Classification and Selection of Consumer Purchase Behaviour Models

R. Decker / W. Gaul
Institut für Entscheidungstheorie und Unternehmensforschung
Universität Karlsruhe (TH)

Summary

Consumer purchase behaviour analysis situations can be characterized by two extremes: On the one side, a magnitude of models for analyzing data of consumer purchase behaviour is available, on the other side, potential users who are not so familiar with details of the corresponding methodology refrain from applying adequate models in day-to-day activities of market research and marketing. In this paper, we will give a classification of consumer purchase behaviour models and describe a selection procedure which - on the basis of the data provided and the market diagnostics desired - helps to find (an) appropriate model(s).
An example based on household panel data provided by a German market research institute is included for illustration.

Motivation

Models for the analysis of consumer purchase behaviour are of great importance for theoretical as well as practical considerations within market and marketing research. In Figure 1 the non-observability of the individual purchase decision process is emphasized by depicting a "black box" approach which is - indeed - unsatisfactory for many reasons.

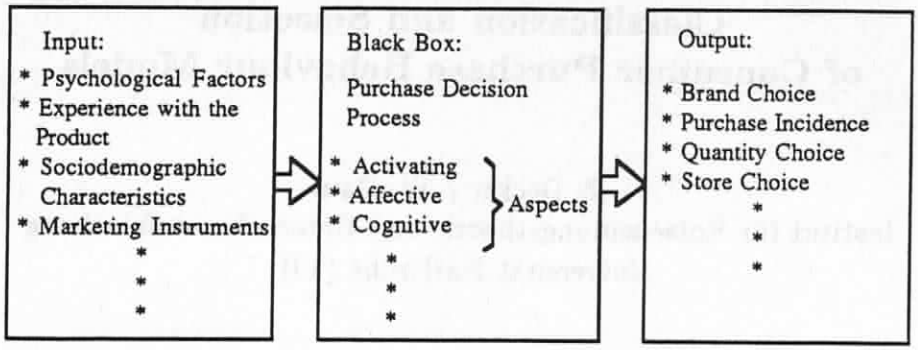

Figure 1: "Black Box" Approach of Individual Purchase Decision Process

Of course, efforts have been undertaken to design models which endeavour to illuminate this "black box" approach. Among others, we mention **simulation models** (see e.g. Amstutz (1967), Klenger/Krautter (1972) and Lavington (1972)), **structural models** (see e.g. Andreasen (1968), Engel/Blackwell/Miniard (1986), Howard/Sheth (1969) and Nicosia (1966) for system approaches, and Bettman (1970) as an example for a decision net model), and **stochastic models** which can be divided into econometric and "purely" stochastic models (see the next section for classification). **Store choice models** (see e.g. Fotheringham (1988) for a logit approach) and **quantity choice models** (see e.g. Paull (1978)) also belong to the set of "purely" stochastic models, however, in the next sections we will stress **brand choice** and **purchase incidence models** as well as models which are a **combination of both.** Distinctions can be made between homogeneous and heterogeneous, stationary and nonstationary, univariate and multivariate models of different order (see e.g. Hauser/Wisniewski (1982) for a nonstationary Semi-Markov Model, Jeuland/Bass/Wright (1980) for a Gamma Erlang/Dirichlet Multinomial Model, Jones/Zufryden (1980) for a nonstationary Gamma Poisson/Beta Binomial Logit Model, Wagner/Taudes (1986) for a nonstationary multivariate Polya Model and Zufryden (1977) for a condensed Gamma Possion/Beta-distributed Linear Learning Model).

This by far not complete list (in alphabetical order) of some references for brand choice/ purchase incidence models shows the need of some classification and selection technique which - based on appropriate classification criteria - aids potential users who are not so familiar with consumer purchase behaviour methodology in finding (an) adequate model(s) (for German language introductions at different levels into the area of consumer behaviour see e.g. Bänsch (1986), Berndt (1982), Böckenholt/Decker/Gaul (1989), Topritzhofer (1974) or Wagner (1985)).

Classification Structure and Selection Procedure for Consumer Purchase Behaviour Models

In this section we make an attempt to classify the different types of models mentioned above using a tree chart (see Figure 2).
We will restrict our discussion to the class of stochastic models and have a more detailed look at these models which often show considerable differences with respect to theoretical formulation and degrees of complexity.
To get a general impression of the procedure of determining (an) appropriate model(s) - for underlying data - we use a kind of flow chart (see Figure 3) which should help to find that (those) model(s) which fit(s) best and deliver(s) the most meaningful market diagnostics under given conditions. In the next section we will show an exemplary application of the selection procedure described in Figure 3.
The actual proceeding for the empirical example within this classification and selection approach is already depicted in Figure 2 in bold lines and rectangles.

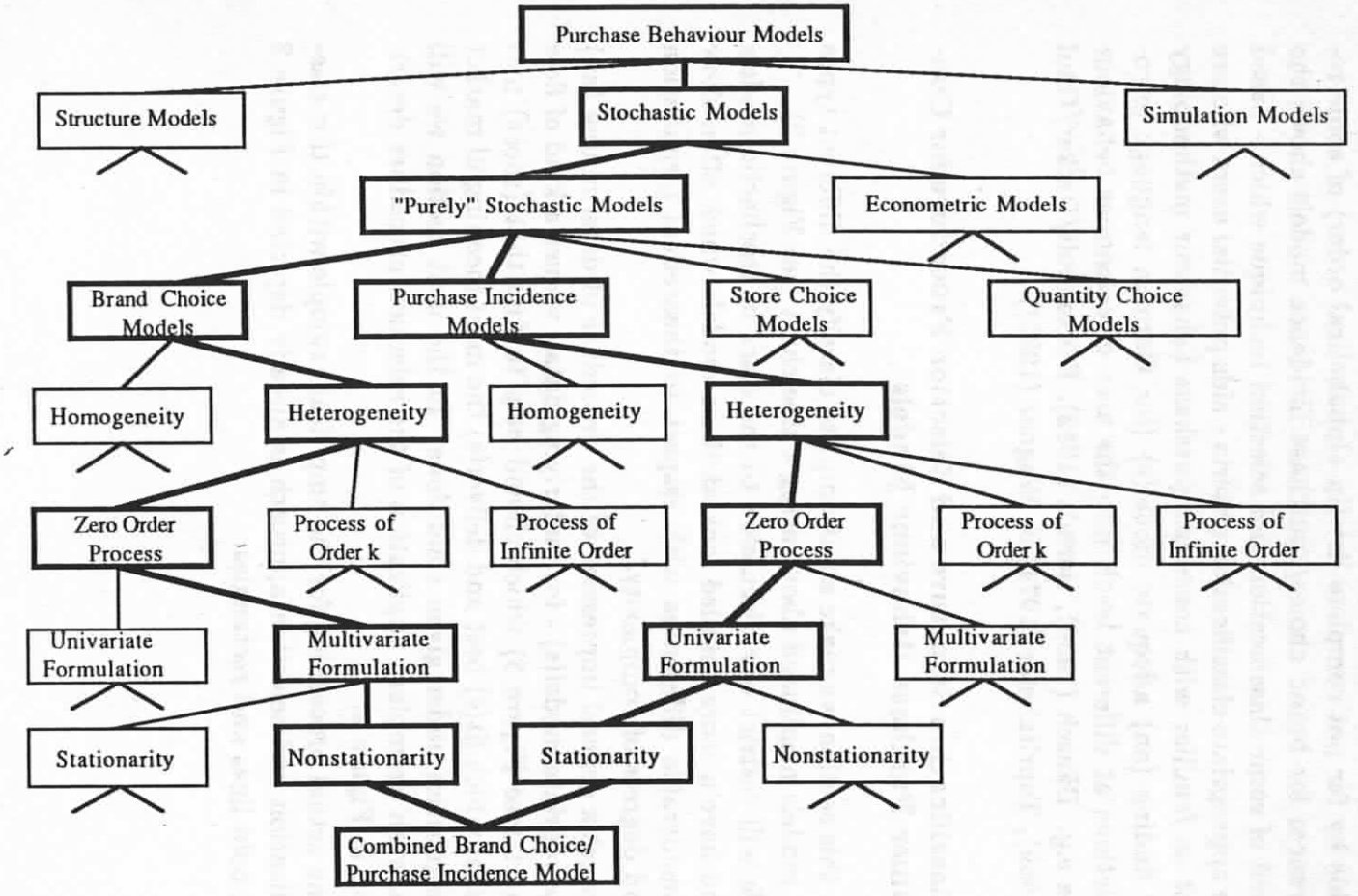

Figure 2: Part of the Classification Structure for Consumer Purchase Behaviour Models (actual proceeding for empirical example is depicted in bold lines and rectangles)

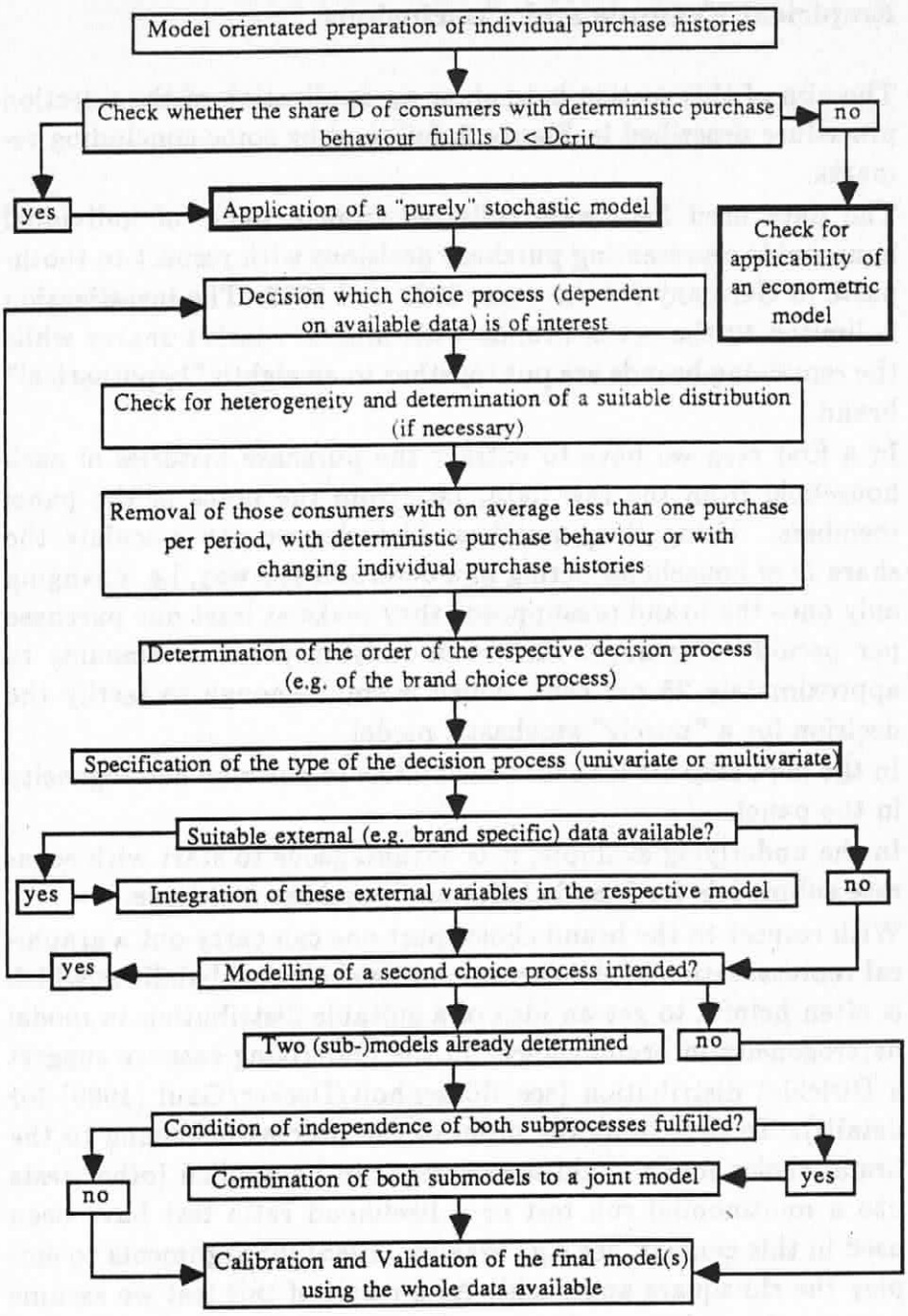

Figure 3.: Part of the Flow Chart for the Model Selection Procedure

Empirical Example and Conclusions

The aim of this section is to show an application of the selection procedure described in Figure 3, followed by some concluding remarks.

The data used here were collected from a panel of individual households representing purchase decisions with respect to toothpaste in Germany for the years 1981 and 1982. The investigation is limited to the seven brands with highest market shares while the remaining brands are put together to an eighth "hypothetical" brand.

In a first step we have to extract the purchase histories of each household from the raw data, i.e. from the notes of the panel members. Using these purchase histories we can calculate the share D of households acting in a deterministic way, i.e. changing only once the brand presupposed they make at least one purchase per period on average. In the underlying panel D amounts to approximately 25 per cent, which is small enough to justify the decision for a "purely" stochastic model.

In the next step we look for possibilities to describe heterogeneity in the panel.

In the underlying example, it is advantageous to start with separate submodels for brand choice and purchase incidence.

With respect to the brand choice part one can carry out a graphical representation of the corresponding choice probabilities which is often helpful to get an idea of a suitable distribution to model heterogeneity in brand choice. In the underlying case we suggest a Dirichlet distribution (see Böckenholt/Decker/Gaul (1989) for details). To determine the order of the process belonging to the brand choice model a chi-square test can be applied (other tests like a multinomial run test or a likelihood ratio test have been used in this context, see e.g. Wagner (1985) for arguments to employ the chi-square approach). As a result of this test we assume a zero order process and - as we have a multibrand approach - a Multinomial distribution.

In the underlying case, additional external variables (e.g. time-

dependent prices, display and advertising expenditures) are available. This information can be used for a reparametrization of the brand choice submodel leading to a nonstationary formulation (see Böckenholt/Decker/Gaul (1989) for details).
Using the same proceeding with respect to the purchase incidence part we suggest a Beta-distribution for heterogeneity and, again, a zero order process and - as we have a binary decision for the purchase situation - a Binomial distribution.
Assuming independence of both subprocesses a combination of the submodels to a joint "Brand Choice/Purchase Incidence Model" is possible.
To sum up, we get a nonstationary Beta Binomial/Dirichlet Multinomial (BB/DM) Model.
Without the described classification and selection procedure a potential user could have chosen one of the (adequate) models from the literature, e.g. the Negative Multinomial (NM) Model following an approach of Chatfield/Goodhardt (1975).

	Brand 1			Brand 5			Brand 8		
Number of purchases	Observ Freq.	Nonstat. BB/DM	Stat. NM	Observ Freq.	Nonstat. BB/DM	Stat. NM	Observ. Freq.	Nonstat. BB/DM	Stat. NM
0	1401	1393.8	1396.3	1764	1781.1	1772.7	979	961.3	967.4
1	240	245.9	269.8	76	78.7	114.0	302	310.6	378.6
2	117	125.9	127.4	60	49.2	46.5	229	225.4	216.3
3	74	76.2	72.0	24	27.8	24.3	147	147.9	136.5
4	42	46.5	44.1	22	17.0	14.1	105	108.8	90.3
5	46	32.3	28.2	16	10.8	8.6	73	76.9	61.3
6	22	21.5	18.6	11	7.0	5.5	63	53.1	42.4
7	15	14.3	12.5	4	4.5	3.6	31	33.4	29.7
8	10	9.5	8.5	4	2.9	2.4	17	22.9	21.0
9	5	6.2	5.9	1	1.9	1.6	15	15.6	14.9
10	5	4.0	4.1	2	1.2	1.1	9	10.4	10.7
11	1	2.6	2.9	0	0.7	0.8	5	6.8	7.7
12	6	5.9	7.0	0	1.1	1.8	9	10.9	20.3
chi-square		8.7	18.7		12.6	37.5		5.5	40.1

Table 1: Observed and Estimated Frequencies of Purchases for some Representative Brands

In Table 1, the frequencies of the purchases (observed vs. estimated by the BB/DM and NM models) are shown. While we leave it to the reader to recognize the discrepancies of Table 1, in Table 2 some goodness-of-fit measures (chi-square (DF), p-value, Theil inequality coefficient) are displayed. Table 1 and 2 are showing in

Model / Goodness-of-fit	Nonstat. BB/DM Model	Stat. NM Model
chi-square value (DF)	86 4 (72)	158.2 (87)
p-value	0.12	0.0
Theil.I.C.	0.05	0.1

Table 2: Goodness-of-fit of the Models

an impressive manner the very different fit of both models. With respect to other diagnostic indicators similar differences can be observed. Of course, if we would have used a more inadequate model for comparison (more inadequate than the NM Model) the discrepancies would have been still more obvious. However, already this very short example is able to demonstrate the importance of a systematic model classification and selection approach in order to get (an) adequate model(s) and corresponding model output from which we have only shown estimated purchase frequencies. Therefore, before the application of a consumer behaviour model - regardless of which type - the user should specify in an accurate way the data available and the diagnostics desired.

Using an adequate stochastic model he is in a position to calculate a multitude of interesting market diagnostics like **market shares, penetrations, switching** or **repeat buying probabilities** and **elasticities.**

Additionally, these models allow to predict **future market situations** and to carry out **sensitivity analyses** in order to deduce **promising marketing strategies.**

At the moment, we try to implement the classification and selection procedure on workstations.

References

Z, A. E.: Computer Simulation of Competitive Market Response, M.I.T Press, Cambridge (Mass.) London, 1967.

ANDREASEN, A. R.: Attitudes and Customer Behaviour: A Decision Model, in: Kassarjian, H.H./Robertson, T. S. (Eds.) Perspectives in Consumer Behaviour, , Scott, Foresman & Co., Glenview (Ill.), 1968, 498 - 510 .

BÄNSCH, A.: Käuferverhalten, 3. Aufl., Oldenbourg, München, 1986.

BERNDT, R.: Stochastische Modelle des Käuferverhaltens als Grundlage für absatzwirtschaftliche Entscheidungsmodelle, in: Operations Research Proceedings, Springer, Berlin Heidelberg, 1982, 149 - 160.

BETTMAN, J. R.: Information Processing Models of Consumer Behaviour, Journal of Marketing Research, Vol. 7, 1970, 370 - 376.

BÖCKENHOLT, I./DECKER, R./GAUL, W.: Ein Modell zur Erfassung nichtstationären, heterogenen Kaufverhaltens bei regelmäßig gekauften Konsumgütern, Working Paper, Institute of Decisions Theory and Operations Research, University of Karlsruhe (TH), 1989.

CHATFIELD, C./GOODHARDT, G. J.: Results Concerning Brand Choice, Journal of Marketing Research, Vol. 12, 1975, 110 - 113.

ENGEL, J. F./BLACKWELL, R. D./MINIARD, P.W.: Consumer Behavior, Dryden Press, CBS Publishing, Chicago, 1986.

FOTHERINGHAM, A. S.: Consumer Store Choice and Choice Set Definition, Marketing Science, Vol. 7, 1988, 299 - 310.

HAUSER, J. R./WISNIEWSKI, K. J.: Dynamic Analysis of Consumer Response to Marketing Strategies, Management Science, Vol. 28, 1982, 455 - 486.

HOWARD, J. A./SHETH, J. N.: The Theory of Buyer Behavior, John Wiley & Sons, New York, 1969.

JEULAND, A. P./BASS, F. M./WRIGHT, G. P.: A Multibrand Stochastic Model Compounding Heterogeneous Erlang Timing and Multinomial Choice Processes, Operations Research, Vol. 28, 1980, 255 - 277.

JONES, M. J./ZUFRYDEN, F. S.: Adding Explanatory Variables to a Consumer Purchase Behavior Model: An Exploratory Study, Journal of Marketing Research, Vol. 17, 1980, 323 - 334.

KLENGER, F./KRAUTTER, J.: Simulation des Käuferverhaltens, Teil I: Werbewirkung und Käuferverhalten; Teil II: Analyse eines Kaufprozesses; Teil III: Computermodell des Käuferverhaltens, Gabler, Wiesbaden, 1972.

LAVINGTON, M. R.: Ein Mikrosimulationsmodell der Nachfragereaktionen beim Konsumgütermarketing, in: Kroeber-Riel (Hrsg): Marketingtheorie - Verhaltensorientierte Erklärungen von Marktreaktionen, Kiepenheuer & Witsch, Köln, 1972, 332 - 358.

NICOSIA, F. M.: Consumer Decision Processes - Marketing and Advertising Implications, Prentice Hall Inc., Englewood Cliffs, 1966.

PAULL, A. E.: A Generalized Compound Poisson Model for Consumer Purchase Panel Data Analysis, Journal of the American Statistical Association, Vol. 73, 1978, 706 -713.

TOPRITZHOFER, E.: Absatzwirtschaftliche Modelle des Kaufentscheidungsprozesses unter Berücksichtigung des Markenwahlaspektes, Verlag der Österreichischen Akademie der Wissenschaften, Wien, 1974.

WAGNER, U.: Vollstochastische Kaufverhaltensmodelle - Ihr Beitrag zur Analyse realer Märkte, Anton Hain, Königstein/Ts., 1985.

WAGNER U., TAUDES, A.: A Multivariate Polya Model of Brand Choice and Purchase Incidence, Marketing Science, Vol. 5, 1986, 219 - 244.

ZUFRYDEN, F. S.: A Composite Heterogeneous Model of Brand Choice and Purchase Timing Behaviour, Management Science, Vol. 24, 1977, 121 - 136.