GARY L. GRAHN*

This article describes a model for predicting repeat-purchase loyalty based on consumer panel data. A modification of the original model is offered that permits its application to areas in which brands are not purchased as a multiple of a single-unit pack size. Results of the original and modified models are shown.

NBD Model of Repeat-Purchase Loyalty: An Empirical Investigation

INTRODUCTION

The design of this article and the work that preceded it has four stages. The beginning portion is devoted to explaining the data and the meaning of repeat purchasing. The second portion introduces the model and discusses its attributes.[1] Next, results are presented from application of the model in its original form. This section will attempt to demonstrate and discuss why the original model failed to fit the specific data.[2]

The fourth portion describes a modification of one of the model's parameters (from amounts purchased to purchasing occasions) that attempted to test the model on aggregate data. Results of this alteration yielded an improved fit for these data. This expansion of the model's capabilities to deal with a wider range of data by a modified parameter is the main contribution of this article.

THE DATA

The data used in this study are from a continuous consumer panel of 1,041 families over a 24-week period. The specific product group is that of a nondurable consumer good and shall be called Product X. Ehrenberg worked with products from sausages to shampoos [3, p. 26], all of which were marketed in prepackaged and branded form. Panel data usually yield the kind of information needed in such a study, i.e., the frequency distribution of purchases and what each consumer bought in different periods.

* Gary L. Grahn is marketing analyst, Owens-Illinois, Inc., Toledo. The author expresses his appreciation to Associate Professors George H. Haines, Jr., and Marcus Alexis, University of Rochester, for their suggestions and critiques of the results.

There are at least two ways in which time periods may be analyzed—successive or nonsuccessive. Four basic time periods were chosen and analyzed under both conditions. However, Period 1 is the basis for prediction of subsequent periods. Four-, six-, eight-, and twelve-week time periods were analyzed. With the four-week data, Weeks 1–4 were used to analyze Weeks 5–8 and 13–16. Data of Weeks 1–6 were used to analyze Weeks 7–12, 13–18, and 19–24. Data of Weeks 1–8 were used to predict Weeks 9–16 and 17–24, and Weeks 1–12 were used to predict Weeks 13–24. These variations were made in order to discover if the time between initial analysis and prediction would have any consistent effect on the results.

Loyal customers of each brand were obtained by comparing the identification numbers appearing in Period 2 also appearing in Period 1. Implicit in this method is the definition of loyal buyer. To some (such as Kuehn and Rohloff [7]), loyalty is measured by the probability of buying the same brand now as the one purchased most recently. Because a customer is only required to purchase a specific brand once in each of the two time periods, it is possible for a consumer (under Ehrenberg's assumptions) to be loyal to more than one brand—as in the Product X data analyzed here. Nonloyal buyers from Period 1 are "lost buyers" who

[1] The model has been applied to consumer panel data and has been developed, to a large extent, by Ehrenberg [2, 3, 4, 6]. Much of this article is dependent on published works by Ehrenberg.

[2] With respect to this failure of the model, it must be recognized that the data under consideration differ from that to which Ehrenberg has previously subjected the NBD model. Although previous work with the model has considered purchases of one particular pack size of a given brand, the data this author dealt with contains an aggregation of different pack sizes of a given brand.

Journal of Marketing Research,
Vol. VI (February 1969), 72-8

bought in the first but not in the second; in Period 2 there is a compensating proportion of new buyers who bought in this period but not in the first one.

THE MODEL

The model used here is the Negative Binomial Distribution (NBD).[3] The NBD has a variety of previous practical applications, e.g., in the study of accident statistics, in certain ecological birth, death, and contagious processes, and in operations research theory.

The NBD is a two-parameter distribution for the nonnegative integers 0, 1, 2, 3, 4 and, generally, r. If the two parameters are taken as the mean m and the exponent k, the probability Pr of observing a number r is [3, p. 26]:

$$Pr = \left\{1 + \frac{m}{k}\right\}^{-k} \frac{\Gamma(k + r - 1)}{\Gamma(r)\Gamma(k - 1)} \left(\frac{m}{m + k}\right)^r.$$

These probabilities derive from the expansion of the binomial expression $[1 - m/(m + k)]^{-k}$ in which the exponent's sign is negative. Ehrenberg suggests that it is often convenient to use the parameter $a = m/k$ instead of k. The NBD is always positively skewed. It has one mode at zero for the fairly small values of m and k which occur with consumer purchasing data; hence, the distribution is then reversed, J, shaped. The distribution variance is:

$$m\left(1 + \frac{m}{k}\right) = m(1 + a).$$

"In the fitting of a negative binomial distribution to empirical data, the best estimate of the mean is the sample mean, since it is the maximum-likelihood estimator and is also unbiased, but the maximum-likelihood equations for a or k are very cumbersome to solve" [3, p. 27]. Ehrenberg suggests two alternate methods of solution. One way, the method of moments, is to estimate a by equating the observed sample variance to its expected value $m(1 + a)$. But this method of estimation is not particularly efficient, and it is laborious to compute the sample variance especially since in market research the basic frequency distribution is often not tabulated.

The second method is to equate the observed proportion of zero readings P_0 to its expected value,

$$P_0 = \left(1 + \frac{m}{k}\right)^{-k} = (1 + a)^{-m/a}.$$

This equation can easily be solved by iteration, especially if written in the form suggested by Evans [5],

$$a - c \ln (1 + a) = 0,$$

where

$$c = \frac{-m}{\ln P_0}.$$

This estimation method is convenient for market research data, where the mean m and the proportion of nonbuyers P_0 are often all the figures tabulated.

Probability estimates (Pr), which may be required, tend to be tedious to compute if the effective range of r is large. Ehrenberg suggests using the iterative formula:

$$Pr = \left(\frac{a}{1 + a}\right)\left(1 - \frac{a - m}{ar}\right) Pr - 1.$$

The goodness of fit of the estimated distribution can be tested by calculating the value of χ^2 for the observed and theoretical frequencies, done here with both Ehrenberg's and the present results.

Turning from the generalized distribution to the specific model under consideration, note that the model has two dimensions: one is time, the other (an unordered one) is consumers. Following is a description of requirements of these two dimensions:

1. *Poisson Distribution in Time.* "For the purchases of any particular consumer in successive periods of time, e.g., purchases of 2, 0, 1, 1 units and so on, the model requires that these purchases behave like independent random samples from a Poisson distribution" [3, p. 34].

This appears reasonable under two conditions which we suspect would normally be fulfilled: (a) that successive time periods are not only of equal length but also similar to each other (one week being basically similar to the next), and (b) that periods are not too short so that purchases made in one period do not directly affect those made in the next. The Poisson distribution has one parameter, the mean μ, i.e., the average rate of purchasing in the long run.

2. *A Chi-Square Distribution of Consumers.* The second part of the model specifies that the distribution of average purchasing rates μ of different customers should be proportional to a chi square or Type III distribution with $2k$ degrees of freedom. "This is also plausible, since such a distribution is fairly flexible (having two adjustable parameters) and of the right shape (i.e., a continuous distribution for non-negative values, reversed J-shaped or humpbacked and always positively skewed)" [3, p. 34].

The importance of the preceding discussion is that individual consumers should follow a Poisson distribution, but aggregate values during any time period should follow the NBD [3, p. 35].

All deductions from the NBD can be subsumed by the general and powerful reformulation of the model as a multivariate NBD. Mathematically this can be represented by the probability generating function (pgf) of the distribution of people buying r_i units in the ith period out of T periods of length T_i which is:

---

[3] For much of this section, the author is indebted to Ehrenberg's previous work on this subject.

Table 1
ANALYSIS OF PREDICTED AND OBSERVED PERCENTAGES OF LOYAL BUYERS FOR FOUR-WEEK DATA, WEEKS 5–8[a]

Brand	*Percentage of buyers = 100B*	*Mean buying rate = 100M*	*Loyal buyers predicted*	*Loyal buyers observed*	*Discrepancy*	*Discrepancy as percentage of observed value*
H	21.80%	47.35	.13	.12	.01	8%
B	20.26	59.75	.14	.10	.04	36
A	13.64	41.01	.09	.05	.04	61
G	12.43	33.71	.07	.05	.02	25
C	10.27	22.38	.06	.05	.01	17
D	3.93	8.74	.02	.01	.01	54
E	2.97	8.35	.02	.01	.01	10
F	1.44	3.84	.01	.00	.01	394

[a] Sample size is 1,041.

NOTE:
Averages:

All Brands	.016	76.1%
Without Brand F	.015	26.8
"+ and −"	.016	76.1
"+ and −" without Brand F	.015	26.8.

$$\left[1 + a \sum_{i=1}^{t} T_i(1 - u_i)\right]^{-k},$$

where

$a = m/k$ and m is the average amount bought in some time period of "unit" length

μ_i are dummy variables.

Expanding these pgf's in terms of *rth* powers of the dummy variable μ_i gives as coefficients the probabilities of observing purchases of amount r_i in the relevant time-period. "As a particular illustration, consider stationary purchasing in two equal time-periods I and II, with mean m and NBD exponent k in each time-period. Given all the consumers who bought exactly r units in period I (r being any whole, nonnegative integer), the 'conditional' distribution of their purchases in period II then turns out to be itself negative binomial, with mean $(k + r)\,\{a/(1 + a)\}$ and with the k type of parameter or exponent taking the value $(k + r)$" [2, p. 326].

The NBD model's two parameters are usually denoted by the mean amount m of purchases made by all consumers in any one time period, and by either the exponent k or the quantity $a = m/k$. The mean m must be constant in equal stationary time periods, but its value clearly depends on length of the time period. Denoting more generally by m_T the mean in a period of length T relative to some "unit" time period, we have under stationary conditions:

$$m_T = T_m,$$

where

m is the mean purchasing rate in the unit period.

The parameter m (or equivalently $a = m/k$) acts therefore as a scale factor, reflecting length of the analysis period.

In contrast, the parameter k in the NBD models should be constant for any one stationary brand or pack size, irrespective of period over which purchases are analyzed.

The empirical behavior of the parameter k for any given brand can be illustrated by estimating its value in periods of different lengths T. According to the NBD, the relationship between the proportion b_T of consumers who buy and the mean quantity m_T bought per consumer in a given time period of length T is:

$$b_T = 1 - \left(1 + \frac{m_T}{k_T}\right)^{-k} T.$$

Given observed values of b_T and m_T for time periods of different lengths, this equation allows you to estimate a value of k_T for each period. If the NBD model holds, the value of k_T must be constant, i.e., independent of T [2, p. 327].

Besides the restrictions placed on the parameters discussed with regard to the NBD, three additional conditions must be met:

1. *Stationarity.* Purchasing behavior must show no overall trend from one period to the next. Most frequently bought consumer goods are approximately stationary most of the time. Though complete nonstationarity is not fulfilled with respect to each brand and for each pair of time periods of Product X data analyzed, no overall trend was found to exist.
2. *One brand at a time.* The analysis deals with purchasing one brand at a time. No account is taken whether a consumer has also bought other brands in the product field and, if so, how much he has bought.
3. *One pack size at a time.* The brand or product must essentially be bought as a multiple of a single-unit pack size (12-ounce bottle, half-pound package, etc.) because the statistical distributions and models used deal with integral values only. When a brand or product is marketed in two or more different pack sizes, each pack size must be analyzed separately.

As mentioned earlier, this last restriction could not be guaranteed by the kind of data used here since purchases of different pack sizes of the same brand were aggregated.

ORIGINAL RESULTS

This section discusses (a) results obtained with Ehrenberg's model as applied to Product X data, (b) why the model did not yield a good fit to the actual results, and (c) various model attributes shown by the data.

A computer calculated the intermediate value a in the model. "Mean Rate of Purchasing" is the value m in the model, and the expression $-m/\ln\,(1 - B)$ is used to develop the value a in the model. These values and

Table 2

ANALYSIS OF PREDICTED AND OBSERVED PERCENTAGES OF LOYAL BUYERS BY TIME PERIOD, AVERAGED OVER ALL BRANDS

Method	All brands		Without Brand F		"+ and −"		"+ and −" without Brand F	
	Discrepancy	Discrepancy as percentage of observed	Discrepancy	Discrepancy as percentage of observed	Discrepancy	Discrepancy as percentage of observed	Discrepancy	Discrepancy as percentage of observed
Successive period								
4 weeks 5–8	0.016	76.1	0.015	26.8	0.016	76.1	0.015	26.8
6 weeks 7–12	0.021	65.4	0.020	28.4	0.021	65.4	0.020	28.4
8 weeks 9–16	0.015	33.9	0.014	16.9	0.011	31.2	0.010	14.3
12 weeks 13–24	0.032	42.3	0.030	22.3	0.024	37.8	0.022	19.5
Nonsuccessive								
4 weeks 13–16	0.012	67.3	0.011	18.0	0.012	67.3	0.011	18.0
6 weeks 13–18	0.017	58.4	0.010	24.5	0.017	58.4	0.019	24.5
6 weeks 19–24	0.026	58.4	0.025	47.1	0.026	58.4	0.025	47.1
8 weeks 17–24	0.023	43.3	0.021	30.1	0.019	40.6	0.021	27.0

the converted a value are then used to develop the logarithmic (Naperian) value of the predicted percentage of repeat purchases. Finally a computation is made of the actual observed loyalty, the difference between observed and predicted loyalties, this difference as a percentage of the observed value, and the chi-square value for the total period.[4]

Table 1 shows discrepancies between the observed and predicted loyalties for a typical time period and is important in several respects. It is the kind of analysis that Ehrenberg uses, and it indicates the poor fit that the model yields with the present data.

Table 2 summarizes results over the eight time periods examined. Several features of the results presented in this table need further discussion. First, an adequate appreciation of either the model or the results cannot be obtained from a percentage analysis (subsequently an alternative chi-square analysis will be discussed). Second, the results represent averages of each brand in the eight time periods analyzed. The breakdown by period and brand indicated that Brand F (newly introduced) accounted for a large part of any discrepancy. Because it held such a small market share, a change of even a few customers in its loyalty produced very large effects. Results were therefore analyzed with this brand both included and excluded. It is also relevant to look at both absolute differences and differences that will have cancelling effects if the sign is allowed to remain, as it should for an aggregate model. Because Ehrenberg's model almost consistently overpredicted, this did not improve the results greatly. Note that the data appear to validate the belief that neither length of the original time period, i.e., 4, 6, 8, or 12 weeks, nor successiveness or nonsuccessiveness appears to have any consistent bearing on the results obtained.

[4] Copies of these programs are available on request.

At this point of analysis it became apparent that the NBD as used by Ehrenberg did not fit these data. To understand this discrepancy the requirements and assumptions of the generalized NBD and the specific Ehrenberg model were reevaluated since it is also possible that Product X data is not applicable to this model because of one or more reasons to be discussed.

As indicated earlier, Ehrenberg mentions that the parameter k in the NBD model should be constant for any one stationary brand or pack size, irrespective of the period length over which purchases are analyzed. Table 3 shows that this is not completely true for the present data; however, the values obtained do not vary enough to be held solely accountable for the lack of fit. In this regard note that while k varies by about 10 percent, the time period involved increases 300 percent.

Table 3

RELATIVE CONSISTENCY OF NBD PARAMETER K AND OF OTHER STATISTICS WITH INCREASING LENGTH OF TIME PERIOD

Exemplary brand (H)	Time period (weeks) 4	6	8	12	12 weeks/ 4 weeks
Mean amount bought per household, m_T	0.474	0.705	0.895	1.345	2.8
Proportion of buyers, b_T	0.218	0.267	0.290	0.348	1.6
Average amount per buyer, $m/b = w_T$	2.2	2.6	3.1	3.9	1.8
NBD parameter K_T	0.206	0.212	0.203	0.186	0.9
Average K_T for 8 brands	0.080	0.086	0.088	0.094	1.2

Table 4
ANALYSIS OF PREDICTED AND OBSERVED PERCENTAGES OF LOYAL BUYERS BY TIME PERIOD, AVERAGED OVER ALL BRANDS

Method	*All brands*		*Without Brand F*		*"+ and −"*		*"+ and −" without Brand F*	
	Discrepancy	*Discrepancy as percentage of observed*	*Discrepancy*	*Discrepancy as percentage of observed*	*Discrepancy*	*Discrepancy as percentage of observed*	*Discrepancy*	*Discrepancy as percentage of observed*
Successive period								
4 weeks 5–8	0.007	39.5	0.008	17.0	0.005	12.7	0.005	13.8
6 weeks 7–12	0.009	25.0	0.010	9.2	0.001	17.1	0.001	.1
8 weeks 9–16	0.006	14.0	0.006	7.0	0.001	9.7	0.002	2.1
12 weeks 13–24	0.007	17.2	0.008	7.1	0.007	17.2	0.008	7.1
Nonsuccessive								
4 weeks 13–16	0.009	42.1	0.010	8.0	0.008	8.0	0.010	20.1
6 weeks 13–18	0.008	25.1	0.008	9.3	0.004	12.3	0.004	5.3
6 weeks 19–24	0.008	5.1	0.009	11.4	0.006	12.7	0.007	8.6
8 weeks 17–24	0.006	13.4	0.007	7.8	0.006	12.8	0.006	7.1

Another discrepancy between the model and the data concerns the "one pack size at a time" limitation discussed earlier. The requirement that a brand be bought as a multiple of a single-unit pack size could not be guaranteed with these data, nor could conversion into equivalent units of a single "basic" size be made because of data inadequacies. Because of the pricing and packaging techniques used, the model is not using integral values as required.

Consistency of the discrepancy and belief in the value of the kind of analysis which the model provides led to the conclusion that the model could be meaningfully altered to deal with situations in which the unit pack size requirement could not be satisfied. As evident from the alteration, aggregation of pack sizes is critical to the application being investigated.

PROPOSED ALTERATION OF EHRENBERG'S MODEL

In Ehrenberg's model the parameter m denotes the mean purchases by all consumers of a brand in any one time period and is related to the second NBD parameter, the exponent k through the quantity $a = m/k$, m acting as a scale factor and reflecting the length of the analysis period. It is the definition of this parameter that is now altered. Though Ehrenberg defines it as the mean rate of *purchases*, this author defines it as the mean rate of *purchase occasions*. That is, if a consumer purchases Brand A as follows during a given four-week period (0, 2, 1, 1), Ehrenberg would establish an m value of $4/4 = 1.0$, but my definition yields a value of $3/4 = 0.75$, i.e., only the number of times and not the quantity purchased is now relevant. Before discussing this parameter further, compare the results in Table 4 and those obtained earlier in Table 2 with the straight Ehrenberg model. As an indication of improvement, this author's method produces an average discrepancy over time periods in the "+ and −" without Brand F column of 8 percent, but the Ehrenberg model yields 26 percent.

Even more impressive than these results are those in Table 5, which shows the typical improvement redefinition of m yields. This table represents an improvement in the percentage-type analysis used by Ehrenberg and discussed in connection with Table 1. This author chose to look at the number of loyal consumers observed and predicted and to analyze their difference by chi-square techniques. The null hypothesis is that no significant difference exists between sample frequency distribution and actual population distribution. The significance or nonsignificance of the observed differences may be found by determining from a chi-square table the probability that a χ^2 larger than that computed is likely to occur as a result of sampling variations. At the .05 level of significance, with six degrees of freedom, each time period had a sample χ^2 greater than χ_0^2 (equal to 12.6). Therefore, the hypothesis is rejected at the .05 level of significance [1, p. 175].

With redefinition of m, in every time period analyzed the computed chi-square value was less than 12.6, indicating that the difference is not significant at the level chosen. The signs of the successive numerical differences would usually be expected to alternate in some erratic fashion. That is, first, the sample value might exceed the population value for one or two class intervals; then the population value might exceed the sample value; then the sample value might exceed the population value, etc. If the alternation in signs does not follow some such erratic pattern, presence of a factor other than random sampling variations is usually suspected. For example, the completely "one-signed" differences in the Ehrenberg results would not generally be attributed to random differences in the two distributions.

As with Ehrenberg's results, no consistent discrepancy could be attributed to either time period length or whether the prediction was made for successive or non-successive periods. It is true, however, that both predictions for the four-week periods (Weeks 5–8 and 13–16) are considerably less accurate than the other periods of 6-, 8-, and 12-week durations. This suggests that the model's modification is more sensitive to short time periods, as can be expected when consequences of the redefinition of the mean rate are considered. Certainly over a short period the value of m, as redefined, could be unusually low, and this could have strange effects on the total value predicted by the model because m is in both the base and exponent of the expression.

It was hoped that redefinition of m would produce a more stable value of the NBD exponent k and therefore better fit the distribution. Unfortunately, this was not the case, and the difference is not resolved as simply as desired. A downward rather than an upward trend was obtained for the average value of k over time.

The reason why my modification of Ehrenberg's model is a better fit to these data is first, the parameter m has been redefined so that it is intrinsically meaningful in the traditional concept of market analysis. Just as Ehrenberg's mean rate over purchases has marketing meaning, the mean rate of purchase occasions is a meaningful concept instead of simply an isolated numerical approximation forced on the NBD model to obtain a better fit in specific areas.

Second, using purchase occasions will broaden the scope of the NBD model. The modification should allow the NBD to be used in areas where restrictions on unit pack size previously limited its applicability. The kind of data represented by Product X, i.e., aggregated pack sizes, is important and needs the kind of analysis which the NBD model offers.

A problem that the modification is as unable to solve as the unmodified model is that which Ehrenberg calls "shelving"—"the shelving phenomenon may contain the main, if not the entire, explanations of the various discrepancies" of the fitted theoretical NBD [4, p. 335]. For each brand there is a sudden drop or discontinuity in the frequencies of heavier purchases at or about the number of units equal to the number of weeks (or a fraction of this) in the analysis period.

Table 5

CHI-SQUARE ANALYSIS OF FOUR-WEEK DATA (WEEKS 5–8) UNDER EHRENBERG'S DEFINITION AND AS REFINED

Brand	*Families purchasing brand at least once in Period 1*	*Total purchases of brand in Period 1*	*Observed families repurchasing same brand in Period 2*	*Predicted families repurchasing same brand in Period 2*	*Difference between observed and predicted*
H	227.0	493.0	133.0	144.3	−11.3
B	211.0	622.0	108.0	147.6	−39.6
A	142.0	427.0	61.0	98.2	−37.2
G	130.0	351.0	61.0	76.6	−15.6
C	107.0	233.0	56.0	65.7	−9.7
D	41.0	91.0	16.0	24.7	−8.7
E	31.0	87.0	21.0	30.7	−9.7
F	15.0	40.0			
m is mean amount of purchases (χ^2 = 53.4)					
H	227.0	359.0	133.0	114.6	18.3
B	211.0	328.0	108.0	103.6	4.3
A	142.0	221.0	61.0	68.3	−7.3
G	130.0	192.0	61.0	47.7	13.2
C	107.0	172.0	56.0	52.6	3.3
D	41.0	50.0	16.0	11.4	4.5
E	31.0	44.0	21.0	18.4	2.6
F	15.0	20.0			
m is mean rate of purchase occasions (χ^2 = 8.3)					

APPLICATIONS

The model is valuable in several ways previously developed [2, 3, 4, 6]. First, in regard to this discussion, it can estimate loyalty over two time periods based on collection of purchasing data for only one. This is true regardless whether the periods are successive or non-successive. With respect to the product data analyzed, this means that had a company purchased data for only one half or even less of the full 24-week period, loyalty and repeat-purchase information for the full period could perhaps be estimated from only one quarter of the data. The goal here is not necessarily to drastically reduce a marketing research budget, but rather to get more information for the amount spent.

A second use is more fundamental; the model can pinpoint abnormal loyalty levels. Because the model is dependent on a stationary market, it can examine results of a bouyant market. That is, if a brand's market share increases, it is worthwhile to know whether this resulted from new buyers or an increased proportion of loyal buyers. Similarly important would be the rate and extent of establishing an adequate loyalty core. It should be recalled here that Brand F was indeed a new product and that it had one of the lowest loyalty percentages. We would not necessarily expect a new brand to capture a large market share initially. However, if it is to exist on a small share, it must have loyal consumers.

Also, a marketing group should be better able to formulate promotional budgets based on loyalty information. That is, a brand with a low loyalty factor must continuously seek new customers to maintain its market share. However, a brand with a high loyalty factor should be promoted occasionally on a loss leader philosophy to attract new buyers who might remain with it.

CONCLUSION

This article attempted to describe a model for predicting repeat-purchase loyalty and to measure this model's fit to consumer panel data. The article was divided into the chronological and logical stages in this author's work. First an explanation of the data and of repeat-purchasing was presented. Ehrenberg's model was then developed from the NBD, and statistical attributes of both were discussed. The third section presented the results obtained when Ehrenberg's model was applied to Product X data. Because the fit of these results was unsatisfactory (especially considering the market's relative stability), a meaningful modification of one of the NBD model parameters was proposed. The results closely approximated those observed from the data. The modification proposed is meaningful in a marketing and statistical sense, and can be used in other product areas where there is some question of the reasonableness of the unit pack-size restriction on which Ehrenberg's model relies. The modification proposed here enables the NBD model to be broader in scope and to handle products which formerly it could not.

REFERENCES

1. Richard S. Burlington and Donald C. May, *Handbook of Probability and Statistics*, Sandusky, Ohio: Handbook Publishers, Inc., 1958, 175–86.
2. C. Chatfield, A. S. C. Ehrenberg, and Gerald J. Goodhardt, "Progress on a Simplified Model of Stationary Purchasing Behavior," *Journal of Royal Statistical Society*, (Series A), 129 (Part 3, 1966), 317–60.
3. A. S. C. Ehrenberg, "The Pattern of Consumer Purchases," *Applied Statistics*, 8 (March 1959), 26–41.
4. ———, "Estimating the Proportion of Loyal Buyers," *Journal of Marketing Research*, 1 (February 1964), 56–9.
5. D. A. Evans, "Experimental Evidence Concerning Contagious Distributions in Ecology," *Biometrika*, 40 (June 1953), 186–211.
6. Gerald J. Goodhardt and A. S. C. Ehrenberg, "Conditional Trend Analysis: A Breakdown by Initial Purchasing Level," *Journal of Marketing Research*, 4 (May 1967), 155–61.
7. Alfred E. Kuehn and Albert C. Rohloff, "New Dimensions in Analysis of Brand Switching," *New Directions in Marketing*, Chicago: American Marketing Association, 1965, 297–308.