different models.

The statistic S determines – for a table of given dimensions – the value of W. S is defined as follows:

$$S = \sum_{j=1}^{N} \left(T_j - \frac{\Sigma T_j}{N} \right)^2 ,$$

where N is the number of objects to be ranked (here: 4). It is clear, that the nearer the respective total scores are to each other – i.e., the greater the disagreement between the judges – the smaller will be the value of S.

Now W is defined as:

$$W = \frac{S}{\frac{1}{12} k^2 (N^3 - N)} ,$$

where k is the number of judges (here: 6). W lies between 0 (complete disagreement) and 1 (complete agreement).

Now S is used to test the null-hypothesis that all judges perform the ranking at random against the alternative that there is agreement between their judgments. For a 6×4 table, such as we have here, the critical value for S ($\alpha = .05$) is 75.7.

For Table 4.28 we get: $S = 110$ and $W = .61$. So for this case the following is concluded: the models which produce simulated brand choice processes that are best in accordance with the original process, are in the following rank order: HEBM, LLM, s.o. HOMM, f.o. HOMM, and the various comparison points agree on this rank order.

Of course the fact that the different comparison points get equal weights in the above comparison-procedure is somewhat arbitrary. Further there is the aspect that it is possible that certain comparison points favor certain models; e.g., because the learning parameters were derived from the relative frequencies of purchase sequences of length 4, it can be expected that the LLM generally reproduces c_4 rather well. Insofar as this is true, it makes it more difficult to reach the conclusion that all comparison points agree on the rank order of the different models. In that sense the procedure followed here is a conservative one.

In the procedure discussed, one realisation of the simulated process is compared with one realisation of the original process. In this respect the following can be remarked.

The result of a simulation is stochastic; it depends on the starting value of the random number generation process. For a large number of households and many purchases, the influence of the starting value on the resulting values of the comparison points will not be great. The sensitivity of the outcomes, with respect to the starting values, is examined in section 4.6.5.

The brand choice process, originally observed, is a specific realisation of a stochastic process. When only one such a realisation is considered, accidental events can have a big influence. When more processes are studied, as is the case here, this danger is considerably diminished.

4.6.4.2 *General results*

In Table 4.29 the results of the simulation are given in the form of the general rankings of the 4 models for the different brand choice processes. In the previous section we saw in detail how this general ranking is effected in the case of fopro, brand 1 = F1.

Table 4.29 General rankings of models for different brand choice processes

Product	*Brand* 1 =	*Model*					
		HEBM	HOMM (*f.o*)	HOMM (*s.o.*)	LLM	*S*	*W*
Fopro	F1	1	4	3	2	110	.61
	F2	1	4	3	2	148	.82
	F3	2½	4	2½	1	132	.73
	F4	3	4	1	2	56	.31
Beer	B1	2	4	3	1	26	.14
	B2	2½	4	2½	1	76	.42
	B3	1	4	3	2	106	.59
	B4	3	4	2	1	114	.63
Marg	M1	2	4	3	1	185	.10
	M2	1	4	3	2	116	.64
	M3	1	4	3	2	116	.64
	M4	1	4	2	3	67	.37
Total score		21	48	31	20	506	.70
Total general ranking		2	4	3	1		

As can be inferred from this table: in 4 out of 12 cases the *S*-value lies below the critical level of 75.5. In these cases, the hypothesis that there is no agreement between the outcomes of the different comparison points is not rejected.

A general picture can be obtained by adding the rankings of the models for the 12 brand choice processes studied. From these total scores a total general ranking can be derived.

It is seen that – based on this total general ranking – the rank order of the models with respect to the ability to reproduce the original process is: LLM, HEBM, s.o. HOMM, f.o. HOMM. The difference in performance between the LLM and HEBM is very small, the f.o. HOMM performs the worst, the s.o. HOMM takes an intermediate position.

For judging the agreement on the rank orders of the models among the different brand choice processes, Kendall's Coefficient of Concordance can again be computed. We find $S = 506$ and $W = .70$. Because the critical value of S for this (12×4) table is smaller than 269.8 ($\alpha = .01$), it can be concluded that the individual brand choice processes agree on the total general rank order for the models just given.

4.6.5 *Sensitivity analysis*

As seen above, given the brand choice model, the whole sequence of imaginary purchases constituting a simulated brand choice process is determined by the starting value of the random number generation process. Therefore, it is useful to examine the influence of this starting value on the resulting simulated process, especially with respect to the values produced for the comparison points. For this sensitivity analysis, the brand choice process of margarine, brand 1 = M4, was used. The reason this process was chosen was based on the practical point that for this process, which was the last one simulated, the parameters, distribution functions of the beta distribution, etc., were still present in computer programs and on files.

For each model, the simulation was run 5 times, using 5 different starting values in the process of random number generation. These starting values themselves were obtained from a table of random numbers. The values of the comparison points were computed for each of the 5 simulated processes. So for each model there were 5 figures for each comparison point. Of those values the lowest and highest could be determined. These are given in Table 4.30.

Table 4.30 Lowest and highest values of comparison points resulting from 5 different starting values in the random number generation process. Product: margarine; brand 1 = M4

Comparison point \ Model	HEBM		HOMM (*f.o*)		HOMM (*s.o*)		LLM		*Original process*
	Min	Max	Min	Max	Min	Max	Min	Max	
c_1	7.98	13.31	1837.3	2055.3	549.7	608.4	11.0	36.4	47.8
c_2	11432.9	12261.1	11763.7	12021.6	11739.9	12126.0	11152.3	11513.1	11840.9
c_3	2926.8	3397.9	.94	12.37	3142.4	3492.4	2579.9	3068.2	3688.3
c_4	16	28	141	163	33	53	33	50	0
c_5	9.65	10.42	9.67	10.06	9.64	10.25	8.97	9.21	9.88
c_6	8.69	8.76	6.47	6.69	7.90	8.07	8.10	8.27	8.62

It can be observed that in some cases the value-ranges of two different models for one comparison point partially overlap. This means that for such a comparison point the rank order of the models depends on the particular simulated processes – with respect to the starting values – chosen for the comparison.

The influence of the starting values on the ultimate general ranking was examined as follows.

With the 5 simulations, for each of the 4 models $5^4 = 625$ different combinations of values for the comparison points can be formed. For each of these combinations the rank ordering procedure for the models can then be performed. So every combination produces a general ranking of the 4 models in the way described in 4.6.4.1.

Generally $4! = 24$ different general rankings are possible. Now for each of the 625 combinations the resulting general ranking is determined. If there were no influence of the starting values on the general ranking, each of the 625 combinations would produce the same general ranking, viz., the ranking given for margarine (brand 1 = M4) in Table 4.29, which is: 1–4–2–3 for the order of the models given there. The relative frequency of the different general rank orders resulting from the 625 combinations is given in Table 4.31.

Table 4.31 Frequency of different general rankings in 625 combinations of comparison points

General Ranking[1]	Absolute Frequency	Relative Frequency
1423	543	.87
2413	60	.10
1432	16	.03
1324	6	.01
others	0	0
Total	625	1.00

1. For the models in the order: HEBM, f.o. HOMM, s.o. HOMM, LLM

We see that the rank order 1–4–2–3 occurs by far the most frequently. The other rankings, which also appear, are only slight variants of the order 1–4–2–3, viz., with 2 models interchanged. It can be concluded that the results are rather stable and not very sensitive to the particular starting values used in the random number generation process. This strengthens confidence in the simulation results given in Table 4.29.

4.6.6 *Conclusions from the simulation study*

The brand choice processes for the first 10(20) purchases were simulated with 4 different brand choice models. From a comparison of the processes resulting from the simulation, with the corresponding original processes, it is concluded that the performance of the first order HOMM was the worst. The second order HOMM gave a somewhat better reproduction of the original process, while the LLM and the HEBM produced the best results, the LLM having a minor lead. The superiority of the HEBM over the HOMM suggests that population heterogeneity is a heavier weighing aspect of the brand choice process than purchase feedback.

These simulation results will be discussed further in section 4.7.

4.7 CONSIDERATION OF THE VARIOUS RESULTS IN RELATION TO THE MODELS

4.7.1 *Test and simulation results for the* HEBM: *a contradiction?*

In sections 4.2 to 4.5 we tested if the empirical brand choice processes could be satisfactorily described by various brand choice models. With respect to the HOMM, it appeared that higher order Markov models are needed to describe the brand choice process. The LLM, which is, generally speaking, a higher order model, was found to fit the data satisfactorily. The assumption of no purchase feedback, which underlies the HEBM, was rejected for our data, although the coefficients of fit were higher than for the HOMM.

The simulation study produces a partially different picture. Again the LLM described the brand choice process under study the best. As was to be expected from the test results, the lower order Markov models did not reproduce the original processes very well. The fact that the second order model performed much better than the first order, justified the presumption that still higher order models would give better results. However our data made it impossible to estimate transition matrices of an order higher than two.

The divergent result in the simulation study is produced by the HEBM. Although the essential feature of this model, i.e., no purchase feedback, was found to be not in agreement with the data in all the tests performed, in the simulation study the model generated the brand choice processes approximately as well as the LLM, and much better than the lower order Markov models.

So we reach the conclusion, which seems contradictory, that the HEBM, which is of order zero, reproduces the brand choice processes almost as well as higher order models such as the LLM. This phenomenon was observed earlier (see Frank, 1962). In the next section we will show that this is not so contradictory as it seems to be.

4.7.2 *Temporary seeming zero order behavior of higher order brand choice processes*

4.7.2.1 *The* LLM

Let us assume that the brand choice process can actually be described by the LLM. Here we will examine if it is possible then for the process just to look as if it were zero order.

4.7.2.1.1 *Boundary behavior*

In the LLM, the probability p can never go below a lower bound p_L or above an upper bound p_U. Take the case of fopro (brand 1 = F1, first 10 purchases) as an example. Here $p_L = .0114$ and $p_U = .9765$ (see 4.5.6.1.2). Now, because .9765 is the upper bound for p, a consumer who has a p-value of .9765 and then purchases brand 1 still has the same p-value after that purchase.

In the same way a consumer with $p = .0114$, who purchases brand 0, also keeps the same p-value.

A consumer with $p = .9765$ purchases brand 1 with probability .9765. So with a probability of .9765 his p-value will not be changed by the next purchase. Because of this great chance that p will not change, a consumer who once reaches the upper bound for his p-value may be expected to remain in this state for a large number of purchases. In fact the number of purchases it will take until this upper boundary is left – i.e., a brand 0 purchase is made – is a geometric variate with parameter-value $1-.9765 = .0235$. So the expected value of the number of times that a consumer remains in the state $p = .9765$ is $1/.0235 \approx 43$. In the same way, a consumer who reaches the state: $p = .0114$, i.e. with the probability of a brand 1 purchase of .0114, is expected to remain at this lower bound of p during $1/.0114 \approx 88$ purchases.

So for this fopro brand choice process, when p has reached one of the boundary values, a big number of purchases will generally follow during which p is constant.

As seen in section 4.5.5.1.2, the lower and upper bounds of the p-values for all LLM-parameters estimated are generally near to 0, respectively 1.

Therefore these processes all show the same 'constant-p-behavior' at the boundary values as found for fopro above.

4.7.2.1.2 *Importance of boundary behavior*

The next question is: is boundary behavior important? i.e., does it often happen that consumers reach the boundary values for p? It is useful here to look at the parameters of the fitted beta distributions for the distribution of p_0 (Table 4.26). For the fopro case, treated above, we have $\alpha = .0734$ and $\beta = .1217$. These small values of the beta parameters mean that the probability mass is strongly concentrated at $p = 0$ and $p = 1$. The p.d.f. of this distribution is given in Figure 4.3.

In fact for the beta distribution with parameters .0734 and .1217, the probability of a value lower than .0114 is .4554 and the probability of a value greater than .9765 is .2420.

So when we assume that the truncated beta distribution, as it was called in section 4.6.3.2, fits the distribution of p_0, it can be concluded that for the fopro case about 70% of the consumers start with a p-value at one of the boundaries. From Table 4.26 it can be seen that all the estimated beta parameters lie between 0 and 1. Therefore, all beta distributions are U-formed, which means that for all the brand choice processes considered there are many consumers, who start with p-values equal to the upper or lower bound.

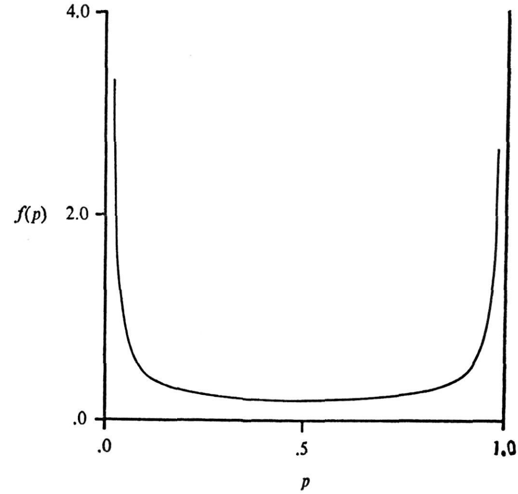

Fig. 4.3 P.d.f. of the beta distribution with parameters .0734 and .1217

4.7.2.1.3 *Consequences for brand choice processes*

We saw that for those LLM-brand choice processes studied, a consumer who has a boundary p-value will generally keep this p-value over a large number of purchases. In addition to this, there are many consumers starting with such a p-value. Let us examine the consequences of this for the brand choice process, using the example of the fopro case discussed earlier.

In this case ten purchases for each household were observed. The probability that a household starting with $p = .9765$, i.e. after a large number of brand 1 purchases, remains in this state over 10 purchases is: $.9765^{10} = .7883$. A household, starting with $p = .0114$, i.e. after a big number of brand 0 purchases, remains in this state over 10 purchases with probability $(1-.0114)^{10} = .8915$. Now when, as seen in the previous section, .4554 of the consumers start with $p = .0114$ and .2420 of the consumers with $p = .9765$, the expected portion of consumers having a constant p-value over the 10 purchases is .5968. So it can be expected that about 60% of the 672 fopro households will keep the same p-value during all 10 purchases.

In addition, there are a number of households for which the starting values for p approach p_L or p_U (this can be inferred from Fig. 4.3). These households also have a high probability of buying brand 0, respectively brand 1, so that there will be only minor changes in p. So in these cases p is almost constant. Of course this constant and almost-constant p-behavior with p at or near a boundary value implies that the same brand (0 or 1) is always bought. Actually it was observed that in the fopro case under study, 549 of the 672 households (82%) bought only brand 1 or only brand 0 during the first 10 purchases. As expected, this is more than the 60% of households expected to keep their p-values exactly constant.

With respect to the seeming zero-order character of the process, the following can be remarked: when a brand choice process is produced by a consumer who buys according to the LLM, and during the purchases observed p happens to remain at the boundary value of .9762, this process cannot be discriminated from that in which the consumer buys according to the HEBM, where his p-value is .9762. The non-zero order of the process would be shown only if a brand switch was made.

Reasoning along the same lines, as was done in detail for the fopro case, the following can be stated for the brand choice processes studied. A consumer, who is once at a boundary value for p, is expected to keep the same p-value over a large number of subsequent purchases. Moreover, there are many consumers with a p-value equal to a boundary value. Therefore, many of the purchases are made by consumers who have constant p-values during those purchases. So for the parameter-values in the processes studied here, the LLM builds in the idea of a large portion of brand choice processes being effected under conditions of constant p. In these circumstances it is not surprising that the HEBM, a model having this constant p as an underlying feature, reproduces the brand choice process rather well.

In this connection, it can be remarked that for the tests on zero order, which led to rejection of the zero order hypothesis presented in section 4.3, only the purchases of those households purchasing brand 0 as well as brand 1 were present. Purchases of those households buying only brand 0 or brand 1, were not included, whilst it was seen in this section that such households may represent the greater part of the purchases. Obviously, for the brand choice processes we studied, 'constant-p-buying' was such an important part of the process that neglecting the purchase feedback of consumers who were not in a constant-p-state, was not too harmful for reproduction of the process.

4.7.2.1.4 *Influence of the value of the* LLM-*parameters*

The effect of the parameter-values on the seeming zero order behavior of an LLM brand choice process can be demonstrated by a comparison with the brand choice process studied by Kuehn, which was discussed in section 4.5.1. For this process we have $a = .015$, $b = .305$ and $c = .612$.

The boundary values computed from these parameters are: $p_L = .0387$ and $p_U = .8247$. Here a consumer whose p-value once reaches the upper bound is expected to remain in this state for $1/(1-.8247) \approx 6$ purchases, which is much less than for the processes studied above. The 'constant-p-buying' occurs much less in this process, therefore it will behave much less than the processes studied earlier as if it were zero order.

To verify this, this process was simulated with the HEBM and the LLM. According to the information given in Kuehn (1962), purchase histories of 600 households were simulated with 25 purchases per household. The parameters of the HEBM were obtained from the given relative frequencies of purchase sequences of length four, in the way described in section 3.4.3.1.

The only comparison point which could be used here to test the reproduction of the original process, was the difference between simulated and original relative frequencies of purchase sequences of length 4, which was indicated as c_4 in section 4.7.

The result of the simulation was that the mean absolute difference between simulated and original relative frequencies was .0026 for the LLM and .0073 for the HEBM. So, as was to be expected for these parameter values, the HEBM here gives an evident worse reproduction of the brand choice process than the LLM.

In addition to the situation where p is equal to a boundary value, the brand choice process also looks very much zero order when the LLM-parameter c is near to one and b near to zero. In section 4.5.6.1.1 it was seen that the probability p then changes very slowly, which may also lead to a zero order appearance. Of the brand choice processes studied, this situation is most nearly reached for margarine. This may have resulted in the particularly good simulation results for the HEBM in the case of margarine, which are shown in Table 4.29.

4.7.2.2 *Higher order* HOMM*'s*

The LLM is not unique in the feature that it – being a higher order model – can produce brand choice processes that sometimes look zero order; the same is true for a higher order HOMM. For example, let us look at the estimated 3rd order transition probabilities for fopro, brand $1 = F1$, given in Table 4.6. We see that the probability of choosing brand 1 after purchase history 111 is .9813. So a consumer starting with this purchase history, makes a brand 1 purchase with probability .9813, in which case the next purchase history is again 111 and the consumer's p-value remains unchanged. A consumer starting in the state: $p = .9813$ is expected to remain in this state during $1/(1-.9813) \approx 53$ purchases. So, quite analogous to the LLM, a higher order HOMM can produce temporary 'constant-p-behavior', which may mean that the process can hardly be discriminated from an HEBM brand choice process.

From the discussion in section 4.7.2, it can be concluded that it is not so contradictory after all for a brand choice process that it can be generated by a higher order model, as well as by the zero order HEBM. For specific parameter-values, the higher order process can exhibit seeming zero order behavior for rather long series of purchases.

4.7.3 *General conclusions about the models*

Of the models applied, the LLM is the only one, which produced good results, both in the test procedures and in the simulation study. Therefore, it is concluded that this model describes the brand choice processes for forpro, beer and margarine the best.

The HEBM was rejected in the relevant tests but showed a rather good performance in the simulation study; this is compatible with the good results for the LLM, because, as seen in section 4.7.2.1, the LLM brand choice processes in question exhibit rather a lot of temporary 'constant-p-behavior'.

The HEMM's, of which the BLM and the LPLM were used, seem to be not so very promising; the LPLM in particular did not perform very well.

When the processes are to be described by HOMM's, these models should be higher order. When the HOMM becomes higher order, however, the difference with the LLM becomes smaller in the sense that an individual consumer may be in more different states with respect to his p-value. In the LLM this number is infinite. But for higher order Markov models the number of parameters, the transition probabilities, increases very quickly, which means that many data are needed to estimate them, whereas in the LLM the number of parameters is independent of the order of the process. So for higher order brand choice processes, such as we have in this study, the LLM is to be preferred over the Markov Model.

The brand choice processes for the three products are not very different. The only point to be distinguished is that the b.c.p.'s for margarine are somewhat more zero order in character than those for the other products.

For the outcomes of the various model tests it does not make much difference if the first 10(20) purchases or all purchases are used as the data base. Generally, the conclusions in the case of all purchases are, because of the bigger sample size, somewhat more powerful than in the case of the first 10(20) purchases, but the directions are the same.

With respect to the LLM the following can still be added. As seen in section 4.7.2.1, the LLM can describe the course of somebody's brand choice process towards a certain equilibrium point, corresponding with a boundary value for p. Once such a point is reached, it is generally not abandoned very quickly, but with probability 1, this will ultimately occur.

This seems to be quite compatible with what can be expected from a consumer. After having tried out different brands, it is likely that he will show subsequent routinized behavior for some time, but not forever.

In this connection Frank (1962, p. 389) can be cited: 'A learning model may be relevant for explaining how consumers arrive at equilibrium. But, once there, one of the sequences of choices for individual families might appear to be consistent with a simple model, that assumes no learning.' We have seen that this 'equilibrium behavior' can be built into the LLM; it is not necessary to switch to another model to describe the brand choice process after an equilibrium point is reached.

For the rest it would not be justified to consider an LLM brand choice process, once an equilibrium is reached, as a Bernoulli process, because in the latter p is assumed to remain always the same. This appears to be a too simple concept of consumer behavior. In the LLM, also in an equilibrium situation, p is expected to change ultimately, which seems to be more realistic.

This aspect of the LLM, viz., that in the brand choice process periods of brand switching activity are alternated with periods of repose, together with the fact that – as seen in section 3.6.1 – in the LLM the influence of former purchases diminishes gradually, makes the LLM intuitively appealing for the description of consumer brand choice processes. This intuition is confirmed by the test and simulation results of this chapter.

Another point in favor of the LLM is that, as observed in section 3.6.4, it contains two other models, the HEBM and the first order HOMM, as special cases. So if a brand choice process in fact can be described by one of these simpler models, and the LLM is used, the parameter values will indicate that one can switch to such a less complicated model.

Because of the good possibilities offered by the LLM, this model will be examined in somewhat more detail and learning models in general will be discussed in the next chapter.

5. Learning models for brand choice: a closer examination

5.1 INTRODUCTION

In section 3.6 the formulation and some properties of the Linear Learning Model (LLM) for brand choice processes were discussed. It was seen that the LLM has a number of features which make it intuitively appealing as a brand choice model. Moreover, from the results given in chapter 4, it was concluded that of all the models used the LLM gave the best description of the empirical brand choice processes investigated.

The LLM has its origin in mathematical learning theory. But the LLM is only one specific type of a great variety of learning models and it seems useful, because of the good results for the LLM obtained in chapter 4, to consider other types of learning models as possible candidates for the description of brand choice processes. This is what is done in the present chapter, where also some additional properties of the LLM, especially with respect to equilibrium behavior, and some possible generalisations of this model are treated.

In section 2 a brief summary of some important types of learning models, developed in mathematical learning theory, is given and the possibilities of applying them to brand choice processes discussed.

In section 3 linear operator models are treated in some detail and in section 4 some generalisations of the LLM are proposed, 2 types of which are applied to empirical data.

In section 5 we discuss a special kind of stimulus sampling model for brand choice, viz., the so-called Probability Diffusion Model, developed by Montgomery (1969). This model is applied to the fopro, beer and margarine data. The chapter ends with some concluding remarks.

5.2 A SURVEY OF LEARNING MODELS

5.2.1 *General*

The learning models discussed here can be divided into operator models and stimulus sampling or state models. The operator models – in turn – are sub-divided into linear and non-linear operator models. In section 5.2 we will briefly discuss these different types of learning models; it will appear, that the LLM is a special case of the linear operator models.

For a more comprehensive and complete treatment of learning models, the reader is referred to Bush & Mosteller (1955), Luce et. al. (1963, chapters 9 and 10) and Coombs et. al. (1970, chapter 9).

A common feature of all the models to be discussed is their probabilistic nature. Learning is considered to be stochastic. In all models the learning process is conceived of as a sequence of discrete trials. At every trial or response occasion there is a certain stimulus situation and the subject chooses one from a number of possible responses. At any given trial the subject has an initial probability distribution over the possible responses. After one of these responses is made, an outcome follows and subsequently a change in the probability distribution over the possible responses may occur, so that the resulting probability values differ from the initial ones. In that case it is said that a probability flow has occurred and then the next trial starts with a probability distribution different to that of the previous one.

As an example let us consider a very common type of experiment in psychological learning research, namely a so-called *T*-maze experiment with a rat. In this design a rat walks through a corridor, at the end of which he has to choose whether turn right or left. He is rewarded or not rewarded by finding or not finding a piece of food in the direction chosen. By manipulating the frequency with which the food is placed right and/or left, the experimenter can study the way in which the rat learns from former experience under different conditions. The mathematical learning models, discussed in this chapter, have for a major part been constructed to describe the learning processes in this kind of experimental situations.

In the *T*-maze experiment each walk of the rat constitutes a response occasion. The response possibilities are turning left and turning right. The possible outcomes are 'being rewarded' and 'being not-rewarded'. The probability of a right, respectively left, turn may change as a consequence of experience. For example, it can be imagined, that the response: turning right, followed by the outcome: finding a piece of food, results in an increase of the probability of a right turn. Thus at the next trial the chance that the rat turns right is greater than at the previous one, so the probability distribution over the responses has been changed. Similar statements can be made for other combinations of responses and outcomes.

Summarizing, the following definition of a learning process can be given. A learning process is a sequence of trials – response occasions – with a stimulus situation, a response, an outcome and a resulting probability flow, which may be zero. This is the definition which will be used in this chapter. Essentially the various models differ in their description of this probability flow from trial to trial.

Now the brand choice process can be conceived of as a learning process in the sense just described. Here a consumer (subject) makes subsequent purchases (trials) of a certain product. At every purchase moment a special brand (response) is chosen. The experience with the brand after the purchase can be conceived of as the outcome of the trial, which can change the probability that the brand in question will be bought again. When the brand chosen uniquely determines the resultant probability flow, as is the case in the LLM, no distinction can be made between response and outcome; in this case they coincide.

In section 3.6.1 it was seen, that the concept 'learning' has a very broad interpretation in learning models. Every change in the probability of a certain response can be called learning. For the treatment of the different learning models in the next subsections, the classification given by Coombs et. al. (ibid) is followed.

5.2.2 *Operator models*

5.2.2.1 *Events and operators*

We saw that in a trial – given the initial probability distribution over the responses – the response and/or outcome determine the resulting probability flow. Now a combination of response and outcome, is called an experimental event. If different experimental events bring about the same probability flow, they are equivalent for the mathematical model, therefore they are said to constitute the same model event. When we use the word event we mean 'model event', which implies that the occurrence of such an event can be caused by different response/outcome combinations. But when, for example, in the T-maze experiment mentioned above the combination turning left/rewarded has the effect of increasing the probability to turn left, while the combination turning left/not-rewarded brings about an opposite probability flow, these two combinations represent different events. In a brand choice process the purchase of a specific brand may be conceived of as an event, which transforms the probability distribution over the different brands in a specific way.

Now in the case of operator models, with every event an operator is associated, which transforms the probabilities in a way characteristic for that event.

5.2.2.2 *Linear operator models*

The theory of linear operator models in learning theory is for the greater part due to Bush and Mosteller. In their comprehensive work, Bush & Mosteller (1955), further be referred to as B & M, occupy themselves with this type of learning model only.

Here we only treat the case of two responses, indicated as 1 and 2. We call the probability of a response 1: p. The operators are denoted by the symbol Q. We assume that only two different events, type 1 and type 2 are possible, which means that there are two different operators, indicated as Q_1 and Q_2. In the case of two responses it is sufficient to specify only the way in which the probability of response 1 is transformed by an operator. Since the probabilities of response 1 and response 2 sum to unity, the latter is determined by the first.

Now in the linear operator model an event affects the probability p in the following way. When a type 1 event occurs p is transformed by the operator Q_1 according to:

$$p' = Q_1 p = \alpha_1 p + a_1 , \tag{5.1}$$

where p' is the probability of a type 1 response after the transformation. For a type 2 event the transformation is:

$$p' = Q_2 p = \alpha_2 p + a_2 . \tag{5.2}$$

In both cases we are concerned with a linear relation, where the slope and the intercept depend on the type of operator.

We assume non-negative learning, so

$$\alpha_i \geq 0 \; (i = 1, 2) \tag{5.3}$$

Since p and p' are probabilities, which lie between 0 and 1, we have the restrictions:

$$\left.\begin{array}{l} 0 \leq a_i \leq 1 \\ 0 \leq \alpha_i \leq (1 - a_i) \end{array}\right\} (i = 1, 2). \tag{5.4}$$

The operator Q_i is not a linear operator in the mathematical sense. For this to be true, the following relation must hold:

$$Q_i(\lambda p_1 + \mu p_2) = \lambda Q_i p_1 + \mu Q_i p_2 ,$$

which is clearly not fulfilled here. The name 'linear operator model' indicates that in the expression for $Q_i p$, p only appears with the powers 0 and 1. Of course it is conceivable for higher powers of p to be included in this expression, but as B & M put it: 'Even with this easy (linear) transformations the mathematics becomes complicated and leads to many unsolved problems and there is little hope of solving these problems with non-linear transformations.'

In this survey the linear operator model will not be treated in detail. This will be done to some extent in section 5.3. In section 3.6.1 the phenomenon of boundary values for p, which generally are not equal to zero and one were discussed. Here only one further property of the linear operator model will be discussed, namely the non-commutativity of the operators, which is useful for comparison with other operator models.

When a series of events occurs consecutively, this is expressed by a sequence of corresponding operators. For example, $Q_2 Q_1 p$ denotes the resulting probability when first Q_1 works on p, resulting in the new probability $Q_1 p$, whereafter Q_2 works on this new probability. We have:

$$Q_2 Q_1 p = \alpha_2(\alpha_1 p + a_1) + a_2 = \alpha_2 \alpha_1 p + \alpha_2 a_1 + a_2$$
$$Q_1 Q_2 p = \alpha_1(\alpha_2 p + a_2) + a_1 = \alpha_1 \alpha_2 p + \alpha_1 a_2 + a_1$$

So $Q_2 Q_1 p = Q_1 Q_2 p$ if

$$\alpha_2 a_1 + a_2 = \alpha_1 a_2 + a_1,$$

which is generally not true. Therefore, we can conclude that in general the operators do not commute or:

$$Q_2 Q_1 p \neq Q_1 Q_2 p$$

So, when the initial probability is p, the resulting probabilities are different for the cases when first event 1 and then event 2 occurs and when first event 2 and then event 1 occurs.

In brand choice processes the events can be conceived of as the brands bought. Let, in a two-brand-market, Q_1 correspond with a purchase of one brand, Q_2 with a purchase of the other brand. Now the non-commutativity of the linear operator model seems to be in agreement with reality of the brand choice. If the operators were commutative, to give a good description of the brand choice process the following would be required. If, after a certain starting value for p, during the next 10 purchases brand 1 is bought one time and brand 0 is bought nine times, it does not make any difference to the probability p after these ten purchases if the brand 1 purchase is the most recent one or, if brand 1 has been bought ten purchases ago with nine brand 0 purchases in between.

This would mean that purchases cannot be forgotten, which is contrary to the findings made in chapter 4, where it was seen that the influence of a brand bought at a certain purchase occasion diminishes with the number of purchases made since that occasion. So the property of non-commutativity of operators is an attractive feature in relation to the possibilities of describing brand choice processes by linear operator models. The same holds for the property of the linear operator model, i.e., that p will never exactly reach the values 0 or 1, because if this occurred it would mean that from a certain point in time a consumer would always buy the same brand, which is unrealistic.

5.2.2.3 *Non-linear operator models*

Under this heading the model known as Luce's Beta-model and the Urn model will be treated.

5.2.2.3.1 *The Beta model*

Contrary to the linear operator models of the Bush-Mosteller type, also called Alpha models, in Luce's Beta Model, the operator associated with an event does not work directly on a response probability, but on a more fundamental variable called response strength. The response probabilities are functions of the response strengths. The Beta model is based on the response strength scale derived from Luce's choice axioms (see Luce, 1959). In the model, to every response i there corresponds a response strength, indicated as v_i. When there are n different responses, the chance that response i will be chosen is:

$$\text{Prob (response } i) = \frac{v_i}{\sum_{j=1}^{n} v_j}$$

For this probability only the ratio of the strength of response i to the other response strengths is important, not the values of the response strength itself. In this sense it can be said, that the response strength is more basic than the probability.

An operator associated with an event has the effect of multiplying the response strengths by a constant. Suppose there are two responses 1 and 2 with response strengths v_1 and v_2, and event k occurs, with which operator L_k is associated. Generally the effect of event k may be different for v_1 and v_2 and both effects should be specified separately. This can be done by letting L_k operate on the vector (v_1, v_2). Then the effect of event k can be expressed as follows:

$$L_k(v_1, v_2) = (a_k v_1, b_k v_2) \qquad (5.5)$$

a_k and b_k are positive constants, which are specific for event k. Note that response strength is unbounded.

When p_1 is the initial probability of a response 1, and p_2 is this probability after event k has occurred, we have:

$$p_1 = \frac{v_1}{v_1 + v_2}$$

and

$$p_2 = \frac{a_k v_1}{a_k v_1 + b_k v_2}$$

If we let $v = v_1/v_2$ and $\beta_k = a_k/b_k$,
we have for the transformation of the probability of response 1, which is indicated by the symbol Q:

$$p_2 = Q_k p_1 = \frac{\beta_k p_1}{1 - p_1 + \beta_k p_1} \tag{5.6}$$

which is evidently non-linear.

Two properties of this Beta model are worth mentioning. The first refers to the repetitive application of the same operator. For example, when L_k is applied to the initial vector of response strengths (v_1, v_2) twice in succession we get:

$$L_k L_k(v_1, v_2) = L_k^2(v_1, v_2) = (a_k^2 v_1, b_k^2 v_2) \text{ and generally:}$$
$$L_k^t(v_1, v_2) = (a_k^t v_1, b_k^t v_2)$$

Therefore p_t, the probability of choosing response 1 after t successive applications of L_k to the initial response strengths vector (v_1, v_2), is:

$$p_t = \frac{a_k^t v_1}{a_k^t v_1 + b_k^t v_2} = \frac{\beta^t v}{\beta^t v + 1}$$

$$\text{So for } \beta > 1 \quad \lim_{t \to \infty} p_t = 1$$

$$\text{So for } \beta = 1 \quad \lim_{t \to \infty} p_t = \frac{v}{v+1} = p_1$$

$$\text{So for } \beta < 1 \quad \lim_{t \to \infty} p_t = 0$$

We see that the limit points of p are generally 0 or 1. It could be imagined that the brand choice process can be described by the Beta model. Then in a two-brand market a brand 1 purchase can be conceived of as being one of the two possible responses, say a response of type 1. In the case of positive learning, for the operator associated with a brand 1 purchase, generally the multiplying constant a will be greater than b, so $\beta > 1$. This means that after many purchases of brand 1, the probability of repurchasing brand 1 will ultimately reach the value 1. As remarked earlier, this is unrealistic for a brand choice process.

The second property of the Beta model to be considered is that of the commutativity of operators. With operators L_k and L_j we have:

$$L_j L_k(v_1, v_2) = L_k L_j(v_1, v_2) = (a_k a_j v_1, b_k b_j v_2)$$

and

$$Q_j Q_k p_1 = Q_k Q_j p_1 = \frac{\beta_j \beta_k v_1}{\beta_j \beta_k v_1 + 1} \tag{5.7}$$

So the order in which the operators are applied i.e., the order in which the events occur, is not important for their influence on p. For a brand choice process this would mean that the influence of former purchases does not diminish with the number of purchases made later. Referring to the discussion in section 5.2.2.2, it can be said that this is not in agreement with reality.

The conclusion to be drawn from the two properties discussed is that the Beta model does not seem to be very suitable as a model for brand choice processes.

5.2.2.3.2 *The Urn Model*

In the Urn model (see Audley and Jonckhere, 1956), the strength of a response is represented by a physical equivalent, viz., the number of balls of a certain kind in an urn. For example, when there are two possible responses, 1 and 2, their respective strengths can be represented by the number of white and red balls in an urn. Let the number of white balls be w and the number of red balls r. Then p, the probability of response 1, is:

$$p = \frac{w}{w+r}$$

An event works out as a change in the respective numbers of balls in the urn. For example, a type k event means that the number of white balls increases with w_k and the number of red balls with r_k. Of course w_k and/or r_k may be negative. Now the effect of event k on the response strengths, which can be indicated by the operator U_k, is:

$$U_k(w, r) = (w + w_k, r + r_k) \tag{5.8}$$

Note via comparison of (5.8) with (5.5), that in the Urn model the effect of an operator on the response strengths is additive, whereas in the Beta model it is multiplicative. In both cases the response strengths can grow without limit, in the Beta model this growth is generally faster.

When p_1 is the initial probability of response 1, and p_2 is this probability after an event of type k, we have:

$$p_2 = Q_k p_1 = \frac{w+w_k}{w+r+w_k+r_k}$$

The recursive formula for p, analogous to (5.6), is complicated and will not be given here. It is clear that the transformation of p is not linear.

With respect to the properties of boundary values for p and commutativity of operators, the following can be noted for the Urn model:

$$Q_k^t p_1 = \frac{w+tw_k}{(w+r)+t(w_k+r_k)} = \frac{w}{(w+r)+t(w_k+r_k)} +$$

$$+ \frac{1}{\left(\frac{w+r}{tw_k}\right)+\left(\frac{w_k+r_k}{w_k}\right)}$$

$$\text{So} \lim_{t\to\infty} Q_k^t p_1 = \frac{w_k}{w_k+r_k},$$

which means, that generally the limit point of p is not equal to 0 or 1, as is the case in the Beta model. This latter point is attractive with respect to the application of the Urn model in brand choice processes. In addition we have:

$$Q_j Q_k p_1 = Q_k Q_j p_1 = \frac{w+w_k+w_j}{w+r+w_k+w_j+r_k+r_j} \tag{5.9}$$

So we see that, as in the Beta model, the operators commute. Because this means that former purchases cannot be forgotten, the Urn model does not appear to be a suitable model for brand choice processes.

One general remark which can be made here is that – as can be seen from the treatment given above – in the non-linear operator models the operator works indirectly on the response probability p, while in the linear models the influence on this probability is direct.

5.2.3 *Stimulus sampling models*

In stimulus sampling models, initially developed by Estes (1950), the stimulus situation at a response occasion is conceived of as a set of hypothetical 'molecular' stimulus elements or stimulus components. Every stimulus element is connected or associated with one and only one response. Further, these stimulus elements are left undefined, they have no equivalents in reality. We will consider the case where only two different responses are possible: response 1 and response 2. At a certain moment a number of the stimulus elements in the set are associated with or conditioned to response 1. Let this number be n. The other stimulus elements are associated with response 2. Their number is $N-n$, where N is the total number of hypothetical stimulus elements in the set. Stimulus models are also called State Models. The number of an individual's stimulus elements associated with response 1 can be called the state of that individual.

In the stimulus sampling model it is assumed that at a response occasion a sample of all the stimulus elements in the set is drawn. Let the sample size be S. Generally, the sample will contain stimulus elements associated with response 1 and with response 2. Let these numbers be s and $(S-s)$ respectively. Then the probability of a response 1 is s/S, the proportion of stimulus elements associated with response 1 in the sample. Of course, when the sample is randomly drawn from the set, the expected value of s/S is n/N. The effect of the response made and/or the consecutive outcome may be that the associations of the stimulus elements present in the sample change. Elements initially associated with response 1 may become associated with response 2 and vice versa. This means that the probability of a response 1 at the next response occasion may be different from the probability before.

In these few sentences we have roughly outlined the idea of stimulus sampling models, an idea about which much recent literature can be found. There is a great variety of types of these models. The models differ, for example, with respect to the number of stimulus elements N assumed to be present in the set, which can be one, very small, or can go to infinity. Also the number of stimulus elements sampled S is different in different models, sometimes it is only one. Another aspect in which the models differ is the way in which the effect of response and/or outcome works out on the association of the sampled elements. For an extensive survey of various types of stimulus sampling models see Luce, et. al. (1963, chapter 10) or Coombs, et al. (1970, section 9.3).

Because of this variety, there is no one unique mathematical formulation for the flow of probability from trial to trial in the case of stimulus sampling models. Here an example for a specific case will be given. Let the effect of a response 1 be such that all stimulus elements in the sample associated with response 1 remain in that condition, while those elements in the sample initially associated with response 2 become conditioned to 1. In this case response 1 is said to be reinforced. When the initial number of stimulus elements in the set associated with response 1 is n_1, and after a response 1 this number is n_2, the effect of the latter response can be expressed as:

$$n_2 = n_1 + (S - s) = n_1 + S\left(1 - \frac{s}{S}\right),$$

or

$$\frac{n_2}{N} = \frac{n_1}{N} + \frac{S}{N}\left(1 - \frac{s}{S}\right) \tag{5.10}$$

Let the probability of response 1 before the response was made be p_1 and after the response p_2. Then we have:

$$p_1 = \frac{n_1}{N} \quad \text{and} \quad p_2 = \frac{n_2}{N}$$

When s/S is set equal to its expected value n_1/N, (5.10) becomes:

$$p_2 = p_1 + \frac{S}{N}(1 - p_1) \tag{5.11}$$

With $\alpha = 1 - \frac{S}{N}$, we get:

$$p_2 = \alpha p_1 + (1 - \alpha) \tag{5.12}$$

From a comparison with (5.1) it can be seen that this transformation is of the Bush-Mosteller type. In fact we have a special case here, because there is only one parameter α.

Bush and Mosteller (1955, chapter 2) show that a stimulus sampling model can also be constructed which generates their general linear operator model. So we see that stimulus sampling models may mathematically be equal to linear operator models, which, of course, does not hold for all types of stimulus sampling models. Because of their great flexibility, it may be that stimulus sampling models can be fruitfully used for brand choice processes. In section 5.5 a brand choice model of the stimulus sampling type will be treated.

In this section some further properties of the linear operator model, given in section 5.2.2.2 will be discussed. Throughout this section the case of two responses is assumed.

5.3.1 *The general case*

For the general case, i.e., with no specific assumptions about the slope parameters in the transformation (as opposed to that in section 5.3.2), some alternative expressions in which the operator can be written will first be given. Then the effect on p of the occurrence of a number of consecutive events, i.e., the consecutive application of a sequence of operators, will be examined. After that the origin of events, i.e., how which event occurs and how which operator applies in a given situation are determined, will be discussed, and a classification of events according to their origins given.

Finally some remarks will be made about the equilibrium distribution of p, the probability of a response 1.

5.3.1.1 *Alternative formulations for the operator*

The transformation of p by the ith operator associated with an event of type i, was given in section 5.2.2.2 as

$$Q_i p = \alpha_i p + a_i \tag{5.13}$$

with the restrictions:

$$\left.\begin{aligned} &0 \leq a_i \leq 1 \\ &0 \leq \alpha_i \leq (1-a_i) \end{aligned}\right\} \tag{5.14}$$

Now (5.13) is known as the so-called slope-intercept form of the operator. Here we will give two other forms in which the operator is often written. When we define:

$$b_i = 1 - a_i - \alpha_i \tag{5.15}$$

we can write (5.13) as:

$$Q_i p = p + a_i(1-p) - b_i p, \tag{5.16}$$

which is the so-called gain-loss form of the operator. Here the new probability is written as the old probability p, plus a term proportional to the maximal possible gain in probability, minus a term proportional to the maximal loss in probability. This way of writing facilitates interpretation of the values of the learning parameters.

$$\text{Of course: } 0 \leq b_i \leq 1, \tag{5.17}$$

as can be derived from the restrictions (5.14).

By defining:

$$\lambda_i = \frac{a_i}{1-\alpha_i}, \tag{5.18}$$

$$\text{with, of course, } \quad 0 \leq \lambda_i \leq 1 \tag{5.19}$$

we can write (5.13) as:

$$Q_i p = \alpha_i p + (1-\alpha_i)\lambda_i \tag{5.20}$$

which is known as the fixed-point form of the operator. It is immediately clear that for $p = \lambda_i$, we have

$$Q_i p = p,$$

so in this case p remains fixed and is not changed by the operator $Q_i \cdot \lambda_i$ is the limit point or boundary value of p for the operator Q_i. The fixed-point formulation is useful when the effect of repetitive application of the same operator is studied, as is done in the next section.

In Figure 5.1 the relationship between p and $Q_i p$ is shown graphically. The geometric meaning of the parameters of the different formulations is also indicated there.

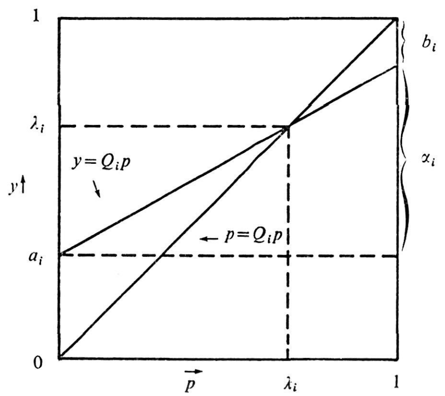

Fig. 5.1 Relationship between p and $Q_i p$ with the geometric interpretation of the different parameters

5.3.1.2 *A sequence of events*

Here we consider the effect on p of a number of consecutive transformations, i.e., of the occurrence of a sequence of events. When two events i and j occur in succession the sequence of operators is $Q_j Q_i$ and the corresponding transformation of p is

$$Q_j Q_i p = \alpha_j \alpha_i p + \alpha_j a_i + a_j \tag{5.21}$$

In section 5.2.2.2 it was seen that generally:

$$Q_j Q_i p \neq Q_i Q_j p, \quad \text{for} \quad i \neq j,$$

so the operators do not commute. This implies that the algebraic expressions for the probability after a number of different events are generally very complicated. Except for sequences in which a small number of operators alternate in a very systematic way, it is impracticable to consider these cumbersome expressions.

In what follows, we treat only the case of repetitive application of the same operator, i.e., the case where a series of events of the same type occurs. In brand choice terms, this means the consecutive buying of the same brand.

The transformation of p by the operator Q_i, written in the fixed-point form (5.20), is:

$$Q_i p = \alpha_i p + (1-\alpha_i)\lambda_i \tag{5.22}$$

So: $Q_i Q_i p = Q_i^2 p = \alpha_i^2 p + \alpha_i(1-\alpha_i)\, \lambda_i + (1-\alpha_i)\, \lambda_i = \alpha_i^2 p + (1-\alpha_i^2)\lambda_i$,

and, via the method of mathematical induction, it can be proven that generally:

$$Q_i^t p = \alpha_i^t p + (1-\alpha_i^t)\lambda_i \tag{5.23}$$

When $0 \leq \alpha_i < 1$ – which is mostly the case, because α_i must satisfy the restrictions (5.14) and $\alpha_i = 1$ would imply the extreme case of constant p without learning – we have:

$$\lim_{t\to\infty} \alpha_i^t = 0$$

and

$$\lim_{t\to\infty} Q_i^t p = \lambda_i \tag{5.24}$$

So we see that the fixed point λ_i, also called the limit value or boundary value, is the value which the probability of a response 1 approaches, after a certain number of successive transformations by the operator Q_i. The rate of approach depends on the value of α_i. A value near to one gives a slow movement, when α_i is small the approach to λ_i is much faster. Note that the initial probability p has no influence on the boundary value λ_i.

As an example let us take the brand choice processes for fopro and margarine analyzed in chapter 4. These processes can be described by the LLM, which is a special case of a linear operator model. Then p is the probability of a brand 1 purchase and an event 1 is defined as a purchase of brand 1. We take for fopro: brand 1 = F1 and for margarine: brand 1 = M1. The parameter values given in Table 4.23 for these processes are used. Translated into fixed-point form parameters they are: $\alpha = .5448$ and $\lambda = .9765$ for fopro, while $\alpha = .8829$ and $\lambda = .9898$ for margarine. Now in Figure 5.2 $Q_1^t p$, computed according to (5.23), is drawn for fopro and margarine for $t = 1, \ldots, 60$. The initial value of the probability p is set equal to .5. The curves are indicated in the figure as curve a and curve b respectively.

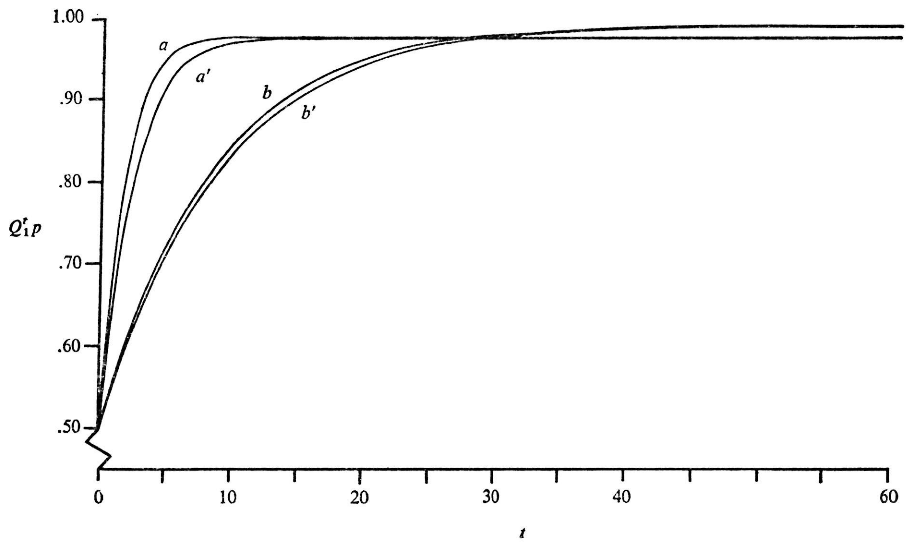

Fig. 5.2 $Q_1^t p$ *for fopro*$_1$ *(brand 1 = F1) and margarine (brand 1 = M1): Exact (a resp. b) and exponentially approximated (a' resp. b'); p = .5*

The effect of the higher α-value for margarine becomes immediately clear from the diagram. The movement of the probability to the boundary value is much slower than for fopro. Since the slope parameters for beer are not very different from those for fopro, the picture for beer would be very similar to that for fopro, therefore it is not given here.

The gradually diminishing growth of $Q_i^t p$, when t increases, which is shown by Figure 5.2, raises the question if it is possible to describe $Q_i^t p$ by an exponential function. As shown by B & M, this is possible as an approximation.

Let $p(t) = Q_i^t p$, where $p(0) = p$

Then from (5.22):

$$p(t+1) - p(t) = (1-\alpha_i)(\lambda_i - p(t))$$

When $\Delta p = p(t+1) - p(t)$

and $\Delta t = 1$, we have

$$\frac{\Delta p}{\Delta t} = (1-\alpha_i)(\lambda_i - p(t))$$

When as an approximation t is thought to be continuous, we can write down the differential equation:

$$\frac{\mathrm{d}p}{\mathrm{d}t} = p' = (1-\alpha_i)(\lambda_i - p(t)) \tag{5.25}$$

When we define

$$f(t) = \lambda_i - p(t),$$

(5.25) can be written as

$$f'(t) = -(1-\alpha_i)\, f(t),$$

with the well-known solution:

$$f(t) = f(0)\, \mathrm{e}^{-(1-\alpha_i)t} \tag{5.26}$$

Substituting back, we get:

$$p(t) = \mathrm{e}^{-t(1-\alpha_i)}\, p(0) + (1-\mathrm{e}^{-t(1-\alpha_i)})\, \lambda_i \tag{5.27}$$

So, as an approximation, the growth of $p(t)$ to the upper bound λ_i can be described by an exponential function.

In Figure 5.2 we have also drawn these 'exponential approximations' of $Q_i^t p$, computed according to (5.27), for the fopro and margarine case, exact functions for which have already been pictured. The exponential curves are indicated in the figure as curve a' and curve b' respectively.

It appears that for margarine the approximation is almost perfect, while for fopro it is somewhat less. This is a consequence of the higher α-value for margarine, as can be seen from what follows.

Let $\alpha_i' = e^{-(1-\alpha_i)}$, then (5.27) becomes:

$$p(t) = (\alpha_i')^t\, p(0) + (1 - (\alpha_i')^t)\lambda_i \tag{5.28}$$

So the approximation is perfect if

$$\alpha_i' = e^{-(1-\alpha_i)} = \alpha_i \tag{5.29}$$

It can be inferred a series expansion of $e^{-(1-\alpha_i)}$, that the nearer α_i is to one, the better equation (5.29) is fulfilled. So a high value for α results in a better approximation by the exponential function. In this connection it can be remarked that the estimates for the slope parameter α, which were found for the empirical brand choice processes in section 4.5, were all well above .5.

5.3.1.3 *Types of events*

Until now we have gone no further than stating that the response and/or outcome of a trial determine the event which occurs, with the corresponding operator which brings about a transformation of p. Here a classification of types of events according to their origin will be given. The terminology used, with such terms as experimenter and subject, has been taken from psychological learning experiments, the origin of linear learning models.

a. *Experimenter controlled events*

Here it is exclusively the outcome of a trial which determines the occurring event. The subject's response has no influence. An example is a T-maze experiment in which a rat is rewarded for a fixed portion of the times that it turns right and in which a trial is terminated only after a right turn, so left turns are disregarded. Because of the fixed probability that an operator is applied, mathematically this case is the most tractable. But due to the fact that the subject's response does not influence the subsequent choice probabilities at all, this type of event is not likely to fit many real learning processes, brand choice processes included.

b. *Subject controlled events*

For this type of event it is exclusively the response made by the subject which determines the event occurring. As an example a T-maze experiment can be mentioned, where going one direction is always rewarded and going the other direction is never rewarded. In fact response and outcome can be said to coincide for this case.

The LLM for brand choice processes, as defined in section 3.6.1, is a case of subject controlled events. In this model the consumer, who can be indicated as the subject, makes purchases of a product, where the response is defined as the brand bought. When brand 1 is bought, the purchase operator is applied, otherwise the rejection operator works, see section 3.6.1. Here the response made, i.e., the brand chosen, exclusively determines the type of event and with it the operator to be applied. Each purchase of brand 1 constitutes the same type of event. In *T*-maze terminology, it can be said that a brand 1 purchase, for example, always constitutes a reward and a brand 0 purchase a non-reward. This assumption, inherent in the LLM, can not always be satisfied because it would not then be possible for experience with a brand to also influence the probability of buying it again. In section 5.4 it will be shown that the linear operator model can also be applied to brand choice processes without this assumption.

c. *Experimenter-subject controlled events*

For this type of event the response and the outcome together determine which type of event occurs and which operator applies. An example is the *T*-maze experiment, where every combination of (1) going left or going right and (2) being or not being rewarded constitutes an event. In this case there are 4 different events possible.

A generalisation of the LLM, presented in section 5.4, is a brand choice model with experimenter-subject controlled events.

5.3.1.4 *Equilibrium distribution of p*

When a large number of subjects perform the same learning process, i.e., with the same parameters, it is interesting to know the distribution of p over the subjects after a great many trials.

In B & M, chapter 5, it is shown for the case of 2 experimenter controlled events, that an asymptotic or equilibrium distribution exists. This means that the distribution of p at trial t and trial $(t+1)$ are equal for $t \to \infty$. This equilibrium distribution is independent of the initial distribution of p-values over the subjects, i.e. when the learning process began.

For our purposes it is important to know how the probability p in the LLM behaves after a large number of purchases. The first moment of this equilibrium distribution takes then the interpretation of the long-term market share of brand 1.

It should be emphasized that a different equilibrium situation from that mentioned in section 4.7.3 is meant here. Here what is referred to is the distribution of p in the population after a great number of purchases;

in section 4.7.3 the situation where an individual consumer purchases the same brand during a great number of consecutive purchases and reaches the boundary value for p is, after Frank (1962), called an equilibrium state for that consumer.

Now for the subject controlled event case, which is the LLM, it has been proven for the 2-response situation (B & M, p. 199) that an equilibrium distribution of p exists which is independent of the initial distribution of p, provided that the slope parameters α_i are not negative and the absolute value of the difference between the two limit points λ_i is less than unity. Note that the latter conditions are met for all the empirical processes analyzed in section 4.5.

Contrary to the case of experimenter-controlled events, it is not possible to give explicit expressions for the moments of the equilibrium distribution of p. The difficulty inherent here can be shown as follows. We have from (5.13):

$$Q_1 p = a_1 + \alpha_1 p$$
$$Q_2 p = a_2 + \alpha_2 p.$$

When the starting value of the probability of a response 1 is p_0, the probability after one trial, denoted as p_1, is:

$Q_1 p_0$ with probability p_0

and $Q_2 p_0$ with probability $(1-p_0)$.

So the expected value of p_1 is:

$$\begin{aligned} Ep_1 &= p_0(a_1+\alpha_1 p_0) + (1-p_0)(a_2+\alpha_2 p_0) \\ &= a_2 + (a_1 - a_2 + \alpha_2)\, p_0 + (\alpha_1 - \alpha_2)\, p^2 \end{aligned} \tag{5.30}$$

Thus for the first moment of p_1, p_0^2 is required. And when p_0 itself is not fixed, but stochastic, Ep_0^2 must be known to find Ep_1.

As is shown by B & M, this extends to higher moments. Generally, it can be stated that for the mth moment of p_t the $(m+1)$th moment of p_{t-1} is required. So only recurrent relations for the moments can be found and no explicit expressions.

From (5.30) it is clear that this difficulty is resolved when:

$$\alpha_1 = \alpha_2, \tag{5.31}$$

i.e., when the slope parameters of the two operators are equal.

Therefore (5.31) is often introduced as an assumption in the case of subject controlled events; this is known as the equal α-condition. Generally this is not done to make the learning model agree more with reality, but just to make the model mathematically tractable. The LLM used for brand choice processes in the preceding chapters is an example of subject controlled events, with the equal α-condition satisfied. This is clear from the formulation given in section 3.6.1. So the model, which – in brand choice literature – is called the Linear Learning Model is in fact a special case of those models which, in mathematical learning psychology, are called linear operator models of the subject controlled event type.

In section 5.3.2 we consider only linear operator models in which the equal α-condition satisfied.

5.3.2 *The case of equal slope parameters*

In the application of learning models to brand choice processes the equal slopes parameters of purchase and rejection operator have been accepted without further preface. It is desirable, however, to examine the severity of this limitation for real brand choice processes. This is done first in the present section. After that the computation of the moments for the equilibrium distribution of p for the LLM will be discussed.

5.3.2.1 *Severity of the restriction*

The LLM satisfies the equal α-condition because the slope parameters of purchase and rejection operators are equal. Apart from the better mathematical tractability so obtained, there is no reason why this should be so. It is important to know if this equal α-condition constitutes a major limitation for real brand choice processes.

To obtain some impression, the learning model parameters were computed with the slope parameters being allowed to differ for the brand choice process of fopro described and analyzed in chapters 2 and 4 of this book. The estimation procedure is quite similar to that given in section 3.6.2, except that now the relative frequencies of sequences of length 4 are functions of 8 parameters instead of 7. The way in which the expressions for these relative frequencies are formed is analogous. The results can be found in Table 5.1, where the figures from Table 4.23, the case of equal slope parameters, with which a comparison should be made, are also given. In this table c_1 is the slope of the purchase operator, c_2 is the slope of the rejection operator, and a and b are the LLM parameters as defined in section 3.6.

From Table 5.1 it can be concluded that when c_1 and c_2 are allowed to differ they remain rather near to the original c-values. In all cases the original value for the common slope parameter c lies between the estimated values for c_1 and c_2. Also the parameters a and b do not change very much, while the fit of the model for the case of unequal slope parameters is not evidently better.

Table 5.1 Estimates of the parameter of the LLM *for the case of equal slopes (es) of purchase and rejection operator, resp. unequal slopes (us) for these operators. Product fopro, first 10 purchases*

Brand 1 =	a_{es}	a_{us}	b_{es}	b_{us}	c_{es}	c_1	c_2	$\chi^2_{8(es)}$	$\chi^2_{7(us)}$
F1	.0052	.0031	.4393	.4795	.5448	.5042	.5699	2.67	3.11
F2	.0022	.0034	.4138	.3383	.5652	.6458	.5216	5.67	3.53
F3	.0033	.0029	.2625	.2900	.7302	.7004	.7338	5.20	5.71
F4	.0058	.0022	.3844	.4103	.5925	.5646	.6742	10.57	10.33

Generally one would expect the discrepancy between model and reality to be less in this case, because unequal slope parameters provide one possibility more to fit the model to the data. In this connection it should be noted that the pattern search method for minimization does not guarantee an absolute minimum and that the minimization procedure is complicated by the parameter added. So for the fopro brand choice process examined here, it is concluded that the condition of equal slope parameters for the purchase and rejection operators, as is implicit in the LLM, does not represent a severe restriction. Of course it would have been useful to also carry out this analysis for beer and margarine, but because of the long computation times this was left for later research. It should be observed that the linear operator model with unequal slope parameters, as applied here, constitutes a generalisation of the LLM. In practical situations it may happen that the usual LLM does not fit an empirical brand choice process, but that the unequal-slopes-LLM does. In the rest of this section only the case of equal slope parameters will be treated.

5.3.2.2 *Moments of the equilibrium distribution of p in the* LLM

In this section we are concerned with the moments of the equilibrium distribution of p for the LLM.

5.3.2.2.1 *The first moment*

The LLM is a case of subject controlled events, so with regard to the first moment, the starting point is (5.30). Because of the equal slope condition:

$$\alpha_1 = \alpha_2 = \alpha,$$

$$Ep_1 = a_2 + (a_1 - a_2 + \alpha)\, p_0$$

This holds when p_0 is fixed. When p_0 is stochastic with an expected value Ep_0, we get:

$$Ep_1 = a_2 + (a_1 - a_2 + \alpha)\, Ep_0 .$$

This can be extended to the general relation:

$$Ep_{t+1} = a_2 + (a_1 - a_2 + \alpha)\, Ep_t \qquad (t = 0, 1, 2, ...) \tag{5.32}$$

Let $\beta = (a_1 - a_2 + \alpha)$ and $\mu = \dfrac{a_2}{1-\beta}$, then (5.32) becomes:

$$Ep_{t+1} = \beta Ep_t + (1-\beta)\, \mu \qquad (t = 0, 1, 2, ...)$$

This is a recurrent relation of the same mathematical form as the equation for the linear operator in the fixed-point form (5.22) and the solution is:

$$Ep_t = \beta^t Ep_0 + (1-\beta^t)\, \mu \tag{5.33}$$

From restrictions (5.14) we derive that:

$$-1 \leq \beta \leq 1,$$

Therefore, generally:

$$Ep_\infty = \lim_{t\to\infty} Ep_t = \mu = \frac{a_2}{1-(a_1 - a_2 + \alpha)} \tag{5.34}$$

Expressed in the LLM-parameters a, b and c, used in sections 3.6 and 4.5 with Q_1 as purchase and Q_2 as rejection operator, (5.34) becomes:

$$Ep_\infty = \lim_{t\to\infty} Ep_t = \frac{a}{1-b-c} . \tag{5.35}$$

In marketing terms Ep_∞ can be interpreted as the long term market share of brand 1, when it is assumed that the same brand choice process remains in operation.

In the first column of Table 5.2, the equilibrium market shares are given for the parameter values from Table 4.23, i.e., for the brand choice processes of fopro, beer and margarine, first 10(20) purchases. From Table 5.2 we can infer that the long term market share for brand F1 is .3245. Compared with the current value of .3762 in Table 4.23 this means a decrease, so it can be said that when this brand choice process remains in operation brand F1 will decrease in market share. The other brands can be considered in analogous ways. The long term market shares for margarine do not seem to make much sense. This can be explained by the low a-values for these brand choice processes, which are given in Table

4.23. When computing the long term market share by means of (5.35) the parameter a plays a crucial role. When a is very small, as is the case for margarine, small estimation errors have a relatively great influence so that long term predictions are not then reliable.

Table 5.2 Moments for the equilibrium distribution of p in the case of the LLM *for the parameter values of Table 4.23*

Product	*Brand* 1 =	Ep_∞	β	Ep_∞^2	σ_{p_∞}	Ep_∞^3 (*exact*)	Ep_β^3 (*beta distr.*)
Fopro	F1	.3245	.9841	.2937	.4340	.2788	.2788
	F2	.1050	.9790	.0867	.2750	.0785	.0783
	F3	.4509	.9927	.4081	.4525	.3870	.3869
	F4	.2503	.9769	.2061	.3787	.1858	.1855
Beer	B1	.4841	.9650	.4104	.4196	.3746	.3737
	B2	.1224	.9755	.0824	.2597	.0663	.0658
	B3	.0574	.9791	.0473	.2098	.0429	.0427
	B4	.2757	.9699	.2214	.3813	.1967	.1962
Marg	M1	.1643	.9986	.1405	.3369	.1295	.1294
	M2	.3947	.9962	.3735	.4666	.3630	.3630
	M3	.0680	.9950	.0519	.2174	.0450	.0449
	M4	.0074	.9980	.0060	.0771	.0053	.0053

Apart from the long term market share itself, it is important how quickly this share will be reached. As can be derived from (5.33), this is determined by the value of β. The higher the value of β, the slower the convergence of Ep_t. Table 5.2 also gives the β-values for the brand choice processes in question. It can be seen that these values are generally high, which means that the processes do not quickly go to equilibrium. Apparently the markets behave in a very stable way which is in agreement with the relatively small variations in market shares, as is shown by Table 2.3.

5.3.2.2.2 *Second and third moment; approximation by the beta distribution*
In an analogous way, as was done above for the first moment, recurrent relations can be written for the second moment of p, from which Ep_∞^2 can be derived. We will not give this derivation here, but directly state the expression as it can be found in B & M (equation 5.47):

$$Ep_\infty^2 = \frac{a_2^2 + (a_1^2 - a_2^2 + 2a_2\alpha)\, Ep_\infty}{1 - \alpha(2a_1 - 2a_2 + \alpha)} \tag{5.36}$$

In Table 5.2 the asymptotic second moments and standard deviations for p are given for the brand choice processes mentioned earlier. It appears that the standard deviations are relatively great, which can be brought into relation with the fact that these processes, as seen in section 4.7.2, show a lot of boundary behavior, i.e., there are many consumers with p either near to zero or near to one.

From the general recurrent formula given by B & M (1955, p. 113, eq. 5.40), an expression for the third moment as a function of the first and second moment can be derived. For the expectation of p^3 at trial $t+1$ we get:

$$\begin{aligned} Ep_{t+1}^3 &= a_2^3 + (a_1^3 - a_2^3 + 3a_2^2\alpha)\, Ep_t + \alpha(3a_1^2 + 3a_2^2 + 3a_2\alpha)\, Ep_t^2 \\ &\quad + \alpha^2(3a_1 - 3a_2 + \alpha)\, Ep_t^3 . \end{aligned} \tag{5.37}$$

Now when $t \to \infty$, $Ep_{t+1}^3 = Ep_t^3 = Ep_\infty^3$, which, substituted in (5.37), gives:

$$\begin{aligned} Ep_\infty^3 &= a_2^3 + (a_1^3 - a_2^3 + 3a_2^2\alpha)\, Ep_\infty + 3\alpha(a_1^2 + a_2^2 + a_2\alpha)\, Ep_\infty^2 \\ &\quad + \alpha^2(3a_1 - 3a_2 + \alpha)\, Ep_\infty^3 . \end{aligned} \tag{5.38}$$

When Ep_∞ and Ep_∞^2 are known from (5.35) and (5.36), Ep_∞^3 can be derived as the solution of (5.38). The numeric values for Ep_∞^3 for the brand choice processes of Table 5.2 were calculated; they are given in the last but one column of that table.

This asymptotic third moment can be used to examine if the asymptotic distribution of p can be fit by a beta distribution. When it is assumed that this is the case, then by means of the relations (3.3) and (3.4) the parameters of the beta distribution can be derived from the first and second asymptotic moment and with these parameters the third moment can be computed. The agreement between the third moment of the asymptotic distribution, calculated by means of the beta distribution, and the third moment derived from (5.38) is an indication of the fit of the beta distribution for the asymptotic distribution of p. The numerical values of Ep_∞^3, calculated via the beta distribution, are given in the last column of Table 5.2. Comparing these values with the exact values in the last but one column of this table, we see that the agreement is strikingly good. So it appears that for the brand choice processes investigated, the distribution of p in the equilibrium situation can be fit very well by a beta distribution. This point was referred to in section 4.5.5.3.

In section 5.3.1.3 it was seen that the LLM is a linear learning model of the subject controlled events type. This means that the response made exclusively determines the operator that is applied. So in the LLM every time brand 1 is bought the purchase operator works and after each brand 0 purchase the rejection operator is applied. Generally speaking, this means that every time brand 1 is purchased the probability of buying brand 1 again is increased, while each brand 0 purchase has the effect of decreasing this probability. This is a rather rigid assumption because it can be imagined that after a brand 1 purchase the probability of buying this brand again will not increase at all, e.g., as a consequence of disappointing experience with the brand. Generally, every brand 1 purchase will not always constitute a 'reward' for every consumer. The same can be said with regard to brand 0. Of course, it can also be imagined that the parameter b of the LLM is negative. In that case a purchase of brand 1 would always imply a decrease in the probability p to choose brand 1, which is still less realistic than the assumption that each brand 1 purchase increases p.

Because of these limitations of the LLM a linear operator model was considered in which the purchase c.q. rejection operator is not assumed to always work when brand 1 c.q. brand 0 is purchased, but where this occurs only with probability π. We call this the π-model. In this model, when the operator does not work the probability remains unchanged, so it can be said that then the unit operator is applied. So the application of purchase and rejection operator is contingent on the working of the learning mechanism. It can be imagined, for example, that in the case of a good experience with brand 1 the learning mechanism works but that at another occasion the product is consumed as a matter of course, with

	Learning mechanism works	*Learning mechanism does not work*
Brand 1 bought	$p' = a+b+cp$	$p' = p$
Brand 0 bought	$p' = a+cp$	$p' = p$
Probability	π	$(1-\pi)$

Fig. 5.3 Different possibilities for the transformation of p in the π-model

the consequence that p does not change at all. The different possibilities for the transformation of p are given in Figure 5.3. Here p' denotes the probability after transformation.
Here the brand bought is the response and the working/not working of the learning mechanism can be conceived of as the outcome of the trial (purchase). As is the case in experimenter-subject controlled events, response and outcome together determine the operator that is applied.

The parameters for this π-model for the fopro brand choice processes, first 10 purchases, were estimated. The estimation method is the same as that for the LLM given in section 3.6.2, except that now every purchase sequence of length 4 can be brought about in 8 different ways. After each of the first, second and third purchases of such a sequence, the learning mechanism may work or not work. For a specific purchase sequence the probabilities for each of these 8 possibilities, which individually can be expressed in a, b, c, π, μ_1, μ_2, μ_3 and μ_4 – added together – constitute the theoretical relative frequency. It will be clear that this complication increases considerably the computing time required. This is the reason that the estimation was not carried out for beer and margarine.

The results are given in Table 5.3.

Table 5.3 Estimated parameters of the Linear Operator Model, where the learning mechanism works with probability π, for fopro, first 10 purchases, and for Kuehn's snow crop data

Brand 1 =	a	b	c	μ_1	μ_2	μ_3	μ_4	π	χ^2_7	N
F1	.0086	.5930	.3828	.3684	.3324	.3179	.3158	.5677	4.81	4704
F2	.0028	.4802	.4961	.1515	.1322	.1264	.1263	.7688	6.86	4704
F3	.0084	.4579	.5348	.2403	.2048	.1862	.1750	.2858	6.97	4704
F4	.0080	.4707	.4986	.2196	.1874	.1727	.1652	.5158	9.21	4704
Snow Crop (Kuehn)	.0369	.5478	.2497	.2129	.1195	.0848	.0704	.4029	2.14	13519

It is concluded that the π-model gives a very satisfying description of the brand choice processes for the product fopro. The probability of the working of the learning mechanism is above .5 for F1, F2 and F4 and only .25 for F3. Comparing the parameter values with those for fopro in the case of the LLM, given in Table 4.23, we see that for the π-model the c-values are generally lower and the b-values higher. This means that the probability change, induced by the application of purchase or rejection operator is greater if such an operator is really applied. But of course in the π-model this application only occurs with probability π.

The parameters of the π-model for the frozen orange juice data (brand 1 = Snow Crop), which were used by Kuehn, were also estimated. These data were discussed in section 4.5.1. The results for these data are also presented in Table 5.3. From a comparison with the results for the LLM, given in Table 4.21, we conclude that the parameters changed in the same directions as observed for fopro above. The fit of the π-model is excellent for this orange juice data, being almost as good as that of the LLM.

In fact, the π-model is a generalisation of the LLM. When $\pi = 1$, we have the LLM again. Of course, as a consequence of this greater generality, we would expect the π-model to give a lower χ^2-value as a result of the estimation procedure, but in fact this is only the case for F4. Analogous to the remarks made in section 5.3.2.1 this will be due to the minimization procedure applied, which does not guarantee finding the absolute minimum and may be expected to produce worse results when the number of parameters increases.

The χ^2-values in Tables 5.3 and 4.23 for fopro are of the same order of magnitude. Therefore it can be said that the π-model meets the criterion for the fit of the empirical brand choice processes as well as the LLM; the same holds for the snow crop data. So the rigid assumption that the probability p will increase after every brand 1 purchase and decrease after a brand 0 purchase, is not essential for the use of linear operator models for brand choice processes.

The model can be generalized further. For example, it can be assumed that in the case of a brand 1 purchase the purchase operator is applied with probability π_1, while after a brand 0 purchase the rejection operator is applied with probability π_2. In the π-model above we have $\pi_1 = \pi_2 = \pi$.

Another possibility is to assume that when the purchase c.q. rejection operator does not work, other operators, which are not unit operators, are applied, as was assumed in the π-model. Then the model can be made such that it is possible for the probability of a brand 1 purchase to decrease after brand 1 is bought. Such an operator, which decreases the probability of a brand 1 purchase, can be assumed to work when a consumer has had a bad experience with this brand. Another kind of generalization of the LLM was given in section 5.3.2. So the general conclusion is that it is not necessary to restrict the application of linear operator models in brand choice processes to the LLM, but that more general models can also be used. In other words: the possibilities of linear operator models in brand choice processes are richer than those of the LLM alone.

Computer time for the estimation of the parameters, which is already considerable for the LLM, is still more for the generalized models. However faster minimization procedures than the pattern search method used reduces this problem.

5.5 THE PROBABILITY DIFFUSION MODEL (PDM)

This section treats the so-called Probability Diffusion Model (PDM), developed by Montgomery (1969), which is a special kind of stimulus sampling model, a type of learning model dealt with in section 5.2.3. We will briefly sketch the formulation of the model and the way in which its parameters can be estimated. After that the results of the application of the PDM to the fopro, beer and margarine data will be presented and these outcomes discussed at some length. Finally some attention will be paid to an extension of the PDM.

5.5.1 *Formulation*

First it should be emphasized that only an outline of the model is given here. An extensive description can be found in Montgomery (1969) or MM & M, chapter 6.

In the PDM every consumer is assumed to have a set of stimulus elements, each element of which is uniquely associated with one of two possible responses, indicated as A and B. In brand choice terms these responses are respectively a purchase of brand 1 and a purchase of brand 0. Let every consumer have N stimulus elements (s.e.) in his set and let the number of stimulus elements associated with response A be indicated by i. The number of s.e. N is the same for all consumers, while different consumers may have different i-values at the same moment. It can be said that i is the state variable for a consumer. A consumer in state i has a probability of an A response equal to i/N. As noted in section 5.2.3 the s.e. are hypothetical constructs, with no direct relationships to real observable quantities.

There is a mechanism which results in a consumer not always being in the same state. Stimulus elements first associated with A may become associated with B and vice versa. In the PDM these transitions of association are not a consequence of purchase feedback, but are assumed to be brought about by the environment of the consumer, an environment in which advertising effects play an important role. So the PDM is essentially a zero-order model in that it assumes that there is no purchase feedback. It is postulated that a consumer who is in state i at response occasion t, will make a transition $i \rightarrow i+1$ in the interval $(t, t+\Delta t)$ with probability $\lambda_i \Delta t + O(\Delta t)$, where $O(\Delta t)$ denotes a function of the order Δt, i.e. which tends to zero faster than Δt. In an analogous way, the probability that a consumer being in state i at response occasion t will make a transition to state $(i-1)$ in the interval $(t, t+\Delta t)$ is $\mu_i \Delta t + O(\Delta t)$. The probability of a transition to something other than a neighbouring state is $O(\Delta t)$. In this way a so-called linear birth-death process, see Cox & Miller (1965, chapter 4), has been formulated, where a transition of association of an s.e. from $B \rightarrow A$ is conceived of as a birth and a transition in the opposite direction is a death.

In the formulation t is conceived of as being a continuous variable, having a unit interval between subsequent response occasions (purchase moments). The transitions of association do not take place at discrete moments, i.e., at response occasions, as was indicated for the general stimulus sampling models in section 5.2.3, but is a continuous process here. Therefore the PDM is a special kind of stimulus sampling model. Montgomery gives two different specifications of the PDM. These specifications refer to the so-called transition intensities λ_i and μ_i.

1. *The Independent Elements Specification* (IES)

Here each stimulus element associated with B has a transition intensity α toward becoming associated with A, and each s.e. associated with A has a transition intensity β of becoming associated with B, while the N stimulus elements of a consumer are assumed to behave independently from one another.

Thus:

$$\lambda_i = (N-i)\alpha$$
$$\mu_i = i\beta.$$

It can be shown that such a birth-death process will ultimately reach a steady state with:

$$E\left(\frac{i}{N}\right) = \frac{\alpha}{\alpha+\beta} \tag{5.39}$$

$$\text{Var}\left(\frac{i}{N}\right) = \frac{1}{N}\left(\frac{\alpha}{\alpha+\beta}\right)\left(1-\frac{\alpha}{\alpha+\beta}\right). \tag{5.40}$$

Now the magnitude of N, the number of hypothetical stimulus elements of an individual consumer, is important. When N is small, the number of possible states in which a consumer may be – which is 1, 2, 3 ..., N – is small, i.e., an individual's response probability can only take a limited number of values. Montgomery argues that this is not realistic and that therefore N should be assumed to increase without limit. But then, as can be seen from (5.40), Var(i/N) in the steady state goes to zero, which implies that in the steady state all consumers will have a probability of choosing brand 1 equal to $\alpha/(\alpha+\beta)$ and that there is no variation around that value. Because all consumers have the same parameters α and β and their state variable i behaves according to the same stochastic process, in the IES with $N \to \infty$ all consumers will ultimately reach a probability of purchasing brand 1 equal to $\alpha/(\alpha+\beta)$ and remain in that state. So in the steady state the population will be homogeneous in the sense that each consumer will have the same probability. Concluding that this consequence of the IES is not in agreement with reality, Montgomery rejects this specification and turns to what he calls: the cohesive elements specification.

2. *The Cohesive Elements Specification* (CES)
This specification assumes that the stimulus elements do not behave independently, but attract one another. It is postulated that the transition intensity of each s.e. is increased by an amount γ for each element associated with the opposite response. Thus, with α and β defined as in the case of IES:

$$\lambda_i = (N-i)\,(\alpha+i\gamma)$$

$$\mu_i = i(\beta+(N-i)\gamma)\,.$$

For this case, it can be derived for the steady state, when $N \to \infty$, that:

$$E\left(\frac{i}{N}\right) = \frac{\alpha}{\alpha+\beta} \tag{5.41}$$

$$\text{Var}\left(\frac{i}{N}\right) = \frac{\alpha\beta\gamma}{(\alpha+\beta)^2\,(\alpha+\beta+\gamma)}\,. \tag{5.42}$$

So for the CES the steady state value of $E(i/N)$ is the same as for the IES, but now the variance of that quantity is not equal to zero. Therefore, an individual's value of i/N now remains varying around the expected value and in the steady state different individuals may have different probabilities of a brand 1 purchase.

Because this seems to be a more satisfactory description of reality, Montgomery decided to work further with the CES and derived estimation and testing procedures for this specification.

5.5.2 *Estimation and testing*

After formulating a model, the next task is to develop estimation and testing procedures to examine the performance of the model in empirical situations. We will briefly outline the procedure developed by Montgomery for the CES, for a more detailed treatment the reader is once more referred to the references mentioned.

Let $X(t)$ denote an individual's probability of response A at response occasion t. So $X(t)=i/N$, when i is the number of stimulus elements associated with response A at response occasion t. Now for the CES the following can be derived.

When the expected proportion of respondents making response A at t is denoted as $P(t)$, we have:

$$P(t) = E(X(t)) \tag{5.43}$$

and $$P(t) = P(0)\exp(-(\alpha+\beta)t) + \frac{\alpha}{\alpha+\beta}(1-\exp(-(\alpha+\beta)t)). \tag{5.44}$$

When $P(0,t)$ is the expected proportion of the respondents making response A at time 0 and A at time t again, then

$$P(0,t) = P(0,0)\exp(-(\alpha+\beta)t) + P(0)\frac{\alpha}{\alpha+\beta}(1-\exp(-(\alpha+\beta)t)) \tag{5.45}$$

where

$$P(0,0) = E(X^2(0)). \tag{5.46}$$

Equations (5.44) and (5.45) constitute the basis of the estimation and testing procedure, a procedure for which purchase sequences of individual families are needed. Let us assume that we have the purchase sequences of individual families during $(k+1)$ subsequent purchases. The first purchase is assumed to have taken place at $t=0$. Then the following quantities can be computed:

1. $Q(t)$ = the fraction of families purchasing brand 1 at time t, for

   $t = 0, 1, 2, \ldots, k.$

2. $Q(0,t)$ = the fraction of families purchasing brand 1 at time 0 and brand 1 again at time t, for $t=1, 2, \ldots, k$.

Now $P(0)$ can be set equal to its estimator $Q(0)$ and then with (5.44) and

(5.45) the theoretical values $P(t)$ and $P(0, t)$ can be computed for $t = 1, 2, \ldots, k$ and compared with the corresponding values of $Q(t)$ and $Q(0, t)$. For this to be possible it is necessary to assume values for the unknown α, β and $P(0, 0)$. The fit between observed and theoretical fractions, the Q-variables and the P-variables, can be expressed in a chi-square quantity. Now the unknowns α, β and $P(0, 0)$ are estimated by finding those values for them for which the corresponding P-variables are as close as possible to the observed Q-values, as measured by the chi-square quantity. Such minimum chi-square estimators were encountered earlier in sections 3.4.3 and 3.6.2. The resulting minimum chi-square quantity, which has $(2k-4)$ degrees of freedom, can be used as a test statistic for the fit of the model.

5.5.3 *Application to brand choice data*

The PDM was applied to the empirical brand choice processes of fopro, beer and margarine used throughout this study. Because it is necessary to have the same number of purchases for all households, the data for the first 10(20) purchases were used.

This means that for fopro and beer $k = 9$, while for margarine $k = 19$. In Table 4.5 the results for the goodness of fit measures and the estimated parameter values are given. The minimization was performed by the pattern search method discussed in section 3.7.2.

Table 5.4 Results of the PDM *for fopro, beer and margarine, first 10(20) purchases*

Product	*Brand* 1 =	χ^2	*df*	*p-level*	$P(0)$	$P(0, 0)$	α	β	$\alpha/(\alpha+\beta)$
Fopro	F1	4.13	14	>.99	.3750	.3541	.0046	.0084	.3538
	F2	10.99	14	.69	.1533	.1429	.0020	.0154	.1149
	F3	4.37	14	>.99	.2426	.2194	.0051	.0092	.3566
	F4	5.02	14	.99	.2292	.2059	.0031	.0134	.1879
Beer	B1	7.68	14	.91	.3923	.3351	.0155	.0077	.6681
	B2	5.27	14	.98	.1643	.1210	.0000	.0147	.0000
	B3	5.21	14	.98	.1292	.1085	.0001	.0074	.0133
	B4	11.48	14	.65	.3142	.2697	.0056	.0215	.2066
Marg	M1	14.83	34	>.99	.2774	.2545	.0016	.0032	.3333
	M2	22.48	34	.94	.0401	.0356	.0019	.0000	1.0000
	M3	18.16	34	.99	.0366	.0324	.0006	.0101	.0561
	M4	14.73	34	>.99	.6458	.6202	.0022	.0048	.4143

From Table 4.5 it can be inferred that the fit of the PDM is extremely good. The p-level, i.e., the probability of finding by chance a higher χ^2-value than the one given is never below .5 and is often near to one. These results will be discussed further in the next section.

5.5.4 *Sensitivity of the test with respect to the* PDM-*assumptions*

As was observed in the previous section, measured by the value of the χ^2-statistic used, the fit of the PDM is extremely good for fopro, beer and margarine. Montgomery, who applied the PDM to brand choice data for toothpaste, with special reference to the brand Crest, also found values for the χ^2-quantities which were very favorable for the PDM, see MM&M, Table 7.3. The number of households from which the toothpaste purchase data for the different cases considered by Montgomery were acquired range from 480 to 894, which is the same order of magnitude as for our data where the numbers for fopro, beer and margarine are respectively 672, 627 and 847. Montgomery used purchase sequences of individual households of length 7, while we have sequences of length 10 for fopro and beer and of length 20 for margarine. So the total numbers of purchases on which the PDM-results are based here are slightly higher.

Now it is a striking fact that the fit of the PDM is so good, that for the toothpaste data in only 2 out of 8 cases was the p-level associated with the χ^2-value found below .50, while for our data this occurred in none of the 12 cases considered. Such an excellent fit is very unusual for brand choice models and it arouses some suspicion regarding the power of the test. Moreover, it is surprising that a model assuming no purchase feedback gives such a good fit for the brand choice processes of fopro, beer and margarine, which were found to be of an order higher than zero, and for which a definite purchase feedback was established in chapter 4. Hence the question arises to what extent does the estimation and testing procedure used deliver statistics which are sufficient, in a statistical sense, with respect to the PDM-assumptions. In other words: how sensitive is the procedure with respect to departures from the PDM?

It should be realized that the whole estimation and testing procedure is based on equations (5.44) and (5.45). These equations describe the development of $P(t)$ and $P(0, t)$ in time as exponential functions with parameters α and β. The estimation procedure tries to fit the points of $Q(t)$ and $Q(0, t)$ to those functions as well as possible.

Now (5.44) and (5.45) describe very general curves. (5.44) is the solution of the differential equation:

$$P'(t) = \alpha - (\alpha + \beta)\, P(t), \tag{5.47}$$

and (5.45) is the solution of the differential equation:

$$P'(0,t) = \alpha P(0) - (\alpha + \beta)\, P(0,t). \tag{5.48}$$

So the only condition imposed by (5.44) and (5.45) is that the rate of growth of the probabilities $P(t)$ and $P(0,t)$ is a constant minus an amount proportional to the level already reached.

Therefore it is questionable if, with the non-rejection of the PDM, the detailed assumptions of the PDM are also verified. One point to be mentioned here is that the basic equations (5.44) and (5.45) are based on an expression for $E(X(t))$ which is the same for the IES as for the CES. This is demonstrated by the fact that in (5.44) and (5.45) γ is not present. Obviously these equations hold for all values of γ, inclusively $\gamma = 0$, which constitutes the IES case. So the non-rejection of the PDM by the test applied can be taken as a confirmation of the CES as well as of the IES, of which the latter was found to be very unlikely. Here we have an indication that the detailed assumptions of the PDM are not tested in the procedure used.

As observed earlier, the estimation procedure fits curves of functions (5.44) and (5.45) through the points $Q(t)$ and $Q(0,t)$, for $t = 1, 2, 3 \ldots$ The long term equilibrium $\alpha/(\alpha+\beta)$, which can be conceived of as a long term market share, is a kind of extrapolation of the observed values of $Q(t)$. We have here the same experience as Montgomery, who found equilibrium market shares, which in a number of cases could hardly be taken seriously. In Table 5.4, for example, it can be seen that the market share of F3 would rise by more than 10 points, that of B1 by about 27 points, B2 and B3 would go back to zero or almost zero, while M2 would absorb the market. These landslides are very unlikely in practice. This again is an indication that (5.44) and (5.45) are too general to represent the essential features of the brand choice processes.

The conclusion is then that the test procedure applied appears to be insufficient to check the detailed PDM-assumptions in empirical data. So the fact that the PDM was not rejected is no sufficient reason to state that the specific PDM- and especially the CES-assumptions are confirmed by the test results.

5.5.5 *An extension of the* PDM

Jones (1970a, 1970b, 1971) extended the PDM to, what he calls, the Dual-Effects model. In this model, not only the influence of the environment (advertising, etc.) is built in, but also the effect of the purchases made. This is done by making the parameters α and β of the PDM variable, so that they can change during the brand choice process. So it can be imagined that after a brand 1 purchase α is increased, while after a brand 0 purchase β is increased. Jones gives several specifications for the way in which these changes of α and β occur. In one of these specifications the change in α and β is brought about in a way analogous to the linear transformation of p in the LLM.

The dual-effects model is theoretically more general than the PDM and the LLM, because the former incorporates only the influence of the environment, while the latter describes only the effects due to purchase feedback. However, in an application of all 3 models to the same toothpaste data used by Montgomery (mentioned in section 5.5.4), Jones found that the dual-effects model, tried with several specifications, gave a worse fit for the data than the PDM and LLM separately. He claims, however, that the dual effects model gives important insights into the actual mechanisms at work in the market place.

5.6 CONCLUDING REMARKS ABOUT LEARNING MODELS

This chapter considered several learning models from the viewpoint of applying them to brand choice processes.

We saw that the LLM treated in chapters 3 and 4 of this book is only a special case of linear operator models and that other models of this type – generalisations of the LLM – can be applied to brand choice processes. Moreover, some further properties of the LLM, with special attention to the asymptotic behavior of p, were treated.

The non-linear learning models discussed, viz., Luce's Beta model and the Urn model, because of their property of commutativity of operators, do not seem to offer great possibilities for brand choice processes. The stimulus sampling model, of which a great variety in types exist, seems to be worth further attention when brand choice models are being built. It was seen that the LLM can be conceived of as a very specific version of a stimulus sampling model, but that the possibilities of this stimulus sampling model are much wider. Another special type of this model, the Probability Diffusion Model, was found not to disagree with the empirical data, but the question arose if the test procedure used was sensitive enough with respect to the PDM-assumptions.

6. An alternative way of studying the brand choice process: the poolsize approach

6.1 INTRODUCTION

In the chapters 3 to 5 of this book we treated rather precisely defined mathematical models which can be used to describe brand choice processes and applied those models to observed processes for fopro, beer and margarine. When one tries to describe a process by a mathematical model, one must ordinarily content oneself with a model that gives a simplified picture of reality. Now the present state of affairs with respect to brand choice models is such that it is almost always necessary to condense the brand choice process to a 0–1 process. Here a one means a purchase of a specific brand, while a zero stands for all other brands. When it is found that such a simplified brand choice process can be described by a mathematical model – for example the LLM – information is provided about the brand choice behavior in question, but it should be kept in mind that only a summarized picture of the underlying brand choice behavior is then considered.

In this chapter an attempt will be made to supplement the findings with respect to the brand choice processes of fopro, beer and margarine by looking at these processes without limiting the number of brands to just two. This is done by studying a variable, called *poolsize*, in the brand choice processes of individual consumers. This variable, which bears a relationship to the number of different brands from which a consumer makes his choice, will be defined in the next section. The connections with the theory of buyer behavior of Howard & Sheth will be discussed in section 6.3.

By means of the poolsize variable, we will try to get more insight into the brand choice behavior of individual consumers. An attempt will be made to find answers to such questions as: how many brands are involved in a brand choice decision? do consumers have a constant or a varying level of brand switching activity? do consumers switch rather straightforwardly from one brand to another or do they exhibit search behavior, etc.? These types of questions are treated in the subsequent sections of this chapter, the last of which contains some conclusions.

Throughout this chapter the empirical brand choice processes studied are the observed processes for fopro, beer and margarine discussed in chapter 2. For each household all purchases are included. For margarine all households are present in the analyses, mono-loyal as well as multi-loyal.

6.2 DEFINITION OF POOLSIZE AND SOME ILLUSTRATIONS OF THE CONCEPT

A variable was looked for with which individual brand choice processes could be characterized without limiting the number of different brands involved. By inspecting a number of observed brand choice processes it was seen that consumers seem to alternate buying periods, during which only one or a small number of different brands are bought, with periods in which transitions are made between a greater number of different brands. This led to the definition of the variable 'poolsize', which is as follows: *the poolsize of a consumer is the number of different brands bought during his last 10 purchases.*

So at purchase moment t the poolsize of a consumer is the number of different brands bought at purchase occasions $(t-9)$, $(t-8)$, ..., $(t-1)$ and t. The word pool is used because the set of different brands bought over the last ten purchases can be conceived of as a pool from which choices were made during the last 10 purchases. By studying this variable 'poolsize' more insight into the nature of the brand choice process can be obtained.

When S_t denotes the poolsize at purchase moment t, the sequence $S_t, S_{t+1}, S_{t+2}, \ldots$ for an individual consumer constitutes a stochastic process, with poolsize being the state variable. S_t is not defined before the tenth purchase of a consumer, i.e. for $t<10$. Of course S_t, S_{t+1}, ..., etc. are not mutually independent. S_t is the number of different brands in the $(t-9)$-th to the t-th purchase, S_{t+1} is the number of different brands in the $(t-8)$-th to the $(t+1)$-th purchase. So S_{t+1} can only take the values S_t-1, S_t and S_t+1.

The number of 10 purchases used in the definition of poolsize, which implies that the sequence of the 10 most recent purchases are considered, is somewhat arbitrary; we could also have taken a sequence length of 5 or 15 for example. It was tried to fix this sequence length such that all those previous purchases were included which markedly influence current brand choice. In this respect it can be mentioned that in chapter 4 the parameter c of the Linear Learning Model, which is the best fitting model for the fopro, beer and margarine brand choice processes, was found to be about .7 on average (Table 4.23). In section 3.5.1 it was seen that in the LLM the contribution of former purchases to the probability p is proportional to the powers of c. Thus the influence of the brand chosen 10 purchases ago is proportional to c^{10}. Now $(.7)^{10} \approx .03$, which means that the influence of the 10th purchase on the empirical processes studied is very small. Therefore in our opinion, with the 10 previous purchases that part of the purchase history of a consumer relevant to his current brand choice is generally grasped.

In section 7 of this chapter it will be seen that the influence of the exact sequence length chosen on the results obtained is not great.

To illustrate the concept poolsize, the realisation of the stochastic process $\{S_t\}$, i.e., the development of the poolsize during the purchase process, has been drawn in Figure 6.1 for 3 empirical processes of individual households for fopro, margarine and beer respectively. Here the time scale has been shifted such that $t = 0$ in the figure corresponds with the tenth purchase of the consumer, i.e., the first purchase for which poolsize is defined.

From the poolsize curves pictured it is clear that the households show periods of little brand switching activity alternated with periods in which quite a number of different brands are bought in a short period. We will discuss the fopro process at some length. The household starts with poolsize 2, but after a short time the poolsize diminishes to one, which means that during 10 purchases only one brand was bought. This remains so during 28 purchases; after that a brand other than the one bought until then is purchased. This implies that during the next 9 purchases the pool will contain at least 2 brands. It is conceivable that a consumer in one time completely switches to the new brand. In that case the situation of the poolsize equalling 2 lasts exactly for 9 purchases. We observe that after 10 purchases the poolsize returns to one, which indicates that no further purchases of the new brand were made, while no other brands were tried in the meantime. After the 'interlude', just discussed, again a period of absence of brand-switching begins, which lasts for 23 purchases. Then again an incidental purchase of a different brand is made, which is followed by a short period of rest. After that, a long period of intensive brand-switching starts, during which the poolsize even increases to 4. Finally the household seems to have made its choice and continues for a long time buying the same brand. For the poolsize curves of the beer and margarine households similar comments can be made. So the development of the poolsize curve of a consumer gives information about his brand choice behavior. It can be said that when S_t is high the consumer shows

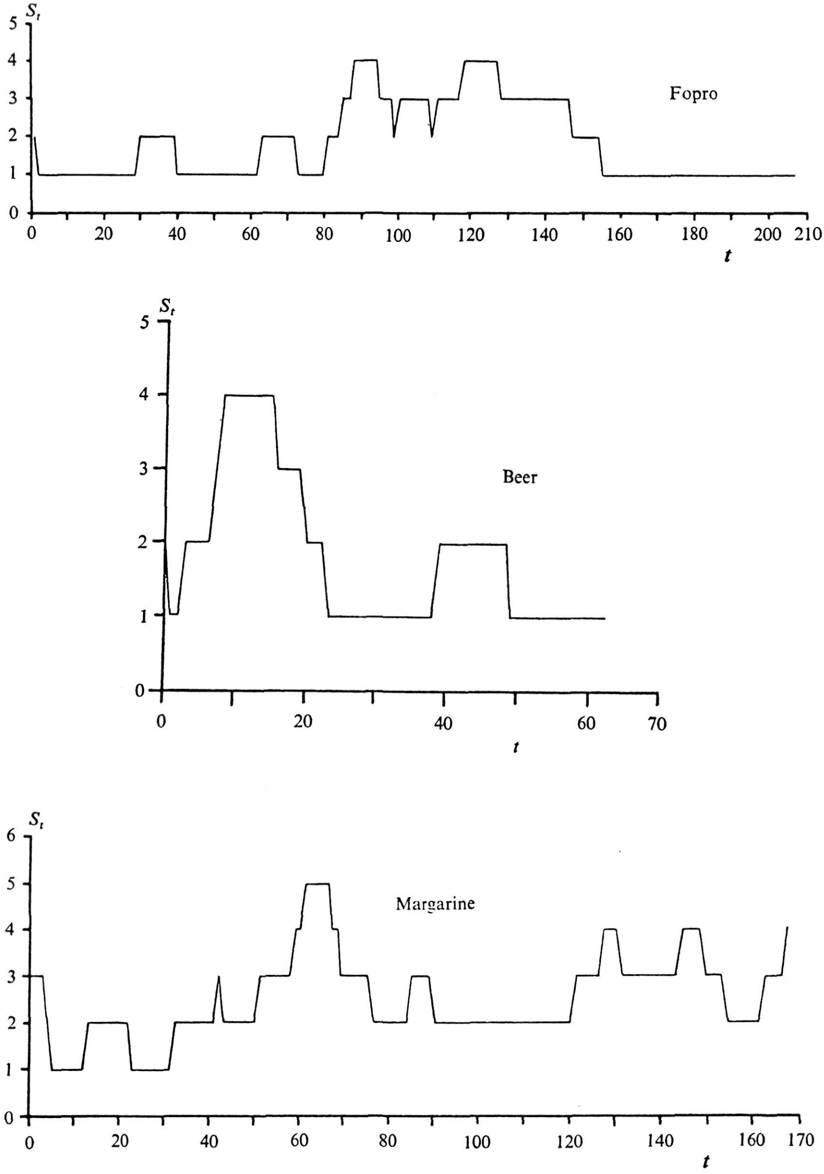

Fig. 6.1 The development of poolsize (S_t) during the purchase process for 3 households

intensive search behavior, while a poolsize of one indicates that the consumer has a favorite brand which he buys continuously. The 3 households, for which the poolsize curves are drawn, show an evident variation of search intensity during the purchase process. The following sections examine, inter alia, how far this is a general picture for all households.

6.3 CONNECTIONS WITH HOWARD AND SHETH'S THEORY OF BUYER BEHAVIOR

The theory of buyer behavior, developed by Howard and Sheth (1969), contains some elements which are relevant to the approach followed in this chapter. They will be briefly discussed in this section.

An important concept in Howard & Sheth's theory is the notion of the 'evoked set' of brands. Generally, a consumer will not consider all available brands of a certain product as potential candidates for his choice. In the first place, a consumer will not be aware of all these different brands; also, not every brand the consumer is aware of will be acceptable to him for his next purchase. Now the brands that form alternatives to a buyer's choice decisions, which generally constitute a smaller number than all brands in the market, are called by Howard & Sheth the 'evoked set' of that buyer (ibid, p. 26). The conditions for a brand to belong to the evoked set of a consumer are that: 1) the consumer must be aware of the brand, and 2) the consumer must consider the brand as a choice possibility.

The existence of an evoked set is very important for a consumer, especially in markets with many brands of the same product. Then the evoked set means that the consumer considerably simplifies his choice situation by choosing from only a few brands instead of from all the brands available. Howard and Sheth mention some empirical evidence for the existence of the evoked set, as reported in an unpublished work by Campbell (1969) in which it was found that for toothpaste and detergents no buyer had an evoked set larger than seven, with the mean numbers being 3.1 and 5 respectively.

Of course, for the same product different consumers may have different magnitudes for their 'evoked sets'. Also, the number of brands in the evoked set of a consumer can change during the purchase process.

The concept of a pool, i.e., the set of different brands during the last 10 purchases of a consumer, can be brought into connection with this evoked set. When it is assumed that a consumer remembers his last 10 purchases, he is aware of the brands in the pool. Furthermore these brands have been shown to be acceptable choice candidates, because they have actually been chosen. So the brands in the pool belong to the evoked set of the consumer. On the other hand, it may be possible that there are still other brands, which are also considered as choice possibilities by the consumer, but which happened not to be purchased during the last 10 purchases. These brands also belong to the evoked set, but are not present in the pool. So it can be said that the pool represents a part of the evoked set, it can be called a truncated evoked set. The brands present in the evoked set but not in the pool, will generally not be very salient brands, i.e. brands with high purchase probabilities. If they were, it is very unlikely that they would not have been bought during the last 10 purchases. So the pool coincides with the most salient part of the evoked set and the poolsize is the number of brands in this subset of the evoked set. Therefore, the development of the poolsize of a consumer during the purchase process gives information about the development of the evoked set. For the rest it is very difficult to exactly specify for a certain consumer at a certain moment which brands belong to his evoked set and which do not. To make the concept evoked set operational in practical situations, it will always be necessary to apply a kind of approximation. The poolsize approach constitutes such an approximation.

Another element of the theory of Howard and Sheth, relevant to this study of poolsize, is the way in which Howard and Sheth look at the buying process and the changes that occur in it over time. In section 2.1 of their book a summary of this view is given. Here a consumer's decision-making is divided into three stages:

1. extensive problem solving

This refers to the early stages of repetitive decision making, in which the buyer has no strong predispositions towards any brand.

2. limited problem solving

This is the next stage, in which the consumer has moderately strong predispositions towards a number of brands, but does not have a particular preference for any one brand.

3. routinized response behavior

Here the buyer has only one or two brands in mind as the most probable choice alternatives. This process of reducing the complexity of the buying situation is called 'the psychology of simplification'. The more the buying situation is simplified, the less the consumer tends towards active search behavior.

According to Howard and Seth, there is a surprising phenomenon which occurs in many instances of frequently purchased products:

'The buyer, after attaining routinization of his decision process, may find himself in too simple a situation. He is likely to feel monotony and boredom associated with such repetitive decision making. He therefore feels a need to complicate his buying situation by considering new brands and this problem can be called "the psychology of complication". The new situation causes him to search for identity with a new brand and so he begins again to simplify.'

According to this view, the brand choice process will show cycles: periods of intensive search will be alternated with periods of continuous buying of the same brand or alternation between two brands. Now this is exactly the phenomenon, observation of which was mentioned in section 6.2 and which was demonstrated by the poolsize curves for the individual households, drawn in Figure 6.1. So at first glance this view of Howard and Sheth on the buying process seems to be in agreement with our empirical data. In the next sections this will be examined further.

6.4 SHAPE OF THE POOLSIZE CURVES

6.4.1 *The variables* MEANPLZ *and* RANGPLZ

In this section we study the form of the poolsize curves for the fopro, beer and margarine brand choice processes. Of course, it is impracticable to analyse the curves for all individual households, as was done for the fopro household in section 6.2. It is necessary to use quantities which are derived from the observed poolsize curves and which characterize these poolsize curves.

Two aspects of a poolsize curve for an individual household are especially interesting, viz., the height and the measure in which the poolsize curve goes up and down, i.e. the variation in poolsize. Therefore we define the following quantities:

1. MEANPLZ = the average value of the poolsize over the whole purchase history of an individual consumer.
2. RANGPLZ = the difference between the maximum and minimum values of the poolsize, taken over the whole purchase history of an individual consumer. Note that the poolsize curve of each individual household produces one observation for MEANPLZ and one observation for RANGPLZ (S_t, the poolsize, is a continuous measure of the level of brand switching activity of a household). With MEANPLZ the general level of brand switching activity, which can be taken as an indication of the search intensity, is measured. A consumer with a high value for MEANPLZ shows a high level of brand switching activity.

Of course S_t, and therefore MEANPLZ, gives a somewhat simplified picture of brand switching. For example, if there are two brands, *A* and *B*, the sequences *ABABBAABAB* and *AAAAABBBBB* will both produce a poolsize of 2, although the switching patterns are different. When one tries to grasp the level of brand switching in one figure, such simplifications cannot be avoided. RANGPLZ, the range over which the poolsize varied, is a measure of the variation in the level of brand-switching activity. When periods of constant buying of the same brand are alternated with periods in which a large number of different brands are tried, the value of RANGPLZ will be high. On the other hand, both when a consumer always chooses the same brand and when a consumer constantly keeps the same (possibly high) level of brand switching activity the value of RANGPLZ will be zero. A consumer showing the cyclical behavior with respect to his search activity postulated by the psychology of simplification and complication of the previous section, will have a value of RANGPLZ which is at least greater than zero and generally greater than one.

6.4.2 *Average values over all households*

We have computed the values of MEANPLZ and RANGPLZ for the brand choice processes of the individual fopro-, beer- and margarine households. The average values over all households are given in Table 6.1. The numbers in parentheses are the corresponding estimated standard deviations of the averages of the MEANPLZ and RANGPLZ.

Table 6.1 Average values with standard deviations for MEANPLZ *and* RANGPLZ *over all households*

Product	MEANPLZ	RANGPLZ	*Mean number of brands per household*	*Number of households*
Fopro	1.553	1.268	2.88	672
	(.026)	(.046)		
Beer	1.745	1.022	2.57	627
	(.031)	(.040)		
Marg	1.778	1.756	4.26	1059
	(.031)	(.042)		

For completeness in Table 6.1 the mean number of brands per household (from Table 2.2) and the number of households for each product are also given.

As can be seen from the figures in the first column of Table 6.1, the general level of brand-switching activity is lowest for fopro. With a Student *t*-test it was verified that MEANPLZ for fopro is significantly lower than for beer and margarine. It may be somewhat surprising that beer, with a number of different brands per household not greater than for fopro, has a higher value for the MEANPLZ. Here it should be remembered (Table 2.1) that the purchase histories for beer are much shorter than for fopro. To attain the same number of different brands in a shorter purchase sequence, a household must show a heavier switching activity during that sequence.

The MEANPLZ for margarine is about equal to the MEANPLZ for beer and only slightly greater than that for fopro, although the number of brands per household is much bigger for margarine. Obviously, the different brands used by a margarine household in the 2 years observed were not bought interchangeably during this whole period; a household buys certain brands in certain periods and other brands in other periods. So it can be said that not all the brands used were simultaneously in the pool, but that the pool contained different brands at different times. If all the different brands had been constantly in the pool during the whole 2-year period, then the MEANPLZ would have been equal to the number of brands for an individual household and this would also have been the case for the averages over all households.

Margarine has significantly higher values for the RANGPLZ than fopro and beer. This means that for this product the variation in the level of brand switching activity is the most evident. So at some moments in a purchase history poolsize was low and at other moments it was high, which again indicates that not all brands bought were present in the pool during the whole purchase history. The RANGPLZ for beer is significantly lower than for fopro. So compared with the other products, the level of brand switching activity for beer does not vary much during the brand choice process. It was seen that the general level of brand switching activity, as measured by the MEANPLZ, is rather high for beer. This different behavior of the poolsize curves for beer may have to do with the generally lower frequency with which beer is bought, which means longer inter-purchase times. Perhaps in that case the sequence: 'extensive problem solving – limited problem solving – routinized response behavior', as postulated by Howard and Sheth, is not passed through so neatly as for more frequently bought products. An indication of this lies in the fact that longer beer-households (i.e. with more purchases during the 2 years) have lower values for the MEANPLZ ($r = -.2572$) and higher values for the RANGPLZ ($r = .1997$). The different results for beer may also be because of the fact that beer purchases in a household are often made by various members of the family, see section 8.3.3. Of course there is further the aspect that shorter purchase histories simply have less chance to show cyclical behavior, because there is less opportunity to complete a whole cycle from poolsize low to poolsize high or vice versa.

6.4.3 *Average values per brand-class*

After having looked at the averages of the MEANPLZ and RANGPLZ over all households for the 3 products, it is instructive to consider the averages per brand-class. By brand-class is meant the set of households with a certain number of different brands in the 2 years observed. Thus there is the set of households which bought only one brand during the 2 years, this is called brand-class 1; brand-class 2 is the set of households buying 2 different brands during the 2 years, etc. So here we examine the differences in poolsize curves between households buying different numbers of brands. With the differences between margarine and the other products observed in the previous section, we saw one possible effect of the number of brands, because margarine-households generally have more different brands.

In Table 6.2 the average values of the MEANPLZ and RANGPLZ per brand-class are given for fopro, beer and margarine. For each household the highest value of the poolsize during the 2 years was also computed; this quantity is called the MAXPLZ. The averages of the MAXPLZ over all

Table 6.2 Average values for MAXPLZ, MEANPLZ *and* RANGPLZ *per brand-class*

	MAXPLZ			MEANPLZ			RANGPLZ		
Brandclass	*Fopro*	*Beer*	*Marg*	*Fopro*	*Beer*	*Marg*	*Fopro*	*Beer*	*Marg*
1	1.00	1.00	1.00	1.00	1.00	1.00	.00	.00	.00
2	2.00	2.00	2.00	1.28	1.50	1.31	.96	.81	.93
3	2.62	2.81	2.57	1.66	2.04	1.59	1.46	1.36	1.39
4	3.19	3.49	3.07	1.90	2.44	1.72	2.03	1.88	1.91
5	3.74	4.05	3.45	2.19	2.85	1.98	2.43	2.14	2.22
6	4.11		3.92	2.30		2.25	2.89		2.64
7			4.42			2.30			3.09
8			4.65			2.56			3.22
9			5.37			3.10			3.87
10			5.35			2.98			3.77
others	5.24	4.93	6.06	2.91	3.03	3.35	3.84	3.19	4.57

households per brand-class are also given in Table 6.2. In the higher brand-classes the number of households is not always high enough to compute a reliable average. Here only for brand-classes with at least 20 households were the average values for fopro, beer and margarine computed. The households in the other brand-classes are taken together in the group which is indicated as 'Others'.

Let us first consider the MAXPLZ. We see that the rate of increase of the MAXPLZ, associated with the increase in the number of brands, levels-off for higher brand-classes. This is also shown by Figure 6.2, where the average MAXPLZ has been plotted against the number of brands. The graph suggests that there is a maximum level of about 7 for MAXPLZ. This is the most evident for margarine, for which product the curve could be drawn furthest, because more households had many brands. So the poolsize, even for households which bought 10 or more brands during the 2 years observed, generally does not become higher than 7, so there appears to be a tendency to limit the poolsize, i.e. the number of brands which can be considered as candidates from which a choice is made. This may have to do with general capacity limits with respect to the human capability to discriminate between alternatives. Campbell (1969), as reported by Howard and Sheth (1969, p. 98) also detected such a limit.

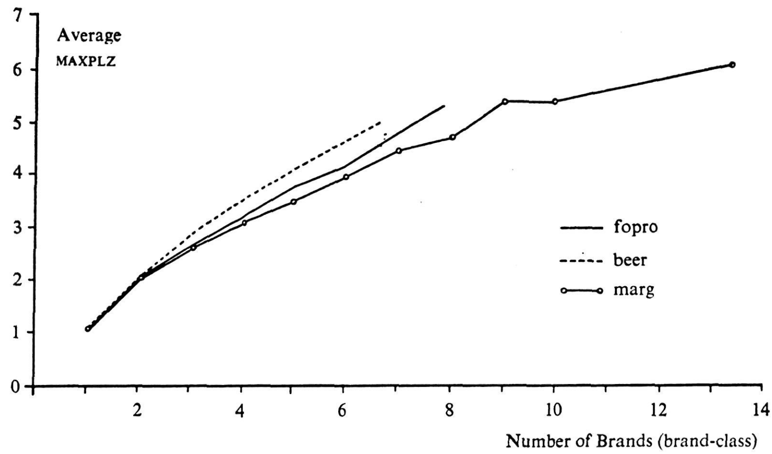

Fig. 6.2 Average values of MAXPLZ *(maximum poolsize) for households with different numbers of brands during the 2 years observed*

He found that the maximum number of brands in the evoked set for toothpaste and detergents was 7. In fact, the number 7 was found as the highest value ever attained for the poolsize over all households for fopro and beer while for margarine there was one household with MAXPLZ = 9. In this connection the classic article of Miller (1956) with the title: 'The magical number seven, plus or minus two: some limits on our capacity for processing information', may be mentioned. In this article Miller discusses a number of experiments in which persons were tested on their capacity to discriminate between alternatives. Among others, he mentions experiments with different tones of music and different salt-concentrations in salt solutions. He concludes that for this kind of judgment people can only discriminate among as many of about 7 alternatives. It is striking that we see this phenomenon again in relation to the number of brands which are simultaneously considered by a consumer as potential choice possibilities.

Considering the MEANPLZ, we see that the tendency to have not too many alternatives at a time is also demonstrated by the values of the

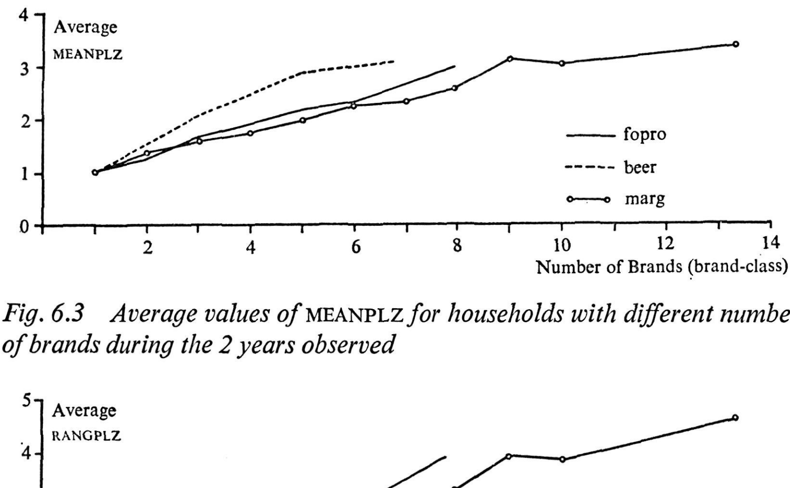

Fig. 6.3 Average values of MEANPLZ *for households with different numbers of brands during the 2 years observed*

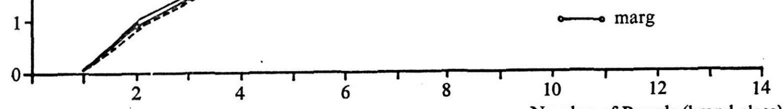

Fig. 6.4 Average values of RANGPLZ *for households with different numbers of brands during the 2 years observed*

average MEANPLZ for the different brand-classes. This is shown by the figures from Table 6.2, which are plotted in Figure 6.3.

We see that the MEANPLZ does not rise very fast and shows a tendency not to move higher than about 3, even for households which bought more than 10 brands during the two years observed. This points to the phenomenon that consumers using many different brands, did not buy them all interchangeably, which kept the number of alternatives at a time limited.

With respect to the RANGPLZ, it can be seen from the figures in Table 6.2, which are plotted in Figure 6.4, that the variable RANGPLZ increases roughly linearly with the number of brands. Households with a large number of brands also have a high value for the RANGPLZ, which means that these households in particular show variation in the poolsize.

6.4.4 *Regression of the* MEANPLZ, *respectively the* RANGPLZ, *on the number of brands*

The relationship between the number of brands and the MEANPLZ, respectively RANGPLZ, can be considered in another way. We define:

X_i = number of different brands of household i bought during the 2 years.

Y_i = the MEANPLZ of household i, respectively RANGPLZ of household i.

Every household produces an observation of Y and X.
Now we can specify the linear relation:

$$Y_i = \alpha + \beta X_i + U_i$$

where U_i is the disturbance term, with expected value zero, constant

Table 6.3 Estimated slopes from the linear regression of the MEANPLZ *and* RANGPLZ *on the number of brands*

Product	*Dependent variable*		*N*
	MEANPLZ	RANGPLZ	
Fopro	.2811	.5635	672
	(.0090)	(.0110)	
Beer	.4126	.5569	627
	(.0139)	(.0162)	
Marg	.2022	.3796	1059
	(.0042)	(.0066)	

variance and with subsequent U-values independent. The coefficients α and β can be estimated by the method of least squares.

Of course the linear relationship postulated here is an approximation; from Figure 6.2 and Figure 6.3 it can be inferred that the assumed linearity does not differ too much from the truth.

Here we are particularly interested in the slope β. We give the values estimated for the slope with the corresponding standard deviations for the different cases in Table 6.3.

In all cases the slopes are significantly greater than zero. The standard deviations are very small, relative to the values of the coefficients, which is due to the large sample sizes.

The important finding here is that the slope of the RANGPLZ is always greater than the corresponding slope of the MEANPLZ. For example, in the case of fopro an increase of one in the number of brands is associated with an average increase in the MEANPLZ of .2811, while the corresponding increase in the RANGPLZ is .5635, which is twice as much. Also, for beer and margarine an increase in the number of brands has a greater impact on the RANGPLZ than on the MEANPLZ. So households with many brands differ from households with few brands more with respect to the RANGPLZ than to the MEANPLZ. Since the MEANPLZ is a measure of the general level of brand switching activity and RANGPLZ for the variation in that level, it can be said that many-brand-households, as compared to few-brand-households, are especially different on the point of demonstrated variation in their level of brand-switching activity. This is a more striking point of difference than the fact that the general level of brand-switching activity of the many-brand-households is higher. So the additional brands of the many-brand-households are used more to intensify the search during search periods – resulting in higher poolsizes and corresponding higher values for the RANGPLZ – than to increase the general level of brand-switching activity.

6.4.5 *Relative frequency of* RANGPLZ-*values*

Some more information about the importance of cyclical behavior in the poolsize is provided by Table 6.4. In that table the relative frequency of occurrence of the different values for the RANGPLZ is given. Of course, households which always bought the same brand during the two years observed have a poolsize value of 1 during their whole purchase history, which implies that for these households RANGPLZ always $= 0$; these households did not switch at all.

Table 6.4 Relative frequency of RANGPLZ*-values for households with more than one brand*

RANGPLZ	*Fopro*	*Beer*	*Marg*
0	.0269	.1204	.0190
1	.4669	.5055	.3814
2	.3223	.2582	.2975
3	.1157	.0875	.1745
4	.0475	.0241	.0738
≥5	.0207	.0044	.0537
Number of Households with ≥2 brands	484	457	894

These households were left out in calculating the relative frequencies, as given in Table 6.4. Using the Howard and Sheth terminology of section 6.3 it can be said that during the purchases observed these households were continuously in the phase of Routinized Response Behavior.

From Table 6.4, we see that for fopro and margarine the RANGPLZ is 2 or greater for more than 50% of the households. For 18, respectively 30% of the households the RANGPLZ is even 3 or more for these 2 products. For beer, the RANGPLZ is generally lower, being 2 or greater in about 37% of the households. This different behavior of the poolsize for beer has already been discussed in connection with Table 6.1.

The general conclusion from Table 6.4 is that a considerable number of households making brand switches show a substantial variation in their level of brand switching activity. Because the presence of cycles in the purchase history, as assumed by Howard and Sheth, implies variation in the level of brand-switching activity, the phenomenon observed is in agreement with that cyclical behavior of search activity hypothesized by Howard and Sheth. Because the observed waves have long periods, as is shown by Figure 6.1, it was not possible to observe for each household a reasonable number of them over the two years. So we can infer from the RANGPLZ-values that brand-switching activity varies, which agrees with Howard and Sheth's theory of cyclical behavior, but the observed period was not long enough to establish if the behavior of individual households was exactly cyclical, i.e. with ups and downs of the same size and with intervals of constant length.

6.4.6 *General conclusions regarding the shape of the poolsize curves*

Section 6.4 tried to grasp the general characteristics of the poolsize curves for fopro, beer and margarine. It should be realized, that every household has its own poolsize curve and that it is difficult to find 2 households for which the poolsize curves are exactly identical. Yet it is possible to draw some tentative conclusions, which are generalisations because they hold for the 'average household' and not necessarily for each individual household. These conclusions, which are the most obvious for fopro and margarine and somewhat less obvious for beer, are interesting, because they tell us something about the way in which a consumer makes his choices during the brand choice process.

The conclusions are:

- The poolsize, which has a mean value over the whole purchase history of 1.5 to 1.8 does not become higher than about 7 during the brand choice process. This means that the number of alternative brands, simultaneously considered as potential choice candidates by a consumer, is limited.
- The poolsize is not constant during the brand choice process but shows periods of low values, and periods of high values. This is in agreement with the theory of the cyclical behavior of a consumer's search activity.

These conclusions support the theory of Howard and Sheth with respect to the limited number of brands in the evoked set of a consumer and the cyclical behavior of the search activity. In the following section we examine the search behavior more closely.

6.5 THE SEARCH BEHAVIOR OBSERVED

In this section we study the importance of search behavior during the brand choice process. We say that search behavior occurs, when not a straight-forward switch from one favorite brand to another is made or a tried brand is rejected immediately, but when during a certain period a number of different brands are tried – sometimes taking the form of an alternation between two brands – before a certain brand is definitively continued with. We also pay attention to the possible causes or inducements of the search behavior.

6.5.1 *Importance of search behavior*

The first approach to this subject is to study the development of poolsize after an increase from 1 to 2, i.e., after the purchase of a new brand, while during the 10 preceding purchases always the same brand (other than the new brand) was bought. It is interesting to see if after such an initial increase the poolsize increases further, which means that still other brands are tried before a decision is made, or if after some time the poolsize decreases to one without having been higher than 2. This was examined for the fopro, beer and margarine brand choice processes by observing what happened after an increase of the poolsize from 1 to 2. From such increases on the poolsize curves were followed in the phase of increase until the first decrease. This was done as often as such an initial increase of the poolsize occurred in the brand choice processes observed. Sometimes a consumer's purchase history did not contain such a point at all; in that case the household was discarded for this analysis. Other households delivered more than one occasion for observation, their poolsize went more often than once from 1 to 2. The results are given in Table 6.5. Of course, a subsequent increase or decrease was not observed after every increase of the poolsize, there is also the possibility that the poolsize remained at the level reached during the further purchase history observed. For example, of the 682 times an increase of the poolsize from 1 to 2 for fopro was observed, in 591 cases the poolsize was seen to afterwards go to 1 or 3, in 91 cases the poolsize remained at the level 2 during the remainder of the observed purchase history.

Table 6.5 Development of poolsize after an increase from one to two

	Fopro	*Beer*	*Marg*
a. Number of times an increase from 1 to 2 was observed	682	322	672
b. Number of times a change to 1 or 3 was observed after an increase as meant in a.	591	227	523
c. Number of times the poolsize increased to 3			
(1) absolute	218	103	291
(2) as % of b.	36.9	45.4	55.6
d. Number of times a change to 2 or 4 was observed after an increase as meant in c.	188	82	243
e. Number of times the poolsize increased to 4			
(1) absolute	50	26	89
(2) as % of d.	26.6	31.7	36.6
f. Number of times a change to 3 or 5 was observed after an increase as meant in e.	47	22	74
g. Number of times the poolsize increased to 5			
(1) absolute	18	6	28
(2) as % of f.	38.3	27.3	37.8

Considering only the cases in which subsequent changes in the poolsize occurred, we see that for fopro in 36.9% of the cases, once the poolsize increased to 2 it rose further to 3, and in 26.6% of these latter cases a further increase to 4 was observed, while in 38.3% of the times poolsize 4 was reached an increase to 5 occurred. For beer the corresponding percentages are 45.4, 31.7 and 27.3, while for margarine they are 55.6, 36.6 and 37.8.

This gives an indication of the importance and magnitude of search behavior. We see that in 1/3 to 1/2 of all cases in which an initial increase of the poolsize occurs, the search is not limited to 2 brands, but more brands are tried, which in 4–8% of all cases goes as far as at least 5 brands.

When the poolsize, after having been at level 2 for a number of purchases, goes ultimately back to one without having been at level 3, this does not imply that the consumer in question did not exhibit search behavior. Whether this is the case or not can be derived from the number of purchase occasions during which the poolsize was at level 2. Both when the new brand bought is accepted at once and all subsequent purchases are purchases of this brand and when the new brand is rejected at once, implying that the consumer immediately continues with his old brand, the poolsize remains at level 2 not longer than for 10 purchases. When the decision is not immediately taken, but the consumer hesitates between the two brands, (the old and the new one), which means that both brands are bought alternately for some time, poolsize remains at level 2 for more than 10 purchases. In this case, it can be said that the consumer practises search behavior with respect to 2 brands. Theoretically it is even possible that during such a period of constant poolsize 2 more than 2 brands were bought, namely when a new brand enters the pool exactly at the moment that an old one leaves it.

In the case of fopro $591-218=373$ times poolsize went back from 2 to 1. It was observed that in 126 of these cases poolsize was at the level 2 during more than 10 purchases. From Table 6.5 we know already that of the 591 cases considered search behavior was shown 218 times because a third brand was tried. Now we can add the 126 times that search behavior occurred because of alternation between two brands and conclude that the percentage of cases in which search behavior was observed after an initial increase of poolsize from 1 to 2 for fopro is $(218+126)/591 \times 100\% = 58\%$. In the same way it was found that for beer in 50 cases of an initial poolsize increase search behavior between two brands occurred. This implies that for beer in total, in 67% of the cases search behavior was shown after an initial increase in poolsize. For margarine, in all cases that the poolsize went back from 2 to 1 the poolsize had been at level 2 for more than 10 purchases. So for margarine there was always search behavior after an initial increase of the poolsize.

So we conclude that mostly consumers do not straight-forwardly switch from one brand to another or immediately reject a tried brand, but that such events are generally accompanied by temporary search behavior. For fopro this occurred in 58%, for beer in 67% and for margarine in 100% of the observed cases.

The results with respect to the search behavior for fopro, beer and margarine can be compared with the results found by Lawrence (1969). Lawrence examined what he calls the 'switch patterns' of 953 American households buying toothpaste. He studied the brand choice behavior during 4 purchases after a switch, where a switch is defined as the purchase of a new brand, after a consumer had bought the same brand (other than the new one) during at least 5 consecutive purchases. 1607 of these switch patterns were studied. Lawrence found that in 35.1% of the cases the consumer returned to the old brand immediately after the switch and remained with that brand. In 12.5% there was a complete conversion to the new brand. So in 47.6% of the cases a direct decision was made. In 22.1% of the switch patterns there was an alternation between the old and the new brand, called 'vacillation' by Lawrence, while in 30.3% of the cases there at least one additional brand was tried. This means that search behavior was observed in 52.4% of cases.

The figure 30.3, found by Lawrence, should be compared with the figures in row c(2) of Table 6.5, while the figure 52.4 for the search behavior in total, should be compared with the figures of 58, 67 and 100 for fopro, beer and margarine respectively. So it can be concluded that although Lawrence's figures on the importance of search behavior are of the same order of magnitude – which is in itself a striking fact, because of the different products and different consumer populations – they tend to be somewhat smaller. This may be due to the different method for establishing search behavior used by Lawrence. In our method more purchases after a switch were considered, which implies that we also observed search behavior manifested after the 4th purchase after a switch. Notwithstanding this the similarity of Lawrence's results and ours may lead to the hypothesis that there is a general kind of search behavior after a switch, such that rather constant percentages of consumers respectively return to the old brand immediately, make a complete conversion to a new brand or show search behavior.

A second approach for examining the importance of search behavior is to consider those households which obviously made a complete transition from one favorite brand to another during their purchase history. For this purpose those households were selected for which the first 10 poolsize values were 1 and the value of the poolsize during the last 10 purchases of their buying process was also constantly 1, while the brand bought at the start of their purchase history was different from the brand bought at the last purchases of their purchase history. It can be said that, during their purchase history, these households went from one stable period to another, with in the meantime a change in the favorite brand. Now it is interesting to know what happened between both stable periods. Was the switch made straight-forwardly or was there manifest search behavior before the new brand was accepted? To answer this question we looked at the behavior of the poolsize between the 2 stable periods; in fact, we examined what was the maximum value of the poolsize in that interval. The results are given in Table 6.6.

Table 6.6 Search behavior of households changing their favorite brand

	Fopro		*Beer*		*Marg*	
	abs	*perc*	*abs*	*perc*	*abs*	*perc*
Number of households considered	35	100	21	100	74	100
Number of households reaching poolsize ≥ 4	5	14.3	1	4.8	6	8.1
Number of households with maximum poolsize $= 3$	13	37.1	10	47.6	23	31.1
Number of households with maximum poolsize $= 2$, but during ≥ 10 purchases	10	28.6	4	19.1	29	39.2
Number of households showing search behavior	28	80.0	15	71.4	58	78.4

We see that, taken over all 3 products, 40 to 50% of the households reached a poolsize of 3 or 4 in the time between the 2 stable periods, which means that they also tried a third and sometimes a fourth or a fifth brand before the decision fell on the new favorite brand. 20–40% of the households showed search behavior between the old and the new brand at some time, which can be inferred from the fact that their poolsizes have as maximum value 2, a level which was maintained over 10 or more purchases.

To summarize: 70–80% of households showed search behavior between both stable periods, which indicates that consumers generally do not straight-forwardly switch from one brand to another, but that such a switch is usually accompanied by a period of search.

6.5.2 *Possible causes or inducements for search behavior*

An interesting question is: what causes or induces search behavior on the part of a consumer? Is it possible to influence this search behavior with external stimuli, e.g., by means of marketing activities? In this section attention will be paid to this type of question.

First the influence of deals will be considered. Deal purchases are defined here as purchases with a temporary price reduction or an associated free gift. Section 6.5.1. reported how we followed the development of the poolsize of individual consumers after an increase from 1 to 2 to establish the importance of search behavior. During this analysis we classified every purchase, which caused an increase in the poolsize – which means that a brand not present in the pool thus far was chosen – according to whether it was a deal purchase or not. In this way it was determined which percentage of the purchases causing an increase in the poolsize were deal purchases. This percentage can be compared with the figure for the deal purchases in all purchases, as given in Table 2.2. The results, thus obtained, are presented in Table 6.7.

Table 6.7 Relative importance of deal purchases in purchases that increase poolsize

	Fopro	*Beer*	*Marg*
(1) Number of poolsize-increasing purchases considered	1345	695	3971
(2) Percentage of deal purchases in the purchases of (1)	7.2	4.6	17.7
(3) Percentage of deal purchases in all purchases	2.5	3.0	6.4

It appears that deal purchases occur more often at pool-increasing purchases than at purchases in general. The big sample sizes mean that the differences in observed fractions are significant, even for $\alpha = .01$. For fopro and margarine the deal-fraction is almost 3 times bigger for the poolsize-increasing purchases than for purchases in general. So the purchase of a new brand, not thus far in the pool, is relatively often associated with a deal, from which it might be deduced that deal purchases can induce or stimulate search behavior. But the difficulty here is to establish the direction of cause and effect. It is conceivable that in the cases considered above, the consumer already intended buying a new brand and that a brand with a deal-offer was the one first considered for a trial.

Another approach followed to obtain information about possible inducements to search behavior, was to look at the start of periods of intensive search. Purchases made at such starting-points were considered, i.e., purchases made at the start of periods of increasing poolsize, during which periods the poolsize attained at least the value 3 before it went back to 1. For example, for the 3 households for which poolsize curves were plotted in Figure 6.1, such search-initiating purchases occurred at $t = 81$, $t = 3$ and $t = 32$ respectively. So purchases were considered, which resulted in the poolsize going from 1 to 2 at the beginning of search periods.

For these search-initiating purchases it was seen if there were concomitant events associated with the change in brand. Hence, it was noted for each of these purchases if the new brand had a different price compared with the old brand, if it was a deal purchase, if the shop in which the purchase was made was different to the shop at the previous purchase and if the size of the unit bought was different from the unit-size bought at the previous occasion.

A price change was said to occur when the price of the new brand differed more than 5 ct from the price of the old brand. The change in unit-size only refers to fopro and beer, because the margarine purchases were almost 100% made in the same unit-size. For fopro there are 5 major size classes and for beer 3. When a change in unit-size occurred, the price-change was not counted as such.

The relative occurrence of the different types of concomitant events is indicated by the figures in Table 6.8.

Table 6.8 Relative occurrence of different types of concomitant events of search-initiating brand switches

	Fopro		*Beer*		*Marg*	
	abs	*perc*	*abs*	*perc*	*abs*	*perc*
Number of brand-switches observed	76	100	34	100	211	100
With concomitant event:						
Price change	29	36.7	5	14.7	141	66.8
Deal	6	7.6	0	0	18	8.5
Shop change	50	63.3	18	52.9	115	54.5
Unit-size change	33	41.8	22	64.7	—	—
None of these	4	5.1	5	14.7	26	12.3

We see from this table that in most cases the search-initiating brand switches are accompanied by some change in other variables, sometimes by more than one at a time. The most general for all 3 products is a change in shop. We come back to the relationship between shop choice and brand choice in the next chapter. In addition, for fopro and beer a change in unit-size often occurs at the same time as the brand switch, while for margarine and fopro price changes are also important. Deal purchases are not frequently observed for these search-initiating purchases, at least for beer and margarine, this in contrast to the purchases which increase poolsize in general, which are relatively often associated with a deal, as seen above. A tentative explanation of this difference might be that the increases of the poolsize considered here, i.e., those which occur at the start of search periods are more planned by a consumer and therefore less influenced by incidental deal-offers.

The term 'concomitant events' was used intentionally in Table 6.8, because it is again difficult to establish the direction of cause and effect. For example, from the fact that a search-initiating brand switch is often associated with a change in shop, it might be deduced that the visit to another shop is an important inducement to the initiation of a search period. But it might also be stated that a consumer's intention to try another brand often leads to a purchase in a different shop. Similar statements can be made for changes in price and unit-size.

So it appears that a brand switch at the start of a search period is often associated with changes in price, shop and/or unit-size.

6.6 AVERAGE POOLSIZE OVER ALL CONSUMERS

Up until now we have considered the development of the poolsize during the purchase history of individual consumers. It is interesting to know if the poolsize curves of different consumers exhibit coherence in time, i.e., if poolsize curves of different consumers are low and high at the same points in time. If this were true, then the average poolsize curve over all consumers would go up and down, indicating periods of intensive brand-switching alternating with periods of routine buying for the whole market. Because, to some extent, all consumers are confronted with the same variation in the level of prices, advertising, promotion, etc., at the same time, such a parallel development of poolsize curves of individual consumers is not, a priori, unlikely. As Howard and Sheth (ibid p. 28) remark, if this is the case it would be useful to a marketing manager to know in which phase the consumers in the market are. If he wants to introduce a new brand, for example, a situation where many consumers are at a level of routinization (low poolsize) and are close to feeling satiated with their brand, may constitute a favorable occasion for the introduction.

Average poolsize values per period were computed to test if this cyclical behavior of the brand-switching activity of the market as a whole really existed. For this purpose we had to transform the time axes of individual consumers to real calendar time. Application of the time axis, where the equidistant points are formed by the purchase moments of individual consumers, as was done in Figure 6.1 for example, was then not possible. The two years, for which the purchases are known, can be divided into 24 periods of 4 weeks. Now for each period for each consumer it was established at which poolsize level that consumer was at the end of that period. This is the poolsize reached at the last purchase before the end of the period concerned. This quantity, averaged over all consumers, is the average poolsize for that period. This was done for all 24 periods; with the resulting 24 average poolsize values, the average poolsize curve could be drawn. As observed earlier, the poolsize of an individual consumer is only defined after a consumer has made at least 10 purchases. So a household can only be used for computing the average poolsize if it has completed the first 10 purchases. For this reason not all households are present in the computation of average poolsize for the first few of the 24 periods; for fopro and beer it did not even make sense to compute the average poolsize for period 1, because of the small number of households. As the period number increases, the number of households which can be used also becomes greater.

In Table 6.9 the resulting figures for the average poolsize per period are given, while the corresponding poolsize curves are drawn in Figure 6.5. The last 2 columns of Table 6.9 will be discussed in section 6.7.

From the figures and curves for the average poolsize it can be deduced that only for fopro is there a tendency to cyclical behavior in the brand-switching activity. We see that there are peaks in the 9th and 21st period with a nadir in the 15th period, but the amplitude of the cycles is very modest. For the other products no regularity at all can be established. On the whole, the average poolsize curves are rather flat, which means that no clear distinction can be made between periods with great and periods with little brand-switching activity of the whole market. Obviously, in every period there are consumers with high and consumers with low brand-switching activity, which in the computation of average poolsize balance each other out. This leads us to the conclusion that the poolsize curves and the moments at which they rise and fall are primarily matters of individual consumers, without such common experiences as advertising, promotions, etc., having much influence.

Table 6.9 Average poolsize-values over all households per 4-weekly periods of fopro, beer and margarine

Period	*Fopro*	*Beer*	*Marg*	*Fopro sequence-length* = 20	*Marg sequence-length* = 20
1	—	—	1.96	—	1.50
2	1.42	1.72	2.03	1.50	2.08
3	1.42	1.54	1.78	1.47	2.22
4	1.54	1.65	1.76	1.69	2.35
5	1.59	1.57	1.75	1.81	2.09
6	1.59	1.59	1.80	1.73	2.09
7	1.67	1.69	1.86	1.83	2.14
8	1.72	1.80	1.86	1.89	2.17
9	1.75	1.86	1.80	2.00	2.18
10	1.75	1.90	1.89	2.05	2.27
11	1.70	1.89	1.85	2.06	2.22
12	1.65	1.88	1.82	2.10	2.19
13	1.67	1.96	1.87	2.08	2.23
14	1.63	1.97	1.87	1.99	2.20
15	1.62	1.94	1.90	1.99	2.25
16	1.70	1.91	1.95	2.02	2.31
17	1.74	1.91	1.92	2.03	2.33
18	1.74	1.97	1.95	2.03	2.35
19	1.81	2.02	2.04	2.11	2.42
20	1.87	2.02	2.03	2.21	2.50
21	1.93	2.07	2.01	2.26	2.50
22	1.90	2.06	2.05	2.29	2.54
23	1.81	2.04	1.98	2.23	2.47
24	1.81	2.07	2.03	2.23	2.49

A striking feature of the average poolsize curves in Figure 6.5 is that they all show an upward trend during the 24 4-weekly periods. For beer and margarine the figures of the first periods seem to be somewhat deviant, but it should be remembered that these figures are the least reliable because of the small number of households used in the computation for these periods. This slight upward trend in the average poolsize might point to a general increase in brand-switching activity during the two years observed.

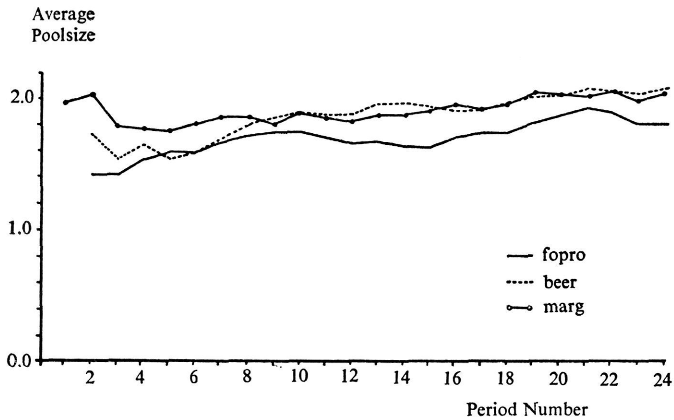

Fig. 6.5 Average poolsize per period for fopro, beer and margarine

One point should be mentioned, which somewhat weakens the statement just made. For fopro and beer it was established that households with fewer purchases (shorter households) have higher values for the MEANPLZ ($r = -.3249$, respectively $r = -.2572$). For the computation of average poolsize these shorter households, because of their lower purchase frequency, can only be used for the later periods and when the shorter households have higher poolsize values, this means that the average poolsize is higher for the later periods, apart from a possible general increase in brand-switching activity.

6.7 EFFECT OF THE SEQUENCE LENGTH CHOSEN

In section 6.2 we defined poolsize as the number of brands bought during the last 10 purchases. The results derived in the previous sections are based on a poolsize with this sequence length 10 and it is necessary to test in how far these results are dependent on the sequence length chosen. For this purpose we examine here how the poolsize curves change when the sequence length of the purchases considered is not 10 but 20. In this case poolsize is defined as the number of different brands bought during the last 20 purchases. It can be directly seen, that this latter poolsize is always greater than or equal to the poolsize with sequence length = 10.

With sequence length = 20 a consumer must have made at least 20 purchases before the poolsize is defined. Since for beer the mean number of purchases is 43, this would result in very short poolsize curves. For this reason the effect of sequence length for beer was not examined. It was assumed that the results for fopro and margarine would give sufficient information.

The effect of the sequence length is considered in 2 ways. One approach is to consider the effect of taking sequence length 20 instead of 10 on the poolsize curves of individual households. We are interested in knowing if the essential characteristics of the poolsize curves then change. Because the mechanism by which a poolsize curve with sequence length 10 is transformed into a poolsize curve with sequence length 20 is the same for all households the effect for 2 specific households is already very informative. Therefore Figure 6.6 gives the poolsize curves with sequence length 20 for those fopro and margarine households for which the poolsize curves with sequence length 10 were drawn in Figure 6.1.

From a comparison of the corresponding curves it can be concluded, that although the curves are not identical of course, the general features are very similar. In particular the feature of the poolsize going up and down is maintained with sequence length = 20. Both types of curves show ups and downs at roughly the same moments. In comparing the curves it should be realized that the curves for sequence length = 20 are shifted 10 purchases to the left, compared with those for sequence length = 10, because in the first case the poolsize is defined 10 purchases later.

A second approach is to study the effect of the sequence length on poolsize values averaged over households. In the last 2 columns of Table 6.9 we find the average poolsize per period for fopro and margarine based on sequence length = 20. These figures should be compared with the corresponding average poolsize curves with sequence length 10, as given in the first, respectively third columns of Table 6.9. We see that the corresponding series of the poolsize values are roughly parallel. Considering the second half of the table, i.e. periods 13 to 24, for which the majority of the households were constantly present in the computation of average poolsize, we see that there is a small, rather constant difference between the average poolsize values based on the two sequence lengths. For fopro this difference ranges from .29 to .41, for margarine from .33 to .49. This parallel development means that analyses such as those made in the previous sections, where differences in poolsize are the important phenomena studied, are not much influenced by the sequence length chosen.

We conclude that the results in this chapter regarding the brand switching behavior of consumers, obtained by studying the poolsize, do not seem sensitive to the specific sequence length chosen.

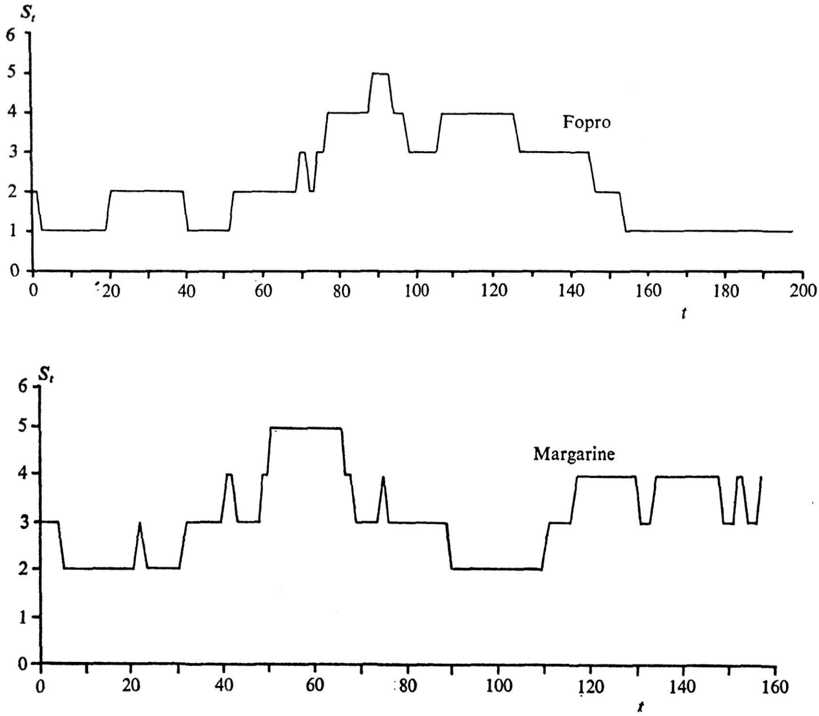

Fig. 6.6 Poolsize curves with sequence length = 20 for the same fopro- and margarine-households as in Figure 6.1

6.8 CONCLUSIONS TO BE DRAWN FROM THE POOLSIZE APPROACH

In this chapter we studied the brand choice process by means of the variable poolsize. The results obtained concern the brand choice process of the 'average consumer' and not necessarily that of an individual consumer.

Here some global findings will be mentioned. More specific results have been reported in the various sections of this chapter.

In section 6.4.6 some conclusions, derived from the general shape of the poolsize curves, were given. The most important findings were that in the brand choice process periods of routinized buying alternate with periods of brand-switching and that a consumer simultaneously considers a limited number of brands as potential choice candidates.

It was also found in this chapter that search behavior is important in the brand choice process. A consumer does not often switch straightforwardly from one brand to another, such a switch is mostly accompanied by more-or-less extensive search behavior.

It was found that the start of search behavior is often accompanied by changes in price, shop and unit-size bought, while increases in search activity are relatively often associated with deal-purchases.

The behavior of the brand-switching activity of individual consumers is not coherent in the sense that different consumers simultaneously exhibit periods of great, or little switching activity.

It appears that the specific sequence length chosen in the definition of poolsize does not strongly influence the results.

These conclusions support the theory of Howard and Sheth with respect to their concept 'evoked set' and the assumed cyclical behavior of search activity during the brand choice process.

The reader will be aware of the fact that the poolsize approach was used for the same brand choice data to which, in chapter 4, a number of different brand choice models was applied. Therefore it is useful to compare the results obtained here with those of chapter 4. In that chapter it was found that the model which gave the best description of the brand choice processes is the Linear Learning Model. In section 4.7.3 we noted that the LLM, with the parameter values for the observed fopro, beer and margarine processes, implies a brand choice process with rather long periods of routinized buying alternating with brand switching periods. This is quite in agreement with the finding of the varying level of brand switching activity in the present chapter.

Because the LLM requires the brand choice process to be condensed to a 0–1 process, in chapter 4 we could only observe switching behavior between two brands. In this chapter a more detailed picture of this search behavior of consumers was obtained.

7. The brand choice process and its relations to environmental variables

7.1 INTRODUCTION

The subject of research in this book is the brand choice process. For this study, the purchase histories of individual consumers are examined. Thus far, we have almost exclusively considered the sequences of brands purchased by consumers. Such sequences were called brand choice processes. For this approach the empirical purchase histories were reduced to the aspect of the brands chosen at the subsequent purchase occasions. It should be realized, however, that brand choices are not made in a vacuum but in a living environment.

Important elements of this environment of the brand choice process are the shops in which purchases are made, and marketing variables such as price, advertising and deal-offers. In addition, the length of the time interval between subsequent purchases is an interesting variable. These inter-purchase times determine the time between two successive purchases during which a consumer is exposed to environmental influences such as advertising, group influences, etc. For this reason the variable inter-purchase time is also treated under the heading of environmental variables.

In this chapter we examine the relationship between the brand choice process and environmental variables. This is done for the empirical purchase histories of fopro, beer and margarine, as discussed in chapter 2. In section 7.2 we study the relationship between brand choice and store choice, while in 7.3 the influence of the marketing variables price, advertising and deal-offers is considered. In section 7.4 the influence of inter-purchase times on brand choice is studied. These 3 sections of this chapter are thus rather independent of each other.

7.2 BRAND CHOICE AND STORE CHOICE

7.2.1 *General*

When a consumer purchases a product in a certain shop, his choice possibilities with respect to brands are limited to the different brands of the product which are available in that shop. Ordinarily a shop will not carry all the brands of a product that are in the market. Generally the national brands will be present and in addition some regional and/or private brands may be available. So a consumer cannot buy any brand in any shop and with the choice of a shop for the purchase of a product the number of possible alternatives is thereby limited. This mere 'limited availability' effect of a shop makes it likely that there is a relationship between brand choice and shop choice. But it is interesting to know if, apart from this, there is also an 'autonomous' relationship between brand choice and shop choice, i.e. if consumers who often change shop also often change brand, without this being necessarily a consequence of the limited availability effect of a shop.

The following first examines in how far there is a relationship between brand choice and shop choice in the empirical purchase histories. The results obtained are compared with findings obtained from the literature. After that we try to get more insight into the nature of the interdependence of shop choice and brand choice. Finally, by means of factor analysis an attempt will be made to unravel the influences of brand and store.

7.2.2 *The relationship observed*

In this section we examine in how far a relationship exists between brand choice and store choice in the empirical fopro, beer and margarine purchase histories.

The following variables will be considered:

NUMBR = number of different brands bought;
NUMSH = number of different shops in which the product was bought;
SFAVBR = share of favorite brand by volume of purchases;
SFAVSH = share of favorite shop for the product in question by volume of purchases.

All these variables refer to the whole 2-year-period observed. Each household produced an observation for each of the above 4 variables. With these values the coefficient of correlation ρ between NUMBR and NUMSH and between SFAVBR and SFAVSH was estimated for all 3 products.

These estimates of the correlation coefficients, represented by the symbol r, are given in Table 7.1.

Table 7.1 Estimated correlation coefficients for the relationship between brand choice and store choice

	Fopro	*Beer*	*Marg*
Number of observations	672	627	1059
r (NUMBR, NUMSH)	.691	.469	.623
r (SFAVBR, SFAVSH)	.577	.349	.358

It appears that the correlation coefficients are quite considerable. From the large number of observations it is immediately clear that all coefficients are significantly greater than zero, even for $\alpha < .01$. So it can be concluded that households with many brands also have many shops and households with a big share for the favorite brand also have a big share for the favorite shop. For the test on difference between two correlation coefficients the z-transformation:

$$z = \tfrac{1}{2}\ln\frac{1+r}{1-r}$$

can be used (see Kendall & Stuart, II, 1967, pp. 292–295). The hypothesis that two correlation coefficients ρ_1 and ρ_2 are equal can be tested by considering:

$$d = \frac{z_1 - z_2}{\dfrac{1}{N_1-3}+\dfrac{1}{N_2-3}},$$

where z_1 and z_2 are the z-transforms of r_1 and r_2 and N_1 and N_2 are the numbers of observations from which r_1 and r_2 were respectively computed. For N_1 and N_2 large d is distributed approximately standard normal when $\rho_1 = \rho_2$. With this test it was verified that ρ(NUMBR, NUMSH) and ρ(SFAVBR, SFAVSH) for fopro are greater than for beer and margarine. This means that the relationship between brand and shop is closer for fopro than for the two other products. Further ρ(NUMBR, NUMSH) for margarine is greater than for beer but this does not hold for ρ(SFAVBR, SFAVSH).

Another approach which can be followed to establish the relationship between shop choice and brand choice is to examine how often brand switches or brand changes go together with shop changes and vice versa. For this purpose for all households we considered the purchases for which the previous purchases were known and noted, if the brand was the same as the brand bought at the previous purchase and if the shop was the same as the shop at which the previous purchase of the product was made. These figures are presented in Table 7.2.

For our purposes the results under d(3) and d(4) are especially interesting. We see that for fopro, of all purchases with a brand change 65.1% also have a shop change, compared with 9.5% for fopro purchases in general. So purchases with a brand change more often have a shop change than purchases in general. On the other hand, for all shop changes 60.8% go with a brand change, which is much more than the 8.9% for all purchases. Also for beer and margarine there is strong evidence that brand changes and shop changes 'stick together': brand changes are, relatively, very often associated with shop changes and vice versa.

Table 7.2 Association of brand change and shop change

	Fopro	*Beer*	*Marg*
a. Number of purchases considered	55482	20364	111540
b. Number of purchases with a brand change			
(1) absolute	4925	2325	29374
(2) as % of a.	8.9	11.4	26.3
c. Number of purchases with a shop change			
(1) absolute	5279	2814	15289
(2) as % of a.	9.5	13.8	13.7
d. Number of purchases with brand and shop change			
(1) absolute	3209	1198	10961
(2) as % of a.	5.8	5.9	9.8
(3) as % of b(1)	65.1	51.5	37.3
(4) as % of c(1)	60.8	42.5	71.7

Because of the higher percentage values for d(3) and d(4) the relationship between shop choice and brand choice is again narrower for fopro than for beer. The same holds for fopro as compared to margarine, when d(3) is considered but not when d(4) is considered.

From these results and the rather high correlation coefficients found above it is clear that brand choice and shop choice are rather strongly related. In sections 7.2.4 and 7.2.5 this interdependence is studied more closely. First the findings of this section will be compared with some results from the literature.

7.2.3 *Some results from the literature*

Here we give some results from the literature with which the above findings can be compared.

Cunningham (1961), who analysed the purchases of 18 food products made by 50 families during one year, found for 10 of the 18 products that households with a big share for the favorite brand also had a big share for the favorite shop.

Massy, Frank and Lodahl (1968), who considered coffee purchases of 670 families during one year, found a correlation coefficient between the number of different brands and the number of different shops of .38.

Rao (1969) analysed purchases of 3 consumer products during 2 years and concluded that the more a household changes stores, the more it changes brands. Carman (1970) studied the purchases of housewives of members of the Berkeley Marketing Faculty during a 15-week period. He found that the number of different food chains visited explained 63.5% of the total variance in brand loyalty for coffee, 66% for canned fruit and 63% for frozen juice.

So in all these results a clear relationship between brand choice and store choice is demonstrated, which is quite in agreement with our findings.

7.2.4 *The nature of the relationship between brand choice and store choice*

The crucial point with respect to the relationship between brand choice and store choice is: does this relationship only exist because the choice of a shop limits the set of brands from which a choice can be made – we called this the limited availability effect of a shop – or is there an autonomous underlying factor, which, for example, might be called the 'general proneness-to-change factor' which means that certain consumers tend to show great variation in store as well as in brand, while others show routine behavior with respect to store as well as to brand? A related question is, if in addition to a general proneness-to-change factor there exist an independent proneness-to-brand-change factor and an independent proneness-to-shop-change factor. If all these 3 factors exist, there will be households which show a lot of brand change without much shop change, there will be households which show a lot of shop change without much brand change and there will be households showing a lot of brand change and shop change.

This section will show that not the whole relationship between brand change and shop change can be traced back to the mere limited availability effect of a shop and that the 3 factors, mentioned above, can indeed be distinguished. In the next section a factor analysis will be applied to get still more insight into the nature of the interdependence of brand choice and shop choice.

First it will be shown that there is a general proneness-to-change factor, i.e., that there are consumers who often change brand and shop, without this being necessary because of the limited availability effect of a shop. To this aim we examine more closely the association of brand change and shop change. For the brand choice figures in Table 7.2 any change between two brands, insofar they were differently coded, was considered as a brand change. In this connection it can be mentioned that for fopro 52 brands were distinguished, for beer 50 and for margarine 87. In addition there was for every product a remaining brand, which stood for 'all others'. To draw conclusions with respect to the impact of the limited availability effect of a shop, it is necessary to know the distribution of the different brands, i.e. in which percentage of all shops the different brands are available.

This is not known for all the numbers of brands just mentioned. Therefore a more limited concept of brand change will be applied by distinguishing in every market only 4 different brands, which are brands F1 to F4 for fopro, B1 to B4 for beer and M1 to M4 for margarine, as mentioned in section 2.5. Now only changes between these brands are conceived of as brand changes. Because we have national brands here, together with a category 'all others' it is to be expected that their distribution will be rather good, so that each of these brands can be bought in the majority of shops. So in many cases it will not be necessary for a consumer to change shop in order to be able to change brand, in the sense just defined. When shop change and brand change go together, only because of the limited-availability effect of a shop, it is to be expected that these brand changes will not often be accompanied by shop changes. Yet the following figures with respect to the degree to which shop change goes together with brand change were found: for fopro, in 58% of the cases of a brand change a shop change also took place; for beer this figure was 41% and for margarine 17%. So we see here that shop change and brand change also often go together, although this will not be necessary because of distribution limitations.

We can consider this in some more detail. For brands F1, B1 and M1 we know the distribution figures; these were obtained from a retailers panel and a company in the branch concerned. There is one difficulty, viz., the distribution figures refer only to grocery shops; we use them here for all outlets together. Now it is known that brand F1 was available in 85% of shops during the 2-year period observed, brand B1 in 91% and brand M1 in 99% of shops. Consider a transition for fopro from another brand (not F1) to F1. Assuming that the shops in which consumers buy before making such a transition, 'carry brand F1', respectively 'do not carry brand F1' in the same proportions as all shops, we can expect that in 100% − 85% = 15% of the cases such a brand change will be accompanied by a shop change. It was found, however, that 44% of this type of transition goes with a shop change. Quite analogously, when considering transitions for beer of the type: other brand → B1, we would expect, when only the limited availability aspect is taken into account, 9% of such transitions to be associated with a shop change. In reality, however, the figure was 38%. For margarine the expected percentage of shop change for transitions of the type: other brand → M1 would be 1%, while in fact it was found to be 32%. So we find that shop change is much more often associated with brand change than can be explained by limitations in the availability of the different brands. The fact that brand change and shop change often go together points to the existence of a general proneness-to-change factor, which has reference to brand as well as to shop.

Next we consider the question of an autonomous proneness-to-brand-change factor, respectively, proneness-to-shop-change factor. We first define 2 new variables:

BRPSH = mean number of brands per shop. The value of this variable was computed for every household by counting for each shop in which purchases were made the number of different brands bought. The results for all shops of a household were then averaged, which produced the mean number of brands per shop.

SHPBR = mean number of shops per brand. This variable was computed as follows: for every household it was counted for each brand purchased in how many different shops that brand was bought. The average of the results for all brands of a household is the mean number of shops per brand.

Now consider the relationship between NUMBR and BRPSH. If the fact that certain households have many brands, compared to other households, is completely due to the fact that they have many shops and therefore more choice alternatives, households with many brands will also have many shops (which was already confirmed above), but there is no reason to suppose that those households also have many brands per shop. So then we expect no relationship between NUMBR and BRPSH. When, on the other hand, there would be a clear positive correlation between NUMBR and BRPSH, indicating that households with many brands also have many brands per shop, this would point to the existence of an autonomous proneness-to-brand-change factor.

The correlation coefficients between NUMBR and BRPSH have been estimated for the fopro, beer and margarine households and are given in Table 7.3.

Table 7.3 Estimated correlation coefficients, indicating the existence of distinct proneness-to-brand-change and proneness-to-shop-change factors

	Fopro	*Beer*	*Marg*
Number of observations	672	627	1059
r (NUMBR, BRPSH)	.582	.584	.609
r (NUMSH, SHPBR)	.440	.554	.528

We see that the magnitude of these coefficients is considerable. For all 3 products, consumers with many brands also have many brands per shop, which indicates that there is an autonomous proneness-to-brand-change factor.

The relationship between NUMSH and SHPBR is considered in an analogous way. When the limited availability effect exclusively governs the relationship between brand choice and shop choice, it is not expected that consumers with many shops will also have many shops per brand, i.e. that consumers with many shops will buy each of their brands in a relatively large number of shops. Just because the shop is then the limiting factor, this is not likely. But in Table 7.3 we see that the NUMSH and SHPBR are evidently positively correlated. So consumers with many shops generally also buy each of their respective brands in many different shops, which indicates the existence of an autonomous proneness-to-shop-change factor. It can be further mentioned that the correlation coefficients between NUMBR and SHPBR, respectively between NUMSH and BRPSH, appeared to be very small for all 3 products.

So we conclude that in addition to the general proneness-to-change factor already observed, a separate proneness-to-brand-change factor and proneness-to-shop-change factor can be distinguished. In the next section we will try to unravel the interdependence of shop choice and brand choice somewhat further.

7.2.5 *A factor analytical approach to the interdependence of shop choice and brand choice*

7.2.5.1 *Introduction and definition of the variables*

To obtain more insight into the interdependence of shop choice and brand choice a factor analysis was carried out on a number of variables, which all refer to the relationship between shop and brand.

For extensive treatments of factor analysis the reader is referred to Lawley & Maxwell (1971) or D. F. Morrison (1967). For the results, which follow, we used a maximum likelihood estimation procedure and applied the iterative algorithm described in Morrison (ibid., chapter 8). The rotation of factors was done according to the varimax procedure, also given there.

The factor analysis was performed with 12 variables. These variables, the first six of which were defined earlier in this chapter, are;

1. SFAVBR;
2. NUMBR;
3. SFAVSH;
4. NUMSH;
5. SHPBR;
6. BRPSH;
7. BRCHWSHCH = brand change with shop change = portion of brand changes with an associated shop change;
8. SHCHWBRCH = shop change with brand change = portion of shop changes with an associated brand change;
9. NOCH = no change = portion of purchases with neither a shop change nor a brand change;
10. BRCH = brand change = portion of purchases with a brand change;
11. SHCH = shop change = portion of purchases with a shop change;
12. BRSHCH = brand change and shop change = portion of purchases with a shop change and a brand change.

All variables refer to the whole 2-year period observed for each household.

For every household an observation for each of these 12 variables was obtained. Because it is impossible to calculate the value of variable 7 for a household not changing its brand and of variable 8 for a household not changing its shop, for such households the average value for variable 7 and variable 8 over all remaining households were taken as observations.

For every product a (12 × 12) correlation matrix for the above variables was computed and these correlation matrices formed the starting points of the factor analysis.

7.2.5.2 *Results of the factor analysis*

For all 3 products 4 uncorrelated factors were extracted, which were afterwards rotated by means of the varimax procedure. The number of 4 for the extracted factors is somewhat arbitrary. In all 3 cases the appropriate test statistic, given in D. F. Morrison (1967, p. 269) indicated that by adding more factors the fit of the factor model could still be improved. In this connection it should be realized, however, that the height of the test statistic is especially due to the large sample sizes (672, 627 and 1059 for fopro, beer and margarine respectively). If additional factors, even though they are significant, contribute little to the explanation of the correlations observed, they are difficult to interpret. It was experienced that the improvement in the explanation of the correlation matrix resulting from the additional fourth factor was only slight. This led to the decision to stop the factor analysis after 4 factors.

The matrices of factor loadings obtained are given in Table 7.4. These factor loadings are correlation coefficients between original variables and factors. The factors are indicated as shown in the line under the matrices of factor loadings. The last row of Table 7.4 gives the portion in the sum of variances of all variables which is 'explained' by the respective factors. This portion is equal to the sum of squares of the loadings in the column concerned, divided by 12. In the following we discuss the results obtained and try to find suitable names for the factors. For the interpretation of a factor it is important to know on which variable(s) it has (a) high loading(s) in an absolute sense. Therefore, the discussion of the factors concentrates on the big loadings; i.e., loadings with an absolute value greater than .4 are mainly considered.

For *fopro* we see that the first factor FFA is positively related with NOCH and SFAVBR, and negatively with SHCH, BRCH, BRSHCH, NUMBR and NUMSH. So high FFA-values go together with little change, with respect to brand as well as to shop, and with few different brands and shops. Therefore FFA can be called a general variability factor or *general proneness-to-change* factor.

Table 7.4 Matrices of factor loadings

	Fopro				*Beer*				*Marg*			
1. SFAVBR	.553	−.189	.174	−.469	.104	−.670	.355	−.097	.592	−.137	−.069	−.499
2. NUMBR	−.462	.164	−.013	.869	−.237	.904	−.164	.213	−.306	.247	.034	.916
3. SFAVSH	.570	.223	−.186	−.314	.616	−.091	.371	−.043	.143	−.452	.283	−.256
4. NUMSH	−.497	−.154	.239	.562	−.952	.203	−.172	.150	−.196	.417	−.258	.510
5. SHPBR	−.199	−.471	.078	−.116	−.688	−.424	−.166	−.290	−.090	.244	−.431	−.130
6. BRPSH	−.263	.097	−.401	.504	.289	.721	−.186	−.235	−.392	−.037	.076	.539
7. BRCHWSHCH	−.236	.092	.637	−.021	−.258	−.214	−.112	.678	.144	.671	−.001	.090
8. SHCHWBRCH	−.191	.607	.098	.272	.142	.391	−.141	.706	−.328	.275	.382	.348
9. NOCH	.945	.142	.109	−.274	.238	−.300	.786	.038	.924	−.289	.144	−.204
10. BRCH	−.883	.319	−.185	.290	−.027	.490	−.667	.296	−.948	.214	.108	.209
11. SHCH	−.918	−.197	.281	.198	−.416	−.009	−.746	.213	−.319	.909	−.214	.160
12. BRSHCH	−.870	.350	.277	.210	−.130	.199	−.653	.661	−.375	.888	.189	.184
Factor	FFA	FFB	FFC	FFD	FBA	FBB	FBC	FBD	FMA	FMB	FMC	FMD
Portion in Variance	.38	.09	.08	.16	.19	.22	.21	.15	.23	.23	.05	.17

Measured after the portion in variance, the second factor in importance for fopro is FFD. FFD is positively correlated with NUMBR, NUMSH and BRPSH. So FFD also refers to a kind of general variability, although people with high FFD-values tend to be somewhat more brand change-prone than shop change-prone. Unlike FFA, the other general variability factor, FFD is not strongly correlated with the change variables 9 to 12. So consumers with high FFD-values, having many different brands and shops, do not show many changes in brand and/or shop. On the other hand, consumers who score low on FFA also have many brands and shops but these consumers also show many changes, i.e. these consumers vacillate more. It can be said therefore, that the changes of consumers scoring high on FFD are more deliberate. Therefore we call FFD the *deliberate-change* factor.

The factors FFB and FFC are less important because of their generally lower correlations with the variables. Without trying to give them names, some tentative remarks can be made with respect to these factors. Both have to do with the limited availability effect of a shop, but in a different sense. High scores on FFB mean that many of the shop changes are also brand changes (var 8), while a brand is bought in relatively few shops. Yet the changes within a shop are normal (there is no correlation with BRPSH) and not especially many of the brand changes are also shop changes (var 7). So consumers with high FFB-values seem to be rather brand-change-prone, but the fact that not many shops were visited (and perhaps especially shops with few different brands) limits their change possibilities. On the contrary, households with high FFC-values have few different brands per shop (var 6) and for them brand change mostly goes together with shop change (var 7), but when the shop changes the brand does not very often change with (var 8). So high FFC-values point to little proneness-to-brand-change, which is somewhat hidden because limited availability sometimes means that shop changes force brand changes.

The first factor for *beer*, FBA, is negatively correlated with NUMSH, SHPBR and SHCH and positively with SFAVSH. So high values of FBA are associated with few shops, few shops per brand, little shop change and a big share in the favorite shop, while there are no strong correlations with number of brands, share of favorite brand, etc. Therefore FBA can be called an autonomous *proneness-to-shop-change* factor. FBB shows about the same with respect to brand as FBA does with respect to shop. High values for FBB mean many brands (var 2), a small share of brand 1 (var 1), many brands per shop (var 6) and a lot of brand change (var 10). So FBB can be called an autonomous *proneness-to-brand-change* factor. FBC has high loadings only on the change variables 9 to 12. High values of FBC mean few changes with respect to brand as well as to shop, without especially small numbers of brands and shops. Therefore FBC can be called the *vacillation* factor.

The factor FBD is positively correlated with BRCHWSHCH, SHCHWBRCH and BRSHCH. So high values of FBD mean that brand changes and shop changes mostly occur together and occur relatively often. We call FBD the *change-together* factor, because it indicates in how far shop change and brand change go together.

The first factor of *margarine*, FMA, is negatively correlated with BRCH and positively with NOCH and SFAVBR. So high values of FMA mean little brand change, little change in general and a big share of the favorite brand. We call FMA the *proneness-to-brand-change* factor.

High values of FMB mean a lot of shop change (var 4), a lot of shop and brand change (var 12), a low share of the favorite shop (var 3) and many shops (var 4). FMB can be called the *proneness-to-shop-change* factor. The factor FMC, which has only a small portion in the variance, is difficult to interpret.

Because of the same signs and magnitudes of the respective loadings, FMD shows a striking similarity with the factor FFD of fopro. Therefore we also call it the *deliberate-change* factor.

Summarizing the most important results, for fopro a general proneness-to-change factor and a deliberate-change factor were found, for beer a change-together factor and for margarine a deliberate-change factor again, which all point to a basic variability factor referring to brands as well as to shops. Further, for beer and margarine evidently distinct proneness-to-brand-change and proneness-to-shop-change factors were found, while for fopro the proneness-to-brand-change factor seemed to be hidden by the limited availability effect of shops. In this connection it may be mentioned, that about 50% of purchases of the product fopro were made from ambulant shops, i.e., shops which bring products to the consumer's door. Generally the assortments of such ambulant shops as rather limited. The fact that we find here for fopro the strongest general change factor and no such clear distinct shop change and brand change factors as for beer and margarine, is in agreement with the results obtained in section 7.2.2, i.e., that of the 3 products, the relationship between brand change and shop change is strongest for fopro.

In general the factors which were found support the earlier findings, that separate proneness-to-brand-change, proneness-to-shop-change and general proneness-to-change factors can be distinguished.

A remarkable phenomenon is the deliberate-change factor, observed for fopro and margarine. In chapter 8 it will be noted how factor scores for individual households were computed back from the matrices of factor loadings. Now it was found that the scores on the factors FFD and FMD correlated strongly with the poolsize variable RANGPLZ, defined in the previous chapter (section 6.4.1). The correlation coefficient between FFD and RANGPLZ for fopro was .809 and between FMD and RANGPLZ for margarine it was .807. So consumers scoring high on the deliberate-change factor also showed the cyclical behavior in search activity as discussed in chapter 6. This can be understood when it is remembered that both phenomena refer to the extent to which well-considered changes in brand are made. The fact that no such deliberate-change factor was found for beer is in agreement with the finding of the previous chapter that the cyclical pattern of search behavior was less evident for beer than for the 2 other products.

7.2.6 *Conclusions regarding the relationship of brand choice and shop choice*

It was found that brand choice and shop choice are rather closely related. This interdependence cannot be completely traced back to the limited availability of the different brands in a shop, but there is an autonomous general proneness-to-change factor, which means that some consumers often change shop and brand, while others show routine behavior with respect to store as well as to brand. Besides the general proneness-to-change factor, a specific proneness-to-brand-change factor and a proneness-to-shop-change factor can be distinguished.

The brand choice models developed thus far, generally do not take into account the shop in which a purchase is made. In the light of the findings in this section it might be useful to develop brand choice models which incorporate the effect of store choice.

7.3 INFLUENCE OF MARKETING VARIABLES ON BRAND CHOICE

7.3.1 *Introduction*

The marketing variables, also called the elements of the marketing mix, are the means by which a company implements its marketing policy. When a company wants to influence the position of its brand in the market, it can attempt to do this by manipulating one or more of the instruments of the marketing mix. Therefore it is very important to know the effect of the marketing variables on brand choice. With the data here it is possible to study the effect of the price, advertising and deal variables, which is done in the present section.

One possible procedure would be to try to directly establish the effect of the marketing variables on the parameters of the model describing the brand choice process. In chapter 4 it was seen that the model which best fitted the empirical data was the LLM. Parameter estimation of the LLM as such is not a simple affair, but the problems become much bigger if such parameters are conceived of as being functions of the marketing variables, so that the parameters of these functions have also to be established. Therefore not this approach will be followed, but that in which merely the influence of the marketing variables on brand choice is examined, without exactly establishing the effect for the LLM-parameters.

In section 7.3.2 we treat the effect of price and advertising on brand choice, while section 7.3.3 considers the influence of deal-offers.

7.3.2 *Price and advertising*

7.3.2.1 *The method used*

7.3.2.1.1 *Introduction*

In the following we assume that certain brand choice variables are linear functions of price and advertising. Even though, in reality, this may be not the case for all values of price and advertising variables, for a certain range of these variables a linear function can be used as an approximation.

Generally stated:

$$Y_t = \lambda_0 + \lambda_1 X_{1,t} + \lambda_2 X_{2,t} + \ldots + \lambda_p X_{p,t} + U_t \qquad (7.1)$$

Here Y_t denotes a brand choice variable and $X_{1,t}$ to $X_{p,t}$ represent a number of price and advertising variables plus a trend factor; t refers to the time period. Exact definitions of variables X_t and Y_t will be given in the next section. U_t is the disturbance term for which the usual assumptions are made of U_t being normally distributed with expected value zero and constant variance. Further, successive U_t values are assumed to be mutually independent. Now, when observations of the set Y and X_1 to X_p are available for a number of periods, the parameters λ_0 to λ_p of equation (7.1) can be estimated via a multiple regression technique and after that statements about the reliability of these estimates can be made. For an extensive treatment of this technique of multiple regression the reader is referred to Johnston (1972), Draper and Smith (1968) or Wonnacott and Wonnacott (1970).

The next section gives a number of different specifications for the relationship between brand choice and price and advertising, which all have the form (7.1). Four different definitions for the brand choice variable Y will be used.

7.3.2.1.2 *Definition of variables and specification of equations*

The effect of price and advertising was studied for a number of brands, viz., those brands for which advertising figures were available. They are F1 and F2 for fopro, B1, B2 and B3 for beer and M1, M2 and M3 for margarine. In the following, when a certain brand is spoken of and it does not matter which of the ones just mentioned is meant, we indicate that brand as A.

As mentioned earlier, the 2-year period for which purchase data for fopro, beer and margarine were available can be divided into 24 4-weekly periods. Now for every period we computed the values of 4 brand choice variables for each brand studied. These brand choice variables are:

MSV = the market share by volume of brand A;
MST = the market share of brand A by number of purchase times;
$P(A \mid A)$ = the fraction of times that after an A purchase another A purchase was made. We call this the *repeat purchase probability*;
$P(A \mid \bar{A})$ = the fraction of times that the purchase of a brand other than A was followed by an A purchase. We call this the *transition probability*.

Of course the market share variables MSV and MST are more indirect brand choice variables, as compared with variables $P(A \mid A)$ and $P(A \mid \bar{A})$, which refer directly to transitions from one brand to another.

The value of each of these variables for a certain brand in a certain period was computed by considering all purchases of all households in that period. These 4 brand choice variables became in turn the dependent variables in each of the 5 function-specifications which follow.

As independent variables we used a trend factor, the relative price of brand A, i.e. the price of brand A divided by the average price of all other brands in the market, and variables which refer to the advertising expenditures on brand A and other brands in the market. As indicated in section 2.3, for every purchase in the consumer panel the price paid is recorded, so the prices of the different brands per period could be computed from the purchase data. For the advertising figures the data discussed in section 2.6 were used. Those data were available on a monthly basis and were transformed to 4-weekly data by assigning to each 4-weekly period parts of the corresponding monthly expenditures proportional to the distribution of the days of that period over the months in question. This is somewhat arbitrary and there is no complete certainty that the advertising expenditures were exactly spent in the periods to which they were assigned. For the first 4-weekly period there were no advertising data available, so this period was discarded for the regression. Because it is possible that advertising effects show a certain time lag, advertising in the current as well as in the previous period were always used as variables in the regression. This implies that again one period is lost as an observation. Therefore all regressions were performed with $24-2=22$ observations. The exact definitions of the independent variables are as follows. (The index t is omitted for simplicity.)

X_1 = Trend: period $1=1$, period $2=2$, etc.

X_2 = Price of brand A divided by the average price of all other brands in the market (relative price)

X_3 = Advertising expenditures on brand A in the current period

X_4 = Advertising expenditures on all other brands in the market in the current period

X_5 = Advertising ratio of brand $A = X_3/X_4$

X_6 = As X_3, but for the previous period

X_7 = As X_4, but for the previous period

X_8 = X_6/X_7

X_9 = Advertising difference = advertising expenditures on all other brands − advertising expenditures on brand A in the current period

X_{10} = Idem, but for the previous period

X_{11} = X_3 for periods in which X_5 is greater than its average value
= 0, otherwise

X_{12} = X_6 for periods in which X_8 is greater than the average value
= 0, otherwise

X_{13} = X_5 for periods in which the total advertising expenditures in the current period are above average
= 0, otherwise

X_{14} = X_8 for periods in which the total advertising expenditures in the previous periods are above average
= 0, otherwise

All variables defined here refer to 4-weekly periods.

The following specifications were used for the relationship between brand choice and price and advertising variables, which all have the general form (7.1). In these equations the dependent brand choice variable is indicated as Y, which stands for MSV, MST, $P(A|A)$ and $P(A|\bar{A})$ respectively. The differences in these specifications, which are briefly discussed, all refer to the advertising variables.

a. $Y = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \alpha_5 X_5 + \alpha_8 X_8 + U$

In this specification it is assumed that the advertising ratio influences Y.

b. $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_9 X_9 + \beta_{10} X_{10} + U$

Here the absolute difference in advertising expenditures between the brand in question and all other brands together is taken as a measure of the advertising effect.

c. $Y = \gamma_0 + \gamma_1 X_1 + \gamma_2 X_2 + \gamma_3 X_3 + \gamma_4 X_4 + \gamma_6 X_6 + \gamma_7 X_7 + U$

Now the absolute values of advertising expenditures, both for the brand concerned and all other brands, are taken as advertising variables.

d. $Y = \delta_0 + \delta_1 X_1 + \delta_2 X_2 + \delta_5 X_5 + \delta_{13} X_{13} + \delta_8 X_8 + \delta_{14} X_{14} + U$

This is an extension of specification (a). As in this specification, the advertising ratio is taken as independent variable, but now the possibility is built in for this ratio to have a different effect for 2 different general levels of advertising.

e. $Y = \varepsilon_0 + \varepsilon_1 X_1 + \varepsilon_2 X_2 + \varepsilon_3 X_3 + \varepsilon_{11} X_{11} + \varepsilon_6 X_6 + \varepsilon_{12} X_{12} + U$

This specification is a variant of (c), which was introduced because in some cases, in particular for brand B1 of beer, the advertising expenditures on the brand in question were strongly correlated with the advertising expenditures on all other brands. This implies that X_3 is correlated with X_4 and X_6 with X_7 and estimation of specification (c) involves the well-known problem of multi-collinearity, which makes it difficult to disentangle the separate effects of the independent variables.

Because it is always possible to increase the coefficient of determination R^2 by including additional independent variables, when the fit of 2 specifications is to be compared, it is better to use, instead of R^2 itself, the so-called adjusted multiple correlation coefficient $\bar{R}^2$, which is defined as

$$\bar{R}^2 = R^2 - \frac{k-1}{n-k}(1-R^2), \tag{7.2}$$

see Wonnacott and Wonnacott (1970, p. 311). Here k is the number of independent variables, inclusive of the constant, and n is the number of observations. This function clearly penalizes R^2-values with high values of k, especially when n is small.

7.3.2.1.3 *Specific difficulties*

As mentioned in section 7.3.2.1.1, in the usual regression procedure it is assumed that the disturbance term U has constant variance and that successive U-values are mutually independent. Here we discuss two possibilities of deviation from this situation and how to cope with them.

The fact that the 4 brand choice variables, which we use as dependent variables in the regressions, are all proportions, can mean that the requirement of constant variance of U is violated. Let us consider, for example, the dependent variable MST, i.e. the number of purchases of the brand considered (generally indicated as brand A) divided by the total number of purchases. Now when we write MST as a function of a number of independent variables X in the general form (7.1) we get:

$$\text{MST}_t = \lambda_0 + \lambda_1 X_{1,t} + \ldots + \lambda_p X_{p,t} + U_t \qquad (t = 1, n)$$

Then, because $EU_t = 0$:

$$E\ \text{MST}_t = \lambda_0 + \lambda_1 X_{1,t} \ldots + \lambda_p X_{p,t}$$

Thus when $\theta_t \overset{\text{def}}{=} \lambda_0 + \lambda_1 X_{1,t} \ldots + \lambda_p X_{p,t}$,

θ_t is the expected or true fraction of A purchases in period t. Then, as is well-known from the theory of the binomial distribution, the variance of the fraction MST_t is:

$$\text{Var MST}_t = \frac{\theta_t(1-\theta_t)}{N_t},$$

where N_t is the number of observations in period t. Therefore, because different periods can have different true fractions θ_t and different numbers of observations N_t, the variance of MST, and as a consequence the variance of U, can be different for different periods. This is in contrast with the assumption of constant variance mentioned above.

An additional difficulty with the use of proportions as dependent variables in regression is that the value of this variable must always lie between 0 and 1. It is possible, however, that in working with specifications (a) to (e) above, we find such values for the regression coefficients that the values of MST fall outside the [0, 1] interval in certain cases. When the results of the regression analysis are used to make predictions, predictions outside the [0, 1] interval are difficult to interpret of course.

Cox (1970), chapters 2 and 3, discusses the problems mentioned above. To avoid them he proposes a certain transformation of the observed variables before the least squares estimation is performed. This transformation was applied to the data here, but it was found that the regression results were hardly changed by it. This may be due to the fact that the observed fractions and the sample sizes varied over a limited range, which makes the problems just mentioned less serious.

A second source of difficulty is the assumption regarding the disturbance term U_t, i.e., that subsequent U_t-values are mutually independent. However, in econometric analyses, especially when the time periods to which the observations refer are short, it often happens that the disturbances are serially correlated. This is the phenomenon of auto-correlation (see Johnston, 1972, chapter 8).

With the Durbin-Watson d-statistic it can be tested whether or not auto-correlation is present. The policy was followed that whenever d was lower than 1.25, an attempt was made to remove auto-correlation by the Cochrane-Orcutt procedure (also given by Johnston).

7.3.2.2 *The results obtained*

7.3.2.2.1 *Presentation of results*

As mentioned earlier, the effect of price and advertising was analysed for the brands F1, F2, B1, B2, B3, M1, M2 and M3. For each brand the regression was performed for the 4 dependent variables MSV, MST, $P(A \mid A)$ and $P(A \mid \bar{A})$, as defined in section 7.3.2.1.2. For each dependent variable the regression was run for the 5 specifications (a) to (e), also given in section 7.3.2.1.2. Now for each specification the overall regression result was first judged by considering R^2. Every regression was discarded for which the R^2 was so low that the corresponding F-value was lower than the upper 10% point of the F-distribution. This implied that for some dependent variables for some brands no regression was left for which the result was sufficiently significant.

After that it was verified for each dependent variable which of the remaining regressions gave the best results, measured after the value of $\bar{R}^2$, (equation (7.2)). In Table 7.5 it is indicated for each brand and each dependent variable, with the corresponding R^2-values and Durbin-Watson d-statistics, which specification gave the best result. When a dependent variable is not mentioned for a certain brand in Table 7.5, this means that for that dependent variable none of the 5 specifications gave sufficiently significant results. The regression coefficients of the specifications presented in Table 7.5 are also given, so long as the absolute values of the corresponding t-values were not too small. Here, the qualification was used that the observed t-value should be greater than the upper 10% point or smaller than the lower 10% point of its distribution under the null-hypothesis, the null-hypothesis being that the coefficient is zero. For each regression coefficient of Table 7.5 is given the probability of obtaining – under the null-hypothesis – a bigger value of the regression coefficient, resp. a smaller one, depending on whether the observed coefficient is positive, respectively negative. This is the quantity given in brackets. Here .01 stands for '.01 and lower'. By considering these probabilities the reader can judge for himself the significance of the results.

Table 7.5 gives the regression coefficients for the trend variable and for the price and advertising variables, not for the constant which appears in each of the equations (a) to (e). The regression coefficients were rounded off to 4 decimals, except in cases where resulting value would be zero. In those cases a fifth decimal was also given. The procedure for removing auto-correlation, mentioned in section 7.3.2.1.3, was performed only once, viz., in the case of the dependent variable $P(\text{M1} \mid \text{M1})$.

7.3.2.2.2 *Discussion of the results regarding the observed effect of price and advertising*

In the following the results of the regressions are discussed insofar as they refer to the effect of price and advertising; the estimated values of the constants and the trend coefficients are not discussed.

Fopro

For brand F1 of fopro, we see that for the dependent variable MSV specification (e) gives the best result. There is a negative influence of the relative price (X_2), which means that the higher the relative price of F1, the lower the market share by volume. There is a tendency for brand F1 advertising to have a negative influence on the MSV when the portion of F1 in the total advertising expenditure is high (X_{11}). This points to the possibility of an overdose of advertising for brand F1, which becomes somewhat explicable when it is known that brand F1, taken over all periods together, spent almost 50% of all advertising expenditures in the product field. The probability of making a transition from another brand to F1, $P(\text{F1} \mid \overline{\text{F1}})$, is negatively influenced by advertising expenditures on other brands in the current period (X_4), and positively influenced by its own advertising expenditures (X_6) and by competing advertising expenditures (X_7) occurring in the previous period.

The latter is somewhat surprising and might be explained by the reasoning that when competing advertising is high in the previous period, many consumers attracted to a competing brand come back to F1 in the current period. For brand F2, only the MSV and MST gave significant regression results. In both cases, the relative price (X_2) and competing advertising expenditures in the previous period (X_7) had a negative influ-

Table 7.5 Results of the regression analysis establishing the effect of price and advertising on brand choice; N = 22 for all regressions (Figures in brackets are critical levels)

Product	*Brand*	*Dependent variable*	*Best spec.*	R^2	*d*	*Regression coefficients with t-values above 10%-level*
Fopro	F1	MSV	e	.67	1.58	$\varepsilon_1 = .0032(.01)$ $\varepsilon_2 = -.0081(.01)$ $\varepsilon_{11} = -.0001(.06)$
		$P(F1\mid\overline{F1})$	c	.68	1.91	$\gamma_1 = .0009(.03)$ $\gamma_4 = -.0001(.01)$ $\gamma_6 = .0001(.04)$ $\gamma_7 = .0001(.01)$
	F2	MSV	c	.84	1.87	$\gamma_1 = .0008(.01)$ $\gamma_2 = -.0046(.01)$ $\gamma_7 = -.0001(.06)$
		MST	c	.68	1.89	$\gamma_1 = .0009(.03)$ $\gamma_2 = -.0041(.01)$ $\gamma_3 = .0002(.04)$ $\gamma_7 = -.0001(.02)$
Beer	B1	MST	d	.63	1.44	$\delta_5 = -.0659(.01)$ $\delta_8 = .0369(.04)$ $\delta_{14} = .0303(.01)$
		$P(B1\mid B1)$	d	.51	1.54	$\delta_{13} = -.0315(.02)$ $\delta_{14} = -.0192(.09)$
		$P(B1\mid\overline{B1})$	d	.50	1.80	$\delta_{13} = .0313(.01)$
	B2	MSV	b	.59	1.49	$\beta_2 = -.0048(.01)$ $\beta_5 = .00004(.07)$
		MST	b	.65	1.53	$\beta_1 = -.0008(.07)$ $\beta_2 = -.0022(.06)$ $\beta_9 = .0001(.01)$
	B3	MST	e	.71	1.40	$\varepsilon_2 = .0011(.05)$ $\varepsilon_3 = .0001(.04)$ $\varepsilon_6 = .0001(.01)$ $\varepsilon_{12} = .00004(.05)$
Marg	M1	MSV	b	.48	1.38	$\beta_9 = -.0001(.01)$
		$P(M1\mid M1)$	c	.48	.79	—
		idem after transf.	c	.51	1.65	$\gamma_3 = .0001(.07)$
		$P(M1\mid\overline{M1})$	d	.54	1.32	$\delta_{13} = -.0209(.06)$
	M2	MSV	d	.57	2.02	$\delta_{13} = .0371(.02)$ $\delta_{14} = .0294(.02)$
		MST	d	.70	2.19	$\delta_1 = -.0078(.07)$ $\delta_5 = -.0412(.05)$ $\delta_{13} = .0609(.01)$ $\delta_8 = .0507(.01)$ $\delta_{14} = .0478(.01)$
		$P(M2\mid\overline{M2})$	e	.62	1.68	$\varepsilon_1 = -.0008(.04)$ $\varepsilon_3 = .0001(.07)$ $\varepsilon_{11} = -.0001(.05)$ $\varepsilon_6 = .0002(.04)$
	M3	MST	b	.65	1.41	$\beta_{10} = -.00001(.05)$

ence. For the MST, advertising expenditures on F2 in the current period (X_3) also had an evident positive influence.

Beer

For brand B1, only advertising variables are significant. With respect to the dependent variable MST, it is remarkable that a high advertising ratio of B1 has a negative influence on the MST in the current period (X_5) but a positive influence on the MST in the next period (X_8), which influence is stronger when the general level of advertising is high (X_{14}).

The repeat purchase probability $P(\text{B1} \mid \text{B1})$ is negatively influenced by a high advertising ratio for brand B1 in the current period when the general level of advertising is high (X_{13}); for the transition probability $P(\text{B1} \mid \overline{\text{B1}})$ the opposite holds. So it appears that a high level of advertising for B1 can be harmful with respect to the loyalty of consumers already using B1, while on the other hand it attracts new users. Again this might point to the possibility of an overdose of advertising, especially in relation to the consumers who already use the brand in question. Like F1 in the fopro market, B1 is the biggest brand in the beer market, which does a major part of all advertising (more than 30% over the whole 2 years).

For brand B2 only the two market share variables MSV and MST give significant regression results. In both cases there is a negative influence of price (X_2) and a positive influence of the advertising differences (X_9). The latter implies that as the competing brands spend more on advertising than B2, this is favorable for the market share of B2, which is difficult to explain.

For brand B3 the dependent variable MST is the only one which gives significant regression results. There is a positive influence on MST of price (X_2), a negative influence of current advertising expenditures on B3 (X_3) and a positive influence of advertising expenditures on B3 in the previous period (X_6), which is stronger as the advertising ratio for brand B3 is higher (X_{12}). So a higher price means a bigger market share for B3, which is surprising. It might point to the effect of price as a quality index. The negative effect of current advertising is difficult to explain.

Margarine

For brand M1 of margarine the MSV is higher as the advertising difference (X_9) is bigger, which is what we would expect. The repeat purchase probability $P(\text{M1} \mid \text{M1})$ is positively influenced by current advertising (X_3), while the transition probability $P(\text{M1} \mid \overline{\text{M1}})$ is smaller as the advertising ratio of M1 is higher in periods with a high general level of advertising (X_{13}). Note, that the sign of δ_{13} is opposite to the corresponding case for beer, i.e., for the dependent variable $P(\text{B1} \mid \overline{\text{B1}})$.

For brand M2, the MSV is positively influenced by a high advertising ratio in the current (X_{13}) and previous periods (X_{14}) when the general level of advertising is high. For the MST there is a negative influence of the current advertising ratio when the general level of advertising is low (X_5), but the influence of the advertising ratio in the previous period (X_8) is then positive. When the general level of advertising is high both advertising ratios have opposing influences on the MST (X_{13} and X_{14}).

For the transition probability $P(\text{M2} \mid \overline{\text{M2}})$ there is a positive influence of current (X_3) and previous (X_6) advertising, while current advertising has less effect when the advertising ratio of M2 is high (X_{11}).

For brand M3 the only significant regression result was produced with the MST as dependent variable. Here the greater the advertising difference in the previous period, the higher was the MST (X_{10}).

Summarizing the discussion above, the following can be noted. With respect to the effect of the relative price on brand choice, for 4 brands a significant effect of relative price on market share was found. For brands F1, F2 and B2 that influence was negative, as expected. In the case of brand B3, however, the effect was found to be positive. For none of the margarine brands was a significant effect of price found.

With respect to the effect of advertising on brand choice, it was found that in all the regression results, reported in Table 7.6, there was a significant influence for one or more advertising variables. In most cases the signs of the regression coefficients of the advertising variables were as was expected, which means that advertising expenditures of a particular brand have a positive effect on the probability of choosing that brand, while competing advertising works out negatively. For brands F1 and B1 there seemed to be a possibility of an overdose of advertising; in 2 other cases the direction of the advertising effect was difficult to explain.

Of the 17 regressions reported in Table 7.6, in none of the cases was it found that specification (a) gave the best results. So the relationship with only the advertising ratios as advertising variable seems to be not so satisfactory. When it is assumed, however, that the advertising ratio has a different effect for different general levels of advertising (specification (d), which occurred most often as the best specification, i.e., in 6 out of 17 times) a much better explanation of the brand choice variables is obtained. In contrast with (a) and (d), specifications (b), (c) and (e), which are respectively 4, 4 and 3 times the best specification, use primarily absolute figures for the advertising expenditures, instead of advertising ratios. Because (d), on the one side, as well as (b), (c) and (e), on the other side, are the best specifications in a number of cases, it cannot be said that either absolute advertising figures or advertising ratios generally give the best explanation of brand choice.

7.3.3 *Deal-offers*

Two instruments of the promotion policy of the seller of a product are the possibility of selling the product at a temporarily reduced price and the possibility of offering a free gift to every buyer of his product. These possibilities are called deal-offers, or 'deals'. Here we examine the influence of deals on brand choice.

For the purchases of fopro, beer and margarine used throughout this book, it was recorded for each purchase, whether or not it was a 'deal-purchase'. Therefore figures could be computed giving information about the importance of deal-purchases and the impact of deals on brand choice. Those figures are presented together in Table 7.6.

Table 7.6 Information about deal-purchases

	Fopro	*Beer*	*Marg*
(1) % deal-purchases by number of purchase occasions	2.5	3.0	6.4
(2) % deal-purchases by volume	4.3	4.3	8.3
(3) % deal-purchases in purchases with brand change	5.9	4.6	8.4
(4) % brand changes in deal-purchases	21.2	17.6	34.4
(5) % brand change in all purchases	8.8	11.4	26.3
(6) Total number of purchases	59297	27265	130572

Line (1) gives the percentage of deal-purchases in the total number of purchases. We see that the deal purchases represent only a small minority of purchases: 2 to 3% for fopro and beer and 6% for margarine. When we compare line (2) of Table 7.6 with line (1) we see that the percentages of deal-purchases by volume are bigger than those by number of purchases. For fopro the portion by volume is about 70% higher than the portion by number of purchases, for beer it is 40% higher and for margarine 30%. Because of the very big number of purchases given in line (6), on which the percentages of Table 7.6 are based, even small differences in percentages are significant. So deal-purchases are bigger in size than no-deal purchases; apparently when there is a deal offer more units of the product in question are bought at a time.

From a comparison of the figures of line (3) with those of line (1) it is clear that purchases with brand change, i.e. purchases in which the brand is different from the previous brand, are relatively often purchases with a deal. This can also be concluded from a comparison of line (4) with line (5). For fopro 8.8% of all purchases are purchases with brand change, but 21.2% of all deal-purchases are purchases with a brand change. For beer these figures are respectively 11.4 and 17.6 and for margarine 26.3 and 34.4. So deal-purchases and brand changes go relatively often together.

We conclude with respect to deal-offers that although deal-purchases were relatively unimportant in the fopro, beer and margarine purchases, they do appear to be able to induce brand change.

7.4 BRAND CHOICE AND INTER-PURCHASE TIMES

7.4.1 *Introduction*

In most cases there is a certain time interval, which can vary in length, between subsequent purchases. During this time interval the consumer has experience with the brand he bought at his last purchase and he has also the opportunity of receiving information, from commercial and non-commercial sources, about the brand he purchased most recently and about competing brands. Here we are interested in the influence of the length of the time interval between subsequent purchases, which is called the inter-purchase time, on brand choice.

Thus far these inter-purchase times have not been taken into account. As indicated in section 3.1, in the analyses of the previous chapters the equidistant points on the time scale are the indexing numbers of the purchases. Thus an inter-purchase time of 1 day cannot be distinguished from an inter-purchase time of 2 weeks, for example. So the influence of the inter-purchase times on brand choice was disregarded.

In examining the effect of inter-purchase time here we centre on the question: does the probability of a repeat purchase, i.e. the probability of buying the same brand as the brand bought at the previous purchase occasion, increase or decrease with increasing inter-purchase time or has the inter-purchase time no influence at all? It i sdifficult to hypothesize in advance what, if it exists, the direction of the influence of inter-purchase time will be. On the one side a longer inter-purchase time means more possibilities of forgetting the last purchase and receiving advertising messages, etc. from brands other than the brand last bought. This is an argument in favour of the hypothesis that the probability of a repeat purchase decreases with increasing inter-purchase times. On the other hand a longer inter-purchase time can imply that the brand is longer in use, at least kept longer in the house, which means a greater opportunity of becoming accustomed to it, which is an argument in the other direction.

In the following we will first examine the influence of inter-purchase times for the fopro, beer and margarine brand choice processes. Then these findings will be compared with some results from the literature and finally the possibilities of applying Semi-Markov processes to brand choice, in view of the findings with respect to inter-purchase times,will be discussed.

7.4.2 *The influence of inter-purchase times for fopro, beer and margarine*

For the analysis of the effect of inter-purchase time the brands in the empirical brand choice processes were coded such that in each market 4 brands were distinguished, F1 to F4 for fopro, B1 to B4 for beer, and for margarine M1 to M4. These are the brands discussed in section 2.5. Now for each purchase for which the previous purchase was known, it was noted how many days had elapsed since the previous purchase and whether the brand bought was the same as the brand of the previous

Table 7.7 Repeat purchase fractions for different inter-purchase times

	Fopro		*Beer*		*Marg*	
Inter-purchase time (in days)	*Repeat purchase fraction*	*N = number of purchases considered*	*Repeat purchase fraction*	*N*	*Repeat purchase fraction*	*N*
1	.8853	2275	.7675	400	.6615	14956
2	.9451	5795	.8323	471	.8353	12638
3	.9440	6878	.8377	573	.8503	12224
4	.9337	5596	.8494	571	.8599	8944
5	.8983	3393	.8566	565	.8114	5811
6	.8954	3299	.8945	1242	.7341	9005
7	.9613	17793	.9552	8913	.9673	36647
8	.9083	2398	.9086	1028	.9236	3715
9	.8910	1147	.8349	315	.9173	1426
10	.9026	842	.8584	233	.8892	821
11	.8830	624	.8599	157	.8928	597
12	.8594	384	.8642	162	.8661	433
13	.8637	477	.8768	284	.8418	531
14	.9112	1802	.9278	1496	.9357	1741
15	.8719	367	.8952	248	.8986	296
16	.8439	173	.8537	123	.8516	155
17	.8297	182	.8804	92	.8319	113
18	.8661	127	.8171	82	.8673	113
19	.8571	98	.8833	60	.8854	96
20	.8106	132	.8919	148	.8491	106
21-25	.8406	759	.9005	935	.8869	566
26-30	.8439	346	.8911	606	.8694	245
31-35	.8053	190	.8786	387	.8968	126
>35	.7610	385	.8241	1205	.8579	190
Number of + signs for Cox-Stuart test = T	12		4		6	

purchase occasion. Of course multiple purchases were disregarded, because for these purchases the order in which they were made is not known. The purchases were then classified according to their corresponding inter-purchase times. Then for each inter-purchase time category all purchases in that category were considered and the repeat purchase fraction, i.e. the number of times the brand purchased was the same as the previous brand, divided by the total number of purchases in the inter-purchase time category, was computed. These fractions are estimates of the corresponding repeat purchase probabilities. The repeat purchase fractions computed for fopro, beer and margarine are given in Table 7.7. To verify if the repeat purchase fractions were systematically higher or lower for longer inter-purchase times as compared with short inter-purchase times, Cox and Stuart's sign test for trend in location was used, (see Bradley, 1968, p. 174). According to this test, the repeat purchase fraction for inter-purchase time 1 is compared with the fraction for inter-purchase time 13. When the first fraction is higher than the second we note a plus sign, when the second fraction is higher we note a minus sign. In the same way we compare the repeat purchase fraction of inter-purchase time 2 with the repeat purchase fraction of inter-purchase time 14, etc.

The statistic considered is $T =$ the number of $+$ signs among the 12 signs. If the repeat purchase probability decreases with an increase in the inter-purchase time, we can expect many $+$ signs. If the repeat purchase probability increases with an increase in the inter-purchase times we can expect few $+$ signs. Under the null-hypothesis of no change in the repeat purchase probability with increasing inter-purchase time, in each comparison a $+$ sign and a $-$ sign are equally likely and occur with a probability of $\frac{1}{2}$. So then we expect T to be 6. From a table of the binomial distribution with $p = \frac{1}{2}$ and $n = 12$ it can be derived that for $\alpha = .05$ the null-hypothesis is rejected leftsidely for $T \leq 2$ and rightsidely for $T \geq 10$. In the first case it is concluded that the repeat purchase probability increases with increasing inter-purchase time, in the second case the conclusion is that the repeat purchase probability decreases with increasing inter-purchase time.

The values for T, computed from comparison of the repeat purchase fractions of Table 7.7, are given in the last row of that table. We see that only for fopro is there a significant result, viz., that the repeat purchase probability decreases when inter-purchase time becomes longer. For beer and margarine the null-hypothesis is not rejected.

In the above analysis the purchases of all consumers are included, consumers who buy very frequently as well as consumers who buy with long inter-purchase times. In the computation of the repeat purchase fractions for the short inter-purchase times relatively many purchases originate from households which buy frequently, while in computation of the repeat purchase fractions for the long inter-purchase times infrequently buying households make a relatively large contribution. Now if frequently buying households have repeat purchase probabilities that differ from those of infrequently buying households anyhow, this difference might have caused the significant result for fopro. In this case there is no real change in repeat purchase probability when the inter-purchase time increases; this is only a spurious result of mixing two different groups of consumers. On the other hand it is also conceivable that such a difference hides any real influence exerted by inter-purchase time, viz., when the two effects have opposite directions. This might have happened for beer and margarine.

To remove the possible frequency-of-buying effect, discussed above, the households were divided into 2 groups: frequently buying households, i.e. households which had more than the average number of purchases, and infrequently buying households, i.e., those households which bought less than the average number of times. For each of these sub-groups the analysis was then performed as was done earlier for all households together. For considerations of space we do not give the repeat purchase fractions, but only report the T-values. They are for the frequently buying households 11, 5 and 7 respectively for fopro, beer and margarine, and for the infrequently buying households 10, 6 and 8 respectively. So for both groups of households the repeat purchase probability for fopro decreases significantly with increasing inter-purchase time, but for beer and margarine no significant influence of the inter-purchase time is observed. It appears that the results of Table 7.7 are not spurious outcomes due to the mixing of frequently and infrequently buying households together.

It was concluded that for fopro the repeat purchase probability decreases when inter-purchase time becomes longer. For the 2 other products no influence of the inter-purchase time on the repeat purchase probability was established. From this result it can be tentatively concluded that for fopro the effect of forgetting the last purchase and of being influenced by competing advertising in the time interval between subsequent purchases is stronger than for beer and margarine.

Table 7.7 also gives for each inter-purchase time category the numbers of purchases on which the computed repeat purchase fractions are based. It is obvious that for beer, the less frequently bought product, the longer inter-purchase times are relatively better represented than for fopro and margarine. One remarkable point is that inter-purchase times which are multiples of 7 are relatively strongly represented and that the repeat purchase fractions for these inter-purchase times are high. These are purchases which are made regularly, with time intervals of 1 week, 2 weeks, etc.

7.4.3 *Some results from the literature regarding inter-purchase times*

Here the results of the previous section will be compared with some findings from the literature.

Kuehn (1962), in the study of frozen orange juice purchases referred to in section 3.6.3, also examined the influence of inter-purchase time. He concluded that the probability of a repeat purchase of a certain brand decreases as the inter-purchase time increases. He discusses the effect of inter-purchase times in the framework of the linear learning model (LLM) and states that the slopes of the purchase and rejection operator are greater for high frequency purchases, i.e. purchases with short inter-purchase times, than for low frequency purchases. The LLM was extensively discussed in section 5.3. From Figure 5.1 it can be seen that higher slopes mean that the lines representing the operators are steeper, which implies that the probability of buying the brand in question again, does not change very quickly.

Carman (1966) in the analysis mentioned in section 3.6.3 applied the LLM separately to purchases of 5 different groups of consumers buying toothpaste. The groups were different as to the frequency with which they bought. He found no significant differences among the LLM-parameters of the various groups. This is contrary to the statement of Kuehn just mentioned and indicates that there is no influence of inter-purchase time on repeat purchase probability. D. G. Morrison (1966 b) in an analysis of coffee purchases isolates the frequency-of-buying effect, discussed in the previous section, from the real inter-purchase time effect by studying the relationship between repeat purchase probability and the deviation of inter-purchase time from the average inter-purchase time of a family. He found that inter-purchase time had no significant influence.

It is concluded that no general statements about the effect of inter-purchase time on repeat purchase probability can be made. In two cases, viz., Kuehn's frozen orange juice data and the product fopro, it was found that the repeat purchase probability decreases when inter-purchase time becomes longer. In a number of other cases, however, no influence could be established.

7.4.4 *The application of Semi-Markov processes to brand choice*

Section 3.3.5.1 mentioned R.A. Howard's suggestion (1963) of using the Semi-Markov process for the description of brand choice processes. Because a Semi-Markov process is a Markov process, where the times between successive transitions are allowed to differ, its application to brand choice processes makes it possible to build the real inter-purchase times into the model. Here, where inter-purchase times are treated, some remarks will be made about this possible application.

An essential feature of a Semi-Markov process is that the so-called holding times, which in applications to brand choice processes are the inter-purchase times, are allowed to differ for different types of transition. From the results in the previous sections it cannot be inferred, however, that the relationship between inter-purchase time and type of transition is a very essential characteristic of the brand choice process. On the contrary we observed a significant relationship only for fopro and even then, as can be seen from an inspection of Table 7.7, the differences in repeat purchase probabilities for different inter-purchase times are not very striking. So it seems that there is hardly any need for the generality offered by this feature of the Semi-Markov process.

There are indications that a source of differences more important in inter-purchase times than the type of transition is constituted by the differences between households. The coefficient of variation of the inter-purchase times for individual households and for all inter-purchase times of all households together was computed. For fopro .50 was found as the median value for the individual coefficients of variation over all households, while the coefficient of variation over all inter-purchase times of all households was 1.03. For beer the corresponding figures are .91 and 1.45 and for margarine .48 and .86. (For all 3 products the means of the individual coefficients of variation are near to the median.) So it appears that the variation within households is small relative to the variation among households. Therefore when a model is to be built which describes the brand choice process in real time, such a model should take into account these differences in inter-purchase times among households. It is a question if the Semi-Markov model, with the assumption that the distribution of inter-purchase times for a certain transition is equal for all households, can then give satisfactory results.

In a Semi-Markov process, also called a Markov Renewal Process, it is assumed that successive holding times, which are equivalent to inter-purchase times here, are mutually independent. To examine if this was true for the empirical brand choice processes here, correlation coefficients between successive inter-purchase times were computed. When all inter-purchase times of all consumers were considered, for fopro an estimated correlation coefficient between successive inter-purchase times of .396 ($n = 54540$) was found. For beer $r = .272$ was found ($n = 23104$) and for margarine $r = .485$ ($n = 122416$). So successive inter-purchase times are not independent of each other, but long inter-purchase times are relatively often followed by long inter-purchase times and the same holds for short inter-purchase times. This suggests that in the purchase history of a household, periods of purchases with long inter-purchase times alternate with periods during which purchases are made with short intervals between subsequent purchases. This may have to do with the influence of season or with the fact that in different periods, the product is bought at different outlets (e.g., an ordinary shop versus an ambulant shop).

For the above figures, the inter-purchase times of all households together produced one correlation coefficient. It is also possible to consider the serial dependence of inter-purchase times within individual households. In this way it was found that (with $\alpha = .05$) for 164 of the 672 fopro households (24%) there was a significant positive correlation between successive inter-purchase times. For beer this was the case for 124 of the 627 households (20%) and for margarine for 334 of 1059 households (32%). So the serial dependence of inter-purchase times is considerable. The fact that this has to be neglected when the brand choice process is described by a Semi-Markov model constitutes a drawback to this process.

These few remarks indicate that the Semi-Markov model, although appealing at first glance because of its possibility of taking into account real inter-purchase times, is not without further preface a suitable model for the brand choice process.

8. The brand choice process and its relations to household variables

8.1 INTRODUCTION

In this chapter we will examine the connections between the characteristics of brand choice processes of individual households and certain properties of these households. For this purpose we will study the relationship between a number of brand choice variables – characteristics of the brand choice process – and a number of household variables. These household variables are divided into socio-economic variables and purchase variables, other than brand choice variables. The analysis is carried out for the empirical brand choice processes of fopro, beer and margarine, used throughout this book.

In section 8.2 the variables to be used are defined and in 8.3 the observed relationships are reported.

One question, closely connected with the question of relationships between brand choice behavior and household variables is, whether or not there exists a kind of general brand choice behavior for a household, i.e. whether or not a household's brand choice processes are more or less similar with respect to different products. If this is the case, a household with few brands for one product will also have few different brands for another product. It can then be said that brand choice behavior is transitive over products. To see if this applies to the products fopro, beer and margarine, section 8.4 examines if the corresponding brand choice variables of the same households for different products are narrowly related.

Section 8.5 gives some results from the literature, with which the findings obtained in sections 8.3 and 8.4 are compared, while in the last section of this chapter some conclusions are presented.

8.2 THE VARIABLES USED

In this section the variables to be used further in this chapter are defined. These variables were partly defined earlier, i.e., in chapters 6 and 7.

The variables are divided into brand choice variables, socio-economic variables and purchase variables other than brand choice variables. All the variables refer to purchase histories of individual households during the whole 2-year period observed.

8.2.1 *Brand choice variables*

1. SFAVBR : Share of favorite brand by volume of purchases.
2. SFAVBRN : Share of favorite brand by number of purchase occasions.
3. NUMBR : Number of different brands bought.
4. RANGPLZ: This variable, originating from the poolsize approach discussed in chapter 6, was exactly defined in section 6.4.1. It refers to the extent to which the level of brand switching activity of a household has varied in the 2 years observed.
5. MEANPLZ: This variable is also defined in section 6.4.1. It refers to the general level of brand switching activity of a household, as measured by the number of different brands in the last ten purchases.
6. PRIVBR : Share of private brands by volume of purchases.
7. FIRBR : A nominal variable indicating which was the favorite brand, i.e. the brand most purchased, according to volume of purchases, in the two years observed. For all products 5 categories were distinguished and for each household it was established for each product to which of the 5 categories the most favored brand of the household belonged.
   For fopro these brand codes are as follows:
   1 = F1
   2 = F2
   3 = A private brand of a chain
   4 = A private brand of a voluntary chain
   5 = Other
   For beer:
   1 = B1
   2 = B2
   3 = B3
   4 = A private brand
   5 = Other
   For margarine:
   1 = M1
   2 = M2
   3 = M3

4 = A private brand
5 = Other

So the first brands are always the major national brands, as discussed in section 2.5, while private brands are taken as a separate category. To give the 5 brand codes somewhat sharper profiles the following can be remarked with respect to their prices. For fopro the national brands F1 and F2 (brand codes 1 and 2) are higher in price than brand codes 3, 4 and 5. For beer the same applies to the national brands B1, B2 and B3 as compared with the private brands (brand code 4) and brand code 5. For margarine the national brands M2 and M3 are more expensive than brand codes 4 and 5. Brand M1 (coded as 1), which is also a major national brand is cheaper, its price being of the same order of magnitude as that of brand codes 4 and 5.

In section 7.2.3, where the relationship between brand choice and shop choice was studied, a factor analysis was carried out to arrive at more fundamental brand choice, respectively shop choice, factors than the variables directly observed. Now it is possible to calculate for individual households the factor scores for the brand choice factors found in section 7.2.5 (see, for example, D.F. Morrison, 1967, section 8.10). Such factor scores can then be further used as ordinary variables for which the relationship with other variables is studied. As such some of these factors could have been included as brand choice variables in the list just given. In fact we proceeded in this way and studied the relationships of these factors with socio-economic and purchase variables. It was found, however, that in no case did these factors give a better explanation than the original variables from which they were derived. For this reason, these factors are omitted here and only those results regarding the original brand choice variables are given.

8.2.2 *Socio-economic variables*

8. SIZTOW : Size of town. Coding according to number of inhabitants of the town or village in which the household lives:
   1 = > 500000 inhabitants;
   2 = 80000 – 500000 inhabitants;
   3 = 30000 – 80000 inhabitants;
   4 = < 30000 inhabitants;
   5 = countryside.
9. SOCCL : Social class. Coding on a 4-points scale following the Attwood classification:

1 = high social class;
4 = low social class.

10. FAMSIZ : Family size: number of persons in the household.

11. AGEHOU: Age of housewife, rounded off in multiples of 5 years.

The next 4 variables – 12 to 15 – are scores on attitude scales regarding work in the household and buying behavior. These scores were obtained by Attwood by interviewing the housewives of the households in the panel. They are scores on 4-points scales.

The variables are:

12. REGHOU : Regularity in the household:
1 = very regular;
4 = not regular.

13. TRAD : Traditionalism towards work in the household:
1 = very traditional;
4 = not traditional.

14. BRTIED : Brand-tiedness, the extent to which the housewife feels tied to the brands she uses:
1 = very much brand-tied;
4 = not brand-tied.

15. PRICON : Price consciousness:
1 = very price-conscious;
4 = not price-conscious.

16. DISTR : District, region in which the household lives:
1 = agglomerations of the 3 big cities Amsterdam, Rotterdam and The Hague;
2 = the west of the Netherlands, except the regions mentioned in 1;
3 = the north of the country;
4 = the east of the country;
5 = the south of the country.

17. CHILD : Children, the presence of children:
0 = no children < 15 years;
1 = child(ren) between 5 and 14 years, but not under 5 years;
2 = Child(ren) < 5 years.

18. REFRI : Presence of a refrigerator:
0 = no refrigerator present;
1 = refrigerator present;
This variable was included because a refrigerator can be

used to preserve fopro, beer and margarine, and this can influence the buying process.

19. TELEV : Presence of a television set:
0 = no television set present;
1 = television set present.
The variable TELEV was included, because tv-commercials are a major form of advertising for the 3 products under study.

8.2.3 *Purchase variables other than brand choice variables*

20. PRICE : Average price paid per standard weight unit.
21. DEAL : Share of deal purchases in total purchases, by volume.
22. SELSUP : Share in total purchases by volume of purchases in self-service shops and supermarkets.
23. CHAINSH: Share in total purchases by volume of purchases in shops belonging to chains and voluntary chains.
24. NUMPU : Number of purchases, the number of times a purchase of the product was made (= purchase occasions).
25. VOLPU : Volume of purchases, the total volume of the purchases of the product.
26. VOLPP : Volume per purchase, the average volume of the product bought per purchase occasion.
27. VARVOL : Coefficient of variation of the volume bought per purchase occasion.
28. VARIPT : Coefficient of variation of inter-purchase time.
29. NUMSIZ : Number of different package sizes of the product bought. For fopro 5 standard sizes in which the product is sold were distinguished; for beer there are 3 such standard sizes. Practically all margarine purchases were of the 250 gram package size, so for this product the variable NUMSIZ is not relevant.
30. NUMPUR: Number of different purchasers in the household. This variable is only defined for beer, because for this product it was recorded for each purchase by whom this purchase was made: Housewife, husband, son, daughter or someone else.
31. NUMVAR: Number of different varieties. This variable is also only defined for beer. In total 4 varieties were distinguished and for each purchase the variety was recorded.

Note that among the 'purchase variables other than brand choice variables' no variables regarding store choice are included here, because the relationship between brand choice and store choice was extensively treated in section 7.2.

For every household the value of each of these 31 variables (as far as the variables are defined of course) was determined. It was then examined per product if there were relationships between the brand choice variables on the one hand and the socio-economic and purchase variables on the other hand. The results of this are reported in section 8.3. Next it was examined if in general a household behaved in the same way – with respect to brand choice – towards different products. For this purpose the relationships between corresponding variables (e.g. NUMBR) of the same households for different products were studied. The results of this analysis are presented in section 8.4.

8.3 THE RELATIONSHIP OF BRAND CHOICE VARIABLES TO SOCIO-ECONOMIC AND PURCHASE VARIABLES

8.3.1 *The method used*

The aim was to study the relationships between the brand choice variables, i.e. variables 1 to 7 as defined in section 8.2, and
a. the socio-economic variables (variables 8 to 19)
b. the purchase variables (variables 20 to 31).
Variables 1 to 6, 10, 11 and 20 to 31 are interval-scaled variables. Variables 8, 9 and 12 to 15 have an ordinal scale and variables 7 and 16 to 19 are nominal variables.

To study the mutual relationship between interval-scaled variables the ordinary (Pearson-)correlation coefficient can be used. For the relationship between an ordinal variable and an interval-scaled variable or another ordinal variable, computation of the Spearman rank correlation coefficient or the Kendall rank correlation coefficient is appropriate. This implies, however, that for the variables concerned all observations must be replaced by their rank numbers. For fopro and beer this means that more than 600 numbers must be rank ordered for each variable and for margarine more than 1000. This is a very cumbersome and time-consuming affair, even when a computer is used. Because – as will be seen presently – the effect of taking original values instead of ranking numbers does not seem great, it was decided to consider the ordinal variables as if they were interval-scaled and to compute ordinary correlation coefficients for these variables.

When an ordinary correlation coefficient is used, the variables should be normally distributed, to enable inferences to be made. If this is not the case a distribution-free technique, such as the method of rank correlation, with the computational disadvantages just mentioned, is more appropriate. This normality condition was not explicitly verified for the data used here, but the effect of a possible non-normality and the effect of the policy of treating ordinal variables as if they were interval-scaled were checked in the following way. For the product fopro we computed for the relationship between brand choice variables 1, 2, 3, 6, on the one hand, and variables 8 to 15 and 20 to 29, on the other hand, the ordinary correlation coefficients as well as the Spearman rank correlation coefficients. (The interval-scaled brand choice variables 4 and 5 were left out, because their values were not yet available at the moment this check was carried out.) It appeared that for the 72 relationships studied in this way, in only 7 cases did the values of the ordinary and Spearman rank correlation coefficients lead to a different conclusion with respect to the significance of the correlation. (In both cases an α of .05, two-sided, was applied.) In these 7 cases the ordinary correlation coefficient led to the conclusion of no significant correlation, while the Spearman coefficient concluded a significant correlation or vice versa. In all these 7 cases the signs of both the coefficients were equal, for 6 of them the difference between the values of the 2 coefficients was smaller than .04, and in one case it was .11. So the results of this comparison suggest that the use of the ordinary correlation coefficient instead of the theoretically more appropriate rank correlation coefficient does not lead to divergent results.

To study the relationship between nominal and ordinal or interval-scaled variables, the analysis of variance technique leading to the F-statistic was used, and for tests on the mutual dependence of nominal variables contingency tables were made for which χ^2-statistics were computed.

8.3.2 *Brand choice variables and socio-economic variables*

In Table 8.1 the results of the tests on dependence between brand choice and socio-economic variables are given.

To the extent that correlation coefficients were computed to examine the relationship between two variables, the resulting coefficients are given whenever they were found to be significant when a two-sided confidence region of 5% in total was applied. Because of the big sample sizes, rather small correlation coefficients are still significant. For fopro all correlation coefficients with an absolute value greater than .076 are significant, for beer this lower boundary is .078 and for margarine .060. It should be realized that when a correlation coefficient between two variables has the

Table 8.1 Relationships between brand choice variables and socio-economic variables (for significant relationships the value of the correlation coefficient or an × is shown)

Product	*Socio-econ. variable* / *Brand choice variable*	SFAVBR	SFAVBRN	NUMBR	RANGPLZ	MEANPLZ	PRIVBR	FIRBR
Fopro ($n = 672$)	SIZTOW						−.11	×
	SOCCL						.09	×
	FAMSIZ							×
	AGEHOU				−.09		−.08	×
	REGHOU							
	TRAD							
	BRTIED	−.09	−.08	.13	.10	.11		×
	PRICON	.17	.17	−.23	−.22	−.20	−.16	×
	DISTR			×		×	×	×
	CHILD				×		×	
	REFRI						×	×
	TELEV							
Beer ($n = 627$)	SIZTOW				−.08			×
	SOCCL							
	FAMSIZ	−.09	−.10	.12	.09	.15		
	AGEHOU							
	REGHOU							
	TRAD							
	BRTIED							
	PRICON	.15	.15	−.16	−.16	−.14	−.10	×
	DISTR			×		×		×
	CHILD			×	×			
	REFRI							
	TELEV							
Marg ($n = 1059$)	SIZTOW	.15	.15	−.17	−.15	−.13		×
	SOCCL						.10	×
	FAMSIZ	.08				−.06	.06	×
	AGEHOU							×
	REGHOU							
	TRAD	−.09	−.09	.13	.11	.10	.06	×
	BRTIED	−.07	−.08	.19	.17	.16		
	PRICON	.14	.13	−.20	−.18	−.15	−.11	
	DISTR	×		×	×	×		×
	CHILD			×				×
	REFRI	×		×	×	×		×
	TELEV	×		×	×	×	×	

order of magnitude of these lower bounds, one variable explains only a very small percentage of the variance of the other, even though a significant relationship might exist. A quick inspection of Table 8.1 makes it clear that there are not many significant correlations between brand choice variables and socio-economic variables and that, if the correlations are significant, they are generally rather weak.

The significance of relationships between nominal variables mutually and between nominal variables and other variables is indicated as follows. An × means: a significant relationship; when there is no ×, the relationship was found to be not significant. For the chi-square tests and Fisher test used here, a 5% confidence level was also applied. Of the analyses of variance, for the significant cases the F-value found is given and the rank order of the classes of the nominal variables according to increasing values of the class-averages of the corresponding interval-scaled variables. Of the chi-square test results, for the significant cases the χ^2-value is reported and the most striking features of the contingency tables. This is done at the appropriate places in the following text, where the results with respect to the various socio-economic variables are discussed.

SIZTOW. For fopro, households in smaller towns buy fewer private brands. (This discussion of results often contains statements of the type just made; when it is said that households with a certain characteristic do something more or have something less in relation to a certain variable, this should always be read as: 'compared with households without this characteristic'. For reasons of conciseness the latter qualification is usually omitted.) For beer the correlation with RANGPLZ means that households in smaller towns have somewhat less variation in their level of brand switching activity. For the brand choice behavior in relation to margarine SIZTOW appears to be a more important variable than it does for fopro and beer. Households in smaller towns have higher shares of their favorite brand, less different brands, less variation in their level of brand switching activity and a lower general level of brand switching activity.

With respect to the relationship between SIZTOW and FIRBR, the rank order of the 5 brand codes of FIRBR by increasing values of SIZTOW is for fopro 3, 2, 5, 1, 4 ($F = 18.86$), for beer 1, 5, 4, 3, 2 ($F = 8.14$) and for margarine 3, 2, 4, 5, 1 ($F = 7.09$). So for fopro private brands of a chain (brand code 3) are relatively often the favorite brand in big towns, while private brands of voluntary chains are relatively often the favorite brand in smaller places. For beer and margarine the private brands (brand code 4) take an intermediate position with respect to size of place.

SOCCL. For fopro, households of lower social class buy more private brands. The same applies to margarine. For beer SOCCL has no influence on any brand choice variable.

With respect to FIRBR, the rank order of the 5 brand codes is 2, 1, 3, 5, 4 ($F = 5.39$) for fopro, and 2, 3, 1, 5, 4 ($F = 7.90$) for margarine. This means that for both products the major national brands (codes 1 and 2 for fopro, codes 1, 2 and 3 for margarine) are relatively often the favorite brand of households of higher social class while, at least for margarine, the private brands (code 4) occur more in households of lower social class.

FAMSIZ. For fopro there are no significant correlations. For beer bigger families have more different brands, a higher general level of brand switching activity, more variation in this level, and their shares of the favorite brand are smaller. This greater variability of bigger families may be due to the fact that in these households beer purchases are made by more different persons (see also the results with respect to NUMPUR in section 8.3.3). For margarine there is some evidence of the opposite: bigger families have a somewhat bigger share of the favorite brand and a lower general level of brand switching activity. Also bigger families buy more private brands.

Fopro and margarine have significant relationships between FAMSIZ and FIRBR. For fopro the rank order of the brand codes is 2, 1, 3, 4, 5 ($F = 4.14$) and for margarine 2, 3, 5, 4, 1 ($F = 12.13$). So for fopro the national brands F1 and F2 are relatively often the first brand in small households. For margarine this applies to brands M2 and M3 while M1 occurs more in big households.

AGEHOU. For fopro, households with older housewives have somewhat less variation in the level of brand switching activity and buy fewer private brands. For beer and margarine there are no significant correlations. With respect to FIRBR the rank order of the brand codes is 4, 5, 3, 1, 2 ($F = 6.01$) for fopro and 4, 5, 1, 3, 2 ($F = 7.13$) for margarine. So it appears that the more expensive brands – F1 and F2 for fopro, M2 and M3 for margarine – which were already found to occur more often in smaller households, also occur more often in households with older housewives.

REGHOU. For none of the 3 products did this attitude variable have a significant relationship with any of the brand choice variables.

TRAD. This variable is successful to some extent for margarine. For this product: the less traditional the attitude of the housewife toward work in the household, the larger the number of different brands, the higher the general level of brand switching activity, the greater the variation in this level, the bigger the share of private brands and the smaller the share of the favorite brand.

With respect to FIRBR, the rank order of the brand codes is 5, 1, 4, 2, 3 ($F = 3.76$) which means that brands M2 and M3 occur more in the least traditional households.

BRTIED. This variable, intended to measure the extent to which a housewife feels tied to her brands, has a relationship with brand choice variables for fopro and margarine. For both products it holds that, as the housewife is less brand-tied, the household has more different brands, a higher general level of brand switching activity, more variation in this level and a smaller share of the favorite brand. At first glance this may appear somewhat trivial, because the variable BRTIED just measures the extent to which one sticks to a brand, but it should be noticed that the BRTIED-score of a household refers to stated brand-tiedness in general, while here we are concerned with demonstrated brand loyalty for special products. For fopro there is, moreover, a relation with FIRBR, with 2, 4, 5, 1, 3 ($F = 3.90$) as rank order of the 5 brand codes, implying that households with F2 as the favorite brand are the most brand-tied.

PRICON. This variable is the most successful of all the socio-economic variables used here to explain differences in brand choice behavior. There are significant correlations with all interval-scaled brand choice variables of all 3 products and the computed correlation coefficients are high, compared with those for the other socio-economic variables.

The picture for the 3 products is very similar: As housewives are less price-conscious, households have fewer brands, a lower general level of brand switching activity, less variation in this level, fewer private brands and a bigger share of the favorite brand. With respect to FIRBR, the rank order of brand codes is 3, 4, 5, 2, 1 ($F = 6.45$) for fopro, and 5, 4, 1, 3, 2 ($F = 2.66$) for beer. So it appears that households with the more expensive national brands as their favorite brand (codes 1 and 2 for fopro, codes 1, 2 and 3 for beer) are less price-conscious, which is quite understandable.

After this discussion of the results for the ordinal and interval-scaled socio-economic variables we now arrive at the nominal variables DISTR, CHILD, REFRI and TELEV. As mentioned earlier, the relationships with the interval-scaled brand choice variables SFAVBR to PRIVBR were examined by means of analysis of variance, while for the relationship with FIRBR a chi-square test was used.

DISTR. For all 3 products there are significant relationships with NUMBR and MEANPLZ. For fopro the rank order of the 5 districts by increasing NUMBR is: 2, 5, 3, 4, 1 ($F = 3.19$), for beer 2, 1, 3, 4, 5 ($F = 2.61$) and for margarine 2, 3, 5, 4, 1 ($F = 4.90$). So for fopro and margarine, households in districts 2, 3 and 5 have the smallest numbers of brands and households in districts 4 and 1 have more brands. For beer the situation with respect to districts 1 and 5 is different. In contrast with the 2 other products, households in district 1 have relatively few and households in district 5 have relatively many brands.

By increasing MEANPLZ the rank order of the 5 districts is: 2, 3, 5, 4, 1 ($F = 2.45$) for fopro, 1, 2, 4, 3, 5 ($F = 2.43$) for beer and 3, 2, 5, 4, 1 ($F = 3.65$) for margarine, which rank orders are very similar to those corresponding with NUMBR.

In addition to the relationships with NUMBR and MEANPLZ for fopro there is a significant relationship of DISTR with PRIVBR. Here the rank order of the 5 districts is 2, 3, 5, 1, 4 ($F = 3.69$), which is rather similar to the rank order by NUMBR, just given. For margarine there are significant relationships of DISTR with SFAVBR and RANGPLZ, with as corresponding rank order of the 5 districts 1, 4, 5, 3, 2 ($F = 5.49$) and 2, 3, 5, 4, 1 ($F = 4.08$) respectively. The first is the inverse of the rank order by NUMBR, which is quite understandable, while the rank order by RANGPLZ is very similar to that by MEANPLZ. In the chi-square tests on the relationship between DISTR and FIRBR, the following χ^2_{16} values were found: 90.37 for fopro, 250.27 for beer and 73.76 for margarine, which all are clearly significant. For fopro, in district 1 brand code 4 occurs relatively infrequently as the favorite brand and brand codes 2 and 3 relatively often. In districts 2 and 3 brand code 1 occurs relatively often however. For beer in district 1 brand code 1 and in district 4 brand code 3 is relatively often the favorite brand. For margarine it was found that in districts 1 and 2 brand code 3 is relatively often the favorite brand and in district 5 this holds for brand code 1.

One general conclusion from the results with respect to DISTR is that, apparently, there are regional differences in brand choice behavior. Because of the modest F-values these differences are not very great; perhaps the differences with respect to the favorite brand (FIRBR) are the most striking ones. These regional differences in brand choice behavior may be due to differences in consumers (taste, buying habits, etc) and to differences in distribution patterns.

CHILD. For fopro the rank order of the 3 categories of the variable CHILD by RANGPLZ and PRIVBR is for both cases 0, 1, 2 ($F = 3.31$, resp. 4.42). So households without children (category 0) have the smallest variation in the level of brand switching activity and buy the least private brands. Households with little children (cat. 2) take the opposite position, while households with bigger children (cat. 1) fall in between.

For beer the rank order of the 3 CHILD-categories is 0, 1, 2 ($F = 3.39$) by NUMBR and 0, 2, 1 ($F = 3.19$) by RANGPLZ. For margarine the rank order is 0, 1, 2 ($F = 3.03$) for NUMBR. So we have the interesting result that households with children, especially with little children, show relatively a lot of brand switching. One possible cause might be that certain promotional activities, which induce brand changes, have a special appeal to children.

For margarine there is a significant relationship between CHILD and FIRBR ($\chi_8^2 = 37.48$). It appears that households with little children more often have brand code 1 as the favorite brand, households with bigger children brand code 5 and households without children brand code 2. Of course housewives without little children are generally older and this result compares quite well with that found for AGEHOU.

REFRI. For fopro, households with a refrigerator buy more private brands ($F = 11.48$). For beer there is no significant relationship with REFRI. For margarine, households with a refrigerator have a smaller share of the favorite brand ($F = 22.46$), a bigger number of brands ($F = 23.73$), a higher general level of brand switching activity ($F = 25.89$) and more variation in this level ($F = 20.30$). For fopro and margarine there are also significant relationships with FIRBR, with χ_4^2-values of 37.08 and 13.00 respectively. For fopro, households with brand codes 2 and 3 have relatively often a refrigerator, and for margarine this applies to households with brand code 4.

It is striking that for margarine REFRI is a more important variable with respect to brand choice behavior than for the other products. In this connection it may be mentioned that margarine is the only product of the 3 for which a refrigerator or another cooling facility is absolutely needed to preserve the product for any length of time. So households with a refrigerator are more flexible with respect to their margarine purchases, which may explain the greater variation in brand choice observed for these households.

TELEV. For fopro and beer there are no significant relationships between TELEV and brand choice variables. For margarine, households with a television set have a smaller share of the favorite brand by volume ($F = 9.34$), more different brands ($F = 9.82$), a higher general level of brand switching activity ($F = 7.42$), a greater variation in this level ($F = 7.63$) and these households buy more private brands ($F = 4.22$). So as for the REFRI variable, TELEV has evident relationships with brand choice variables for margarine but not for fopro and beer. This may be connected with the relative tv-advertising activities of the 3 products in the years 1967 and 1968, to which the data refer. From the figures of the 'Bureau voor Budgetten Controle', mentioned in section 2.6, it can be derived that total expenditures on the broadcasting of tv-commercials for margarine in the years 1967 and 1968 are about 4 times as high as the corresponding expenditures for fopro and more than twice as high as for beer. In fact in 1967 there were no beer tv-commercials at all, as a consequence of the exclusion of alcoholic beverages from tv-advertising in that year. Taken over 1968 alone, the expenditures on margarine tv-advertising were twice as high as those for beer and 5 times as high as those for fopro. Because expenditures on broadcasting of tv-commercials are roughly proportional to the total television time obtained, it can be concluded that tv-advertising for margarine was much more intensive than for fopro and beer. This may contribute to the explanation of the bigger influence of TELEV for margarine.

There is also the point that the possession of a refrigerator or a television set can be used as a general indication of a household's progressiveness. If this plays a role here, in the sense that it influences brand choice behavior, it is curious that this is only the case for margarine and not, or almost not, for the other 2 products.

Here a brief summary of the results regarding the relationships between socio-economic and brand choice variables is given.

With respect to the effect of size of town, it can be remarked that households in smaller towns generally show less brand switching. This is most evident for margarine.

Social class has not much influence on brand choice behavior. The only effect worth mentioning is that for fopro and margarine the major national brands occur relatively often in households of higher social class and private brands in households of lower social class.

The larger the number of persons in the household, the more brand switching occurs for beer. For margarine there is some evidence of the opposite; for fopro there is no influence of family size.

Only for fopro is there an influence of the age of the housewife, viz., in the sense that older housewives show less brand switching. It can be further mentioned that for fopro and margarine the more expensive national brands (F1 and F2, respectively M2 and M3) occur relatively often in smaller households with older housewives.

With respect to the attitude variables, the variable measuring regularity in the household showed no influence at all, traditionalism had some influence for margarine, brand-tiedness for fopro and margarine, while price-consciousness – the most successful socio-economic variable of all those included – had an evident effect for all 3 products. For these variables it was always so, that the less traditional a housewife, the less brand-tied, and the more price-conscious, the more inclined to brand switching was the behavior of the household.

From the results regarding region it is clear that for all 3 products regional differences in brand choice behavior exist. Further, it appeared that households with children, especially those with little children, showed relatively a lot of brand switching.

Only for margarine did the possession of a refrigerator and a television set appear to bear evident relationships to brand choice behavior.

It should be noted that although the statements just made are all based on relationships significant at the 5%-confidence level, generally the contribution to the explanation of the variation of brand choice variables by the respective socio-economic variables is modest.

8.3.3 *Brand choice variables and purchase variables*

Here the results of the tests on the relationships between brand choice variables and other purchase variables are presented. The estimated significant correlation coefficients and the indications with respect to the results of the analyses of variance are given in Table 8.2, which should be read in the same way as Table 8.1. Note that all purchase variables are at least interval-scaled, so that no chi-square tests were performed. As seen in section 8.2, not all purchase variables are defined for all 3 products. This explains why the lists of purchase variables for the 3 products in Table 8.2 are not identical.

A first inspection of this table and a comparison with Table 8.1 show that statistically significant relationships occur between purchase variables and brand choice variables much more often than they do between socio-economic and brand choice variables, while the correlation coefficients of Table 8.2 are higher generally.

The following discusses these results in some detail.

PRICE. For fopro and beer the pictures are very similar: higher prices are paid by households having a large share of the favorite brand, few different brands, a low general level of brand switching activity, little variation in this level, and by households which infrequently buy private brands. So for fopro and beer households which are brand-loyal, pay higher prices. For margarine this seems also to be the case, although the evidence here is somewhat less: there are significant correlations only with SFAVBR, SFAVBRN and PRIVBR, and these correlation coefficients are somewhat smaller than for the other products. A weaker relationship between price and brand choice variables for margarine was observed earlier in section 7.3.2.2, where we examined the influence of the marketing variable price.

With FIRBR there are strong relationships, indicating that the brand codes evidently differ in price. For fopro the rank order of the 5 brand codes is 3, 5, 4, 1, 2 ($F = 102.03$), for beer 4, 5, 3, 1, 2 ($F = 170.10$) and for margarine 4, 1, 5, 3, 2 ($F = 71.25$). These differences between the prices of the various brand codes were already referred to in section 8.2.1, when FIRBR was defined.

DEAL. For fopro and margarine Table 8.2 shows that households with many deal purchases have smaller shares of the favorite brand, more different brands, a higher general level of brand switching activity and more variation in that level. For beer there is only one significant result, which also applies to fopro, viz., that households with many deal purchases also buy private brands relatively often.

With respect to FIRBR the rank orders of the 5 brand codes are: 2, 1, 3, 4, 5 ($F = 6.41$) for fopro, 2, 3, 1, 4, 5 ($F = 12.54$) for beer and 1, 4, 5, 2, 3 ($F = 3.82$) for margarine. So for fopro and beer, households with the major national brands as favorite brand (brand codes 1 and 2 for fopro, brand codes 1, 2, 3 for beer) have the least number of deal purchases. This also applies to the first national brand (brand code 1) of margarine but not for the others (brand codes 2 and 3).

SELSUP. For all 3 products, households making bigger shares of their purchases in self-service shops or supermarkets, have smaller shares of the favorite brand, more different brands, a higher general level of brand switching activity, more variation in this level and buy more private brands. There are further clear relationships with FIRBR. The rank order of the 5 brand codes is 2, 1, 5, 4, 3 ($F = 111.83$) for fopro, 2, 3, 1, 5, 4 ($F = 21.49$) for beer and 3, 5, 2, 1, 4 ($F = 48.66$) for margarine. So for fopro and beer, households with a major national brand as the favorite one make a smaller part of their purchases in self-service shops or supermarkets.

CHAINSH. For this variable, which refers to the extent to which households make their purchases in shops belonging to chains (voluntary or not), the picture with respect to the correlation coefficients is very similar to that for SELSUP, with the exception that the correlations for margarine are weaker here. With respect to FIRBR,

Table 8.2 Relationship between brand choice variables and other purchase variables (For significant relationships the value of the correlation coefficient or an × is indicated)

Product	*Purchase variable* \ *Brand choice variable*	SFAVBR	SFAVBRN	NUMBR	RANGPLZ	MEANPLZ	PRIVBR	FIRBR
Fopro	PRICE	.24	.28	−.33	−.32	−.25	−.46	×
($n = 672$)	DEAL	−.19	−.17	.21	.20	.22	.12	×
	SELSUP	−.22	−.25	.28	.26	.30	.65	×
	CHAINSH	−.22	−.25	.27	.26	.28	.88	×
	NUMPU	.17	.16	−.11	−.09	−.19	−.19	×
	VOLPU	.08				−.20		×
	VOLPP	−.09	−.12	.12	.12	.12	.21	×
	VARVOL	−.25	−.17	.20	.21	.19	.11	×
	VARIPT	−.20	−.16	.13	.12	.23		×
	NUMSIZ	−.31	−.31	.50	.44	.44	.15	×
Beer	PRICE	.27	.27	−.26	−.25	−.20	−.59	×
($n = 627$)	DEAL						.21	×
	SELSUP	−.16	−.14	.12	.16	.08	.42	×
	CHAINSH	−.14	−.13	.11	.14	.10	.53	×
	NUMPU	.12	.11		.20	−.26		×
	VOLPU	.13	.09	.08	.12	−.16		×
	VOLPP	.11						×
	VARVOL	−.12	−.12	.15	.12	.14		
	VARIPT	−.18	−.17	.10		.27		
	NUMSIZ	−.29	−.30	.45	.42	.31	.12	
	NUMPUR	−.09	−.10	.14	.11	.10		
	NUMVAR	−.14	−.17	.17	.13	.20		
Marg	PRICE	−.10	−.12				−.26	×
($n = 1059$)	DEAL	−.16	−.16	.26	.24	.25		×
	SELSUP	−.10	−.09	.14	.12	.14	.40	×
	CHAINSH				.08	.08	.60	×
	NUMPU	−.15	−.16	.08			−.16	×
	VOLPU	.10	.08			−.13	.06	×
	VOLPP	.19	.16	−.11	−.09	−.13	.13	×
	VARVOL	−.26	−.29	.32	.32	.36		
	VARIPT	−.48	−.51	.41	.36	.56		×

the rank order of the 5 brand codes is: 5, 1, 2, 4, 3 ($F = 370.75$) for fopro, 2, 1, 3, 5, 4 ($F = 39.67$) for beer and 3, 5, 1, 2, 4 ($F = 138.44$) for margarine.

These rank orders are also very similar to those for SELSUP. The fact that households having a private brand as the favorite one buy more in chain-shops is particularly understandable because these shops carry private brands. The close agreement between the results of SELSUP and CHAINSH is also not surprising, because chains have relatively more self-service shops and supermarkets than others.

NUMPU. For fopro, households which made many purchases (NUMPU refers to purchase occasions) in the two years observed, have a bigger share of the favorite brand, fewer brands, a lower general level of brand switching activity and less variation in this level. For beer there seems to be a similar tendency, at least as appears from the correlation coefficients for SFAVBR, SFAVBRN and MEANPLZ. In contrast with the 2 other products, for margarine households making many purchases show relatively a lot of brand switching. These households have a smaller share of the favorite brand and more different brands. With respect to PRIVBR the results for fopro and margarine are in agreement: households making many purchases buy fewer private brands. With respect to FIRBR the rank order of the 5 brand codes is 3, 4, 2, 5, 1 ($F = 6.55$) for fopro, 2, 5, 1, 4, 3 ($F = 3.64$) for beer, and 4, 2, 1, 5, 3 ($F = 5.76$) for margarine.

VOLPU. For fopro and beer the correlation coefficients with the brand choice variables, as far as they are significant, have the same signs as the corresponding coefficients for NUMPU. For margarine, however, there is a remarkable reversal of signs. Here households buying large volumes have bigger shares of the favorite brand and a lower general level of brand switching activity. So for margarine households which are heavy buyers, in the sense of purchasing frequently, show relatively a lot of brand switching, but households which are heavy buyers in the sense of buying a large volume show relatively little brand switching. With respect to FIRBR the rank order of the 5 brand codes is 2, 3, 4, 1, 5 ($F = 5.00$) for fopro, 2, 1, 4, 5, 3 ($F = 2.37$) for beer and 2, 3, 5, 1, 4 ($F = 12.30$) for margarine. It can be remarked that for margarine apparently the cheapest brand codes occur relatively often as the favorite brand in households buying a large volume.

VOLPP. This variable is related to the two preceding ones because it is the average volume bought per purchase occasion. For fopro, households which buy larger volumes per purchase occasion have a smaller share of the favorite brand, more different brands, a higher general level of brand switching activity and more variation in this level. For beer there is only a correlation with SFAVBR, which points in the opposite direction. For margarine the picture is quite different from that of fopro. Here households which buy larger volumes per purchase occasion have bigger shares of the favorite brands, fewer brands, a lower general level of brand switching activity and less variation in that level. Fopro and margarine agree on the point that households buying larger volumes per purchase buy more private brands.

With respect to FIRBR the rank order of the brand codes is: 2, 1, 5, 3, 4 ($F = 11.96$) for fopro, 1, 4, 3, 2, 5 ($F = 4.37$) for beer and 2, 3, 5, 1, 4 ($F = 10.38$) for margarine.

Summarizing the results for NUMPU, VOLPU and VOLPP the following can be said. For fopro households with many purchase occasions show little brand switching; households buying larger volumes per purchase occasion show relatively a lot of brand switching. For beer, households with many purchase occasions, as well as households buying larger volumes show little brand switching. Households buying large volumes per purchase occasion are not very different in brand choice behavior than other households. For margarine, households with many purchase occasions show a lot of brand switching, households buying large volumes show little brand switching and the same applies to households with large volumes per purchase.

VARVOL and *VARIPT*. These variables refer to the variability in volume bought per purchase occasion, respectively to the variability in inter-purchase time, during the purchase history of a household. The pictures of correlation coefficients corresponding to these 2 variables are very similar per product and over the products. Generally, households with great variability in volume and inter-purchase time have smaller shares of the favorite brand, more different brands, a higher general level of brand switching activity and more variation in this level. So households with more variation in volume per purchase and in inter-purchase time also show more variation with respect to the brands chosen. This phenomenon, which is very evidently demonstrated here, is interesting because it points to a basic variability factor in the buying behavior of a household, which apparently has an effect on a number of different aepects of the purchase process. With respect to FIRBR, the rank order of the brand codes for fopro by VARVOL is 1, 2, 4, 5, 3 ($F = 5.41$), and by VARIPT 1, 4, 3, 2, 5 ($F = 5.21$). The rank order for margarine by VARIPT is 4, 1, 5, 3, 2 ($F = 4.96$).

NUMSIZ. This variable, referring to the different package sizes of the product bought by a household, is only applicable for fopro and beer. It appears that for both products households with many package sizes have smaller shares of the favorite brand, more different brands, a higher general level of brand switching activity, more variation in this level and buy more private brands. So households with more variation in package size also show more variation with respect to brands. This can be partly explained by the fact that not all brands carry the full range of package sizes, which implies that in some cases a household which wants to change its package size is also forced to change its brand. But it is also possible that we encounter here the basic variability factor, just discussed in relation to VARVOL and VARIPT, which is also at work in the choice of package size. An indication of this is the following. For fopro the rank order of brand codes by increasing NUMSIZ is 1, 5, 2, 4, 3 ($F = 6.28$). We know that for brand code 1 (brand F1) the full range of package sizes is available and we see that households with brand code 1 as the favorite brand have the smallest number of package sizes.

NUMPUR and *NUMVAR*. These variables only apply to beer. We see that households in which there are more purchasers of beer and which buy more different varieties of beer have smaller shares of the favorite brand, more different brands, a higher general level of brand switching activity and more variation in this level; in short these households show more variation in brand choice. The result with respect to NUMPUR is quite in agreement with the result for FAMSIZ outlined in section 8.3.2: bigger households simply have more possibilities of having more different purchasers.

Via a similar reasoning as that made for NUMSIZ, the result for NUMVAR can partly be explained by the fact that not all brands have all varieties and partly by the existence of a basic variability factor.

Here the most important findings with respect to the relationships between brand choice variables and other purchase variables are summarized. Households with a lot of brand switching pay lower prices. This is the most evident for fopro and beer, but also for margarine there are clear indications in this direction. One way of thinking about this phenomenon might be to consider the higher price paid by the brand loyal households as being a risk-premium, for which they receive the certainty of a good quality. There are other households for which the different brands are apparently used more as substitutes for each other; these households concentrate on cheaper brands.

For fopro and margarine households with larger proportions of deal purchases show more brand switching. For beer this could not be established.

Households buying a larger part of their purchases in self-service shops and supermarkets and in shops belonging to chains, show relatively a lot of brand switching.

As far as total volume of purchases is used as the relevant indicator, heavy-buying households show less brand switching than light buyers. When the total number of purchase occasions is considered, the picture for margarine is the opposite. Households buying large volumes per purchase are relatively less brand loyal for fopro, but more brand loyal for margarine.

A remarkable result is that households showing great variation with respect to brands, also show great variation in the volume bought per purchase occasion and in inter-purchase time. For fopro and beer these households also have great variation in package size and for beer also in the variety bought. These findings seem to point to the existence of a basic variability factor in buying behavior.

For beer, households with many different beer purchasers show relatively a lot of brand switching.

8.4 BRAND CHOICE BEHAVIOR OF HOUSEHOLDS WITH RESPECT TO DIFFERENT PRODUCTS

8.4.1 *General*

In this section we examine if brand choice behavior is transitive over the 3 products fopro, beer and margarine. In other words we want to see if households show a similar brand choice behavior for these 3 products. For this purpose the observed brand choice behavior for different products of the same households is mutually compared. Of course for this analysis we can only use those households for which the purchase histories of all 3 products are known. When, in chapter 2, the data to be used in this book were discussed, it was observed that the households used in the analyses for the 3 products were not all the same. For 378 households the purchase histories for all 3 products over the whole two years could be used. These 378 households were present in the 672 fopro households, as well as in the 627 beer households and the 1059 margarine households. These 378 households are the ones used in the present section.

8.4.2 *The method used*

When households show the same type of brand choice behavior with respect to different products, there will be agreement between their values for the corresponding brand choice variables referring to different products, so in this case it is to be expected that, for example, households with many different brands for fopro also have many different brands for beer and for margarine. To examine how far this is true for the 378 households these mutual relationships between corresponding brand choice variables for different products for the same households were studied. This was done for the brand choice variables SFAVBR, NUMBR, RANGPLZ, MEANPLZ and FIRBR. In addition, this analysis was performed for a number of purchase variables, viz., PRICE, DEAL, SELSUP, CHAINSH, SFAVSH and NUMSH. The last 2 variables, which refer to shop choice, were defined in section 7.2.2. Not all brand choice and purchase variables as defined in section 8.2 were included; only a number of them which were considered as being the most important.

The variables just mentioned are all interval-scaled, except FIRBR. For the interval-scaled variables mutual (ordinary) correlation coefficients were computed and – which was more practicable here, because of the smaller number of households – also the Spearman rank correlation coefficients.

Correlation coefficients give an indication of how far *two* variables go up and down together. But we are also interested in the question of how far corresponding variables of *all 3* products fopro, beer and margarine show common movements, i.e. how far households with high values of e.g. NUMBR for fopro, also have high values of NUMBR for beer and margarine, etc. To answer this question the so-called Kendall coefficient of concordance W was computed. This statistic was used in chapter 4 of this book and in section 4.6.4.1.2 a brief description was given. For this large number of observations (378), a χ^2-test can test if the rankings by different variables are in agreement or not (see Siegel, 1956, eq. 9.18). Here, by different variables, variables with the same name but referring to different products are meant.

For the nominal variable FIRBR contingency tables with χ^2-tests were used to test the agreement on this variable for different products.

8.4.3 *Results*

Table 8.3 presents the computed ordinary correlation coefficients, Spearman rank correlation coefficients and the values for W with the corresponding χ^2-values for the respective brand choice and other purchase variables. With 378 observations and a 5% one-side confidence level (the null-hypothesis of no correlation is tested against the alternative of positive correlation) all correlation coefficients above .085 are significant. It can be seen that this applies to all ordinary and rank correlation coefficients in Table 8.3. So it can be concluded that for each pair of products there is a positive relationship between corresponding brand choice variables and purchase variables of the same households. One point to be mentioned is the generally close agreement of the values of ordinary and corresponding rank correlation coefficients. With $\alpha = .05$ the critical value of χ^2 associated with the Kendall coefficient W for 378 observations is 426; with $\alpha = .01$ the critical value is 445. All observed χ^2-values are well above these boundaries, so it is concluded that the rankings of households by corresponding variables for the 3 different products are in agreement with each other. This applies to all brand choice and purchase variables used.

Table 8.3 Relationships between corresponding variables of different products for the same households; n = 378
(W = Kendall's coefficient of concordance; W = 0: no agreement, W = 1: complete agreement)

Product combination	*Fopro/Beer*		*Fopro/Marg*		*Beer/Marg*		*All 3 products*	
Variable	*Correlation coefficients*							
	Ordinary	*Rank*	*Ordinary*	*Rank*	*Ordinary*	*Rank*	*W*	χ^2_{377}
SFAVBR	.1420	.1730	.2154	.2475	.1130	.1475	.46	520
NUMBR	.2126	.2572	.4871	.4440	.2726	.2898	.55	626
RANGPLZ	.1950	.2139	.3650	.3604	.2610	.2715	.52	590
MEANPLZ	.2027	.2121	.4388	.4079	.1603	.1703	.51	576
PRICE	.3557	.3152	.2422	.2388	.0963	.1066	.48	543
DEAL	.2322	.2180	.3749	.2276	.1758	.2368	.48	549
SELSUP	.5461	.5583	.4647	.5197	.5525	.6058	.71	800
CHAINSH	.5812	.5962	.4479	.5173	.5162	.5894	.71	805
SFAVSH	.2092	.2431	.2502	.2596	.2893	.3474	.52	591
NUMSH	.3694	.3162	.4256	.4140	.3782	.3833	.58	657

The important conclusion to be drawn from Table 8.3 is that the buying habits of individual households with respect to fopro, beer and margarine are not mutually independent, i.e., that buying behavior is to some extent transitive over these 3 products. Households which have a big share of the favorite brand for fopro, generally also have a big share of the favorite brand for beer and margarine.

Households which have many different brands for fopro generally also have many different brands for beer and margarine. Analogous statements can be made for the other brand choice variables RANGPLZ and MEANPLZ, and the purchase variables PRICE to NUMSH.

When we look at the estimated correlation coefficients, it appears that of the brand choice variables NUMBR is the most transitive; for all 3 combinations the correlation coefficient for this variable is the highest. When coherence between pairs of products is mutually compared, it can be said that on the point of brand choice behavior fopro and margarine show the most cohesion, because for this combination the correlation coefficients between corresponding brand choice variables are highest.

With respect to the other purchase variables, the correlation coefficients for SELSUP and CHAINSH are particularly high. Generally, households buying a lot of their fopro in self-service shops or supermarkets and in shops belonging to chains, also do this for beer and margarine. Further, it appears that households which pay high prices or which often purchase deals for one product generally also do this for the other products.

Next the observed relationships for the nominal variable FIRBR, as referring to the 3 different products of the same households are reported. For a better mutual comparison, the number of different categories of FIRBR, which was initially 5, was reduced to 3, viz., major national brand (MNB), private brand (PRIV) and other (OTHER). For fopro the category MNB contains brands F1 and F2 (former codes 1 and 2), PRIV contains former codes 3 and 4 and OTHER is the same category as former code 5. For beer and margarine MNB contains former codes 1, 2 and 3 (brands B1, B2, B3, respectively M1, M2 and M3), PRIV is equal to former code 4 and OTHER to former code 5.

By means of contingency tables the agreement regarding the type of favorite brand for different products of the same household is examined. In other words, it is verified if households which have a major national brand as favorite brand for fopro, also have relatively often a major national brand as favorite brand for beer, etc.

These contingency tables are given in Table 8.4.

Table 8.4 Contingency table examining the agreement between the types of the favorite brand for different products for the same household; entries within tables: percentages of row totals

	Rows: Fopro/Columns: Beer				*Rows: Fopro/Columns: Marg*				*Rows: Beer/Columns: Marg*			
	MNB	PRIV	OTHER	*N*	MNB	PRIV	OTHER	*N*	MNB	PRIV	OTHER	*N*
MNB	80	6	14	182	45	10	45	182	48	12	40	251
PRIV	46	36	18	115	28	36	36	115	15	46	39	65
OTHER	65	16	19	81	47	15	38	81	37	19	44	62
Total	66	17	16	378	40	20	40	378	40	20	40	378
χ^2_4				48.95				31.70				45.54

This table should be read as follows: of the 182 households which for fopro have an MNB as the favorite brand, 80% also have for beer an MNB as the favorite brand, 6% have for beer a PRIV and 14% an OTHER as the favorite brand. The construction of Table 8.4 should now be clear.

It appears that for all 3 tables the computed χ^2-value is considerable and clearly significant. So the categories to which the favorite brands for fopro, beer and margarine belong for the same households are not mutually independent. From the fopro-beer contingency table (the utmost left one of Table 8.4) it appears that households with an MNB as favorite brand for fopro are relatively often also in the category MNB for beer, and households in category PRIV for fopro are relatively often also in the category PRIV for beer. This latter result, namely with respect to PRIV, is the most striking feature for all 3 contingency tables of Table 8.4. So it appears that households which have a private brand as favorite brand for one of the products, fopro, beer or margarine, relatively often also have a private brand as favorite brand for the two other products.

From the results of this section it appears that buying behavior, and especially brand choice behavior, is to a certain extent transitive over products. As a consequence, it can be expected that there are general characteristics of a household (the housewife in particular) which determine – partly – the type of buying behavior in relation to different products. In this connection it is remarkable that the correlations of socio-economic variables with brand choice variables, as given in section 8.3.1, are so weak. The coefficients of correlation given there are generally much smaller than those given in Table 8.4, where brand choice variables with respect to different products were directly related. Apparently those socio-economic variables used in section 8.3.1, are not the ones which discriminate well between households with different brand choice patterns and perhaps better ones could be found.

In the next section these results are compared with those found in the literature and it is examined in how far other researchers have succeeded in finding general indicators for brand choice behavior.

8.5 COMPARISON WITH RESULTS FROM THE LITERATURE

The questions of how far brand choice behavior is related to socio-economic, personality and other purchase variables and if brand choice behavior is transitive over products have been examined by a number of authors. This section briefly presents a number of findings from these studies which are interesting in comparison with the results presented in sections 8.3 and 8.4. These studies partly use the same variables as were used here; in such cases, the names as defined in section 8.2 are used for these variables.

In his pioneering study in the field of brand choice research, mentioned in section 1.1, *Cunningham (1956)* reports that no relationship between the socio-economic characteristics of households and brand loyalty was found. With respect to the relationship VOLPU-SFAVBR he concluded that for none of the 7 products studied did a significant relationship exist.

For margarine, which was one of the 7 products in Cunningham's research, the estimated rank correlation coefficient was .13, which can be compared with the value for the corresponding correlation coefficient of .10 in Table 8.2. Although significant correlations between VOLPU and SFAVBR were found here for all 3 products, the correlation coefficients are rather small: .08 for fopro, .13 for beer and .10 for margarine.

To examine the transitivity of brand choice behavior over products, Cunningham computed rank correlation coefficients between the SFAVBR-values of each pair of the 7 products he studied. Thus 21 correlation coefficients were obtained, which ranged from – .06 to .30 with a median value of .12, only 2 of which were statistically significant. His conclusion is that there is no such phenomenon as, what he calls, 'loyalty proneness'. In this connection the estimated rank correlation coefficients for SFAVBR, given in Table 8.3 can be mentioned: .17 for fopro and beer,

.25 for fopro and margarine and .15 for beer and margarine, which values are somewhat higher than the median value found by Cunningham.

Farley (1964) applied a multiple regression analysis with NUMBR as the dependent variable and FAMSIZ, VOLPU and income as independent variables. The results are rather disappointing: in 11 of the 17 regressions R^2 was below .04. The only significant result was the positive effect of VOLPU on NUMBR. This implies that households which buy larger volumes show more brand switching, which is in contrast with the results obtained in our research.

In contrast to Farley's hypotheses, income and family size had no significant influence on NUMBR. In our study for family size, a significant influence was found only for beer and the variable SOCCL, which to some extent is an indicator for income, was also found to have not much influence on brand choice behavior.

Frank (1967a) examined by means of a multiple regression study, if certain buying variables (dependent variables) could be conceived of as being a linear function of a number of socio-economic and purchase variables.

In terms of the research performed in this book, the most important buying variables used by Frank as dependent variables were PRIVBR and SFAVBR. For PRIVBR a significant influence for a socio-economic variable was only found for FAMSIZ: bigger households buy more private brands. A similar result was found here for margarine. With respect to the relationship PRIVBR – other purchase variables, the only result reported by Frank is that households buying more in those shops which carry private brands, buy more private brands, which is not surprising.

With respect to the results for SFAVBR Frank states: 'There appears to be virtually no association between household, socio-economic and purchase characteristics included in the analysis and the degree of brand loyalty exhibited by a household'. So his results are more negative than those found here, because for the socio-economic variables a number of significant results could be reported and for the purchase variables many significant relationships. A special point to be mentioned is that Frank included the average price paid as a purchase variable and apparently found no significant relationships between this variable and the SFAVBR and PRIVBR, while we found significant correlations of PRICE with SFAVBR and PRIVBR for all 3 products.

In another article *Frank (1967b)* reviews the major findings of a number of different studies on brand choice. He comes to the conclusion that 'brand loyal customers almost completely lack identifiability in terms of either socio-economic or personality characteristics'. He also concludes that loyal customers do not have economically important differences in their sensitivity to either short-run effects of pricing and dealing and that they do not have different demand levels. Because significant relationships between PRICE, DEAL and VOLPU and brand choice variables were found in this research, Frank's latter statement does not seem to apply to the fopro, beer and margarine households discussed here.

A very thorough investigation regarding the relationship between brand choice behavior and socio-economic and personality characteristics was undertaken by *Massy, Frank and Lodahl (1968)*, for coffee, tea and beer. They confronted a battery of brand choice variables for the 3 products with a battery of socio-economic and personality variables. As stated by the authors themselves the results obtained were disappointing, for all 3 products. Their conclusion is that personality and socio-economic variables have no great influence on brand choice behavior here.

Some results in the study of Massy et al., which can directly be compared with the findings obtained here are the following. With respect to the transitivity of brand choice behavior over products, correlation coefficients between values for different products of the same households for the variables SFAVBR and NUMBR were computed. The resulting correlation coefficient corresponding to SFAVBR for coffee/tea was .28 and for coffee/beer .19. For NUMBR these coefficients were respectively .31 and .24. These values are of the same order of magnitude as the results regarding the transitivity of SFAVBR and NUMBR given in Table 8.3 and indicate that this transitivity is significantly present.

Wind and Frank (1969) computed coefficients of determination R^2 (squares of the simple correlation coefficient R) for the relationship between values for different products of the same households for the variables SFAVBR and PRIVBR. This was done for each pair of the 38 products in their study, which resulted in 703 correlation coefficients for each variable. It appeared that for SFAVBR 30% of the estimated R^2-values were below .01 (corresponds with $R = .1$), 83% below .05 (corresponds with $R = .22$) and 97% below $R^2 = .1$ (corresponds with $R = .32$). For PRIVBR the corresponding percentages were 35%, 73% and 90%. So only small percentages of the variation from household to household in brand loyalty within one product category are associated with variation in brand loyalty in another category. In a large number of cases there is no significant relationship at all between SFAVBR-values for different products of the same household.

For PRIVBR the R^2-values are generally somewhat higher than for SFAVBR, but they are still rather low. It can be noted that in section 8.4, by means of the contingency table made for FIRBR, an evident private-brand-proneness was found in the sense that households, which have a private brand as the favorite brand for one product, relatively often also have this for another product.

When we look over the various results obtained by the different authors and compare them with the findings obtained here, the following can first be noted. All the studies, mentioned above, are analyses of purchasing behavior of consumers in only one country, namely the USA. The results of this book refer to the purchasing behavior of Dutch households in the Netherlands. In this connection the similarities between the results are more striking than the differences.

As for the relationship: brand choice behavior – socio-economic and personality variables, all the results in the literature are rather negative. From the results of Table 8.1 it is clear that also in the research undertaken here no strong relationships could be discovered. Although a number of significant relationships were encountered, the portion in the variation of the brand choice variable explained by the socio-economic variables was generally very small. In this study a relatively good impression is made by the attitude variable price-consciousness. Perhaps this kind of attitude score, directly referring to the buying process itself, can more fruitfully be used to explain brand choice behavior than the very general personality characteristics, obtained from the Edwards personal preference scale for example.

For the relationship between brand choice variables and other purchase variables the results obtained here are more encouraging than those presented in the literature. Most authors conclude that relationships between brand choice variables and other purchase variables are either very weak or non-existent, while – as is clear from Table 8.2 – we found for a number of variables evident relationships with brand choice behavior. As such can be mentioned: price paid, importance of deal purchases, importance of purchases in self-service shops or supermarkets, importance of purchases in shops belonging to chains, volume of purchases, variation in the volume bought per purchase occasion, variation in inter-purchase times and number of package sizes bought.

The question if brand choice behavior is to some extent transitive over products could be answered affirmatively for the fopro, beer and margarine. This diverges from Cunningham's conclusion that brand loyalty proneness does not exist, is somewhat more positive than the findings of Wind & Frank, and agrees with Massy et al.'s results for coffee, tea and beer. Here we would like to point out that, as in Massy et al., the mutual relationships of brand choice variables were examined for only 3 different products. Also, the correlation coefficients in the upper part of Table 8.3 – although evidently significant – are still not close to 1, which means that the transitivity of brand choice behavior is by far not complete and that a lot of room is still left for factors specific to each of the 3 products.

The main interest underlying attempts to relate brand choice behavior to socio-economic and personality variables and the study of transitivity of brand choice behavior over products lies in the fact that, if transitivity exists and indicators for the brand choice behavior of households can be found, brand choice behavior can be used as a basis for market segmentation. It would then be possible to distinguish between consumers who are brand loyal in general and consumers who often change brands in general, which would be very useful from a marketing point of view. From the findings made here with respect to the socio-economic variables it can be concluded that with the variables included in this piece of research, no very successful market segmentation can be carried out. Perhaps it is possible to find more discriminating variables, but – because the transitivity of brand choice behavior over products is far from complete – it is not likely that a very sharp distinction of consumers in the sense mentioned above will ever be possible.

8.6 CONCLUSIONS

In this chapter the relationships between brand choice variables and socio-economic and purchase variables were examined and the question of transitivity of brand choice behavior over products was considered.

It was found that socio-economic variables have only minor relationships with brand choice variables. For all 3 products some influence of size of town, region, children and attitude scores referring to buying behavior was observed. For some products there was an influence of family size, age of housewife, and the possession/non-possession of a refrigerator and a television set.

For the purchase variables there were more evident relationships with brand choice. Significant correlations with brand choice variables were observed for: price paid, importance of deal purchases, importance of purchases in self-service shops or supermarkets, importance of purchases in shops belonging to chains, volume of purchases, variation in volume of purchases, variation in inter-purchase times and number of different package sizes bought.

It was found that brand choice behavior is transitive over products, but not to such an extent that a general brand choice behavior of a consumer could be spoken of, which could serve as an independent basis for market segmentation.

For the effect of socio-economic variables the findings agreed rather well with results obtained from the literature; with respect to purchase variables more evident relationships with brand choice variables were found than were reported in the studies considered.

9. Summary and conclusions

Here a brief recapitulation of the study of brand choice processes is given, and the major conclusions are reported.

In chapter 2 we discussed the empirical brand choice data used throughout the study. We saw that these were purchase histories of members of the Dutch Attwood Consumer Panel for the products fopro (a pseudonym), beer and margarine during the years 1967 and 1968. A brief description of the type of information, available for each purchase, was given, and some important characteristics of the markets for each of the 3 products were also reported. In addition some attention was paid to the special brands of the 3 products which are repeatedly distinguished in the study and the source of information for the advertising figures used in the analysis was mentioned.

In chapter 3, 5 different brand choice models with the corresponding testing and estimation procedures were presented, viz., the:

1. Homogeneous Bernoulli Model (HOBM);
2. Homogeneous Markov Model (HOMM);
3. Heterogeneous Bernoulli Model (HEBM);
4. Heterogeneous Markov Model (HEMM);
5. Linear Learning Model (LLM).

Of these models (1) and (3) assume no influence of former purchases on current brand choice, i.e. no purchase feedback. In these two models it is assumed that a consumer purchases a certain brand with constant probability, which probability – according to (1) – is equal for all consumers, but – in the case of (3) – may vary over consumers.

All the other 3 models assume purchase feedback. In the case of the Markov models (2) and (4) the influence of former purchases is limited to the recent purchase occasions and expressed in so-called transition probabilities. For the HOMM these transition probabilities are equal for all consumers; in the case of the HEMM different consumers can have different transition probabilities. According to the Linear Learning Model a consumer is assumed to have – at a given purchase occasion – a probability p of buying a certain brand. After a purchase this probability is transformed in a way dependent on the brand bought at that purchase. This implies that in the LLM quite a number of former purchases have an influence, but the influence of the brand choice at an earlier purchase occasion diminishes with the number of purchases made since that occasion.

In chapter 4 it was examined how far the brand choice models just mentioned gave a good description of the empirical brand choice processes of fopro, beer and margarine. This was done by carrying out testing procedures and performing a simulation study.

From the test results it became clear that in all cases a definite purchase feedback is present; the Bernoulli models HOBM and HEBM did not give a good fit, and the Markov and Learning models also showed an evident influence of previous purchases. Of the homogeneous Markov models the first order Markov model, which is the brand choice model most discussed in the literature, did not give a good fit. The performance of the second order model was better, but was still not really satisfactory. The Heterogeneous Markov Model (HEMM) also did not give a good description of the brand choice processes observed. The brand choice model which gave the best results, was the Linear Learning Model (LLM). In every case this model was superior to all other models.

In the simulation study, where the ability of the various brand choice models to reproduce the original brand choice processes was examined, the superiority of the Linear Learning Model again appeared, while the homogeneous Markov models (first and second order) offered a much worse reproduction. A curious point is that the HEBM appeared to give a reproduction of the original brand choice processes which was almost as good as that of the LLM. At first glance this seemed contradictory, but a closer examination of the LLM-parameters showed that these parameters were such that the LLM-processes concerned exhibited a lot of seeming zero-order behavior. Because the HEBM is a zero-order model, this explains the phenomenon observed. Taking the results of the testing procedures and the simulation study together, it appears, that – of all brand choice models used – the Linear Learning Model evidently gave the best description of the empirical brand choice processes.

In chapter 5 we briefly discussed a number of learning models from the viewpoint of their application possibilities to brand choice processes. The non-linear operator models treated appeared to offer no great perspectives. The stimulus sampling models, an example of which was applied to the fopro, beer and margarine data looked more promising. Some further properties of the Linear Learning Model were given, in particular relating to equilibrium behavior, which are useful because they can be used to compute long term market shares. It was also shown that the Linear Learning Model can be generalized, so that brand choice processes can be handled, for which not all assumptions of the ordinary LLM hold.

In chapter 6 we analysed the empirical brand choice processes with the aid of the variable 'poolsize' introduced there. Poolsize is defined as the 'number of different brands bought during the last 10 purchases'. An important finding is that – in their purchase histories – consumers show periods of routinized buying alternated with periods of brand switching. This is in agreement with the good fit of the LLM observed in chapter 4, because the parameters of the LLM estimated there are such that in the corresponding brand choice processes there will be long periods during which brand switches are very unlikely, alternated with periods in which the probability of moving to another brand is considerable.

Further conclusions resulting from the poolsize approach, are that a consumer simultaneously considers a limited number of brands as potential choice candidates and that consumers do not often straightforwardly switch from one brand to another, but usually exhibit search behavior, which accompanies a transition to another brand.

In chapter 7 we examined the relationships between brand choice and a number of environmental variables.

With respect to shop choice, it was found that brand choice and shop choice are rather closely related. This interdependence cannot completely be traced back to the fact that the choice of a shop simply limits the set of different brands from which a choice can be made; it seems that an autonomous general proneness-to-change factor exists, which means that some consumers show great variation with respect to shops as well as to brands. Moreover, there can be distinguished specific proneness-to-brand-change and proneness-to-shop-change factors. With respect to the effect of the marketing variables the following can be remarked: by means of a multiple regression analysis it was found that for a number of brands there was a significant influence of price and/or advertising (the latter measured by expenditures made) on market share, repeat purchase probability and on the probability of making a transition from another brand to the brand concerned. Further it was found that deal purchases are relatively often associated with brand switches, so that dealing seems to be an instrument having the ability to induce brand changes.

As for the effect of inter-purchase times on brand choice, it was found for fopro that the probability of purchasing the same brand as the previous one (=repeat purchase probability) decreases as inter-purchase times become longer. For beer and margarine no effect of inter-purchase times could be established.

In chapter 8 we studied the relationship between brand choice behavior and household variables.

It was found that socio-economic variables have only weak relationships with brand choice variables. For all 3 products we observed some influence of size of town, region, children and attitude scores in relation to buying behavior. In incidental cases there was also an influence of family size, age of housewife and the possession of a refrigerator and a television set.

Between brand choice variables and other purchase variables the relationships are stronger. It was observed that households which, relatively, show a lot of brand switching pay a lower price, make more deal-purchases, make more purchases in self-service shops or supermarkets, buy more in shops belonging to chains, have more variation in inter-purchase times and in volume per purchase occasion and buy more different package sizes.

To a certain extent brand choice behavior was found to be transitive over products, i.e. to some extent households showed the same type of brand choice behavior in relation to different products, but not to such a degree that a general brand choice behavior could be spoken of which could serve as an independent basis for market segmentation.

References

(*JMR* = Journal of Marketing Research)
(*JAR* = Journal of Advertising Research)

Abramowitz, M. and I.A. Stegun (eds), *Handbook of Mathematical Functions*, Dover Publications, Inc., New York, 1968.

Anderson, T.W. and L.A. Goodman, "Statistical Inference About Markov Chains", *Annals of Mathematical Statistics*, Vol. 28, 1957, 89-110.

Audley, R.J. and A.R. Jonckheere, "The Statistical Analysis of the Learning Process", *The British Journal of Statistical Psychology*, Vol. 9, Part II, 1956, 87-94.

Bateman, G., "On the Power Function of the Longest Run as a Test for Randomness in a Sequence of Alternatives", *Biometrika*, Vol. 35, 1948, 97-112.

Billingsley, P., " Statistical Methods in Markov Chains ", *Annals of Mathematical Statistics*, Vol. 32, 1961, 12-40.

Boyd, H.W., Jr. and R.L. Westfall, *An Evaluation of Continuous Consumer Panels as a Source of Marketing Information*, American Marketing Association, Illinois, 1960.

Bradley, J.V., *Distribution-Free Statistical Tests*, Prentice Hall, Englewood Cliffs, N.J., 1968.

Brown, G.H., "Brand Loyalty – fact of fiction", *Advertising Age*, Vols. 23–24, 1952-53.

Bush, R.R. and F. Mosteller, *Stochastic Models For Learning*, Wiley, New York, 1955.

Campbell, B.M., *The Existence and Determinants of Evoked Set in Brand Choice Behavior*, unpublished doctoral dissertation, Columbia University, New York, 1969.

Carman, J.M., "Brand Switching and Linear Learning Models", *JAR*, Vol. 6, June, 1966, 23-31.

Carman, J.M., "Correlates of Brand Loyalty, Some Positive Results", *JMR*, Vol. 7, 1970, 67-76.

Coleman, J.S., *Introduction to Mathematical Sociology*, The Free Press of Glencoe, MacMillan Limited, London, 1964.

Coombs, C.H., R.M. Dawes and A. Tversky, *Mathematical Psychology, An Elementary Introduction*, Prentice-Hall, Englewood Cliffs, N.J., 1970.

Cox, D.R., *The Analysis of Binary Data*, Methuen & Co. London, 1970.

Cox, D.R. and P.A.W. Lewis, *The Statistical Analysis of Series of Events*, Methuen & Co. London, 1966.

Cox, D.R. and H.D. Miller, *The Theory of Stochastic Processes*, Methuen & Co., London, 1965.

Cramer, H., *Mathematical Methods of Statistics,* Princeton University Press, Princeton, 1957.
Cunningham, R.M., "Brand Loyalty to Store and Brand", *Harvard Business Review* Vol. 34, no. 1, 1956, 116-128.
Cunningham, R.M., "Customer Loyalty to Store and Brand", *Harvard Business Review* Vol. 39, no. 6, 1961, 127-137.

Draper, J.E. and L.H. Nolin, "A Markov Chain Analysis of Brand Preference", *JAR,* Vol. 4, 1964, 33-38.
Draper, N.R. and H. Smith, *Applied Regression Analysis,* Wiley, New York, 1968.

Ehrenberg, A.S.C., "A Study of Some Potential Biases in the Operation of a Consumer Panel", *Applied Statistics,* Vol. 9, 1960, 20-27.
Ehrenberg, A.S.C., "An Appraisal of Markov Brand Switching Models", *JMR,* Vol. 2, 1965, 347-363.
Engel, J.F., D.T. Kollat and R.D. Blackwell, *Consumer Behavior,* Holt, Rinehart and Winston Inc., New York, 1968.
Estes, W.K., "Toward a Statistical Theory of Learning", *Psychological Review,* Vol. 57, 1950, 94-107.

Farley, J.U., "Brand Loyalty and the Economics of Information", *Journal of Business,* Vol. 37, no. 4, 1964, 370-379.
Feller, W., *An Introduction to Probability Theory and its Applications,* Vol. I, Wiley, New York, 1957.
Frank, R.E., "Brand Choice as a Probability Process", In: R.E. Frank et al. (eds.), *Quantitative Techniques in Marketing Analysis,* Irwin, Homewood, Illinois, 1962, 372-384.
Frank, R.E., "Correlates of Buying Behavior For Grocery Products", *Journal of Marketing,* Vol. 31, Oct., 1967[a], 48-53.
Frank, R.E., "Is Brand Loyalty a Useful Basis For Market Segmentation?", *JAR,* Vol. 7, no. 2, 1967[b], 27-33.

Haines, G.H. Jr., *Consumer Behavior, Learning Models of Purchasing,* The Free Press, New York, 1969.
Harary, F. and B. Lipstein, "The Dynamics of Brand Loyalty: A Markovian Approach", *Operations Research,* Vol. 10, 1962, 19-40.
Herniter, J.D., "A Probabilistic Market Model of Purchasing Timing and Brand Selection", *Manag. Science,* Vol. 18, no. 4, 1971, 102-113.
Herniter, J.D. and R.A. Howard, "Stochastic Marketing Models", in: *Progress in Operations Research,* Vol. 2, Wiley, New York, 1964.
Herniter, J.D. and J.F. Magee, "Customer Behavior as a Markov Process", *Operations Research,* Vol. 9, 1961, 105-122.
Howard, J.A. and J.H. Sheth, *The Theory of Buyer Behavior,* Wiley, New York, 1969.
Howard, R.A., *Dynamic Programming and Markov Processes,* M.I.T. Press, Cambridge, Massachusetts, 1960.
Howard, R.A., "Stochastic Process Models of Consumer Behavior", *JAR,* Vol. 3, no. 3, 1963, 35-42.

Johnston, J., *Econometric Methods,* McGraw-Hill, New York, 1972.

Jones, J.M., "A Dual-Effects Model of Brand Choice", *JMR*, Vol. 7, 1970[a], 458-465.
Jones, J.M., "A Comparison of Three Models of Brand Choice, *JMR*, Vol. 7, 1970[b], 466-474.
Jones, J.M., "A Stochastic Model for Adaptive Behavior in a Dynamic Situation", *Manag. Science*, Vol. 17, 1971, 484-497.

Kemeny, J.G. and J.L. Snell, *Finite Markov Chains*, Van Nostrand Company, Princeton, N.J., 1960.
Kendall, M.G. and A. Stuart, *The Advanced Theory of Statistics*, Vol. I, Charles Griffin & Co., London, 1969.
Kendall, M.G. and A. Stuart, *The Advanced Theory of Statistics* Vol. II, Charles Griffin & Co., London, 1967.
Kuehn, A.A. (1962), "Consumer Brand Choice – A Learning Process?", in: R.E. Frank et. al. (eds.), *Quantitative Techniques in Marketing Analysis*, Irwin, Howewood, Illinois, 1962, 390-403.
Kuehn, A.A. and R.L. Day, "Probabilistic Models of Consumer Behavior", *Journal of Marketing*, Vol. 28, Oct., 1964, 27-31.

Lawley, D.N. and A.E. Maxwell, *Factor Analysis as a Statistical Method*, (sec. ed.), Butterworths, London, 1971.
Lawrence, R.J., "Patterns of Buyer Behavior: Time for a New Approach?" *JMR*, Vol. 6, 1969, 137-144.
Luce, R.D., *Individual Choice Behavior*, Wiley, New York, 1959.
Luce, R.D., R.R. Bush and E. Galanter (eds.), *Handbook of Mathematical Psychology* Vol. II, Wiley, New York, 1963.

Maffei, R.B., "Brand Preferences and Simple Markov Processes", *Operations Research*, Vol. 8, 1960, 210-218.
Massy, W.F., "Order and Homogeneity of Family Specific Brand-Switching Processes", *JMR*, Vol. 3, 1966, 48-54.
Massy, W.F., R.E. Frank and T.M. Lodahl, *Purchasing Behavior and Personal Attributes*, University of Pennsylvania Press, Philadelphia, 1968.
Massy, W.F., D.B. Montgomery and D.G. Morrison, *Stochastic Models of Buying Behavior*, M.I.T. Press, Cambridge (Massachusetts), 1970.
McConnel, D., "Repeat Purchase Estimation and the Linear Learning Model", *JMR*, Vol. 5, 1968, 304-306.
McGloughin, I., "Continuous Purchasing Data can be Collected from Individuals – But is it Valid?", *Papers Esomar – Wapor Congress*, Part II, Amsterdam, 1971, 317-334.
Miller, G.A., "The Magical Number Seven Plus Or Minus Two", *Psychological Review*, Vol. 63, 1956, 81-97.
Montgomery, D.B., "A Stochastic Response Model with Application to Brand Choice", *Manag. Science*, Vol. 15, no. 7, 1969.
Morrison, D.F., *Multivariate Statistical Methods*, McGraw-Hill, New York, 1967.
Morrison, D.G., "Testing Brand Switching Models", *JMR*, Vol. 3, 1966[a], 401-409.
Morrison, D.G., "Interpurchase Time and Brand Loyalty", *JMR*, Vol. 3, 1966[b], 289-291.
Morrison, D.G., R.E. Frank and W.F. Massy, "A Note on Panel Bias", *JMR*, Vol. 3, 1966, 85-88.

Nicosia, F. M., *Consumer Decision Processes; Marketing and Advertising Implications*, Prentice-Hall, Englewood Cliffs, N.J., 1966.

Powell, M.J.D., "Recent Advances in Unconstrained Optimization", *Mathematical Programming*, Vol. 1, 1971, 26-57.

Raiffa, H. and R. Schlaifer, *Applied Statistical Decision Theory*, Boston: Division of Research, Graduate School of Business Administration, Harvard University, 1967.
Rao, T.R., "Are Some Consumers More Prone to Purchase Private Brands?" *JMR*, Vol. 6, 1969, 447-451.

Siegel, S., *Non-parametric Statistics For the Behavioral Sciences*, McGraw-Hill, 1956.
Styan, G.P.H. and H. Smith, Jr. "Markov Chains Applied to Marketing", *JMR*, Vol. 1, no. 1, 1964, 50-55.
Sudman, S., "On the Accuracy of Recording of Consumer Panels", *JMR*, Vol. 1, no. 2 14-20 and no. 3, 1964, 69-83.

Wagner, H.M., *Principles of Operations Research*, Prentice-Hall, Englewood Cliffs, N.J., 1969.
Wilde, D.J., *Optimum Seeking Methods*, Prentice-Hall, Englewood Cliffs, N.J., 1964.
Wind, Y. and R.E. Frank, "Interproduct Household Loyalty to Brands", *JMR*, Vol. 6, no 4, 1969, 434-435.
Wonnacott, R.J. and T.H. Wonnacott, *Econometrics*, Wiley, New York, 1970.

Indices

AUTHOR INDEX

SUBJECT INDEX

(abbreviations relating to brandchoice models and variable names are given in a separate list on pages 265 and 266)

Samenvatting

In deze studie hielden wij ons bezig met merkkeuze processen, waarbij een merkkeuzeproces is gedefinieerd als het achtereenvolgens kopen van bepaalde merken van een produkt door een consument. Doel van het onderzoek was na te gaan hoe deze merkkeuze processen verlopen, waarbij met name de invloed van merken bij voorgaande aankopen op het te kiezen merk bij een bepaalde aankoop is onderzocht. Hierbij werd onder andere gebruik gemaakt van stochastische merkkeuze modellen. Verder is nagegaan in hoeverre bepaalde faktoren zoals winkelkeuze, marketing variabelen en socio-economische kenmerken het merkkeuzeproces beïnvloeden.

De data, die voor dit onderzoek zijn gebruikt, werden beschreven in hoofdstuk 2. Het zijn gegevens over aankopen van leden van het Attwood Consumentenpanel in Nederland, met betrekking tot de produkten fopro (een pseudoniem), bier en margarine gedurende de jaren 1967 en 1968. Voor de 3 bestudeerde produkten zijn een aantal kenmerken van de betreffende markten vermeld, terwijl tevens de in elk van de markten speciaal onderscheiden merken zijn aangegeven.

In hoofdstuk 3 zijn 5 verschillende merkkeuze modellen met de bijbehorende toetsings- en schattingsprocedures behandeld, te weten:

1. Homogeen Bernoulli model
2. Homogeen Markov model
3. Heterogeen Bernoulli model
4. Heterogeen Markov model
5. Lineair Leermodel

De modellen (1) en (3) veronderstellen dat er geen invloed is op de merkkeuze bij een bepaalde aankoop van de gekozen merken bij voorafgaande aankopen. Er wordt aangenomen, dat een consument een bepaald merk kiest met een constante waarschijnlijkheid. Deze kans wordt – volgens het homogeen Bernoulli model – verondersteld gelijk te zijn voor alle consumenten; in het heterogeen Bernoulli model kunnen verschillende consumenten verschillende kansen hebben om een bepaald merk te kopen. De andere 3 modellen veronderstellen dat er wel invloed is van voorgaande aankopen. In het geval van de Markov modellen – (2) en (4) – is deze invloed beperkt tot de recente aankopen en wordt uitgedrukt in zogenaamde overgangskansen. In het homogeen Markov model wordt aangenomen, dat iedere consument dezelfde overgangskansen heeft, in het heterogeen Markov model kunnen verschillende consumenten verschillende overgangskansen hebben. Volgens het lineair leermodel heeft een consument op een gegeven aankooptijdstip een kans p om een bepaald merk te kiezen. Nadat de aankoop gedaan is, wordt deze kans getransformeerd op een wijze, afhankelijk van het gekochte merk. Deze getransformeerde kans is dan de waarschijnlijkheid, dat het betreffende merk gekocht wordt bij de eerstvolgende aankoop. Bij het lineaire leermodel hebben in principe alle voorgaande aankopen invloed op de merkkeuze bij een bepaalde aankoop, maar de invloed van een eerdere aankoop is minder naarmate er meer aankopen sindsdien gedaan zijn.

In hoofdstuk 4 is onderzocht in hoeverre de genoemde merkkeuze modellen een goede beschrijving geven van de waargenomen merkkeuze processen voor fopro, bier en margarine. Hiertoe werden verschillende toetsingsprocedures en een simulatiestudie uitgevoerd.

Uit de toetsingsresultaten bleek, dat er in alle gevallen een duidelijke invloed van voorgaande aankopen op de merkkeuze is; de Bernoulli modellen gaven geen goede beschrijving van de empirische processen en ook de resultaten voor de Markov- en leermodellen toonden de invloed van voorgaande aankopen duidelijk aan. Van de homogene Markov modellen gaf het eerste orde Markov model – het bekendste merkkeuze model uit de literatuur – geen goede beschrijving van de empirische processen. Het resultaat voor het tweede orde Markov model was beter, maar nog niet erg bevredigend. Ook het heterogene Markov model, waarvan 2 speciale typen werden gehanteerd, gaf geen goede beschrijving. Het model met de beste resultaten is het lineair leermodel; in alle gevallen was dit model beter dan de andere.

Bij de simulatiestudie, waarin nagegaan werd in hoeverre met de verschillende modellen de oorspronkelijke merkkeuze processen kunnen worden gereproduceerd bleek opnieuw het lineair leermodel het beste te zijn, terwijl de homogene Markov modellen (eerste en tweede orde) een veel slechtere reproduktie van de oorspronkelijke processen te zien gaven. Merkwaardig is, dat het heterogeen Bernoulli model een reproduktie tot stand bracht die bijna even goed was als die door het lineair leermodel. Op het eerste gezicht leek dit tegenstrijdig, maar bij een nadere beschouwing van de parameters van het lineair leermodel bleken deze parameterwaarden zodanig te zijn, dat de bijbehorende merkkeuze processen veel z.g. nulde-orde gedrag – d.w.z. zonder invloed van voorgaande aankopen – vertonen. Dit verklaart het goede simulatieresultaat voor het heterogeen Bernoulli model.

Wanneer de resultaten van de verschillende toetsen en van de simulatie worden gecombineerd, dan blijkt dat van alle gehanteerde merkkeuze modellen het lineair leermodel duidelijk de beste beschrijving geeft voor de merkkeuze processen van fopro, bier en margarine.

Mede naar aanleiding van de vermelde goede resultaten voor het lineair leermodel, werden in hoofdstuk 5 een aantal andere leermodellen besproken met het oog op de mogelijkheid deze toe te passen op merkkeuzeprocessen. De behandelde niet-lineaire operator modellen bleken niet veel perspektieven te bieden. Voor de z.g. 'stimulus sampling' modellen waarvan een voorbeeld werd toegepast op gegevens van fopro, bier en margarine leken de toepassingsmogelijkheden groter.

Ook werden in dit hoofdstuk nog enkele eigenschappen van het lineair leermodel besproken, met name met betrekking tot de evenwichtsverdeling, welke laatste nuttig is voor het berekenen van lange termijn marktaandelen. Bovendien werd aangetoond, dat het lineair leermodel in bepaalde opzichten kan worden gegeneraliseerd, zodat dit model ook gebruikt kan worden voor merkkeuze processen waarvoor niet alle veronderstellingen van het gebruikelijke lineaire leermodel geldig zijn.

In hoofdstuk 6 werden de waargenomen merkkeuze processen geanalyseerd met behulp van het daar geïntroduceerde begrip 'poolomvang', gedefinieerd als het aantal verschillende merken, gekocht bij de laatste 10 aankopen. Een belangrijk resultaat is, dat consumenten in hun koopprocessen perioden van routinematig kopen afgewisseld met perioden van veel merkverandering vertonen. Dit stemt overeen met de goede resultaten van het lineair leermodel in hoofdstuk 4, omdat de geschatte parameters van het leermodel zodanig zijn, dat in de corresponderende merkkeuze processen lange perioden zullen voorkomen waarin merkwisselingen onwaarschijnlijk zijn, afgewisseld met perioden met een aanzienlijke kans om van merk te veranderen.

Een andere conclusie, resulterend uit de poolbenadering, is, dat een consument slechts een beperkt aantal merken tegelijkertijd als alternatieven voor zijn merkkeuze beschouwt. Verder bleek dat consumenten niet zonder meer van een bepaald merk op een ander overgaan, maar dat ze meestal in meer of mindere mate zoekgedrag vertonen bij de overgang naar een ander merk.

In hoofdstuk 7 werden de samenhangen tussen merkkeuze processen en een aantal omgevingsvariabelen onderzocht. Wat betreft de relatie met winkelkeuze werd geconstateerd, dat merkkeuze en winkelkeuze onderling nauw samenhangen. Deze samenhang blijkt niet volledig te kunnen worden teruggebracht tot het feit, dat de keuze van een winkel de verzameling van merken, waaruit de consument een keus kan maken, beperkt. Er blijkt een op zichzelf staande factor te bestaan die 'neiging tot verandering' kan worden genoemd en die veroorzaakt, dat sommige consumenten zowel veel variatie in merk als in winkel vertonen. Bovendien kunnen als afzonderlijke faktoren worden onderscheiden een specifieke neiging tot merkverandering en een neiging tot winkelverandering.

Over de samenhang tussen merkkeuze en marketingvariabelen kan het volgende worden gezegd. Het bleek, dat voor de meeste merken er een signifikante invloed was van prijs en/of reclame op grootheden als het marktaandeel, de kans op een herhalingsaankoop van hetzelfde merk en de kans om op een ander merk over te gaan. Verder werd gevonden dat speciale aanbiedingen relatief vaak gepaard gaan met merkwisselingen, zodat deze aanbiedingen een middel lijken om kopers van merk te doen veranderen. Met betrekking tot de invloed van de lengte van de tijd tussen 2 opeenvolgende aankopen werd voor fopro gevonden, dat de kans om hetzelfde merk te kopen als bij de voorgaande aankoop afneemt, naarmate de tussentijd langer is. Voor bier en margarine kon geen invloed van de tussentijd worden vastgesteld.

In hoofdstuk 8 werd de samenhang tussen merkkeuzegedrag en een aantal kenmerken van de huishouding bestudeerd. Het blijkt, dat socio-economische variabelen slechts in geringe mate samenhang vertonen met merkkeuze variabelen, zoals aantal verschillende merken, aandeel 1e merk e.d. Voor alle drie produkten werd enige invloed waargenomen van woonplaatsgrootte, regio, het wel of niet hebben van kinderen en attitude scores met betrekking tot koopgedrag. In een aantal gevallen was er ook enige invloed van gezinsgrootte, van de leeftijd van de huisvrouw en van het bezitten van een koelkast en een televisietoestel.

Tussen merkkeuze variabelen en andere koop-variabelen bleken nauwere samenhangen te bestaan. Er werd geconstateerd, dat huishoudingen, die relatief veel merkverandering vertonen een lagere prijs betalen, vaker speciale aanbiedingen kopen, meer in zelfbedieningswinkels of supermarkten kopen, meer in filialen van grootwinkelbedrijven of in vrijwillig filiaalbedrijven kopen, meer variatie vertonen in tussentijden tussen 2 opeenvolgende aankopen en in de per aankoop gekochte hoeveelheid en dat deze huishoudingen meer verschillende verpakkingseenheden kopen.

Tenslotte werd gevonden dat merkkeuzegedrag tot op zekere hoogte transitief is over produkten, d.w.z. tot op zekere hoogte vertonen huishoudingen hetzelfde type merkkeuzegedrag met betrekking tot verschillende produkten. Dit is echter niet in die mate het geval dat er kan worden gesproken van een algemeen merkkeuzegedrag van een huishouding, welke kan dienen als een zelfstandige basis voor marktsegmentatie.

Curriculum vitae

- Geboren te Zoutkamp 28 juni 1945
- H. de Cockschool voor ULO te Ulrum: 1957-60
- 2e Chr. HBS te Groningen: 1960-62
- Studie Landbouwhogeschool, richting Landhuishoudkunde: 1962-68
- Sinds 1 september 1968 werkzaam als wetenschappelijk medewerker bij de Afdeling Marktkunde en Marktonderzoek der Landbouwhogeschool

(B 1544) Printed in Belgium by Ceuterick s.a.
Brusselse straat 153 B 3000-Louvain
A. Struyf Oude Baan 353 B 3040 Korbeek-Lo

List of abbreviations and variable names

NAME	SHORT DESCRIPTION (further information can be found on the pages indicated)	PAGE OF FIRST APPEARANCE

A Abbreviations for models

BLM	Brand Loyal Model	45
CES	Cohesive Elements Specification (of PDM)	150
CMP	Conditional Model Probability	36
EOS	Equal Occurrence of Sequences	39
HEBM	Heterogeneous Bernoulli Model	33
HEMM	Heterogeneous Markov Model	44
HOBM	Homogeneous Bernoulli Model	18
HOMM	Homogeneous Markov Model	19
IES	Independent Elements Specification (of PDM)	149
LLM	Linear Learning Model	47
LPLM	Last Purchase Loyal Model	45
PDM	Probability Diffusion Model	148
ROS	Relative Occurrence of Sequences	40

B Variable names

AGEHOU	Age of housewife	220
BRCH	Brand change	193
BRCHWSHCH	Brand change with shop change	193
BRPSH	Brands per shop	191
BRSHCH	Brand and shop change	193
BRTIED	Brand-tiedness	220
CHAINSH	Chain shops	221
CHILD	Children	220
DEAL	Deal purchases	221
DISTR	District	220
FAMSIZ	Family size	220
FIRBR	First brand	218
MAXPLZ	Maximum of poolsize	165
MEANPLZ	Mean of poolsize	162
MNB	Major national brand	237

NAME	SHORT DESCRIPTION (further information can be found on the pages indicated)	PAGE OF FIRST APPEARANCE
MST	Market share by number of purchase times	200
MSV	Market share by volume	200
NOCH	No change (neither of brand nor shop)	193
NUMBR	Number of different brands	186
NUMPU	Number of purchase occasions	221
NUMPUR	Number of purchasers	221
NUMSH	Number of different shops	186
NUMSIZ	Number of different package sizes	221
NUMVAR	Number of different varieties	221
PRICE	Average price paid	221
PRICON	Price consciousness	220
PRIV	Private brand (nominal category for the first brand of a household)	237
PRIVBR	Private brands (share in purchases)	218
RANGPLZ	Range of poolsize	162
REFRI	Refrigerator	220
REGHOU	Regularity in the household	220
SELSUP	Self-service shops and supermarkets	221
SFAVBR	Share of favorite brand by volume	186
SFAVBRN	Share of favorite brand by number of purchases	218
SFAVSH	Share of favorite shop by volume	186
SHCH	Shop change	193
SHCHWBRCH	Shop change with brand change	193
SHPBR	Shops per brand	191
SIZTOW	Size of town	219
SOCCL	Social class	219
TELEV	Television set	221
TRAD	Traditionalism	220
VARIPT	Variation in inter-purchase times	221
VARVOL	Variation in volume per purchase	221
VOLPP	Volume per purchase	221
VOLPU	Volume of purchases	221

B. Wierenga AN INVESTIGATION OF BRAND CHOICE PROCESSES