Approaches to Analyzing Brand Switching Matrices

Richard Colombo and Darius Sabavala

Introduction

Brand switching matrices are square tables that display customer purchases on two different, usually contiguous, occasions. Marketers have used such tables in at least three ways. First, by examining substitution patterns — those brands to which and from which consumers tend to switch — marketers can identify the major competitors of a brand. Second, consumer loyalty and customer retention can be assessed from the degree of repeat buying as measured by the diagonal elements of the table. Third, the tables can be used for forecasting since they show losses and gains of each brand from every other.

These matrices capture switching behavior in various ways. The most direct way is to observe switching from one purchase occasion to the next. But switching behavior can also be observed from one time period to another. And, if observation is difficult, as it might be for durable goods with long inter-purchase times, brand switching matrices can be constructed by tabulating answers to questions such as "what was the last brand you bought?" and "what brand do you intend to buy next time?"

The flexibility and usefulness of switching matrices have made them familiar to marketers. But to go beyond the data how do, or should, brand-switching matrices be analyzed (or correspondingly how should switching behavior be modeled)? To answer this question, we asked a group of well-known marketing scientists to analyze a set of similar switching matrices using their own preferred technique. But switching matrices have seen uses in many other fields so that statisticians are also familiar with methods for their analysis. We therefore contacted some statisticians as well as marketers. In all, we received responses from twenty-one out of the thirty-five or so people we approached. The responses are compiled here.

This document can be used by the reader in three ways. First, it serves as a catalogue of approaches to analyzing switching data. You will find here a variety of methods, one of which may strike you as particularly suitable for a particular task at hand. The methods are by no means the only ones that can be used for analyzing switching matrices, but we feel they do represent a fair range of approaches. A second use is to assess what we know about purchasing behaviour. Some of the methods are based on more or less explicit assumptions about how consumers behave. Those assumptions are in some cases incompatible with each other. Empirical work is needed to determine which assumptions are reasonable in any particular context and therefore what methods are useful to marketing decision makers. The third use is to serve as a window into the thought processes of the contributors. In fact, this was our primary interest. The way the problem was approached is at least as interesting as the actual methods used and touches on interesting questions about how marketing science, and statistical analysis, should be done. We touch on some of these issues in section our discussion section below, but our views are much more fully set out elsewhere (Colombo, Ehrenberg and Sabavala, 1994).

To help the reader navigate through the different contributions, we have constructed a framework into which we put the different approaches. We also attempt to show that many of the approaches are related (although, at first sight this is not obvious). We then discuss briefly some modeling issues. We hope

that this will help you assess the different approaches. The contributions themselves can be found by following the hyperlinks in the framework. We also explain how we collected the contributions and show the data the contributors were asked to analyze. This is important for understanding the context in which the contributors worked.

How the study was conducted

About 35 experienced statisticians and marketing scientists, including leading authors on statistical contingency tables, data analysis and market structure analysis, in Europe and North America were asked to analyze the data using their own preferred method. We received about 21 responses which we think are a pretty good portrayal of different approaches to analyzing switching matrices. The people we contacted were a judgment sample: we hoped it represented a good cross-section of people interested in these kinds of data and problems. Certainly, they produced a great variety of methods and answers.

We asked the analysts to apply their own preferred technique or analysis/modelling approach (e.g., GLIM, correspondence analysis, cluster analysis, EDA, IDA, market structure/partitioning etc.) to the same "simple" set of data. The aim was that the results could then be compared. Note that we were asking analysts to use their "preferred" method, i.e., a method that they knew well and felt comfortable with. We were not asking them to use what they thought might be the best method.

We suggested that the data we provided (on diskette and on paper) should be able to tell us something about the competitive structure of the car market. Two broad marketing questions that might be asked of these data by a manufacturer might be

i. i. Who are my main competitors?

ii. ii. Should we aim at strengthening the loyalty of our existing customers or increase the chance of competitors' customers switching to us next time they buy the car?

Included with the data were two recent papers that analyzed contingency tables of car switching - one by Van de Heijden et al. and one by Colombo & Morrison. Finally we suggested some questions or criteria that might be used in assessing an approach or technique, viz.

iii. iii. What kinds of dimensions and/or structures emerge.

iv. iv. How far the original data have been "modelled"

v. v. How quantitative or qualitative the answers are.

vi. vi. How comparable across different replications they are.

vii. vii. How explicit and/or deep any "theory" is, and how generalizable.

viii. viii. How parsimonious the theory or model is.

ix. ix. How transparent and communicable the results are.

x. x. What criteria of success are assumed (and/or satisfied)

Our instructions to the analysts we hoped would to orient them to the problem. Because we asked analysts to apply their preferred technique, we may have artificially obtained a greater variety of responses than if we had asked the same analysts to use what they thought was the best method. On the other hand, because we suggested two research questions and possible criteria by which the contributions might be judged we provided a focus for the analysts which could have served to lead them to use a narrower range of methods.

The Data

The data are from surveys of buyers of new cars (in the previous 12 months) with the make of the new car cross-tabulated with the make of the car previously owned. There are eight sets of data, four from France and four from Britain. The data for France 198a are shown below. The entire data are in the appendix.

France 1989

	Alfa	BMW	Cit	Fiat	Ford	GM	Lada	Merc	Peu	Ren	Rov	Saab	Seat	VW	Volvo	Total
Alfa	82	6	16	16	15	11	0	3	37	24	4	1	7	25	1	256
BMW	1	138	1	7	6	1	0	16	22	31	3	2	4	13	3	257
Cit	7	16	1437	96	100	62	19	18	283	310	40	2	22	107	20	2589
Fiat	8	7	46	358	38	33	6	5	67	116	19	1	24	29	5	798
Ford	7	8	67	39	494	42	6	4	71	96	12	1	8	32	10	927
GM	3	5	19	7	28	245	2	2	41	42	4	0	7	20	6	442
Lada	0	0	9	11	7	5	57	0	4	17	2	0	0	3	0	123
Merc	0	11	2	1	3	1	0	89	5	19	0	0	4	12	4	153
Peu	13	43	251	122	239	130	40	29	2310	653	48	2	34	148	30	4176
Ren	21	49	312	219	310	295	44	34	817	3944	88	3	87	241	34	6629
Rov	2	2	3	15	14	12	0	1	24	31	77	1	0	5	1	203
Saab	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Seat	0	0	0	4	1	0	0	0	9	5	0	0	17	1	0	39
VW	8	23	22	26	45	31	4	9	116	78	12	2	19	477	12	917
Volvo	1	4	0	3	3	4	0	2	4	12	2	1	4	0	49	107
Total	177	354	2482	1180	1510	1020	226	250	4397	6058	411	23	262	1279	209	20532

A framework

Although the contributions were quite varied, it seemed to us that they could be put into one of three broad categories according to whether they used (a) Exploratory methods; (b) Formal techniques; or (c) A prior model. Those authors that took an exploratory approach were unwilling at the outset to use a particular technique or to use whatever knowledge about consumer switching that they had. We classified approaches as using formal methods if either standard software was used or if some explicit statistical model was used, even if the analysis seemed to be done in an exploratory spirit. Where an explicit and existing model of consumer behavior guided the analysis, we classified the approach as using a prior model.

The framework is subjective and there is no doubt room to debate whether such and such approach should more accurately be placed in a different category. Our intent is only to provide some order for looking at the contributions; we would probably not take exception to other classifications.

Classification of Approaches

(a) Exploratory		
	Laurent	Graphical
	Chatfield	Tabular
	Gower & Zielman	Exploratory modelling
(b) Formal methods:		
Standard	Everitt	Cluster analysis; MDS
	Aurier	MDS
	Lebart	Correspondence Analysis
	Morgan	Asymmetric analysis
	Nelder	GLIM
	Goodman	Contingency table analysis
	Jain & Vilcassim	Latent Class Analysis
Non-standard	Cooper	Scaling
	Novak	modified quasi-symmetry
	Hoffman & deLeeuw	restricted quasi-symmetry
	Marcati	network analysis
(c) Prior model		
	Ehrenberg & Pouilleau	(initially exploratory)
	Bemmaor	
	Kalwani, Kannan & Lim	with cluster analysis
	Mirkin	
	Phillips	
	Chandrasekharan & Wright	Mover-stayer
	Morrison & Colombo	Mover-stayer

The table above groups analyses by the kind of approach taken. In the table below we attempt to show the similarities among contributions. The models differ largely in the way they handle the interaction terms in a table.

$p_{ij} = p_i p_j \gamma_{ij}$ where $p_{i.}$ and $p_{.j}$ are the column and row marginals and g_{ij} are the interaction terms		
$\gamma_{ij} = 1$	independence	
$\gamma_{ij} = \begin{cases} 1 & i = j \\ \frac{1-D}{p_i} + D & i \neq j \end{cases}$	Duplication Law	Bemmaor, Ehrenberg, Jain, Kalwani, Philips
$\gamma_{ij} = \begin{cases} \theta & i \neq j \\ 1 & i = j \end{cases}$	Quasi-independence	Chandrasekharan & Wright, Gower, Mirkin, Morrison & Colombo
$\gamma_{ij} = 1 \quad \forall i, j \in S$	Block-independence	Nelder
$\gamma_{ij} = \gamma_{ji}$	Quasi-symmetry	Hoffman, Novak

$\gamma_{ij} = \rho\varepsilon_i\delta_j$	RC model; Correspondence analysis	Goodman, Lebart, Aurier
Distance measures		Aurier, Everit, Kalwani, Mirkin
Other		Chatfield, Laurent

Some modeling issues

In this section we suggest a number of issues and criteria that readers may wish to keep in mind as they consider one of the methods or compare two or more methods. Clearly depending on the reader's interests some of these may be more important than others.

i. i. Does the "story" that accompanies the model or the analysis capture the essence of the problem?

ii. ii. Can the assumptions, procedures and results be easily communicated to other analysts and to decision makers. Are the "pictures" of the "story" easily understood. By pictures we mean not just graphs, charts and the like, but also "graphical tables".

iii. iii. Are the parameters and results directly related to the behaviour of consumers, and are they managerially useful? In particular, can the reader discern information about the role of market share, repeat rates, loyalty, propensity to switch, competitive structure among makes, changes over time?

iv. iv. What are the costs and difficulties of implementing the analysis. Are routines readily available or are special programs required?

v. v. Do the methods provide reasonable ways to predict and test? Do the measures of "fit" for the model provide some kind of managerial relevance as opposed to mere statistical significance?

vi. vi. Do the approaches naturally incorporate the data replicates and permit comparison across datasets (countries and years)?

vii. vii. Did the analyst attempt to generalize across the eight datasets?

viii. viii. Is the model well grounded empirically?

ix. ix. Did the analyst consider a "default analysis", i.e. one that on the basis of prior theory or knowledge could serve as a starting point?

x. x. Could simpler methods do as well ["Simplify as much as possible, but no more so". (Albert Einstein)]

xi. xi. Does the chosen method perform better than an available "benchmark"?

Some Aspects of the Contributions

Approach	Contributors	Statistical Aspects	Marketing Aspects
Exploratory	Chatfield	Tabular exploratory Analysis; Row-column ordering; Use of replicates.	Repeat buying; Exceptional switching patters and Market structure.
	Gower & Zielman	Mover-stayer; Log-linear model; SVD; Symmetry and skew-symmetry.	Market structure; Market share dynamics.
	Laurent	Direct graphical analysis; Logarithmic transformation.	Repeat buying patterns; Switching proportional to share.
Formal Methods	Aurier	ALSCAL; Data transformation.	Market structure.
	Everitt	Cluster analysis; MDS Symmetric and asymmetric switching.	Market structure.
	Goodman	Mover-stayer; quasi-independence; Quasi-symmetry; RC models.	
	Jain & Vilcassim	Multinomial Dirichlet; Latent class analysis.	Zero-order choice; Choice-based segments.
	Lebart	Correspondence Analysis; Hierarchical clustering; Asymmetry.	Market structure; Market share changes.
	Morgan	Principal Components; Data transformation; Asymmetry Use of replicates.	
	Nelder	Log-linear model; Analysis of residuals; Use of replicates.	Market structure; Heterogeneity.
	Cooper	SVD; Mover-stayer; Scaling.	Pairwise switching clusters.
	Hoffman & de Leeuw	MDS; Asymmetry.	Perceptual similarity; Choice-based map.
	Marcati	Network analysis.	Market structure; Switching patterns
	Novak	Log-linear model; Quasi-symmetry; Additive trees.	Switching patterns.
Prior Model	Bemmaor	Multinomial-Dirichlet.	Switching-constant; Market structure;
	Chandrasekharan & Wright		Loyals, shoppers and repeaters
	Ehrenberg & Pouilleau	Multinomial-Dirichlet; Replication.	Switching and repeat buying; Double jeopardy; Market structure; Loyalty.
	Kalwani, Kannan & Lim	Cluster analysis; Multinomial-Dirichlet.	Forced switching; Market partitions.
	Mirkin	Correspondence analysis;	Switching patterns.
	Morrison & Colombo	Quasi-independence; Mover-stayer.	Loyals and switchers; Market structure
	Phillips	Log-linear model; Gravity model.	Brand shifting; Hendry switching constant; Market structure