The Dirichlet distribution as a model of brand choice : further testing

Albert C. Bemmaor

CERESSEC

Centre d'Etudes et de Recherche en Sciences Sociales, Economiques et Commerciales

THE DIRICHLET DISTRIBUTION AS A MODEL OF BRAND CHOICE : FURTHER TESTING

Albert C. Bemmaor

DR CERESSEC 81012 October 1981

Avenue de la Grande-Ecole - B.P. 105 - 95021 CERGY-PONTOISE CEDEX - France - Tél. (3) 030.40.57 - Télex ESSEC 697789 F

Association régie par la loi de 1901 - I.N.S.E.E. 972951270002 H

THE DIRICHLET DISTRIBUTION AS A MODEL OF BRAND CHOICE : FURTHER TESTING

Albert C. Bemmaor

DR CERESSEC 81012 October 1981

THE DIRICHLET DISTRIBUTION AS A MODEL OF BRAND CHOICE : FURTHER TESTING

Albert C. Bemmaor

October 1981

Albert C. Bemmaor is Associate Professor, Ecole Supérieure des Sciences Economiques et Commerciales (E.S.S.E.C.), Cergy-Pontoise 95021, France. The author acknowledges the helpful comments of Professors Frank Bass and Gordon Wright during the research, and the assistance of Mr Bernard Pinet and Mr Marcel Dufour of SECODIP Co. in providing consumer panel data. He is grateful to Professor Andrew S.C. Ehrenberg who wrote a long letter giving thorough comments on an earlier draft. He would like to thank Mr David Butler for sending him information on the Hendry system.

ABSTRACT

This paper provides further evidence on the appropriateness of the Dirichlet model for modeling brand choice. A simple estimation procedure is used and the model is tested on consumer panel data for three product categories (margarine, regular coffee and instant coffee). The results show that a) the model tends to underpredict the proportion of light users of a brand, b) compared to the Hendry model and to the switching law, it provides an acceptable fit to aggregate switching data over consecutive choice occasions, and c) it gives an adequate description of the correlation between relative frequencies of choice across consumers. Consistently with previous studies, the Hendry model tends to overestimate repeat purchase whereas the switching law, like the Dirichlet model, does not indicate any systematic deviation in its predictions of switching behavior.

THE DIRICHLET DISTRIBUTION AS A MODEL OF BRAND CHOICE : FURTHER TESTING

Over the past few years, many researchers have assumed that consumers are consistent with a heterogeneous zero-order process in their brand choice decisions (see for example, Herniter 1973, Bass 1974a). Runs tests on the order of the individual choice process generally supports this assumption (Massy 1966, Dodson 1975, Bemmaor 1978, Jeuland, Bass and Wright 1980). Consequently, the multinomial distribution has been used to describe brand choice behavior of individual consumers (Bass, Jeuland and Wright 1976 ; Ehrenberg and Goodhardt 1976) while the Dirichlet distribution has been taken as a model of market heterogeneity for brand choice probabilities. This distribution has been suggested to provide a theoretical rationale to the "switching law" first noted by Goodhardt (1966) for TV programs. He found that over a given time period, the number of viewers of program X who also watched program Y was proportional to the total number of viewers of program Y. Ehrenberg and Goodhardt (1968) tested this result in a brand choice context and corroborated it. The "switching law" has also been applied to choice behavior over consecutive purchase occasions. Formally, it states that aggregate conditional switching from brand i to brand j is proportional to the market share of j : $p_{j|i} = Kp_j$ where K is a constant ($0<K<1$) for all pairs of brands within the same set of competing brands, and p_j is the market share of j. This constant K has been empirically estimated by averaging the ratios $p_{j|i}/p_j$ over all pairs of competing brands. Bass et al. (1976) suggest another estimation method of K which consists of computing the ratio Total Switching/$(1-\sum_1 p_j^2)$. Ehrenberg and Goodhardt (1976) argue that both estimation methods are in fact equivalent. This paper provides further testing of the Dirichlet distribution on a new data set. In the first section, the mathematics of the model already shown in Mosimann (1962), Chatfield and Goodhardt (1975), and Bass, Jeuland and Wright (1976) are reviewed. In the second section, the

method-of-moment estimators of the parameters are used and the Dirichlet model is tested along three criteria :

a) fit to the empirical distribution of relative frequencies of choice ;

b) fit to actual switching data : the Dirichlet predictions are compared to those given by the "switching law" and by the Hendry model. This model is based on a set of premises about individual brand choice behavior which seem analogous to the multinomial assumption. A sketchy description of this model is given by Ehrenberg and Goodhardt (1974) and by the Hendry Corporation (1976) ;

c) fit to the empirical correlations between the relative frequencies of choice across consumers. The two preceeding alternative models, the "switching law" and the Hendry model, do not provide theoretical values for these correlations. Therefore, the Dirichlet predictions are evalued on their own.

Data

The data used in this study are SECODIP consumer panel data for three product categories (margarine, regular coffee and instant coffee). This panel includes households who continuously reported their purchases (date the product was bought, brand chosen, package size, store visited, price paid and total quantity bought) on a weekly basis. This data involves a total of 38,043 purchase occasions for margarine and 17,385 for instant coffee over a two-year period (1974 and 1975) and 19,416 purchase occasions for regular coffee over the first six-month period of 1974. There is a total of 1,207 households for margarine, 1,688 households for regular coffee and 1,257 households for instant coffee who are at least one-time buyers of the product class over the two-year period. Within each product class, the market shares of the major brands varied substantially from month to month but, in most cases, they did not exhibit any trend.

REVIEW OF THE MODELS

Following Bass, Jeuland and Wright (1976), assume that a consumer is characterized by a vector of scale utilities $[v_1,\ldots,v_i,\ldots,v_{N+1}]$ which describes his preferences over a set of (N+1) brands. According to Luce's axiom (1959), the probability that a consumer will purchase brand i on a purchase occasion is given by

$$\theta_i = \frac{v_i}{\sum_k v_k}$$

where the summation is over all values of k (k=1,...,N+1), v_i is positive and $0<\theta_i<1$. This definition of a probability implies that the odds ratio between two brands is *independent* of the presence or absence of non-chosen third alternatives. This is plausible for a market where brands are highly differentiated, but is questionable in the case where brands possess similar characteristics. For our purposes, Luce's theorem can be considered as a first approximation to choice behavior. Consumers vary in their preferences for brands which is reflected in the distribution of the v_i's over the population. Assume that these v_i's are independently distributed and follow a two-parameter gamma distribution with the same scale parameter (a_i, d_i = d) over the population so that the variability ratio (variance/mean) is assumed to be constant across brands (= d). The gamma distribution being quite flexible in its shape, this assumption is not severely constraining. It follows that the vector $[\theta_1,\ldots,\theta_N]$ is distributed Dirichlet over the population (Mosimann 1962 ; Rogers and Young 1973) with parameters $(a_1,\ldots,a_{N+1})$. Rogers and Young treated the general case where the v_i's have a generalized gamma distribution with parameters (a_i,d_i,e_i). But the increase in complexity in the functional form of the probability density function of $[\theta_1,\ldots,\theta_N]$ is rather substantial, and, at the present state of the study, does not seem to be warranted.

The probability density function of the Dirichlet distribution is given by

$$f(\theta_1,\ldots,\theta_N) = \left[\prod_{i=1}^{N} \theta_i^{a_i-1}/\Gamma(a_i)\right] \left(1-\sum_{i=1}^{N}\theta_i\right)^{a_{N+1}-1} \Gamma(c),$$

where $c = \sum_{i=1}^{N+1} a_i$ and all a_i's are positive. Each component θ_i has a standard beta (a_i, b_i with $b_i = c-a_i$) distribution with the following mean and variance

$$E(\theta_i) = \frac{a_i}{c} \tag{1}$$

and

$$\text{Var}\ (\theta_i) = \frac{a_i}{c^2}\frac{b_i}{(1+c)}.$$

Since $E(\theta_i^2) = \text{Var}\ (\theta_i) + E(\theta_i)^2$, the second moment about zero is

$$E(\theta_i^2) = \frac{a_i(1+a_i)}{c(1+c)}. \tag{2}$$

The joint probability that a consumer with given brand choice probability θ_i buys brand i twice over two purchase occasions is θ_i^2. Hence $E(\theta_i^2)$ is the probability that a randomly chosen consumer buys i twice over two purchase occasions. With consumer panel data, this probability may be assessed by measuring the proportion of pairs of purchases of the product class which are made of two consecutive purchases of brand i. The first moment about zero of the joint distribution of θ_i and θ_j is given by

$$E(\theta_i\theta_j) = \frac{a_i a_j}{c(1+c)},\ i \neq j. \tag{3}$$

$E(\theta_i\theta_j)$ represents the joint probability that a randomly chosen consumer buys i and j over two consecutive purchase occasions.

Note that this probability equals the probability of switching from j to i since $E(\theta_i\theta_j) = E(\theta_j\theta_i)$. It can be interpreted as being the proportion of pairs of purchases of the product class which are made of two purchases of brand i and brand j. With $K = \frac{c}{1+c}$, equation (3) becomes

$$E(\theta_i\theta_j) = K\, E(\theta_i)\, E(\theta_j) \qquad i \neq j \qquad (4)$$

where $E(\theta_i)$ and $E(\theta_j)$ are estimated by the market shares of i and j. Letting X_T denote the brand chosen on the T-th purchase occasion, from (4) we derive the aggregate conditional probability of switching from i to j over two consecutive purchase occasions

$$P(X_{T+1} = j \mid X_T = i) = \frac{E(\theta_i\theta_j)}{E(\theta_i)} = K\, E(\theta_j),\ i \neq j. \qquad (5)$$

This probability is proportional to the market share of j. For a given market share, it will be all the larger as K is large which led Bass, Jeuland and Wright (1976) to interpret K as a "brand loyalty" factor. The closer K is to zero, the higher the loyalty. Here the concept of loyalty solely involves the choices over two consecutive purchase occasions. It essentially reflects the distribution of choice probabilities across the population. As K moves closer to zero (c tends to zero), this distribution becomes more U-shaped.

From (2), we derive the conditional probability of repurchasing brand i

$$P(X_{T+1} = i \mid X_T = i) = \frac{E(\theta_i^2)}{E(\theta_i)} = (1-K) + K\, E(\theta_i).$$

The conditional probability of repurchasing brand i is a linear function of the market share of i which implies that large brands will have (in theory) larger conditional probabilities of repurchase ($0<K<1$) than small brands within the <u>same</u> set of competing brands.

Bass et al. (1976) showed that at the aggregate level, "the T-th step transition probabilities will depend on the entire (purchase) history". By definition, the correlation coefficient between θ_i and θ_j, ρ_{ij}, is equal to

$$\rho_{ij} = \frac{E(\theta_i\theta_j) - E(\theta_i)\,E(\theta_j)}{\sqrt{Var(\theta_i)}\ \sqrt{Var(\theta_j)}}.$$

Using (1) and (3), we find

$$\rho_{ij} = -\sqrt{\frac{a_i a_j}{(c-a_i)\,(c-a_j)}}. \qquad (6)$$

Expressed in terms of $E(\theta_i)$ and $E(\theta_j)$, (6) becomes

$$\rho_{ij} = -\sqrt{\frac{E(\theta_i)\,E(\theta_j)}{[1-E(\theta_i)]\,[1-E(\theta_j)]}} \qquad (7)$$

since $Var(\theta_i) = (1-K)E(\theta_i)\,[1-E(\theta_i)]$.

In total, this model involves the estimation of (N+1) independent parameters for a (N+1) - brand market. These parameters are the components of the vector $[a_1, a_2, \ldots, a_{N+1}]$.

ESTIMATION OF THE PARAMETERS AND TEST OF THE MODELS

Assuming that the vector $[\theta_1,\ldots,\theta_N]$ follows a Dirichlet distribution with parameters $(a_1,\ldots,a_{N+1})$, let $E(\theta_i)$ and $E(\theta_i^2)$ be the first and second moments about the origin for θ_i. Then, we have (see proof in the Appendix)

$$a_i = \frac{E(\theta_i)\ [E(\theta_1) - E(\theta_1^2)]}{Var(\theta_1)},\ i = 1,2,\ldots,N \qquad (A.3)$$

and

$$a_{N+1} = \left(1 - \sum_{i<N+1} E(\theta_i)\right) \frac{E(\theta_1) - E(\theta_1^2)}{Var(\theta_1)} \qquad (A.4)$$

Replacing $E(\theta_1)$, $E(\theta_1^2)$ and $E(\theta_i)$ in (A.3) and (A.4) by the sample estimates, we obtain the method-of-moment estimates of $a_i,\ldots,a_{N+1}$. These estimators are consistent but not generally efficient compared to the maximum likelihood (ML) estimators which are consistent, asymptotically normal and asymptotically efficient (Rao 1973, p.351). They are quite simple to compute by comparison with the ML estimators.[2] These parameters were estimated for the main national brands in the margarine, regular coffee and instant coffee markets. These estimates are shown in Table 1. Note that for all brands, these estimates are in the interval]0,1[which implies that the beta distributions are U-shaped. This result is consistent with other studies of purchase probabilities (see, for example, Kalwani and Morrison 1977a). Once the parameters have been estimated, the Dirichlet distribution has been tested along three criteria :

1. *First validation criterion* : Fit to the distribution of the relative frequencies of choice

As shown in Figure 1, the beta distribution consistently underestimates the proportion of light buyers of the major brands. This discrepancy does not seem to be due to the estimation procedure since the ML estimates for each brand are similar to the method-of-moment estimates. Assessing the fit of the Negative Binomial Distribution, Schmittlein and Morrison (undated) also underpredict the proportion of zeros. The discrepancy might be alleviated by including a mass point at zero for the beta distribution (see for example, Jones 1973).

2. *Second validation criterion* : Fit to actual switching matrices

Actual and theoretical conditional switching were compared across all three product categories. Because of space limitations, only the switching matrix for margarine is shown in Table 2. The Dirichlet model underestimates repeat purchase for all six brands studied.[3] These predictions are now compared to those given by the "switching law" and by the Hendry model.

The "switching law" has received strong empirical support (see, for example, Ehrenberg and Goodhardt 1974, Bass, Jeuland and Wright 1976, and Rubinson, Vanhonacker and Bass 1980). Formally, it states that aggregate conditional switching from brand i to brand j is proportional to the market share of j

$$p_{j|i} = K\, p_j, \qquad i \neq j.$$

This relationship implies that the aggregate joint switching probability is

$$p_{i,j} = K\, p_i p_j$$

which is analogous to (4). Using simple algebra, we derive the value of K,

$$K = \frac{\sum_i \sum_{j \neq i} p_{i,j}}{1 - \sum_i p_i^2} \qquad (7)$$

where the summations are over all brands in the market. K may be estimated by using total switching for the numerator and market share data for the denominator. Actual and theoretical switching for margarine are compared in Table 2. The fit is better than that of the Dirichlet model since the model does not exhibit any systematic discrepancy from actual switching.

The empirical implications of the Hendry model are similar to those of the Dirichlet model and of the switching law. In particular, the aggregate switching predictions over two consecutive purchase occasions are equivalent to those of equation (4) where the theoretical K, denoted by K_w in Hendry terminology, is now equal to

$$K_w = \frac{1 - \sum_i \frac{p_i}{1-p_i \ \ell n \ p_i}}{1 - \sum_i p_i^2} \qquad (8)$$

ℓn being the neperian logarithm. Replacing p_i by the market share of i, one only needs market share data to predict switching and repeat purchase. As indicated by Ehrenberg and Goodhardt (1974), the numerical value of K_w depends on two factors, the *number of brands* in the market and the *size of the share* of the leading brand. According to these authors, K_w takes on a maximum numerical value of about 0.54, the largest values occuring when the number of brands is small and the share of the leading brand is large. In theory, K_w varies within the interval]0,1[which might explain some cases where the Hendry model does not fit the switching data.[4]

For the margarine market, K_w computed from formula (8) equals 0.352. Actual and theoretical switching in the margarine market are compared in Table 2. Similarly to the Dirichlet model, the Hendry system underestimates repeat purchase.

The performance of each model across product categories is summarized in Table 3 where the percentage point deviations from actual are shown. As indicated previously, the discrepancies are small, the maximum deviation being 1.85 percentage point. For the Dirichlet model and the switching law, no systematic deviation in the sign of these deviations emerges across product categories which is consistent with previous results (see, for example, Jeuland 1979). On the other hand, the Hendry model tends to overestimate repeat purchase which corroborates the results found by Ehrenberg and Goodhardt (1974).

Overall, all three models give a fairly good fit to actual switching. The performance of the "switching law" compared to that of the two other models is not surprising since it includes one additional independent parameter, that is K. The test which consists of predicting the switching matrix tends to give a slight edge to this "switching law" against the other two models. The reason for including this "law" within the set of models is to assess the marginal improvement in fit due to this parameter K.[5] The Hendry model tends to overestimate repeat purchase. Butler and Butler of the Hendry Corporation might argue that this discrepancy shows that the brands do not belong to the same partition, which might lead someone to wonder about the degree of falsifiability of the model. Given that K_w is computationally bounded to a 0.54 value, cases may occur where the poor fit is simply due to a poor estimator of K_w rather than to a lack of representativeness of the model. The larger the parameter c is (in the limit, as c tends to infinity, K tends to 1), the poorer the fit of the Hendry system is. As c increases, the beta distribution becomes more and more humped. Luckily, this does not seem consistent with most of the evidence (usually a_i and b_i are between 0 and 1 so c is between 0 and 2).

For the Dirichlet model, the test used here is quite stringent since no single figure form the switching matrix is used in estimating the parameters of the model. Therefore, this model performs quite well compared with the Hendry system which is based on market share data taken from the switching matrix, and with the "switching law" which, in addition, uses total switching. Its advantage over this law" is that it provides a theoretical rationale to the empirical observations. Contrary to the Hendry model, the constant K takes on values within the closed interval [0,1]. It makes explicit assumptions about the form of heterogeneity in the population : whether the distribution of the relative frequencies is U-shaped, humped, J-shaped or inverted J-shaped, depending on the values of the parameters a_i and c (Johnson and Kotz, 1970).

3. *Third validation criterion* : Fit to the correlations between relative frequencies of choice

The expression for the correlation between brand choice probabilities shown in (7) is tested on the three product categories. Out of a total of sixty six comparisons between actual and predicted pairwise simple correlation, we found that, using the z-transform, only one difference was significant at the 5 per cent significance level (Rao 1973, p.434). Because of space limitations, these correlations are not shown here. In Table 4, the average correlations between relative frequencies of choice are compared to their theoretical values for the main national brands in each market. Across product categories, the model does not exhibit systematic discrepancies (nine underpredictions vers twelve overpredictions). The fit is fairly good : the average deviation from actual equals 9 per cent of the actual correlation.

Overall, the Dirichlet model provides a fair description of market heterogeneity. It tends to underestimate the proportion of light buyers.

Tested on switching data, it was compared to the "switching law" and to the Hendry model. Both of these models do not make explicit assumptions about market heterogeneity, and therefore were not tested further. For the three product categories under study, the Hendry model tends to overestimate repeat purchase whereas the Dirichlet model and the switching law do not indicate any systematic deviation from actual switching.

CONCLUSION

The fit of the switching law is generally good : for the three product categories under study, the theoretical values were in close agreement with actual switching. The Hendry model seems less flexible than the Dirichlet model to the extent that the switching constant K_W, as estimated in most publications discussing this model, cannot take on numerical values larger than 0.54. Since K_W should theoretically belong to the interval $]0,1[$, the entropy-based estimator does not seem appropriate. This Hendry model indicates systematic discrepancies in its predictions. The Dirichlet model compare favorably to both alternative models when tested on switching data. However, it tends to underestimate the proportion of light buyers of a brand. Its predictions of the correlations between relative frequencies of choice across consumers are adequate.

The main purpose of using such models is to reduce vast amounts of data such as consumer panel diaries to a smaller set and still be able to recover the original data with a fairly good degree of accuracy. Their managerial significance might be argued since policy variables are not included. But, they allow to specify equilibrium market conditions. The Dirichlet model is a simple, flexible, empirically-based choice model. Even though market aggregates did actually fluctuate from month to month, the Dirichlet model -with the assumptions of stationarity and constant individual probabilities of brand choice- still predicted aggregate choice fairly well. Further testing of this parsimonious model under a wider range of conditions is still needed. Decision variables might also be integrated into this stochastic modeling framework to increase its practical relevance. Work along this line is currently under way. We might safely predict that the inclusion of these variables will not add much to the overall goodness of fit of the Dirichlet model. To some extent, purchase probabilities already contain decision variables in themselves. The explicit allowance for managerial control might help the researcher better understand the formation of these purchase probabilities.

FOOTNOTES

1. This derivation follows the proof by Fielitz and Myers (1975) for the bivariate case.

2. Jeuland (1979) suggests a maximum likelihood estimation procedure of the Dirichlet model based on aggregate switching data. This method was applied to this data set and yielded the following parameter vectors : [.145, .051, .021, .004, .037, .097] for margarine, [.074, .038, .012, .014, .013, .015, .024, .010, .435] for regular coffee and [.096, .147, .054, .031, .031, .046, .071, .121, .236] for instant coffee. The estimated K values are 0.332, 0.384 and 0.454 for each market respectively (IMSL 1977). These parameters are of the same order of magnitude as the method-of-moment parameter estimates.

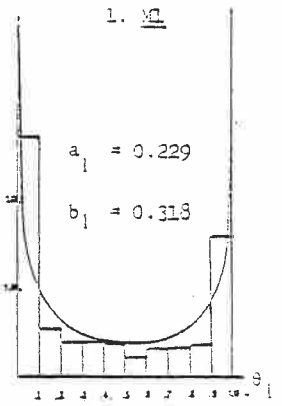
3. The fact that a model exhibits systematic deviations is not by itself a bad thing to the extent that the deviations are <u>really</u> systematic. Estimates of lower (or upper) bounds to repeat purchase might be, in some cases, managerially useful. The hard question is to find out whether the deviations are systematic by nature (or by chance).

4. It might also be argued that formula (8) for K_w is mathematically incorrect since it leads to bounded empirical values by comparison to the theoretical range of values. As no complete information on the Hendry model has been published, this argument is rather difficult to support.

5. One might contend that all three models are in fact the <u>same</u> since they all reduce to the same analytical form (4). Kalwani and Morrison (1977) argue that the Dirichlet model and the Hendry model are analogous. In a comment, Ehrenberg and Goodhardt (1979) showed that the Hendry model is quite distinct from the switching law since no empirical switching intervenes in the computation.

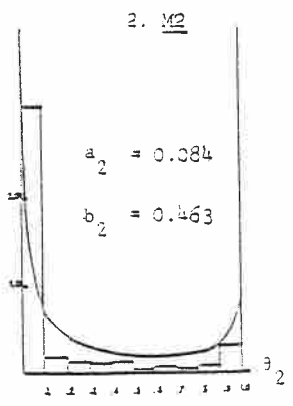
A. Margarine

$f(\theta_1|a_1,b_1)$



χ^2 = 9.095 (d.f.=7, p<.250)

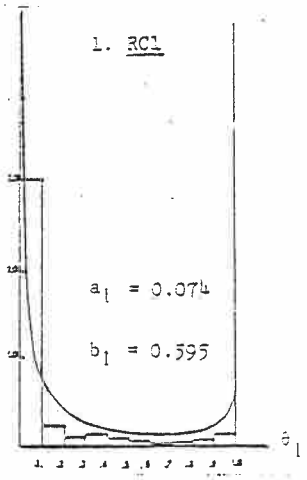
$f(\theta_2|a_2,b_2)$



χ^2= 25.574 (d.f.=7, p<7.6x10^{-4})

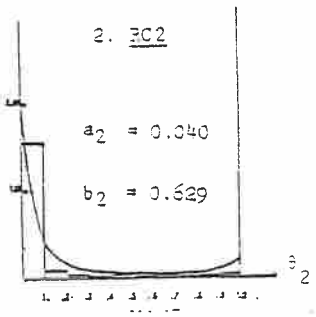
B. Regular coffee

$f(\theta_1|a_1,b_1)$



χ^2= 26.968 (d.f.=7, p<5x10^{-4})

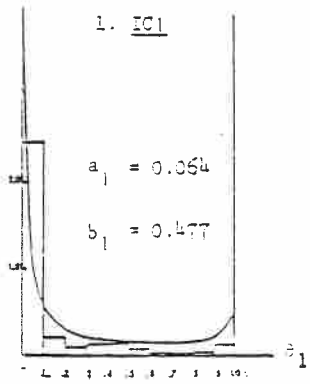
$f(\theta_2|a_2,b_2)$



χ^2= 35.543 (d.f.=7, p<10^{-5})

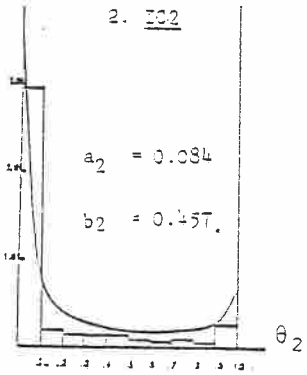
C. Instant coffee

$f(\theta_1|a_1,b_1)$



χ^2= 94.255 (d.f.=7, p<10^{-5})

$f(\theta_2|a_2,b_2)$



χ^2 = 16.399 (d.f.=7, p<.025)

Figure 1
Empirical and Fitted Distributions of Brand Choice Probabilities

Table 1 Parameter Estimates of the Dirichlet Distribution for All Three Product Categories

Product class	Parameter									
	a_1	a_2	a_3	a_4	a_5	a_6[b]	a_7	a_8	a_9[c]	$\sum_i a_i=c$
Margarine[a]	.229	.084	.029	.007	.061	.137	-	-	-	.547
Regular coffee	.074	.040	.011	.017	.010	.019	.025	.014	.460	.669
Instant coffee	.064	.084	.034	.018	.020	.028	.048	.075	.170	.541

[a] Six brands, including the "all other brands" category, were analyzed. For the margarine, regular coffee and instant coffee categories, the "all other brands" category represented a market share of 27, 70 and 29 per cent respectively.

[b] Parameter estimate of the "all the other brands" category for the margarine market.

[c] Parameter estimates of the "all the other brands" category for the regular coffee and instant coffee market.

Table 2 Actual and Theoretical Conditional Switching Within the Margarine Market (38,043 Purchase Occasions)

	M1	M2	M3	M4	M5	M6[a]	Total[b]
M1	.847[c]	.037	.016	.004	.025	.081	1.000
	(.794)[d]	(.054)	(.019)	(.004)	(.039)	(.089)	(1.000)
	[.845][e]	[.038]	[.016]	[.003]	[.027]	[.071]	[1.000]
	(.803)[f]	(.048)	(.020)	(.004)	(.035)	(.090)	(1.000)
M2	.104	.798	.015	.003	.039	.041	1.000
	(.148)	(.700)	(.019)	(.004)	(.039)	(.089)	(1.000)
	[.107]	[.776]	[.016]	[.003]	[.027]	[.071]	[1.000]
	(.136)	(.716)	(.020)	(.004)	(.035)	(.090)	(1.000)
M3	.101	.036	.748	.006	.015	.095	1.000
	(.148)	(.054)	(.665)	(.004)	(.039)	(.089)	(1.000)
	[.107]	[.038]	[.754]	[.003]	[.027]	[.071]	[1.000]
	(.136)	(.048)	(.688)	(.004)	(.035)	(.090)	(1.000)
M4	.130	.050	.015	.693	.000	.1118	1.000
	(.148)	(.054)	(.019)	(.651)	(.039)	(.088)	(1.000)
	[.107]	[.038]	[.016]	[.741]	[.027]	[.071]	[1.000]
	(.136)	(.048)	(.020)	(.673)	(.035)	(.090)	(1.000)
M5	.093	.051	.006	.001	.782	.067	1.000
	(.148)	(.054)	(.019)	(.004)	(.685)	(.088)	(1.000)
	[.107]	[.038]	[.016]	[.003]	[.765]	[.071]	[1.000]
	(.136)	(.048)	(.020)	(.004)	(.702)	(.090)	(1.000)
M6	.122	.022	.020	.004	.027	.805	1.000
	(.148)	(.054)	(.019)	(.004)	(.039)	(.735)	(1.000)
	[.107]	[.038]	[.016]	[.003]	[.027]	[.309]	[1.000]
	(.136)	(.048)	(.020)	(.004)	(.035)	(.758)	(1.000)

[a] All other brands

[b] Some percentages may not add to 1.0 because of rounding errors

[c] 84.7 % of the pairs of purchases which started with brand M1 ended with brand M1

[d] Predictions of the Dirichlet model (K = .354)

[e] Predictions of the switching law (K = .262)

[f] Predictions of the Hendry model (K_w = .332)

Table 3 Mean Percentage Point Deviation from Actual Switching and Repeat Purchase Across Product Categories

Product category	Dirichlet		Switching law		Hendry	
	Repeat purchase	Switching	Repeat purchase	Switching	Repeat purchase	Switching
Margarine	- 1.10	+ .22	- .00	+ .00	- .85	+ .17
Regular coffee	- .22	+ .02	- .01	+ .00	+ .09	- .01
Instant coffee	+ .97	- .12	- .03	+ .00	+ 1.84	- .23
Weighted mean percentage point deviation from actual[b]	+ .01	- .00	- .01	+ .00	+ .51	- .07

[a] The mean deviation from actual repeat purchase is 1.1 percentage point inferior to the actual proportion for the Dirichlet model

[b] The weights are given by the number of brands studied within each product class ("all other brands" included)

Table 4 Average Actual and Theoretical Simple Correlations Between the Relative Frequencies of Choice

Average Correlation	Margarine				
	M1	M2	M3	M4	M5
Actual	-.2526[a]	-.1617	-.1116	-.0583	-.1518
Dirichlet	(-.2396)[b]	(-.1652)	(-.1022)	(-.0523)	(-.1425)

Average Correlation	Regular coffee							
	RC1	RC2	RC3	RC4	RC5	RC6	RC7	RC8
Actual	-.0436	-.0321	-.0221	-.0292	-.0273	-.0316	-.0339	-.0295
Dirichlet	(-.0593)	(-.0364)	(-.0264)	(-.0313)	(-.0339)	(-.0332)	(-.0372)	(-.0293

Average Correlation	Instant coffee							
	IC1	IC2	IC3	IC4	IC5	IC6	IC7	IC8
Actual	-.950	-.1210	-.0732	-.0622	-.0590	-.0798	-.0895	-.1133
Dirichlet	(-.1055)	(-.1193)	(-.0711)	(-.0582)	(-.0613)	(-.0712)	(-.0925)	(-.1136

[a] The average actual correlation across consumers between the relative frequency of choosing brand M1 and the relative frequency of choosing another of the five major margarine brands is -.2526.

[b] Theoretical correlation in parentheses.

APPENDIX

DERIVATION OF THE METHOD-OF-MOMENT PARAMETER ESTIMATORS OF THE DIRICHLET DISTRIBUTION[1]

Assume that the distribution of $[\theta_1,\ldots,\theta_i,\ldots,\theta_N]$ over the population is Dirichlet with parameters $(a_1,\ldots,a_i,\ldots,a_{N+1})$ and $\sum_i \theta_i = 1$.

Let $\sum_{i=1}^{N+1} a_i = c$. The means and second moment about the origin are given by

$$E(\theta_1) = \frac{a_1}{c}$$

$$E(\theta_i) = \frac{a_i}{c} \qquad i = 2,\ldots,N,$$

and

$$E(\theta_1^2) = \frac{a_1(1+a_1)}{c(1+c)}.$$

Therefore

$$c = \frac{a_1}{E(\theta_1)} = \ldots \frac{a_i}{E(\theta_i)} = \ldots = \frac{a_N}{E(\theta_N)} \qquad (A.1)$$

and

$$E(\theta_1^2) = \frac{E(\theta_1)^2 \ (1+a_1)}{\left(E(\theta_1) + a_1\right)}.$$

Therefore

$$a_1 = \frac{E(\theta_1)[E(\theta_1) - E(\theta_1^2)]}{Var(\theta_1)} . \qquad (A.2)$$

From (A.1) and (A.2) we derive

$$c = \frac{a_i}{E(\theta_i)} = \frac{a_1}{E(\theta_1)} = \frac{E(\theta_1) - E(\theta_1^2)}{Var(\theta_1)} .$$

Hence

$$a_i = \frac{E(\theta_i)[E(\theta_1) - E(\theta_1^2)]}{Var(\theta_1)} , \quad i=1,2,\ldots,N, \qquad (A.3)$$

and

$$a_{N+1} = c - (c-a_{N+1}) = \frac{E(\theta_1) - E(\theta_1^2)}{Var(\theta_1)} - \sum_{i<N+1} \frac{E(\theta_i)[E(\theta_1) - E(\theta_1^2)]}{Var(\theta_1)}$$

$$= \left(1 - \sum_{i<N+1} E(\theta_i)\right) \frac{E(\theta_1) - E(\theta_1^2)}{Var(\theta_1)} . \qquad (A.4)$$

Replacing $E(\theta_1)$, $E(\theta_1^2)$ and $E(\theta_i)$ in (A.3) and (A.4) by the sample estimates, we obtain $\hat{a}_1,\ldots,\hat{a}_{N+1}$.

REFERENCES

Bass, Frank M. (1974a), "The Theory of Stochastic Preference and Brand Switching," Journal of Marketing Research, 11, 1-20.

_____ (1974b), "An Empirical Analysis of the Theory of Stochastic Preference", Paper n° 455, Institute for Research in the Behavioral, Economic and Management Sciences, Krannert Graduate School of Management, Purdue University.

_____, Jeuland, Abel, and Wright, Gordon P. (1976), "Equilibrium Stochastic Choice and Market Penetration Theories : Derivations and Comparisons," Management Science, 22, 1051-63.

Bemmaor, Albert C. (1978), "An Empirical Study of the Temporal Aggregation Bias in Stochastic Models of Brand Choice," Proceedings of the 1978 Annual Meeting of the American Statistical Association, Business and Economic Statistics Section, 842-5.

_____ (1981), "A Parsimonious Brand Choice Model with Dependence," Working Paper, Ecole Supérieure des Sciences Economiques et Commerciales, 95021 Cergy-Pontoise, France.

Charlton, P., and Ehrenberg, A.S.C. (1973), "McConnell's Experimental Brand Choice Data," Journal of Marketing Research, 10, 302-7.

Chatfield, Christopher, and Goodhardt, Gerald (1975), "Results Concerning Brand Choice", Journal of Marketing Research, 12, 10-13.

Dodson, Joe Austin (1975), "An Empirical Examination of Buyer Behavior : Individual and Brand Analyses", unpublished Ph.D. dissertation, Purdue University.

Ehrenberg, A.S.C. (1972), Repeat Buying, Theory and Applications, Amsterdam : North Holland Publishing.

_____, and Goodhardt, G.J. (1968), "The Incidence of Brand-Switching", Nature, 220, 5764, 304.

_____ (1974), "The Hendry Brand Switching Coefficient", Admap, 10, 232-38.

_____ (1976), "Decision Models and Descriptive Models in Marketing", Working Paper, Marketing Science Institute, Cambridge, Mass.

_____ (1979), "The Switching Constant", Management Science, 25, 703-705.

Fielitz, Bruce, and Myers, Buddy, L. (1975), "Estimation of Parameters in the Beta Distribution", Decision Sciences, 6, 1-12.

Goodhardt, G.J. (1966), "The Constant in Duplicated Television Viewing" Nature, 212, 5070, 1616.

Herniter, Jerome A. (1973), "An Entropy Model of Brand Purchase Behavior", Journal of Marketing Research, 10, 361-65.

_____ (1974), "A Comparison of the Entropy Model and the Hendry Model", Journal of Marketing Research , 11, 21-9.

International Mathematical Statistical Libraries, Inc. (IMSL 1977), "ZREAL I : ZREAL I Finds Real Zeros of a Real Function and Should Be Used When Initial Guesses are Poor", IMSL Library 02-006 U2 Reference Manual, Houston, Texas, Revised July 1977.

Jeuland, Abel P. (1979), "The Interaction Effect of Preference and Availability on Brand Switching and Market Share", Management Science, 25, 953-965.

_____, Bass, Frank, and Wright, Gordon P. (1980), "A Multibrand Stochastic Model Compounding Heterogeneous Erlang Timing and Multinomial Choice Processes", Operations Research, 28, 255-277.

Johnson, N.J. and Kotz, S. (1970), Continuous Univariate Distributions-2 Boston, Mass. : Houghton Mifflin.

Jones, J. Morgan (1973), "A Composite Heterogeneous Model for Brand Choice Behavior", Management Science, 19, 499-509.

Kalwani, Mahonar U., and Morrison, Donald G. (1977a), "Estimating the Proportions of 'Always Buy' and 'Never Buy' Consumers : A Likelihood Ratio Test with Sample Size Implications", Journal of Marketing Research, 14, 601-6.

_____ (1977b), "A Parsimonious Description of the Hendry System", Management Science, 23, 467-77.

Luce, R.D. (1959), Individual Choice Behavior, A Theoretical Analysis, New-York, N.Y. : John Wiley and Sons.

Massy, William F. (1966), "Order and Homogeneity of Family Specific Brand Switching Processes", Journal of Marketing Research, 3 (February), 48-54.

Mosimann, James E. (1962), "On the Compound Multinomial Distribution, the Multivariate β-Distribution, and Correlations among Proportions," Biometrika, 49, 65-82.

Rao, C.R. (1973), Linear Statistical Inference and Its Applications, Second Edition, New-York : John Wiley and Sons.

Rogers, G.S., and Young, D.L. (1973), "On the Products of Powers of Generalized Dirichlet Components with an Application", Canadian Journal of Statistics, Sections A & B, 1, 159-69.

Rubinson, J.R., Vanhonacker, W.R., and Bass, F.M. (1980), On "A Parsimonious Description of the Hendry System", Management Science, 26, 215-226.

Schmittlein, David C., and Morrison, Donald G. (undated), "Prediction of Future Random Events Based on Past Performance", Graduate School of Business, Columbia University, New-York, New-York 10027.

The Hendry Corporation (1976), Speaking of Hendry.

Wierenga, B. (1974), An Investigation of Brand Choice Processes, Rotterdam University Press.