APPENDICES: CONTENTS

PREFACE TO THE APPENDICES

The Appendices show how to do the calculations in this book. To gain practice this can be tackled in two ways, by working through small-scale numerical examples by hand or by using a computer package. We strongly advise the reader to do both: to "get your hands dirty" and to acquire some helpful computing experience before attempting analyses of larger data sets of this kind.

The worked examples are presented in two Appendices, for an individual brand and for multi-brand buying. In Sections A.1 to A.11 of Appendix A, we show how to calculate the theoretical NBD/LSD repeat-buying statistics. In § A.12 to A.14, which are new to this edition, we comment briefly on the appropriate tabulations of observed data. Appendix B sets out some look-up tables of theoretical statistics of the NBD/LSD model which remain useful even in these days of rapid computing.

In Appendix C, which again is new, the calculation of the Dirichlet parameters is illustrated. This provides alternative predictions of single-brand buying behaviour and, more importantly, it gives us theoretical values for multi-brand buying. Appendix C also outlines how to tabulate the relevant observed data, and how the results can be used to check the assumptions of the Dirichlet model.

Another way to get a feel for the models is to run trials using a computer package. One version of such software is available (for a small fee) on a floppy disc designed for IBM personal computers, though it will transfer to other personal computers and mainframe systems*. On the disc are two programs, written in standard FORTRAN 77; one calculates all the theoretical single-brand statistics and the other is for multi-brand calculations. There are two versions of each program, to allow the user to choose whether to tabulate the raw data and automatically derive the NBD or Dirichlet predictions, or merely to input some summary statistics to obtain the theoretical predictions. Test data are given on the disc, along with some examples of the output.

In the US, consumer panel data are available on a commercial basis from MRCA Information Services, from IRI (Information Resources Inc.), and from NPD (National Purchase Diary), other services are also being developed. In Britain the main source of such data is AGB Ltd. Consumer panels are also operated in a number of other countries in Western Europe,

*The disc is being distributed by Dr M. D. Uncles at the Centre for Marketing and Communication, London Business School, London NW1 4SA, England.

Japan, etc. Most firms are prepared to release small amounts of data for academic research. It is important when asking for data to have a sample of continuous reporters over the whole period concerned (e.g. a year), including non-buyers of the particular product, or at least to know approximately how many there are (see A.14).

APPENDIX A

CALCULATIONS FOR AN INDIVIDUAL BRAND

A.1. Tabulations and Calculations

In this Appendix we mainly give a worked numerical example of the mathematical calculations involved in using the NBD/LSD repeat-buying theory. Application of the theory also involves two other things, namely tabulation of the observed data, and a background of past experience of similar work to help in interpreting the results. In § A.12 to A.14 we give guidance on the tabulations.

The tabulations which are required in repeat-buying studies are in principle quite simple. This is primarily because in studying the observed repeat-buying of any particular brand or other type of item, only purchases of that item are involved and purchasing of other brands can be ignored. For the newcomer to this kind of work it is best to "get his hands dirty" by carrying out some initial tabulations himself, so as to get the feel of the data. Only after that should the work be delegated to a computer. The latter is of value for extensive work on a more or less repetitive basis, but it can take time and effort to get a sufficiently wide and flexible range of programs working; time and money can be saved if you use the programs on our disc to gain initial experience and then adapt them for your own software needs.

The theoretical formulae which are then needed may be tackled in one of three ways. Firstly, using special numerical tables from which the required answers can simply be read off (or estimated by interpolation). A number of such tables are set out in Appendix B, although they are mainly intended to provide a "feel" for the results, and are not always detailed enough for routine use. Secondly, the NBD calculations can be computerised, if and when their routine use for large data files is required, as on our disc.
Thirdly, the theoretical formulae can be worked out by direct hand-calculation in each particular case. This is important to do for the newcomer who wishes to understand how the formulae work (or for those engaged in further research into buyer behaviour). A worked example of such calculations is now given.

A.2. A Worked Example

To familiarize the calculations that are required, a worked numerical example is given of all the theoretical calculations required in Chapter 3.

The most general theoretical formulae are the NBD ones, but for hand-calculation the LSD formulae or the simplifying approximations to them which were set out in Chapters 4 and 8 are easier to use in those ranges of the parameters b and w where they give virtually the same results as the NBD. In what follows, we therefore give for each case the NBD, the LSD and the simplified "approximate" forms of calculation.

Brand E has been chosen for the example because it is in the parameter range where both the LSD and the further simplifications apply, in that they give almost the same results as the NBD. (In working through a further numerical exercise, it may be best to choose Brand A in Chapter 3 to learn to what extent the LSD results *differ* from the NBD ones for high penetration levels.) In what follows some deliberate variation has been included in the time-period chosen to illustrate the *detailed* calculations, but the numerical results for all lengths of time-periods covered in Chapter 3 are given.

A.3. The Basic NBD and LSD Parameters

In most of the calculations, the NBD parameter k (or $a = m/k$) and the LSD parameter q (or $a = q/(1-q)$) are required for the given time-period. These values are obtained now for Brand E in the 4-, 12-, 24- and 48-week periods covered in Chapter 3 *. They are set out in Tables A1 and A2 for ease of subsequent reference.

For the NBD parameter k, a knowledge of the penetration b and the average buying frequency per buyer w is required as input. The observed values for Brand E are given for various time-periods in the "observed" columns of Tables 3.1a and 3.2a respectively. (The values are generally averages of several periods of the stated length, but this does not affect the nature of the calculations here. Results for each separate quarter are for example given in Tables 3.1 and 3.2.)

* In a 1-week period, the observed w is only 1.01 (see Table 3.2a) and no NBD or LSD can usefully be fitted.

Table A1. The Calculation of the NBD Parameters k or a from b and w

Brand E	Period of Length (in weeks)			
	4	12	24	48
Given Data:				
b (Table 3.1a)	.04	.07	.09	.12
w (Table 3.2a)	1.6	3.0	4.9	6.8
$p_0 = 1-b$	.96	.93	.91	.88
$m = bw$*	.064	.21	.441	.816
Derived Statistics				
$c = -m/\ln(p_0)$	1.57	2.89	4.68	6.38
a (Table B3)*	1.32	5.34	12.00	19.18
$k = m/a$ *	.0485	.0393	.0367	.0425

* For footnote see below.

Table A2. The Calculation of the LSD Parameters q or a from w

Brand E	Period of Length (in weeks)			
	4	12	24	48
Given Data:				
w (Table 3.2a)	1.6	3.0	4.9	6.8
Derived Statistics:				
a (Table B3)*	1.40	5.71	12.89	21.0
$q = a/(1+a)$	.584	.851	.928	.955

* Note that the situation is not *quite* stationary (i.e. the values of m and a are not quite proportional to the lengths of the periods, and k is not quite constant).

One case, for the 4-weekly data, will be given in some detail; the calculations for the other time-periods will be shown in summary form only. (This form of presentation will also be followed in other sections, where appropriate.)

The NBD Parameter k. From Tables 3.1a and 3.2a, b (the penetration expressed as a proportion) and w (the buying frequency per buyer) for Brand E in 4 weeks are

$$b = .04, \quad w = 1.6 .$$

The proportion of non-buyers, p_0, is therefore given by

$$p_0 = 1 - b = .96 ,$$

and the mean number of purchases per sample-member, m, by

$$m = bw = .064 \text{ *}.$$

To obtain the value of the NBD parameters k or a we have to solve the equation $p_0 = (1 + m/k)^{-k}$. This is usually best done (as mentioned in §4.2) by calculating a certain quantity $c = -m/\ln(p_0)$ from m and p_0, where "ln" stands for the Natural Logarithm **. We then use Table B2 in Appendix B to read off the corresponding value of a for the given value of c. Thus

$$c = \frac{-m}{\ln(p_0)} = \frac{-.064}{\ln(.96)} = \frac{.064}{.0408}$$

$$= 1.569.$$

From Table B3 (by interpolation)

$$a = 1.32$$

and hence

$$k = m/a = 0.0485 .$$

Table A1 summarises this calculation for the 4-week period and the corresponding ones for the remaining time-periods, of 12, 24 and 48 weeks.

* In practice the value of m would usually be given by direct tabulation of the data.

** Appropriate values of Naperian or Natural Logarithms to base e are given in Table B3 in Appendix B.

The LSD Parameter q. Next, we describe the estimation of the LSD parameter q (or $a = q/(1-q)$). Here only w is required as input and we need to solve the equation $w = -q/(1-q)\ln(1-q)$. The best way is to use Table B2 which was used in the NBD calculation above but which also gives the value of the LSD parameter a for a given w, and hence the LSD parameter q (as $q = a/(1+a)$). In the 4-week period for Brand E, we have (from Tables 3.2a or A2) that

$$w = 1.6 \, .$$

Using Table B2 gives

$$a = 1.402 \, ,$$

and q therefore is

$$q = \frac{a}{1+a} = \frac{1.402}{2.402} = .584 \, .$$

Table A2 summarises this calculation and those for the other time-periods.

A table from which the value of the LSD q can be read off *directly* for a given w is given as Table B1 in Appendix B (see also Table 2.2). Accurate interpolation for relatively high values of q is however relatively difficult, and it is usually better to calculate a first, using Table B2.

More generally, we may note that the values of b and w in the various time-periods here are mostly in the range where the LSD theory tends to give very similar results to the NBD (as will be illustrated by the following sections). The NBD and the LSD versions of the parameter a are however not identical, although Tables A1 and A2 show them to be similar.

A.4. Penetration Growth (Table 3.1a)

The theoretical norms for the change of penetration of Brand E in various length time-periods in Table 3.1a were derived from the observed results in the average *12-week* period. For the NBD, we need to solve the equation $b_T = 1-(1+Tm/k)^{-k}$ as given in §4.4 (Table 4.10) and in §7.5. The lengths of the other time-periods for which the predictions are made are expressed as multiples of the 12-week period, i.e.

by calling this the base-period with $T = 1$. For the other periods, we therefore have

$$\begin{aligned} &1 \text{ week}; && T = 1/12 = .083, \\ &4 \text{ weeks}; && T = 4/12 = .333, \\ &24 \text{ weeks}; && T = 24/12 = 2, \\ &48 \text{ weeks}; && T = 48/12 = 4. \end{aligned}$$

The full NBD calculation is illustrated now for the one-week period. Using the NBD formula $b_T = 1-(1+Tm/k)^{-k} = 1-(1+Ta)^{-k}$, where $m = .2$, $a = 5.34$, $k = .0393$ are the parameters obtained from the 12-week period (Table A1), and $T = .083$, we have

$$\begin{aligned} b_{.083} &= 1-(1+.083 \times 5.34)^{-.0393} \\ &= 1-(1.44)^{-.0393} \\ &= 1-.986 = .014\ , \end{aligned}$$

or just over 1% (the "Theoretical" estimate shown in the 1-week column of Table 3.1a). In other words, given that 7% of the sample bought Brand E on average 3.0 times in the 12-week period (Table A1), the NBD estimate is that about 1.4% of the sample should buy Brand E in the typical *1-week* period (and this compares with the *observed* 1-week penetration of roughly 2%, as was shown in Table 3.1a). This result and these for the longer time-periods are summarised in Table A3.

The corresponding LSD formula (Table 4.10 and §8.5) is expressed in terms of the ratio of b_T to b, the penetration in the 12-week period ($T = 1$),

$$\frac{b_T}{b} = 1 - \frac{\ln(1+(T-1)q)}{\ln(1-q)}\ ,$$

where $b = .07$ and $q = .851$, the LSD parameters in the 12-week period (Table A2).

The quantity $\ln(1-q)$ in this formula enters into the calculation for periods of *any* length and may first be calculated:

$$\ln(1-q) = \ln(.149) = -1.904.$$

Table A3. The NBD Estimates of the Penetration b_T from the 12-Weekly Data ($T = 1$)

Brand E	Period of length (in weeks)			
	1	4	24	48
T	.083	.333	2	4
$1+Ta$	1.44	2.78	11.68	22.36
$(1+Ta)^{-k}$	.986	.961	.908	.885
b_T	.014	.039	.093	.115
$100\,b_T$	1%	4%	9%	11%
Observed $100\,b_T$	2%	4%	9%	12%

Table A4. The LSD Estimates of b_T

Brand E	Period of length (in weeks)			
	1	4	24	48
T	.083	.333	2	4
$(T-1)q$	−.780	−.568	.851	2.553
$\ln(1+(T-1)q)$	−1.514	−.839	.615	1.268
b_T/b	.205	.559	1.323	1.666
b_T	.014	.039	.093	.117
$100\,b_T$	1%	4%	9%	12%

Table A5. Estimates of b_T from the "Approximate" Formula

Brand E	Period of length (in weeks)			
	1	4	24	48
T	.083	.333	2	4
$T^{.82}$	.13	.41	1.76	3.10
b_T/b	.198	.552	1.33	1.67
b_T	.014	.039	.093	.117
$100\,b_T$	1%	4%	9%	12%

For a single week ($T = .083$), we therefore have

$$\frac{b_{.083}}{b} = 1 - \frac{\ln\{1+(.083-1)\,.851\}}{-1.904}$$

$$= 1 - \frac{\ln(1-.917\times .851)}{-1.904}$$

$$= 1 - \frac{1.514}{1.904} = .205, \qquad \text{so that}$$

$$b_{.083} = .07\times .205 = .014 \quad \text{or} \quad 1\%,$$

as for the NBD. The calculations are summarised in Table A4. Note that for the 48-week period the NBD and LSD estimates differ slightly (.115 and .117).

Finally, we can use the approximation to the LSD formulae

$$\frac{b_T}{b} = \frac{Tw}{1+(w-1)T^{.82}}\ ,$$

where w is the buying frequency in the unit period (Tables 4.10 and §8.5). For Brand E in 12 weeks, $w = 3.0$, so that

$$\frac{b_T}{b} = \frac{3T}{1+2T^{.82}}\ ,$$

The estimate of the penetration in a single week ($T = .083$) is therefore

$$\frac{b_{.083}}{b} = \frac{.249}{1+2(.083)^{.82}}$$

$$= \frac{.249}{1+2\times .13} = \frac{.249}{1.26}$$

$$= .198\ , \qquad \text{so that}$$

$$b_{.083} = .198\times .07 = .014,$$

the same value as given by the LSD calculation itself. Table A5 summarises these calculations for all the time-periods.

A.5. The Average Frequency of Buying (Table 3.2a)

We now turn to estimating w_T, the average frequency of purchase per buyer, in time-periods of relative length T, from the b and w results in the "unit" time-period, here 12 weeks (as given in Table 3.2a).

The NBD formula for w_T is

$$w_T = \frac{Tm}{\{1-(1+Tm/k)^{-k}\}},$$

where m and k are the parameters of the 12-week period for Brand E with

$$m = .21, \quad k = .0396.$$

We also have $m/k = a = 5.34$.

To calculate the 4-week value of w_T (with $T = 4/12 = .333$), we therefore have

$$w_{.333} = \frac{.333 \times .21}{\{1-(1+.333 \times 5.34)^{-.0396}\}}$$

$$= \frac{.07}{1-(2.778)^{-.0396}}$$

$$= \frac{.07}{.039} = 1.8.$$

Note that the denominator in the last expression is the estimated value of $b_{.333}$ which can be read off from Table A3 and need not really be recalculated. Table A6 summarises the calculations for the various periods of time and makes use of b_T from Table A3 in this way.

The LSD formula expressed as the ratio w_T/w is

$$\frac{w_T}{w} = \frac{T\ln(1-q)}{\ln(1-q)-\ln(1+(T-1)q)},$$

in terms of q, where $q = .851$ and $w = 3.0$ for Brand E in 12 weeks ($T = 1$).

Table A6. **The NBD Estimates of the Average Purchase Frequency w_T from the 12-Weekly Data ($T = 1$)**

Brand E	Period of length (in weeks)			
	1	4	24	48
T	.083	.333	2.0	4.0
Tm	.017	.07	.42	.84
b_T*	.014	.039	.09	.11
w_T	1.2	1.8	4.6	7.4
Observed w_T	1.0	1.6	4.9	6.8

* The theoretical values as estimated in Table A3.

Table A7. **The LSD Estimates of w_T**

Brand E	Period of length (in weeks)			
	1	4	24	48
T	.083	.333	2	4
b_T/b	.205	.559	1.323	1.666
w_T/w	.405	.596	1.512	2.401
w_T	1.2	1.8	4.5	7.2

Table A8. **Estimates of w_T from the "Approximate" Formula**

Brand E	Period of length (in weeks)			
	1	4	24	48
T	.083	.333	2	4
$T^{.82}$	.13	.41	1.76	3.10
w_T	1.3	1.8	4.5	7.2

For 4 weeks (with $T = .333$ again),

$$\frac{w_T}{w} = \frac{.333 \ln(.149)}{\ln(.149) - \ln(1-.568)}$$

$$= .333 \left(\frac{-1.904}{-1.904+.839}\right)$$

$$= \frac{.333}{.559}$$

$$= .596\ .$$

We therefore have the estimates $w_T = 3.0 \times .596 = 1.8$ in 4 weeks. The numerical value of the denominator in the expression for w_T/w (i.e. .559) need not be directly calculated if the LSD calculations for b_T have been carried out, since it is the value of b_T/b in Table A4. This is the procedure used in the summary Table A.7.

Finally we have the approximation to the LSD formula (Table 4.12)

$$(w_T - 1) = (w-1)\,T^{.82}\ .$$

Substituting $w = 3$ and rearranging gives

$$w_T = 1 + 2T^{.82}\ .$$

In 4 weeks this gives

$$w_{.333} = 1 + 2(.333)^{.82}$$

$$= 1 + 2 \times .406$$

$$= 1.8\ .$$

The values of $T^{.82}$ are already available in Table A5, for the various values of T, and Table A8 shows the approximation for all the time-periods.

It should perhaps be stressed that the reason why the calculations for b_T (Table 3.1a) and w_T (Table 3.2a) have so much in common is that if b_T is known then w_T can easily be calculated from the two simple relationships

$$m_T = Tm \quad \text{and} \quad w_T = m_T/b_T.$$

The calculations set out here are slightly more lengthy to illustrate the various formulae for w_T explicitly.

A.6. Light and Heavy Buyers (Table 3.4)

We now describe the calculations for the theoretical frequency distributions as shown in Table 3.4. Given b and w (and hence k or a) in a given period, we can estimate how many people bought 0, 1, 2, 3 etc. times in this period.

In the NBD, p_r is the theoretical proportion of the population buying the brand r times in the chosen time-period (see §§4.2 and 7.3). The numerical values of p_r are best calculated using the recursive formula,

$$p_r = \left(\frac{a}{1+a}\right)\left(1 - \frac{a-m}{ar}\right)p_{r-1}\,.$$

For Brand E in 48 weeks, we have (Table A1) $a = 19.18$, and $m = .816$.

Therefore,

$$p_r = \left(\frac{19.18}{20.18}\right)\left(1 - \frac{19.18-.816}{19.18r}\right)p_{r-1}$$

$$= 0.9504\left(1-\frac{.9575}{r}\right)p_{r-1}\,.$$

The observed penetration of Brand E in 48 weeks was $b = .12$ (Table 3.1a or Table A1), and so $p_0 = 1-b = .88$. Using this value we can now estimate p_1, and proceed to higher values:

$$p_1 = .9504(1-.9575)\,.88 = .0355,$$

$$p_2 = .9504(1-.4787)\,.0355 = .0176,$$

$$p_3 = .9504(1-.3192)\,.0176 = .0114,$$

$$p_4 = .9504(1-.2394)\,.0114 = .0083,$$

$$p_5 = .9504(1-.1915)\,.0083 = .0064,$$

and so on. To estimate the cumulative "tail" for 6 or more purchases, we have

$$p_{6+} = 1 - p_0 - p_1 - p_2 - p_3 - p_4 - p_5 = .0408.$$

These proportions are of the population as a whole. Expressed as p'_r, the proportion of *buyers* (i.e. the 12% of the population who bought E in the year) who bought r times, they are

$$p'_r = 100\,p_r/12,$$

which is the form given in Table 3.4.

The LSD formula is of course expressed directly in p'_r, i.e.

$$p'_r = \frac{-1}{\ln(1-q)}\,\frac{q^r}{r}\,.$$

The proportion buying *once* is therefore

$$p'_1 = \frac{-1}{\ln(1-q)}\,q = \frac{-.955}{\ln(.045)}$$

$$= .3080 \quad \text{or} \quad 31\%.$$

Values for p'_r for $r > 1$ can best be calculated by the recursive formula

$$p'_r = (r-1)\,q p'_{r-1}/r\,.$$

This gives

$$p'_2 = .3080 \times .955/2 = .2941/2 = .15,$$

$$p'_3 = .2941 \times .955/3 = .2809/3 = .09,$$

$$p'_4 = .2809 \times .955/4 = .2683/4 = .07,$$

$$p'_5 = .2683 \times .955/5 = .2562/5 = .05,$$

$$p'_{6+} = 1 - .31 - .15 - .09 - .07 - .05 = .33.$$

Table A9 compares the observed percentages of light and heavy buyers with these NBD and LSD norms in the form $100p'_r$.

A.7. The Sales Importance of Light and Heavy Buyers (Table 3.5)

Table 3.5 shows in percentage terms what proportion of total sales are accounted for by those consumers who make precisely r purchases each in the analysis period. This amounts to

$$100\, rp_r/m$$

(or $100Nrp_r/Nm$ for a sample of N consumers). Since r and m are given, the theoretical values of this ratio can be simply estimated from the NBD values of p_r in §A.5, and the computational problem is almost trivial. For example, for Brand E in the 48-week period and for $r = 2$, p_2 was .176, so that the share of sales accounted for out of an m-value of .816 (Table A1), is

$$\frac{100 \times 2 \times .0176}{.816} = 4.4\%.$$

The other NBD values follow similarly, with the "tail" (e.g. purchases made by buyers buying 6+ times, say) being obtained by subtraction. The results are set out in Table A10.

In the LSD model, the results can either be obtained by the corresponding calculations in terms of p'_r or through a special formula which gives the answer directly (see §§2.3, 4.4 and 8.4). Thus the proportion of sales accounted for by buyers who make more than r purchases is

$$q^r .$$

The proportion of sales accounted for by those buyers who each make exactly r purchases is therefore

$$q^{r-1} - q^r .$$

Only a tabulation of q^{r-1} for various values of r is therefore required, from which the required estimates can be obtained by subtracting successive terms. This is illustrated for Brand E in Table A10, the value of the 6+ tail being either obtained by subtraction or simply as q^5.

Table A9. The Percentage of Buyers Making 1, 2, 3, etc. Purchases in the Year

Brand E	No. of Purchases in a Year					
	1	2	3	4	5	6+
Obs.	36	12	6	5	6	36
NBD	30	15	10	7	5	34
LSD	31	15	9	7	5	33

Table A10. LSD and NBD Estimates of the Percentage of the Annual Sales of Brand E Accounted for by Buyers Making r Purchases of the Brand E

r	1	2	3	4	5	6+
LSD Estimate						
q^{r-1}	.955	.912	.871	.832	.795	.759
$100(q^{r-1}-q^r)\%$	4.5	4.3	4.1	3.9	3.7	79.5*
NBD Estimate						
$100rp_r/m\%$	4.4	4.4	4.2	4.1	4.0	78.9*

* The 6+ tail is obtained by subtracting previous terms from 1 (and equals q^5 for the LSD).

A.8. The Incidence of Repeat-Buyers (Table 3.6)

The commonest way of calculating the norms for repeat-buying from one period to another is from the values of b and w (or $m = bw$) in the first of the two periods. Sometimes it is also relevant to calculate the results backwards, i.e. given the pattern of buying in the *second* period, how many of the buyers should also have bought in the preceding one? The theoretical calculations which were given in Tables 3.6–3.9 are slightly different still, in that they were based on the b and w values averaged for both periods (and averaged across all the pairs of periods analysed). If we are dealing with a strictly stationary situation, the b and w values are the same in all the periods and none of these variations matters. In practice, some minor fluctuations in the b and w values occur from period to period (as documented in Tables 3.1 and 3.2), and in any specific problem-solving work this needs to be allowed for.

To illustrate the nature of the calculations here, we give the calculation based on the *average* values of b and w for all the periods of each length, which are given in Tables 3.1a and 3.2a. (Some of the resultant theoretical estimates therefore differ slightly from the theoretical values given in the tables in Chapter 3, which were averages of the separate theoretical estimates based on the values of b and w in each individual period.) The data for the average 4-week period are used to illustrate the detailed calculations, for the NBD, the LSD and the "approximate" formulae.

The NBD formula for b_R, the proportion of the population who are repeat-buyers, is (Table 4.6 and §7.6)

$$b_R = 1 - 2(1+a)^{-k} + (1+2a)^{-k}.$$

For Brand E in 4 weeks we have $a = 1.32$ and $k = .0485$, so that

$$\begin{aligned} b_R &= 1 - 2(1+1.32)^{-.0485} + (1+2 \times 1.32)^{-.0485} \\ &= 1 - 2 \times .9599 + .9391 \\ &= .0193\,. \end{aligned}$$

The ratio of those repeat-buying to those who bought in the first period, b_R/b, is the theoretical quantity given in Table 3.6. Thus with $b = .04$ for Brand E in 4 weeks,

$$b_R/b = .0193/.04 = .48, \quad \text{or} \quad 48\%.$$

The NBD calculations are summarised in Table A11.

The LSD formula is expressed directly as the ratio of repeat-buyers to buyers in the first period,

$$b_R/b = 1 + \frac{\ln(1+q)}{\ln(1-q)}.$$

In 4 weeks for Brand E we have $q = .584$ (Table A2) and so

$$\begin{aligned} b_R/b &= 1 + \frac{\ln(1+.584)}{\ln(.416)} \\ &= 1 - \frac{.4600}{.8771} = .48 \quad \text{or} \quad 48\%. \end{aligned}$$

The LSD calculations for all the time-periods are summarised in Table A12.

Table A11. The NBD Estimates of the Incidence of Repeat-Buyers

(The values of a, k and b for each period are from Table A1)

Brand E	Periods of length (in weeks)		
	4	12	24
$1+a$	2.32	6.34	13.00
$1+2a$	3.64	11.68	25.00
$(1+a)^{-k}$	.9599	.9299	.9101
$(1+2a)^{-k}$	.9391	.9080	.8886
b_R	.0193	.048	.068
b_R/b	.48	.69	.75

Table A12. The LSD Estimates of the Proportion of Repeat-Buyers

Brand E	Periods of length (in weeks)		
	4	12	24
q	.584	.851	.928
$\ln(1+q)$	.460	.616	.656
$\ln(1-q)$	−.877	−1.904	−2.631
b_R/b	.48	.68	.75

Table A13. The "Approximate" Estimates of the Proportion of Repeat-Buyers

Brand E	Periods of length (in weeks)		
	4	12	24
w	1.6	3.0	4.9
$2(w-1)$	1.2	4.0	7.8
$2.3w-1$	2.7	5.9	10.3
b_R/b	.45	.68	.76

The approximate formula for the proportions of repeat-buyers in Table 4.6 and §8.6 was

$$\frac{b_R}{b} = \frac{2(w-1)}{2.3w-1},$$

where w is the only input required. For Brand E in the average 4-week period,

$$w = 1.6$$

$$\frac{b_R}{b} = \frac{2 \times .6}{2.3 \times 1.6 - 1} = \frac{1.2}{2.68} .$$

$$= .45 \quad \text{or} \quad 45\%.$$

This and the approximate calculations for the other periods are set out in Table A13.

The observed values of b_R are necessarily the same in both periods which are being analysed (they are the same people), but the observed b_R/b ratios can vary if b varies from the 1st to the 2nd period if there is some non-stationarity. Similarly, the theoretical values of b_R or b_R/b depend on whether the input is that observed in the first or in the second period (or an average of the two). This only matters slightly when dealing with cases of slight non-stationarity, but becomes crucial when using the theoretical norm to interpret major changes from one period to the other (as for example in §§ 2.4 and 6.2).

A.9. The Buying-Frequency per Repeat-Buyer (Table 3.7)

To illustrate the theoretical calculations of the average purchase frequency for repeat-buyers, we take the average 12-week period in Table 3.7.

The NBD formula for m_R, the number of purchases made by repeat-buyers but expressed on a *per informant* basis, gives

$$m_R = m\{1 - (1+m/k)^{-k-1}\}$$

$$= m\{1 - (1+a)^{-k-1}\} .$$

Table A14. The NBD Estimates of m_R or w_R

Brand E	Periods of length (in weeks)		
	4	12	24
m	.064	.21	.44
a	1.32	5.34	12.00
k	.0485	.0393	.0367
$(1+a)^{-k-1}$	.4143	.1466	.0700
m_R	.0375	.1790	.4092
b_R(T.A11)	.0193	.048	.068
w_R	1.9	3.7	6.0

Table A15. The LSD and "Approximate" Estimates of w_R

Brand E	Periods of length (in weeks)		
	4	12	24
q	.584	.851	.928
q^2	.341	.724	.861
$\ln(1-q^2)$	−.417	−1.287	−1.973
$(1-q)\ln(1-q^2)$	−.173	−.192	−.142
w_R	2.0	3.8	6.1
w	1.6	3.0	4.9
$1.23w$	2.0	3.7	6.0

If w_R, the rate of repeat-buying *per repeat-buyer*, is wanted, this is easily obtainable from the relationship

$$w_R = m_R/b_R \,,$$

the theoretical value of b_R being available from Table A.11.

In 12 weeks we have m = .21, a = 5.34, k = .0393 from Table A1, so that

$$m_R = .21 \,\{(1 - (6.34)^{-1.0393})\} \;,$$

$$= .21 \;(1 - .1466) = 0.1790 \,,$$

and $$w_R = \frac{.179}{.048} = 3.7.$$

Table A14 summarises these calculations for all the time-periods.

The corresponding LSD formula for w_R is

$$w_R = \frac{-q^2}{(1-q)\ln(1-q^2)} \quad ,$$

so that in 12 weeks, where $q = .851$,

$$w_R = \frac{-.724}{.149(-1.287)} = \frac{.724}{.192}$$

$$= 3.8$$

Table A15 gives the calculations for 4-, 12- and 24-week repeat-buying. The approximate formula for w_R which can be derived from the LSD is

$$w_R \doteqdot 1.23w,$$

and the results are also set out in Table A15.

Unlike for b_R, the observed values of w_R can differ from the first to the second time-periods because of any non-stationarity (even if only slight), and the theoretical values similarly depend on which period's *observed* values are used as input.

A.10. The Buying-Frequency per "New" Buyer (Table 3.8)

To illustrate the calculations for the theoretical frequency of purchase of "new" or of "lapsed" buyers, we take the two 24-week periods.

The new buyers' purchasing frequency is given on a per informant basis in the NBD theory, as

$$m_N = \frac{m}{(1+a)^{k+1}} \quad .$$

In 24 weeks, $m = .44$, $a = 12.00$, $k = .0367$, so that

$$m_N = \frac{.44}{(13.00)^{1.0367}} = \frac{.44}{14.280} = .0308.$$

The buying-frequency per "new" buyer is given by

$$w_N = m_N / b_N \ ,$$

where the proportion of "new" buyers b_N is given by

$$b_N = b - b_R \ .$$

The total penetration b of Brand E in the 24-week periods being .09 (Table 3.2a) and the theoretical NBD value for b_R at .068 being available from Table A11, we have

$$b_N = .09 - .068 = .022 \ ,$$

so that

$$w_N = \frac{.0308}{.022} = 1.4.$$

Table A16. The NBD Estimates for m_N and w_N

Brand E	Periods of length (in weeks)		
	4	12	24
m	.064	.21	.44
a	1.32	5.34	12.00
k	.0485	.0393	.0367
$(1+a)^{1+k}$	2.417	6.818	14.280
m_N	.0265	.0308	.0308
$b_N = b - b_R$	.021	.022	.022
w_N	1.3	1.4	1.4

Table A17. The LSD and Approximate Estimates for w_N

Brand E	Periods of length (in weeks)		
	4	12	24
q	.584	.851	.928
$\ln(1+q)$	.460	.616	.656
w_N	1.3	1.4	1.4
Approximate Formula	1.4	1.4	1.4

This and the calculations for the average 4- and 12-week periods are given in Table A16.

The LSD formula for w_N is

$$w_N = \frac{q}{\ln(1+q)},$$

and in 24 weeks, $q = .928$, so that

$$w_N = \frac{.928}{\ln(1.928)} = \frac{.928}{.656}$$

$$\doteqdot 1.4.$$

Table A17 shows the calculation of w_N for the average 4- and 12-week periods as well as for the 24-week period, together with the values of the "approximate" formulae, which is simply

$$w_N \doteqdot 1.4.$$

The latter generally holds for $w > 2$ and values of b not too high (see §8.6), and here gives a fractionally different result from the LSD and NBD in the 4-week period, where $w < 2$.

The corresponding buying rates for the "lapsed" buyers (who buy in the first but not the second period) are equal to the above, subject to the effects of any non-stationarity, as already discussed in the previous section.

A.11. Repeat-Buying by Light and Heavy Buyers (Table 3.10)*

In the case of the "Conditional Trend" analysis in Tables 3.10 and 3.10a, only the NBD calculations (§7.6) need to be considered, since the LSD theory does not lead to any effective simplification here.

Tables 3.10 and 3.10a referred specifically to the repeat-buying in Quarters II and III, so that the NBD parameters for Quarter II are required. Their method of determination is that illustrated in §A3 at the

* The theoretical calculations required for Table A9 are not discussed here as they are like those already given in §A7 for Table 3.6a, but as usual, care has to be taken over using the appropriate base-period.

beginning of the Appendix, and here only the numerical results are quoted. For Quarter II, $b = .062$ and $w = 3.2$ *, and this leads to $m = .198$, $a = 6.037$, and $k = .0361$.

The calculations for Tables 3.10 and 3.10a use the conditional formula referred to in §7.6 of Chapter 7. This gives $p_{s/r}$, i.e. the proportion of those buying r times in the first period who buy s times in the second period,

$$p_{s/r} = (1+a')^{-k'} \frac{\Gamma(k'+s)}{\Gamma(k')\Gamma(s+1)} \left(\frac{a'}{1+a'}\right)^s ,$$

where

$$k' = k+r \quad \text{and} \quad a' = a/(1+a) .$$

For Table 3.10, we only require the proportion $p_{./r}$ who buy *at all* in the second period of those who made r purchases in the first period. This is 1 minus $p_{0/r}$, the proportion *not* buying in the second period, i.e.

$$\begin{aligned} p_{./r} &= 1 - p_{0/r} \\ &= 1 - (1+a')^{-k'} \\ &= 1 - (1+a')^{-(k+r)}, \end{aligned}$$

from the above expression for $p_{s/r}$, putting $s = 0$. This is easily evaluated numerically once $(1+a')^{-1}$ and $(1+a')^{-k}$ have been determined. Thus

$$\begin{aligned} a &= 6.037, \\ a' &= \frac{6.037}{1+6.037} = .8579, \\ (1+a')^{-1} &= .5382, \\ (1+a')^{-k} &= (1.8579)^{-.0361} \\ &= .9781, \end{aligned}$$

* These values are given to two significant figures (compared with one in Tables 3.1 and 3.2), as the detailed calculations here are susceptible to rounding errors.

Table A18. NBD Estimates of Repeat-Buying in the Second Period

Brand E	Number of Purchases in 1st Period			Total
	$r=0$	$r=1$	$r=2+$	
Buyers in 2nd Period				
As % of 1st period buyers $p_{./r}$	2.2	47	86	–
As % of total sample *	1.9	1.3	3.0 ***	6.2
Av. Purchases in 2nd Period				
Per buyer of r in 1st period, $m_{./r}$	.031	.889	4.04	–
Per repeat-buyer, $w_{./r}$	1.41	1.88	4.70	–
Per informant	.0268 **	.0237 **	.1475 **	.198

* Prop. of 1st Period Buyers times $p_{./r}$.
** Numerically the same as the proportion of sample buying (r+1) times in the 1st period (Table A19).
*** By subtraction.

and from this we get $p_{./r}$ either directly as $1-(1+a')^{-k}\{(1+a')^{-1}\}^r$, or or by the recurrence formulae $(1-p_{./r})=(1-p_{./r-1})(1+a')^{-1}$.

For Table 3.10a we need $m_{./r}$, the mean of the theoretical distribution of the number of purchases in the second period conditional upon having made r purchases in the first. This is given by

$$m_{./r}=a'(k+r)\,.$$

With these general expressions for $p_{./r}$ and $m_{./r}$ for any r, we can now calculate the specific values for $r=0$, 1, 2, etc.

Starting with non-buyers in the first period, i.e. $r=0$, we have

$$\begin{aligned} p_{./0} &= 1-(1+a')^{-k} \\ &= 1-.9781 \\ &= .0219 \quad \text{or} \quad 2.2\%, \end{aligned}$$

and

$$\begin{aligned} m_{./0} &= a'k \\ &= .8579 \times .0361 = .0310. \end{aligned}$$

Hence the buying frequency,

$$w_{./0} = \frac{m_{./0}}{b_{./0}} = \frac{.0310}{.0219} = 1.41.$$

(This is simply w_N, the buying frequency per "new buyer", as in Table 3.8.) These values are set out in Table A18 for $r = 0$ *.

Next, for once-only buyers in the first period, i.e. $r = 1$, we have

$$\begin{aligned} p_{./1} &= 1 - (1+a')^{-k-1} \\ &= 1 - (1+a')^{-k}(1+a')^{-1} \\ &= 1 - .9781 \times .5382 \\ &= 1 - .5264 \\ &= .474 \quad \text{or} \quad 47\%, \end{aligned}$$

and

$$\begin{aligned} m_{./1} &= a'(k+1) \\ &= .8579 \times 1.0361 = 0.889. \end{aligned}$$

Hence

$$w_{./1} = \frac{m_{./1}}{b_{./1}} = \frac{.889}{.474} = 1.88\,.$$

These values are set out in the $r = 1$ column in Table A18.

And so one can proceed for $r = 2$, $r = 3$, etc. In practice one will stop at some "maximum" value of r. Beyond this we then require the repeat-buying estimates for all buyers who bought more often than this maximum in the first period. These theoretical values of the number of repeat-buyers amongst the "more than r" buyers, and their average purchase frequency in the second period, have to be found by subtracting the appropriate values for $r = 0$, for $r = 1$, for $r = 2$, etc. from the *total* number of buyers and the *total* number of purchases predicted for the second period (expressing all values on a "per informant" basis so as to avoid having explicitly to introduce the sample size). This can

* The derivation of the number of repeat-buyers as a percentage of the total sample (i.e. 1.9% for $r = 0$ in Table A18) is given later in this section.

be illustrated here in terms of the "more than once only" buyers in the first period.

We therefore need to express the repeat-buying values $p_{./r}$ on a "per informant in the total sample" basis. The proportion of informants buying in the first period was .062 (see above) and so the proportion *not* buying is .938. Under the NBD, we have worked out above that 2.2% of these non-buyers ($r = 0$) in the first period buy in the second period (see $p_{./0}$ above). Therefore the theoretical value for the percentage of *all* informants of those who, not having bought in the first period, buy in the second is

$$.02 \times .938 = .019 \text{ or } 1.9\%.$$

Further, they buy at a rate of 1.41 ($w_{./0}$ above), giving a theoretical total of

$$(1.41 \times 1.9)/100 = .0268$$

purchases per informant, as is shown in Table A18.

For $r = 1$, the percentage of households who bought once only in the first period is first of all required. This can of course be calculated directly from the p_r formula in §A6 with $r = 1$, but the calculation is simplified for the present purpose by a certain feature of the NBD, namely that the number of purchase occasions in the second period made by those households who bought r times in the first period is numerically the same as the sheer *number* of households who bought $(r+1)$ times in the first period! It follows that the proportion of households who bought only once in the first period is .0268 (i.e. the numerical value calculated – for a quite different purpose – in the preceding paragraph). This is set out in Table A19.

Table A19. The NBD Estimates of p_r in the First Period

Brand E	r, the Number of Purchases			Total
	0	1	2+	
NBD Proportions p_r	.9380	.0268	.0352 *	1.000

* By subtraction from 1.

Next, of this proportion .0268 of once-only buyers in the first period, 47% were expected to buy in the second period (the $p_{./1}$ value calculated above). The number of these repeat-buyers on a per informant basis is therefore

$$.0268 \times 47 = .0126 \quad \text{or} \quad 1.3\%.$$

These particular repeat-buyers are expected to buy at a rate of 1.88 ($w_{./1}$ above) and therefore account for

$$1.88 \times .0126 = .0237 \text{ purchases on a per informant basis.}$$

We can now calculate the expected proportion of buyers in the second period who bought *more than once* in the first period, namely by subtraction,

$$6.2 - 1.9 - 1.3 = 3.0\%,$$

where 6.2% is of course the total number of buyers in the second period (equal to the number of buyers in the first period). The number of purchases they are expected to make is, similarly,

$$.198 - .0268 - .0237 = .1475 \text{ per informant,}$$

where .198 is the total number of purchases in each period per informant (as given at the beginning of this section).

By subtraction we also find that the proportion of informants who bought more than once in the first period is

$$1.000 - .938 - .0268 = .0352 \quad \text{or} \quad 3.5\%,$$

as shown in Table A19.

Therefore the 3.0% who are expected to buy again in the second period represent

$$\frac{3.0}{3.5} = 86\%$$

of the initial more-than-once buyers. The rate at which they are expected to buy in the second period is

$$\frac{.1475 \times 100}{3.0} = 4.9\ .$$

These results are shown in the 1+ column of Table A18.

A.12. Tabulating Sales, Penetration, and Purchase Frequency

Condensing the raw data from a consumer panel for a single brand into the summary statistics like m, b, and w discussed so far is relatively simple because we need only consider the data for that single brand. It involves identifying the panel member, the brand bought on each purchase occasion and the point in time.

As discussed in § 1.4, all the analyses are in terms of purchase occasions, not amounts bought or paid. (The definition of a purchase occasion may raise problems in some data, where two separate purchases of the same brand made at the same retail outlet in a week may not be distinguishable from two packs of the brand being bought on the same occasion.)

The only step needed beyond simple counts concerns the penetration b of a brand in a given time period; for this, we only count the first time a panel-member buys the brand in the period. In the tabulations we therefore need to allow for whether the buyer has bought that brand before in that time-period.

Data records can be provided in one of several different ways, e.g.

(a) Ordered by panel member. Here all purchases for an individual consumer or household are shown consecutively for a time-period. This time-period (e.g. a year) may be longer than the period we intend to study (e.g. a particular 4 weeks).

(b) Ordered by week. All purchases by different consumers in Week 1 are arranged in some order, then all purchases in Week 2, and so on.

(c) No apparent order.

Especially for hand tabulations it is usually easiest to aim at layout (a) and then order by (i) brand and by (ii) time sequence within brand for each individual. This simplifies the tabulation of penetrations.

To illustrate with a very small-scale hypothetical example, Table A20 sets out the purchases of three brands X, Y and Z over 12 weeks. The total sample is 200 households, of whom 180 did not buy the product-category in that 12-week quarter.* Suppose we want to tabulate summary statistics

*The example will also be used in the new Appendix C. It covers multi-brand buying and therefore is more elaborate than the numerical example used earlier in this Appendix.

for Brand X. We consider statistics mainly for the first 4 week period (similar tabulations will be needed later for the other two months and for the whole quarter), namely:

- Sales, or average purchase occasions for Brand X in the period per individual, denoted m as in the algebraic notation of the main text.
- Penetration, or % buying X at least once in the period, b.
- The frequency distribution of the number of purchases, p_r.
- The average number of purchases per buyer, $w = m/b$.

In doing the counts by hand, we can use five-barred "gates" (e.g. 𝍸), as shown on the right of Table A20. To count "buying at least once" we only note the first purchase by that individual in the time-period.

The tabulation shows that in Weeks 1–4, 6 buyers bought on 8 purchase occasions (with the tally of gate counts and "totals" both summing to 8). Hence with a sample of 200,

$$\begin{aligned} m &= 8/200 = .04, \\ b &= 6/200 = .03 \text{ or } 3\%, \\ w &= 8/6 \quad = 1.33. \end{aligned}$$

The frequency distribution of purchases, counting the number of 0s, 1s, 2s, etc out of 200 in the total column, gives

Number of Purchase Occasions in Weeks 1–4

0	1	2	3	4+
194	5	–	1	–

The figures add up to 200 and give $5 \times 1 + 1 \times 3 = 8$ purchase occasions, calculations which are essential checks on the arithmetic.

The answers for all three 4-week periods and the total 12 weeks from Table A20 should come out to be

For Brand X	1st 4 weeks	2nd 4 weeks	3rd 4 weeks	12 weeks
Total Purchase Occasions	8	9	8	25
Total Buying at least once	6	7	7	13
m	.040	.045	.040	.125
b	.030	.035	.035	.065
w	1.33	1.28	1.14	1.92

Table A20. A Simplified Buying Pattern: Three Brands over 12 Weeks
(Tabulating X in weeks 1 to 4)

Panel-Member	Purchases in Week: 1	2	3	4	5	6	7	8	9	10	11	12	Brand X in Weeks 1–4: Gate Count	Total	Buying at least once
1	X	X	Y	X	X	X	Y	X	.	Y	X	Y	///	3	1
2	.	.	X	.	Z	X	.	.	.	Z	X	X	/	1	1
3	.	.	Z	.	.	Z	.	X	X	.	Z	.		-	-
4	.	.	.	.	Y	.	Y	Y	.	X	.	.		-	-
5	Y	X	.	.	X	.	.	.	.	.	.	Y	/	1	1
6	X	.	.	.	.	.	.	Z	Z.	Y	.	.	/	1	1
7	.	Z	X	.	.	.	.	.	.	.	.	X	/	1	1
8	.	Y	.	Y	.	.	.	.	X	.	.	.		-	-
9	Z	.	.	.	.	.	X	.	.	.	Y	.		-	-
10	.	.	.	X	.	Y	.	.	.	.	.	.	/	1	1
11	.	.	.	Y	.	.	X	.	.	.	.	.		-	-
12	.	.	.	.	.	.	.	.	.	X	.	.		-	-
13	Y	.	.	.	.	.	.	.	.	.	.	.		-	-
14	.	.	.	.	.	.	.	Y	.	.	.	.		-	-
15	.	.	Z	.	.	.	.	.	.	.	.	.		-	-
16	.	.	.	.	.	.	.	.	Y	.	.	.		-	-
17	.	.	.	.	Z	.	.	.	.	.	.	.		-	-
18	.	.	.	.	.	.	.	.	.	.	Z	.		-	-
19	.	.	.	.	.	.	.	.	.	Z	.	.		-	-
20	.	.	.	.	.	X	.	.	.	.	.	.		-	-
21–200	.	.	.	.	.	.	.	.	.	.	.	.		-	-
Totals	2	2	2	2									8	8	6

It is easy to make mistakes and so it is once more essential to do the following checks:

(a) The three 4-week totals should add up to the 12-week total of 25,
(b) The three 4-week *m*s should add up to the 12-week *m* of .125,
(c) The sum of the three 4-week *b*s (.100) should be greater (or possibly equal) to the 12-week *b* (.065).

If the data are almost stationary, then all the 4-week figures will be very similar. (The rather "low" *w* of 1.14 in the 3rd 4-week period is, given the sample size, in effect due to 1 purchase or 1 buyer. Thus $8/7 = 1.14$, whereas $9/7 = 1.29$ or $8/6 = 1.33$, close to the figures for the previous two 4-week periods).

As a further example of the tabulations, the frequency distribution of purchases of X in the whole 12-week period, using gate-counts as an intermediate step, is

Number of Purchase Occasions in Weeks 1-12

0	1	2	3	4	5	6	7	8+
	𝍸 |||	|||		|			|	
187	8	3	–	1	–	–	1	

which correctly adds to 200 individuals again and to 25 purchases

(i.e. $8 \times 1 + 3 \times 2 + 1 \times 4 + 1 \times 7 = 25$).

A.13. Tabulating Period-by-Period Repeat-Buying

To illustrate the tabulation of the repeat-buying statistics from one period to another equal-length period, we consider weeks 1–4 and 5–8 from Table A20. Table A21 shows one possible way of arranging the hand-tabulations. (More elaborate versions are used later.) The table gives the counts that are required for each period, care being needed again in tabulating the penetrations (i.e. counting those who buy at least once in *both* periods, or in only one or the other).

Table A21. Repeat Buying of Brand X, Weeks 1–4 and 5–8

	Purchase Occasions			
	Counts		Totals	
Panel-member	1–4	5–8	1–4	5–8
1	///	///	3	3
2	/	/	1	1
3		/	–	1
4			–	–
5	/	/	1	1
6	/		1	–
7	/		1	–
8			–	–
9		/	–	1
10	/		1	–
11		/	–	1
12			–	–
13			–	–
14			–	–
15			–	–
16			–	–
17			–	–
18			–	–
19			–	–
20		/	–	1
<u>Number of Purchases</u>				
All			8	9
by Repeat Buyers			5	5
by Lapsed Buyers			3	–
by New Buyers			–	4
<u>Number of Buyers</u>				
All			6	7
Repeat Buyers			3	3
Lapsed Buyers			3	–
New Buyers			–	4

The first two columns give the five-barred-gate counts done by hand (which are more helpful than here when one is dealing with a larger number of purchases in longer time-periods). They are to be filled in one row and column at a time (i.e. period-by-period for each panel-member) from Table A20. We then form the totals for each period, as shown in type-set form in the last two columns, and visually differentiate repeat-buyers who bought in each period, by encircling them by hand. We can now form various totals for the total panel as follows, first for the number of purchases and then for the number of buyers:

Purchases

(a) Period totals, i.e. total purchases in Weeks 1–4 (8) and Weeks 5–8 (9).

(b) By repeat-buyers, i.e. purchases in each period by the (encircled) buyers who bought in both (the two sets of numbers happen to be equal, giving a total of 5 in each period).

(c) By lapsed buyers, i.e. (non-encircled) purchases in the first period (which is 3 purchases by those who bought in that period only).

(d) By new buyers, i.e. (non-encircled) purchases in the second period.

Buyers

In the same way we can count the number of buyers (i.e. their sheer occurrence, without taking account of how often they buy), for

(a) Each period (6 and 7),

(b) Repeat-buyers (3 and 3),

(c) Lapsed buyers (3),

(d) New buyers (4).

It is unavoidable to run down each column separately six times to perform these counts. (The first time round this is usually quicker than getting a computer program to work!)

Various checks on the arithmetic can be done. Thus

— The total number of purchases in each period should agree with those recorded in Table A20 (i.e. 8 and 9).

— The two entries in (a) for Weeks 1–4 and 5–8 must each equal the sums of (b) and (c) or of (b) and (d).

— The numbers of repeat-buyers in each of the two periods must be equal, whereas how often they bought need not be equal.

The repeat-buying summary statistics can now be calculated with the sample-size $n = 200$.

Purchase Occasions			Buying at Least Once		
Weeks	1–4	5–8	Weeks	1–4	5–8
m	.040	.045	b	.030	0.35
w	1.33	1.29	–	–	–
m_R	.025	.025	b_R	.015	.015
w_R	1.67	1.67	b_R/b	50%	43%
m_L	.015	–	b_L	.015	–
w_L	1.00	–	b_N	–	.020
m_N	–	.020	b_L/b	50%	–
w_N	–	1.00	b_N/b	–	57%

To check the arithmetic, we note that

(i) $m = m_R + m_L$ in Weeks 1–4, i.e. $.040 = .025 + .015$, and $m = m_R + m_N$ in Weeks 5–8.

(ii) $b = b_R + b_L$ in Weeks 1–4, i.e. $.030 = .015 + .015$, and $= b_R + b_N$ in Weeks 5–8.

(iii) We also have $b_R/b + b_L/b = 100$ and $b_R/b + b_N/b = 100$ in each period.

Extensions of these tabulations, such as separate "conditional" ones for Light, Medium or Heavy buyers (defined as desired) are straightforward in principle. In practice, it is best to mark the panel-members first (e.g. by *L*, *M* and *H*) using their purchases in period 1, say, as the criterion, and then run three separate sets of counts for the summary statistics at the bottom of the table. (As a check, the three sets of figures should add up to the relevant totals of the foot of Table A21). (It may reduce counting errors to make several Xerox copies of Table A20 and block out Light, Medium or Heavy buyers separately on each sheet.)

A.14. Continuous Reporters

The tabulations in A.12 and A.13 help to emphasise the need for basing the analysis on "continuous reporters", sometimes called a "static sample'', over the period analysed. In a diary panel, for example, it is misleading if a panel-member is recorded as a non-buyer of brand *X* in some week or weeks when in fact he or she did not return a diary that week.

Continuous reporting can be difficult to ascertain because in some panel operations no count is kept of whether a usable diary is returned by each panel-member every week, or at least such data are not readily accessible. A partial (but time-consuming) solution is to form a panel of continuous reporters by checking whether each individual recorded any purchases in *any* of the product-categories measured by the panel.

A definition of a ''continuous reporter'' is however somewhat fuzzy. Absence on holiday should presumably be counted as ''non-purchasing but reporting'', if the aim is to cover ''purchases at home''. But absence (or illness) of the normal diary keeper only (with other household members still buying and consuming at home) necessarily introduces ''errors'' in the data, hopefully minor ones.

In general, for the type of studies discussed in this book it seems better to have a relatively small sample of ''more rather than less'' continuous reporters, even if this is potentially biased in other statistical respects, e.g. by age of or even by their amount of purchasing. Repeat-buying patterns for a slightly biased sample are likely to be more valid than ones that are wrong because of discontinuous reporting, even though the sample itself may be more directly representative of the population at large.

APPENDIX B

SOME USEFUL TABLES

This Appendix gives some tables which tend to be useful in the analysis of repeat-buying behaviour along the lines discussed in this book.

Table B1 relates to the LSD and gives the value of the LSD parameter q for a given value of w. The range of values of w is from 1.0 to 10.9 and is therefore much more detailed than the brief extract in Table 2.2 of Chapter 2. However, interpolation in this table (let alone extrapolation for $w > 10$) is relatively difficult because many of the values of q are close to 1. A better approach is to estimate the LSD parameter $a = q/(1+q)$ from w (and then either work with a or calculate q as $a/(1-a)$). A table for reading off a for a given value of w is given in Table B3.

Table B2 essentially gives Naperian or Natural logarithms for numbers running from 0 to 2. This is useful for computing expressions such as $\ln(1-q)$ and $\ln(1+q)$ which appear in the LSD theory.

For convenience, the tables are given in two parts. Table B2a gives $\ln(1-q)$ directly for values of q from .000 to .999 (the values of $\ln(1-q)$ are always negative). This table can also be used to read off $\ln(1-b)$ for a given b, as is required for Table B3. Table B2b gives $\ln(1+q)$ for the same range of q.

Table B3 provides a way of reading off the parameter a in either the NBD or the LSD theory. (The two values will differ for any given data, although only slightly if the parameters are such that the LSD is a good approximation to the NBD.)

For the NBD, we have to calculate first the value of $c = -m/\ln(1-b)$ from the observed values of the mean m and the proportion of buyers b (Table B2a can be used to read off $\ln(1-b)$ for the given value of b). We then read off the value of the NBD parameter a for the calculated value of c from the Table B3. NBD parameter k can be calculated from $k = m/a$.

For the LSD, we can read the LSD values of a directly by entering the table with the observed value of w (where $w = m/b$).

Tables B4–B7 give the NBD and LSD values of four basic repeat-buying statistics in two equal periods under stationary conditions, as follows:
Table B4: $100b_R/b$, i.e. the percentage of the buyers (b) in one time-period who buy the item again in another equal period.
Table B5: w_R, the average purchase frequency per repeat-buyer in one of the periods.
Table B6: $100b_N/b$, i.e. the percentage of buyers in one period who do NOT buy the item in another equal period (where the proportion of "lapsed" buyers is $b_L = b_N = 1-b_R$).
Table B7: w_N, i.e. the average purchase frequency per "new" (or per "lapsed") buyer in the period in which they buy at all ($w_N = w_L$).

Values are given for the NBD and LSD and "approximation" formulae which were set out in Chapters 4, 7 and 8. In all cases, the observed values of b and m (for the NBD) or of w (for the LSD) in *one* of the two time-periods are used to enter the tables.

Only a limited range of values of b and w is covered. The tables are mainly meant to give a feel of how repeat-buying varies numerically with b and w, showing also how for low values of b the NBD values all closely approximate the LSD ones. (For detailed numerical work, direct calculation of the various statistics required rather than reading values off from extensive tables tends to be more useful.)

Tables B8–B9 illustrate the variation with the length of analysis-period of the penetration b_T and of the average purchase frequency w_T in a time-period of length T (relative to some chosen period of arbitrary "unit" length). This has been done by showing b_T and w_T in each case as a fraction of the value of b_1 and w_1 in the unit time-period. To obtain values of $100b_T$ and w_T simply multiply by the values of $100b_1$ and w_1 given in the centre columns of these tables. When interpolating in these tables work in terms of b_T/b_1 and w_T/w_1, and only as a final stage multiply up by the observed values of $100b_1$ and w_1 to obtain $100b_T$ and w_T.

The last block in each table results from the LSD formulation of the model. This is the limiting case as b tends to zero, and the two blocks have been labelled $100b_1 = 0$ accordingly. They may be used either to obtain LSD values for b_T/b_1 and w_T/w_1 (and hence b_T and w_T) when b_1 is small, or to interpolate for the NBD model in cases where the value of $100b_1$ is less than 5%.

Note that NBD's do not exist for high values of b_1 when w_1 is low.

Table B1. Values of q for Values of w from 1.0 to 10.9

w	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1.0	.000	.171	.298	.396	.472	.534	.584	.625	.660	.690
2.0	.715	.737	.757	.774	.789	.802	.814	.825	.834	.843
3.0	.851	.858	.865	.871	.877	.882	.887	.892	.896	.900
4.0	.903	.907	.910	.913	.916	.919	.921	.924	.926	.928
5.0	.930	.932	.934	.936	.937	.939	.940	.942	.943	.945
6.0	.946	.947	.948	.950	.951	.952	.953	.954	.955	.956
7.0	.956	.957	.958	.959	.960	.960	.961	.962	.962	.963
8.0	.964	.964	.965	.965	.966	.967	.967	.968	.968	.969
9.0	.969	.969	.970	.970	.971	.971	.972	.972	.972	.973
10.0	.973	.973	.974	.974	.974	.975	.975	.975	.976	.976

Table B.2a. Values of ln(1−q) for Values of q from .000 to .999

(ln stands for the Naperian or Natural Logarithm)

q	.000	.001	.002	.003	.004	.005	.006	.007	.008	.009
.00	−.000	−.001	−.002	−.003	−.004	−.005	−.006	−.007	−.008	−.009
.01	−.010	−.011	−.012	−.013	−.014	−.015	−.016	−.017	−.018	−.019
.02	−.020	−.021	−.022	−.023	−.024	−.025	−.026	−.027	−.028	−.029
.03	−.030	−.031	−.033	−.034	−.035	−.036	−.037	−.038	−.039	−.040
.04	−.041	−.042	−.043	−.044	−.045	−.046	−.047	−.048	−.049	−.050
.05	−.051	−.052	−.053	−.054	−.056	−.057	−.058	−.059	−.060	−.061
.06	−.062	−.063	−.064	−.065	−.066	−.067	−.068	−.069	−.070	−.071
.07	−.073	−.074	−.075	−.076	−.077	−.078	−.079	−.080	−.081	−.082
.08	−.083	−.084	−.086	−.087	−.088	−.089	−.090	−.091	−.092	−.093
.09	−.094	−.095	−.097	−.098	−.099	−.100	−.101	−.102	−.103	−.104
.10	−.105	−.106	−.108	−.109	−.110	−.111	−.112	−.113	−.114	−.115
.11	−.117	−.118	−.119	−.120	−.121	−.122	−.123	−.124	−.126	−.127
.12	−.128	−.129	−.130	−.131	−.132	−.134	−.135	−.136	−.137	−.138
.13	−.139	−.140	−.142	−.143	−.144	−.145	−.146	−.147	−.149	−.150
.14	−.151	−.152	−.153	−.154	−.155	−.157	−.158	−.159	−.160	−.161
.15	−.163	−.164	−.165	−.166	−.167	−.168	−.170	−.171	−.172	−.173
.16	−.174	−.176	−.177	−.178	−.179	−.180	−.182	−.183	−.184	−.185
.17	−.186	−.188	−.189	−.190	−.191	−.192	−.194	−.195	−.196	−.197
.18	−.198	−.200	−.201	−.202	−.203	−.205	−.206	−.207	−.208	−.209
.19	−.211	−.212	−.213	−.214	−.216	−.217	−.218	−.219	−.221	−.222
.20	−.223	−.224	−.226	−.227	−.228	−.229	−.231	−.232	−.233	−.234
.21	−.236	−.237	−.238	−.240	−.241	−.242	−.243	−.245	−.246	−.247
.22	−.248	−.250	−.251	−.252	−.254	−.255	−.256	−.257	−.259	−.260
.23	−2.61	−.263	−.264	−.265	−.267	−.268	−.269	−.271	−.272	−.273
.24	−.274	−.276	−.277	−.278	−.280	−.281	−.282	−.284	−.285	−.286

Table B2a continued: ln(1−q)

q	.000	.001	.002	.003	.004	.005	.006	007	.008	.009
.25	−.288	−.289	−.290	−.292	−.293	−.294	−.296	−.297	−.298	−.300
.26	−.301	−.302	−.304	−.305	−.307	−.308	−.309	−.311	−.312	−.313
.27	−.315	−.316	−.317	−.319	−.320	−.322	−.323	−.324	−.326	−.327
.28	−.329	−.330	−.331	−.333	−.334	−.335	−.337	−.338	−.340	−.341
.29	−.343	−.344	−.345	−.347	−.348	−.350	−.351	−.352	−.354	−.355
.30	−.357	−.358	−.360	−.361	−.362	−.364	−.365	−.367	−.368	−.370
.31	−.371	−.373	−.374	−.375	−.377	−.378	−.380	−.381	−.383	−.384
.32	−.386	−.387	−.389	−.390	−.392	−.393	−.395	−.396	−.398	−.399
.33	−.401	−.402	−.403	−.405	−.406	−.408	−.409	−.411	−.413	−.414
.34	−.416	−.417	−.419	−.420	−.422	−.423	−.425	−.426	−.428	−.429
.35	−.431	−.432	−.434	−.435	−.437	−.439	−.440	−.442	−.443	−.445
.36	−.446	−.448	−.449	−.451	−.453	−.454	−.456	−.457	−.459	−.460
.37	−.462	−.464	−.465	−.467	−.468	−.470	−.472	−.473	−.475	−.476
.38	−.478	−.480	−.481	−.483	−.485	−.486	−.488	−.489	−.491	−.493
.39	−.494	−.496	−.498	−.499	−.501	−.503	−.504	−.506	−.508	−.509
.40	−.511	−.513	−.514	−.516	−.518	−.519	−.521	−.523	−.524	−.526
.41	−.528	−.529	−.531	−.533	−.534	−.536	−.538	−.540	−.541	−.543
.42	−.545	−.546	−.548	−.550	−.552	−.553	−.555	−.557	−.559	−.560
.43	−.562	−.564	−.566	−.567	−.569	−.571	−.573	−.575	−.576	−.578
.44	−.580	−.582	−.583	−.585	−.587	−.589	−.591	−.592	−.594	−.596
.45	−.598	−.600	−.602	−.603	−.605	−.607	−.609	−.611	−.613	−.614
.46	−.616	−.618	−.620	−.622	−.624	−.626	−.627	−.629	−.631	−.633
.47	−.635	−.637	−.639	−.641	−.642	−.644	−.646	−.648	−.650	−.652
.48	−.654	−.656	−.658	−.660	−.662	−.664	−.666	−.668	−.669	−.671
.49	−.673	−.675	−.677	−.679	−.681	−.683	−.685	−.687	−.689	−.691

Table B2a continued: ln(1−q)

q	.000	.001	.002	.003	.004	.005	.006	.007	.008	.009
.50	− .693	− .695	− .697	− .699	− .701	− .703	− .705	− .707	− .709	− .711
.51	− .713	− .715	− .717	− .720	− .722	− .724	− .726	− .728	− .730	− .732
.52	− .734	− .736	− .738	− .740	− .742	− .744	− .747	− .749	− .751	− .753
.53	− .755	− .757	− .759	− .761	− .764	− .766	− .768	− .770	− .772	− .774
.54	− .777	− .779	− .781	− .783	− .785	− .788	− .790	− .792	− .794	− .796
.55	− .799	− .801	− .803	− .805	− .807	− .810	− .812	− .814	− .817	− .819
.56	− .821	− .823	− .826	− .828	− .830	− .832	− .835	− .837	− .839	− .842
.57	− .844	− .846	− .849	− .851	− .853	− .856	− .858	− .860	− .863	− .865
.58	− .868	− .870	− .872	− .875	− .877	− .880	− .882	− .884	− .887	− .889
.59	− .892	− .894	− .897	− .899	− .901	− .904	− .906	− .909	− .911	− .914
.60	− .916	− .919	− .921	− .924	− .926	− .929	− .931	− .934	− .937	− .939
.61	− .942	− .944	− .947	− .949	− .952	− .955	− .957	− .960	− .962	− .965
.62	− .968	− .970	− .973	− .976	− .978	− .981	− .984	− .986	− .989	− .992
.63	− .994	− .997	−1.000	−1.002	−1.005	−1.008	−1.011	−1.013	−1.016	−1.019
.64	−1.022	−1.025	−1.027	−1.030	−1.033	−1.036	−1.039	−1.041	−1.044	−1.047
.65	−1.050	−1.053	−1.056	−1.059	−1.061	−1.064	−1.067	−1.070	−1.073	−1.076
.66	−1.079	−1.082	−1.085	−1.088	−1.091	−1.094	−1.097	−1.100	−1.103	−1.106
.67	−1.109	−1.112	−1.115	−1.118	−1.121	−1.124	−1.127	−1.130	−1.133	−1.136
.68	−1.140	−1.143	−1.146	−1.149	−1.152	−1.155	−1.158	−1.162	−1.165	−1.168
.69	−1.171	−1.175	−1.178	−1.181	−1.184	−1.188	−1.191	−1.194	−1.197	−1.201
.70	−1.204	−1.207	−1.211	−1.214	−1.218	−1.221	−1.224	−1.228	−1.231	−1.235
.71	−1.238	−1.241	−1.245	−1.248	−1.252	−1.255	−1.259	−1.262	−1.266	−1.270
.72	−1.273	−1.277	−1.280	−1.284	−1.287	−1.291	−1.295	−1.298	−1.302	−1.306
.73	−1.309	−1.313	−1.317	−1.321	−1.324	−1.328	−1.332	−1.336	−1.340	−1.343
.74	−1.347	−1.351	−1.355	−1.359	−1.363	−1.367	−1.371	−1.375	−1.378	−1.382

Table B2a continued: ln(1–q)

q	.000	.001	.002	.003	.004	.005	.006	.007	.008	.009
.75	–1.386	–1.390	–1.394	–1.399	–1.403	–1.407	–1.411	–1.415	–1.419	–1.423
.76	–1.427	–1.431	–1.436	–1.440	–1.444	–1.448	–1.453	–1.457	–1.461	–1.465
.77	–1.470	–1.474	–1.479	–1.483	–1.487	–1.492	–1.496	–1.501	–1.505	–1.510
.78	–1.514	–1.519	–1.523	–1.528	–1.533	–1.537	–1.542	–1.547	–1.551	–1.556
.79	–1.561	–1.566	–1.570	–1.575	–1.580	–1.585	–1.590	–1.595	–1.600	–1.605
.80	–1.610	–1.615	–1.620	–1.625	–1.630	–1.635	–1.640	–1.645	–1.650	–1.656
.81	–1.661	–1.666	–1.672	–1.677	–1.682	–1.688	–1.693	–1.698	–1.704	–1.709
.82	–1.716	–1.721	–1.726	–1.732	–1.737	–1.743	–1.749	–1.755	–1.760	–1.766
.83	–1.772	–1.778	–1.784	–1.790	–1.796	–1.802	–1.808	–1.814	–1.820	–1.827
.84	–1.833	–1.839	–1.845	–1.852	–1.858	–1.865	–1.871	–1.878	–1.884	–1.891
.85	–1.897	–1.904	–1.911	–1.918	–1.924	–1.931	–1.938	–1.945	–1.952	–1.959
.86	–1.966	–1.974	–1.981	–1.988	–1.995	–2.003	–2.010	–2.018	–2.025	–2.033
.87	–2.041	–2.048	–2.056	–2.064	–2.072	–2.080	–2.088	–2.096	–2.104	–2.112
.88	–2.121	–2.129	–2.137	–2.146	–2.155	–2.163	–2.172	–2.181	–2.190	–2.199
.89	–2.208	–2.217	–2.226	–2.235	–2.245	–2.254	–2.264	–2.273	–2.283	–2.293
.90	–2.303	–2.313	–2.323	–2.333	–2.344	–2.354	–2.365	–2.376	–2.386	–2.397
.91	–2.408	–2.420	–2.431	–2.442	–2.454	–2.466	–2.477	–2.489	–2.502	–2.514
.92	–2.526	–2.539	–2.552	–2.565	–2.578	–2.591	–2.604	–2.618	–2.632	–2.646
.93	–2.660	–2.674	–2.689	–2.704	–2.719	–2.734	–2.750	–2.765	–2.781	–2.798
.94	–2.814	–2.831	–2.848	–2.865	–2.883	–2.901	–2.920	–2.938	–2.957	–2.977
.95	–2.997	–3.017	–3.037	–3.059	–3.080	–3.102	–3.125	–3.148	–3.171	–3.195
.96	–3.220	–3.245	–3.271	–3.298	–3.325	–3.354	–3.383	–3.413	–3.443	–3.475
.97	–3.508	–3.540	–3.577	–3.614	–3.651	–3.691	–3.732	–3.774	–3.819	–3.865
.98	–3.912	–3.963	–4.017	–4.075	–4.135	–4.200	–4.269	–4.343	–4.423	–4.510
.99	–4.605	–4.711	–4.828	–4.962	–5.116	–5.298	–5.521	–5.809	–6.215	–6.908

Table B2b. Values of ln(1+q) for Values of q from .000 to .999

q	.000	.001	.002	.003	.004	.005	.006	.007	.008	.009
.00	.000	.001	.002	.003	.004	.005	.006	.007	.008	.009
.01	.010	.011	.012	.013	.014	.015	.016	.017	.018	.019
.02	.020	.021	.022	.023	.024	.025	.026	.027	.028	.029
.03	.030	.031	.031	.032	.033	.034	.035	.036	.037	.038
.04	.039	.040	.041	.042	.043	.044	.045	.046	.047	.048
.05	.049	.050	.051	.052	.053	.054	.054	.055	.056	.057
.06	.058	.059	.060	.061	.062	.063	.064	.065	.066	.067
.07	.068	.069	.069	.070	.071	.072	.073	.074	.075	.076
.08	.077	.078	.079	.080	.081	.082	.082	.083	.084	.085
.09	.086	.087	.088	.089	.090	.091	.092	.093	.093	.094
.10	.095	.096	.097	.098	.099	.100	.101	.102	.103	.103
.11	.104	.105	.106	.107	.108	.109	.110	.111	.111	.112
.12	.113	.114	.115	.116	.117	.118	.119	.119	.120	.121
.13	.122	.123	.124	.125	.126	.127	.127	.128	.129	.130
.14	.131	.132	.133	.134	.134	.135	.136	.137	.138	.139
.15	.140	.141	.141	.142	.143	.144	.145	.146	.147	.147
.16	.148	.149	.150	.151	.152	.153	.154	.154	.155	.156
.17	.157	.158	.159	.159	.160	.161	.162	.163	.164	.165
.18	.165	.166	.167	.168	.169	.170	.171	.171	.172	.173
.19	.174	.175	.176	.176	.177	.178	.179	.180	.181	.181
.20	.182	.183	.184	.185	.186	.186	.187	.188	.189	.190
.21	.191	.191	.192	.193	.194	.195	.195	.196	.197	.198
.22	.199	.200	.200	.201	.202	.203	.204	.204	.205	.206
.23	.207	.208	.209	.209	.210	.211	.212	.213	.213	.214
.24	.215	.216	.217	.217	.218	.219	.220	.221	.221	.222

Table B2b continued: ln(1+q)

q	.000	.001	.002	.003	.004	.005	.006	.007	.008	.009
.25	.223	.224	.225	.225	.226	.227	.228	.229	.229	.230
.26	.231	.232	.233	.233	.234	.235	.236	.237	.237	.238
.27	.239	.240	.240	.241	.242	.243	.244	.244	.245	.246
.28	.247	.248	.248	.249	.250	.251	.251	.252	.253	.254
.29	.255	.255	.256	.257	.258	.258	.259	.260	.261	.261
.30	.262	.263	.264	.265	.265	.266	.267	.268	.268	.269
.31	.270	.271	.271	.272	.273	.274	.274	.275	.276	.277
.32	.277	.278	.279	.280	.281	.281	.282	.283	.284	.284
.33	.285	.286	.287	.287	.288	.289	.290	.290	.291	.292
.34	.293	.293	.294	.295	.296	.296	.297	.298	.298	.299
.35	.300	.301	.301	.302	.303	.304	.304	.305	.306	.307
.36	.307	.308	.309	.310	.310	.311	.312	.312	.313	.314
.37	.315	.315	.316	.317	.318	.318	.319	.320	.320	.321
.38	.322	.323	.323	.324	.325	.326	.326	.327	.328	.328
.39	.329	.330	.331	.331	.332	.333	.333	.334	.335	.336
.40	.336	.337	.338	.338	.339	.340	.341	.341	.342	.343
.41	.343	.344	.345	.346	.346	.347	.348	.348	.349	.350
.42	.350	.351	.352	.353	.353	.354	.355	.355	.356	.357
.43	.358	.358	.359	.360	.360	.361	.362	.362	.363	.364
.44	.364	.365	.366	.367	.367	.368	.369	.369	.370	.371
.45	.371	.372	.373	.373	.374	.375	.376	.376	.377	.378
.46	.378	.379	.380	.380	.381	.382	.382	.383	.384	.384
.47	.385	.386	.386	.387	.388	.388	.389	.390	.391	.391
.48	.392	.393	.393	.394	.395	.395	.396	.397	.397	.398
.49	.399	.399	.400	.401	.401	.402	.403	.403	.404	.405

Table B2b continued: ln(1+q)

q	.000	.001	.002	.003	.004	.005	.006	.007	.008	.009
.50	.405	.406	.407	.407	.408	.409	.409	.410	.411	.411
.51	.412	.413	.413	.414	.415	.415	.416	.417	.417	.418
.52	.419	.419	.420	.420	.421	.422	.422	.423	.424	.424
.53	.425	.426	.426	.427	.428	.428	.429	.430	.430	.431
.54	.432	.432	.433	.434	.434	.435	.435	.436	.437	.437
.55	.438	.439	.439	.440	.441	.441	.442	.443	.443	.444
.56	.444	.445	.446	.446	.447	.448	.448	.449	.450	.450
.57	.451	.452	.452	.453	.453	.454	.455	.455	.456	.457
.58	.457	.458	.458	.459	.460	.460	.461	.462	.462	.463
.59	.464	.464	.465	.465	.466	.467	.467	.468	.469	.469
.60	.470	.470	.471	.472	.472	.473	.474	.474	.475	.475
.61	.476	.477	.477	.478	.479	.479	.480	.480	.481	.482
.62	.482	.483	.483	.484	.485	.485	.486	.487	.487	.488
.63	.488	.489	.490	.490	.491	.491	.492	.493	.493	.494
.64	.494	.495	.496	.496	.497	.498	.498	.499	.499	.500
.65	.501	.501	.502	.502	.503	.504	.504	.505	.505	.506
.66	.507	.507	.508	.508	.509	.510	.510	.511	.511	.512
.67	.513	.513	.514	.514	.515	.516	.516	.517	.517	.518
.68	.519	.519	.520	.520	.521	.522	.522	.523	.523	.524
.69	.525	.525	.526	.526	.527	.527	.528	.529	.529	.530
.70	.530	.531	.532	.532	.533	.533	.534	.535	.535	.536
.71	.536	.537	.537	.538	.539	.539	.540	.540	.541	.542
.72	.542	.543	.543	.544	.544	.545	.546	.546	.547	.547
.73	.548	.548	.549	.550	.550	.551	.551	.552	.553	.553
.74	.554	.554	.555	.555	.556	.557	.557	.558	.558	.559

Table B2b continued: ln(1+q)

q	.000	.001	.002	.003	.004	.005	.006	.007	.008	.009
.75	.559	.560	.561	.561	.562	.562	.563	.563	.564	.565
.76	.565	.566	.566	.567	.567	.568	.568	.569	.570	.570
.77	.571	.571	.572	.572	.573	.574	.574	.575	.575	.576
.78	.576	.577	.577	.578	.579	.579	.580	.580	.581	.581
.79	.582	.583	.583	.584	.584	.585	.585	.586	.586	.587
.80	.588	.588	.589	.589	.590	.590	.591	.591	.592	.593
.81	.593	.594	.594	.595	.595	.596	.596	.597	.597	.598
.82	.599	.599	.600	.600	.601	.601	.602	.602	.603	.604
.83	.604	.605	.605	.606	.606	.607	.607	.608	.608	.609
.84	.610	.610	.611	.611	.612	.612	.613	.613	.614	.614
.85	.615	.615	.616	.617	.617	.618	.618	.619	.619	.620
.86	.620	.621	.621	.622	.622	.623	.624	.624	.625	.625
.87	.626	.626	.627	.627	.628	.628	.629	.629	.630	.630
.88	.631	.632	.632	.633	.633	.634	.634	.635	.635	.636
.89	.636	.637	.637	.638	.638	.639	.639	.640	.641	.641
.90	.642	.642	.643	.643	.644	.644	.645	.645	.646	.646
.91	.647	.647	.648	.648	.649	.649	.650	.650	.651	.652
.92	.652	.653	.653	.654	.654	.655	.655	.656	.656	.657
.93	.657	.658	.658	.659	.659	.660	.660	.661	.661	.662
.94	.662	.663	.663	.664	.664	.665	.666	.666	.667	.667
.95	.668	.668	.669	.669	.670	.670	.671	.671	.672	.672
.96	.673	.673	.674	.674	.675	.675	.676	.676	.677	.677
.97	.678	.678	.679	.679	.680	.680	.681	.681	.682	.682
.98	.683	.683	.684	.684	.685	.685	.686	.686	.687	.687
.99	.688	.688	.689	.689	.690	.690	.691	.691	.692	.692

Table B3. Values of the NBD and LSD Parameters a (Adapted from Chatfield [1969])

NBD: Values of $a = m/k$ for various values of $c = -m/\ln p_0 = -wb/\ln(1-b)$.
LSD: Values of $a = q/(1+q)$ for various values of w.

Values of c for the NBD or w for the LSD

	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1	0.00	0.21	0.43	0.66	0.89	1.14	1.40	1.67	1.94	2.22
2	2.51	2.81	3.11	3.42	3.73	4.05	4.37	4.70	5.03	5.37
3	5.71	6.06	6.41	6.76	7.12	7.48	7.85	8.22	8.59	8.97
4	9.35	9.73	10.11	10.50	10.89	11.29	11.69	12.09	12.49	12.89
5	13.30	13.71	14.13	14.54	14.96	15.38	15.80	16.22	16.65	17.08
6	17.51	17.94	18.38	18.81	19.25	19.69	20.14	20.58	21.03	21.48
7	21.93	22.38	22.83	23.29	23.74	24.20	24.66	25.12	25.59	26.05
8	26.52	26.99	27.46	27.93	28.40	28.87	29.35	29.83	30.30	30.79
9	31.27	31.75	32.23	32.72	33.20	33.69	34.18	34.67	35.16	35.65
10	36.15	36.64	37.14	37.64	38.14	38.64	39.14	39.64	40.14	40.65
11	41.15	41.66	42.17	42.68	43.19	43.70	44.21	44.73	45.24	45.75
12	46.27	46.79	47.31	47.82	48.35	48.87	49.39	49.91	50.44	50.96
13	51.49	52.01	52.54	53.07	53.60	54.13	54.66	55.19	55.73	56.26
14	56.80	57.33	57.87	58.41	58.95	59.48	60.02	60.57	61.11	61.65
15	62.19	62.74	63.28	63.83	64.37	64.92	65.47	66.02	66.57	67.12
16	67.67	68.22	68.77	69.33	69.88	70.43	70.99	71.54	72.10	72.66
17	73.22	73.78	74.34	74.90	75.46	76.02	76.58	77.15	77.71	78.27
18	78.86	79.40	79.97	80.54	81.11	81.67	82.24	82.81	83.38	83.95
19	84.53	86.10	85.67	86.24	86.82	87.39	87.97	88.54	89.12	89.70

	0.0	1.0	2.0	3.0	4.0	5.0	6.0	7.0	8.0	9.0
20	90.3	96.1	101.9	107.9	113.8	119.9	125.9	132.1	138.2	144.4
30	150.6	156.9	163.2	169.6	176.0	182.4	188.9	195.4	201.9	208.4
40	215.0	221.6	228.3	234.9	241.6	248.4	255.1	261.9	268.7	275.5
50	282.3	289.2	296.1	303.0	309.9	316.9	323.9	330.9	337.9	344.9
60	352.0	359.1	366.2	373.3	380.4	387.6	394.7	401.9	409.1	416.3

Table B4. NBD and LSD Values of the Percentage of Buyers of an Item in one Period who Buy it again in Another Period of Equal Length

(For various values of b and w*, as in Table 4.15)

$100 b_R/b$	Average Purchase Frequency per Buyer, w							
	1.1	1.3	1.5	2	3	5	10	20
NBD, for Proportions of Buyers, b =								
.8					85	89	92	93
.6				70	78	84	88	90
.4		42	52	64	74	80	85	88
.2		37	47	60	70	77	83	86
.1	17	35	45	58	69	76	82	85
.05	16	34	45	58	68	76	82	85
LSD: Exact	16	34	44	57	68	75	81	85
Approximate	13	30	41	56	68	76	82	84

* Data for which $w < -\{\ln(1-b)\}/b$ cannot be fitted by an NBD.

Table B5. NBD and LSD Values of the Average Frequency of Purchase in a Period per Repeat-Buyer in an Equal Period

w_R	Average Purchase Frequency per Buyer, w							
	1.1	1.3	1.5	2	3	5	10	20
NBD, for Proportions of Buyers, b =								
.8					3.2	5.4	11	21
.6				2.2	3.4	5.7	11	22
.4		1.3	1.6	2.3	3.5	5.9	11	23
.2		1.4	1.7	2.4	3.7	6.0	12	23
.1	1.1	1.5	1.8	2.5	3.7	6.1	12	23
.05	1.2	1.5	1.8	2.5	3.8	6.1	12	23
LSD: Exact	1.2	1.5	1.8	2.5	3.8	6.2	12	23
Approximate	1.4	1.6	1.9	2.5	3.7	6.2	12	25

Table B6. NBD and LSD Values of the Percentage of Buyers of an Item in one Period who do NOT buy it in Another Equal Period

($100b_N/b = 100b_L/b = 100 - b_R/b$ from Table B4)

$100b_N/b$	Average Purchase Frequency per Buyer, w							
	1.1	1.3	1.5	2	3	5	10	20
NBD, for Proportions of Buyers, b =								
.8					15	11	8	7
.6				30	22	16	12	10
.4		58	48	36	26	20	15	12
.2		63	53	40	30	23	17	14
.1	83	65	55	42	31	24	18	15
.05	84	66	55	42	32	24	18	15
LSD: Exact	84	66	56	43	32	25	19	15
Approximate	87	70	59	44	32	24	18	16

Table B7. NBD and LSD Values of the Average Frequency of Purchase per "New" or "Lapsed" Buyer

($w_N = w_L$)

w_N	Average Purchase Frequency per Buyer, w							
	1.1	1.3	1.5	2	3	5	10	20
NBD, for Proportions of Buyers, b =								
.8					1.9	1.8	1.7	1.7
.6				1.6	1.6	1.6	1.6	1.6
.4		1.3	1.4	1.4	1.5	1.5	1.5	1.5
.2		1.2	1.3	1.4	1.4	1.5	1.5	1.5
.1	1.1	1.2	1.3	1.4	1.4	1.4	1.5	1.5
.05	1.1	1.2	1.3	1.4	1.4	1.4	1.4	1.4
LSD: Exact	1.1	1.2	1.3	1.3	1.4	1.4	1.4	1.4
Approximate	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4

Table B8. Penetration Growth

(NBD and LSD values of b_T/b_1 in time periods of relative lengths T, for different values of $100b_1$ and w_1 in a time-period of unit length, i.e. $T = 1$)

	T: Length of Time-Period as a Fraction of Base Period ($T = 1$)								
	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$	1	2	3	4	12
b_T/b_1	$b_{\frac{1}{12}}/b_1$	$b_{\frac{1}{4}}/b_1$	$b_{\frac{1}{3}}/b_1$	$b_{\frac{1}{2}}/b_1$	$100b_1$	b_2/b_1	b_3/b_1	b_4/b_1	b_{12}/b_1
$100b_1 = 80$									
$w_1 = 20$	.58	.80	.85	.91	80	1.07	1.10	1.11	1.17
$= 10$	.45	.73	.80	.88	80	1.08	1.12	1.14	1.19
$= 5$	.31	.63	.71	.83	80	1.11	1.15	1.18	1.22
$= 3$	.22	.51	.62	.77	80	1.15	1.20	1.22	1.25
$100b_1 = 60$									
$w_1 = 20$	.53	.76	.81	.89	60	1.10	1.15	1.18	1.29
$= 10$	.43	.69	.76	.86	60	1.12	1.18	1.22	1.35
$= 5$	.30	.59	.68	.80	60	1.16	1.24	1.29	1.44
$= 3$	.21	.49	.59	.74	60	1.22	1.32	1.38	1.54
$= 2$	.15	.40	.50	.67	60	1.30	1.43	1.50	1.63
$100b_1 = 40$									
$w_1 = 20$	.51	.73	.79	.87	40	1.12	1.19	1.23	1.39
$= 10$	.41	.67	.74	.84	40	1.15	1.23	1.29	1.48
$= 8$	.29	.58	.66	.79	40	1.20	1.31	1.38	1.62
$= 3$	.21	.48	.58	.73	40	1.26	1.41	1.50	1.80
$= 2$	.15	.40	.50	.66	40	1.36	1.56	1.68	2.05
$= 1.5$	.12	.33	.43	.60	40	1.48	1.76	1.93	2.34
$= 1.3$	.11	.30	.40	.57	40	1.58	1.93	2.14	2.49

Table B8 continued

	T: Length of Time-Period as a Fraction of Base Period ($T = 1$)								
	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$	1	2	3	4	12
b_T/b_1	$b_{\frac{1}{12}}/b_1$	$b_{\frac{1}{4}}/b_1$	$b_{\frac{1}{3}}/b_1$	$b_{\frac{1}{2}}/b_1$	$100b_1$	b_2/b_1	b_3/b_1	b_4/b_1	b_{12}/b_1
$100b_1 = 20$									
$w_1 = 20$	.49	.72	.77	.86	20	1.14	1.22	1.27	1.47
$= 10$	.39	.65	.72	.83	20	1.17	1.27	1.31	1.58
$= 5$	.29	.56	.65	.77	20	1.23	1.36	1.45	1.77
$= 3$	.21	.47	.57	.72	20	1.30	1.47	1.59	2.03
$= 2$	.15	.39	.49	.65	20	1.40	1.64	1.81	2.42
$= 1.5$	.12	.33	.43	.60	20	1.53	1.87	2.12	2.99
$= 1.3$	.11	.30	.39	.56	20	1.63	2.07	2.39	3.54
$100b_1 = 10$									
$w_1 = 20$	.48	.71	.77	.85	10	1.15	1.23	1.29	1.51
$= 10$	.39	.64	.72	.82	10	1.18	1.28	1.36	1.63
$= 5$	.28	.56	.64	.77	10	1.24	1.38	1.47	1.84
$= 3$	.21	.47	.56	.71	10	1.31	1.50	1.63	2.13
$= 2$	.15	.39	.49	.65	10	1.42	1.68	1.86	2.59
$= 1.5$	.12	.33	.42	.59	10	1.55	1.91	2.19	3.28
$= 1.3$	.11	.30	.39	.56	10	1.65	2.12	2.48	3.96
$= 1.1$	.09	.27	.35	.52	10	1.83	2.53	3.13	5.92

Table B8 continued

	T: Length of Time-Period as a Fraction of Base Period ($T = 1$)								
	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$	1	2	3	4	12
b_T/b_1	$b_{\frac{1}{12}}/b_1$	$b_{\frac{1}{4}}/b_1$	$b_{\frac{1}{3}}/b_1$	$b_{\frac{1}{2}}/b_1$	$100b_1$	b_2/b_1	b_3/b_1	b_4/b_1	b_{12}/b_1
$100b_1 = 5$									
$w_1 = 20$	.48	.70	.76	.85	5	1.15	1.24	1.30	1.53
$= 10$	.39	.64	.71	.82	5	1.18	1.29	1.37	1.66
$= 5$	.28	.55	.64	.77	5	1.24	1.39	1.49	1.88
$= 3$	.21	.47	.56	.71	5	1.32	1.51	1.65	2.18
$= 2$	.15	.39	.49	.65	5	1.42	1.69	1.89	2.66
$= 1.5$	.12	.33	.42	.59	5	1.55	1.93	2.22	3.40
$= 1.3$	.11	.30	.39	.56	5	1.66	2.14	2.52	4.15
$= 1.1$	.09	.27	.35	.52	5	1.84	2.55	3.17	6.31
$100b_1 = 0$ (LSD)									
$w_1 = 20$	.47	.70	.76	.85	0	1.15	1.24	1.31	1.55
$= 10$	.38	.64	.71	.82	0	1.19	1.30	1.38	1.68
$= 5$	.28	.55	.64	.76	0	1.25	1.40	1.50	1.91
$= 3$	.20	.47	.56	.71	0	1.32	1.52	1.67	2.23
$= 2$	.15	.39	.48	.65	0	1.43	1.71	1.91	2.74
$= 1.5$	.12	.33	.42	.59	0	1.56	1.95	2.25	3.53
$= 1.3$	.11	.30	.39	.56	0	1.66	2.16	2.55	4.33
$= 1.1$	.09	.27	.35	.52	0	1.84	2.57	3.21	6.63

Table B9. The Average Purchase Frequency in Periods of Different Lengths

(NBD and LSD values of w_T/w_1 in time-periods of relative lengths T, for different values of $100b_1$ and w_1 in a time-period of "unit" length, i.e. $T = 1$)

	T: Length of Time Period as a Fraction of Base Period ($T = 1$)								
	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$	1	2	3	4	12
w_T/w_1	$w_{\frac{1}{12}}/w_1$	$w_{\frac{1}{4}}/w_1$	$w_{\frac{1}{3}}/w_1$	$w_{\frac{1}{2}}/w_1$	w_1	w_2/w_1	w_3/w_1	w_4/w_1	w_{12}/w_1
$100b_1 = 80$									
	.14	.31	.39	.55	20	1.88	2.74	3.59	10.29
	.18	.34	.42	.57	10	1.85	2.68	3.51	10.06
	.27	.40	.47	.60	5	1.80	2.60	3.40	9.80
	.38	.49	.54	.65	3	1.74	2.51	3.29	9.63
$100b_1 = 60$									
	.16	.33	.41	.56	20	1.82	2.61	3.39	9.31
	.20	.36	.44	.58	10	1.78	2.54	3.27	8.91
	.28	.42	.49	.62	5	1.72	2.41	3.10	8.36
	.39	.51	.56	.67	3	1.64	2.27	2.89	7.82
	.54	.62	.66	.74	2	1.54	2.09	2.66	7.35
$100b_1 = 40$									
	.16	.34	.42	.57	20	1.79	2.53	3.25	8.64
	.20	.37	.45	.60	10	1.74	2.44	3.11	8.13
	.29	.43	.50	.63	5	1.67	2.30	2.90	7.42
	.40	.52	.58	.69	3	1.58	2.13	2.67	6.67
	.54	.63	.67	.76	2	1.47	1.93	2.38	5.85
	.69	.75	.78	.83	1.5	1.35	1.71	2.07	5.13
	.79	.82	.84	.88	1.3	1.26	1.55	1.87	4.83

Table B9 continued

	T: Length of Time Period as a Fraction of Base Period ($T = 1$)								
	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$	1	2	3	4	12
w_T/w_1	$w_{\frac{1}{12}}/w_1$	$w_{\frac{1}{4}}/w_1$	$w_{\frac{1}{3}}/w_1$	$w_{\frac{1}{2}}/w_1$	w_1	w_2/w_1	w_3/w_1	w_4/w_1	w_{12}/w_1
$100b_1 = 20$									
	.17	.35	.43	.58	20	1.76	2.47	3.15	8.14
	.21	.38	.46	.61	10	1.71	2.37	2.99	7.57
	.29	.45	.51	.65	5	1.63	2.21	2.77	6.77
	.40	.53	.59	.70	3	1.54	2.04	2.51	5.92
	.55	.64	.68	.76	2	1.43	1.83	2.21	4.96
	.70	.75	.78	.84	1.5	1.31	1.60	1.89	4.01
	.79	.83	.85	.89	1.3	1.23	1.45	1.67	3.39
$100b_1 = 10$									
	.17	.35	.43	.59	20	1.75	2.44	3.10	7.93
	.21	.39	.46	.61	10	1.70	2.34	2.95	7.34
	.29	.45	.52	.65	5	1.62	2.18	2.71	6.51
	.41	.53	.59	.70	3	1.53	2.00	2.45	5.63
	.55	.64	.68	.77	2	1.41	1.79	2.15	4.64
	.70	.76	.78	.84	1.5	1.29	1.57	1.83	3.66
	.79	.83	.85	.89	1.3	1.21	1.42	1.61	3.03
	.92	.93	.94	.95	1.1	1.09	1.18	1.28	2.03

Table B9 continued

	T: Length of Time Period as a Fraction of Base Period ($T = 1$)								
	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$	1	2	3	4	12
w_T/w_1	$w_{\frac{1}{12}}/w_1$	$w_{\frac{1}{4}}/w_1$	$w_{\frac{1}{3}}/w_1$	$w_{\frac{1}{2}}/w_1$	w_1	w_2/w_1	w_3/w_1	w_4/w_1	w_{12}/w_1
$100b_1 = 5$									
	.17	.36	.44	.59	20	1.74	2.43	3.08	7.84
	.22	.39	.47	.61	10	1.69	2.32	2.92	7.24
	.30	.45	.52	.65	5	1.61	2.16	2.69	6.39
	.41	.53	.59	.70	3	1.52	1.99	2.43	5.50
	.55	.64	.69	.77	2	1.41	1.77	2.12	4.51
	.70	.76	.79	.84	1.5	1.29	1.55	1.80	3.52
	.79	.83	.85	.89	1.3	1.21	1.40	1.59	2.89
	.92	.93	.94	.95	1.1	1.09	1.18	1.26	1.90
$100b_1 = 0$ (LSD)									
	.18	.36	.44	.59	20	1.74	2.42	3.06	7.75
	.22	.39	.47	.61	10	1.68	2.31	2.90	7.14
	.30	.45	.52	.65	5	1.60	2.15	2.67	6.29
	.41	.54	.60	.71	3	1.51	1.97	2.40	5.39
	.55	.64	.69	.77	2	1.40	1.76	2.09	4.38
	.70	.76	.79	.84	1.5	1.28	1.54	1.78	3.40
	.79	.83	.85	.89	1.3	1.20	1.39	1.57	2.77
	.92	.93	.94	.96	1.1	1.09	1.17	1.25	1.81

APPENDIX C

CALCULATIONS FOR MULTI-BRAND BUYING

C.1 Calculations and Tabulations

In this Appendix we show how to calculate and model the measures of multi-brand buying discussed in Chapters 9 through 13. The same structure as Appendix A is followed, starting in § C.2 to C.4 with the estimation of the basic parameters of the theoretical model, here the Dirichlet; then showing how to use these in calculating the theoretical multi-brand statistics like w_p, b_s and duplication percentages $b_{X.Y.}$ in § C.5 to C.8; and finally outlining the appropriate tabulations of observed data in § C.9 to C.12. The theoretical treatment of time-periods is somewhat different from Appendix A and a discussion of this is reserved until § C.8.

The theoretical calculations are best done with a computer, using for example the software on the floppy disc mentioned in the Preface. However, by working through a simple example on paper the principles will become clearer. We will use the same hypothetical illustration as in Table A20 of Appendix A which is reproduced as Table C.7 in § C.9.

There are three parts to dealing with the theoretical model before we can calculate values of interest such as w_p. First we have to input values for the distribution of purchases of the product-field as a whole, then estimate the structural parameter S, and thirdly estimate the Dirichlet probabilities. All these steps involve relatively cumbersome numerical iterative calculations and approximations, since the measures we want do not have explicit algebraic formulae in the Dirichlet model. (The difficulties are not due to statistical sampling considerations – it is not for example a question of how to derive maximum likelihood estimates.)

C.2 The Product-Field Distribution

Two methods of dealing with the distribution of purchases of the product-field are available for the Dirichlet model, as noted in Chapter 13. One is the NBD-Dirichlet. This assumes that the distribution of product-field purchases is close to an NBD and estimates this from two summary statistics like B and W in some chosen base-period. Values for any other time-period of length T can then be estimated merely by substituting TA and TM for A and M in what follows (see Section C.8).

If the NBD assumption for the total product-field does not apply, and if the observed distribution is available, this can be used as input for the Empirical-Dirichlet model. This is simpler but less powerful because it does not provide predictions for other time-periods, and it has greater data requirements.

The NBD-Dirichlet. In the NBD-Dirichlet model, B, the percentage buying the product (ie. buying *any* brand) in the time-period, and W, the average purchase frequency of the product per buyer in that period, are used to generate a theoretical NBD. This is fed into subsequent calculations.

We assume that $B = .1$ or 10% and $W = 2.8$ (as in the numerical example of Table C7 later, where for a sample of 200, $B = 20/200 = .1$ and $W = 56/20 = 2.8$, and hence the mean number of purchases per household is $M = 56/200 = .28$ (or $BW = .28$)). Following the procedure in Section A.2 we derive the NBD parameter $A = 4.559$ and hence also the value of the related parameter $K = M/A = .0614$ (here we have obtained A directly rather than by interpolation). This now enables us to follow the guidelines in Section A.6 to calculate the NBD proportion P_n of households who make n purchases of the product-field, where n is any whole number.

$$\begin{aligned} P_n &= \left(\frac{A}{1+A}\right)\left(1-\frac{A-M}{An}\right)P_{n-1}, \quad \text{for } n = 1, 2, 3, \ldots \\ &= \left(\frac{4.559}{5.559}\right)\left(1-\frac{4.559-.28}{4.559n}\right)P_{n-1}, \\ &= .8201 \left(1-\frac{.9386}{n}\right)P_{n-1}. \end{aligned}$$

Given that $P_0 = 1 - B = .9$, we have

$$\begin{aligned} P_1 &= .8201\ (1-.9386)\ .90000 = .04532 \text{ or } 4.5\%, \\ P_2 &= .8201\ (1-.4693)\ .04532 = .01972 \text{ or } 2.0\%, \\ P_3 &= .8201\ (1-.3129)\ .01972 = .01111 \text{ or } 1.1\%, \end{aligned}$$

and so forth (the full distribution is shown later in Table C1). We will use these theoretical proportions buying 1, 2, 3 etc. times in the calculation of the full Dirichlet model.

Dealing with an Infinite Series. However, we need to note that since the NBD is an infinite distribution it is strictly impossible to calculate all these proportions numerically. We therefore have to curtail the calculations at some finite point, n^* say, and develop approximate estimates for the tail of the NBD distribution. For hand calculations, one can generate the individual P values for say .99 of all households, so that we have to deal with a small remainder term of .01. When using computer routines one can easily go further for still greater accuracy, eg. to leave a very small tail of only .0001%. (*These tails are numerically more important than they might seem because they consist of very heavy buyers*).

An approximation procedure, suggested by Professor G. J. Goodhardt, is to divide the sum of the tail probabilities over two values n' and $n' + 1$, determining purchase proportions $P_{n'}$ and $P_{n'+1}$ such that the weighted sum $n' P_{n'} + (n' + 1)P_{n'+1}$ equals the total purchases accounted for in the exact NBD by those buying more than n^* times.

Suppose we have chosen to cut off at a tail of 1%. Calculating numeric values of P_n as above shows that this tail is reached at about $n^* = 6$ and purchases up to this level amount to .18843; ie.

$$\sum_{n=0}^{6} P_n = .99097,$$

and

$$\sum_{n=1}^{6} nP_n = .18843.$$

We represent the tail of the distribution by two proportions at n' and $n' + 1$ by solving the two equations

$$\begin{aligned} P_{n'} + P_{n'+1} &= 1 - \sum_{n=0}^{n^*} P_n = P_R \\ &= 1 - .99097 \\ &= .00903, \end{aligned}$$

$$\begin{aligned} n'P_{n'} + (n' + 1)P_{n'+1} &= M - \sum_{n=1}^{n^*} nP_n = Q_R \\ &= .28 - .18843 \\ &= .09157, \end{aligned}$$

where M is again the mean number of purchases of the product-field per household. Substantively this means .00903 people make .09157 purchases in the tail.

Multiply P_R by n' and subtract from Q_R to give

$$P_{n'+1} = Q_R - n'P_R$$

Divide by P_R to give

$$\frac{P_{n'+1}}{P_R} = \frac{Q_R}{P_R} - n'$$

If Q_R/P_R is an integer (which is unlikely) put $n' = Q_R/P_R$ and then $P_{n'+1} = 0$, $P_{n'} = P_R$. Otherwise, n' must be the integral part of Q_R/P_R. Then

$$P_{n'+1} = P_R \times \text{(non-integral part of } Q_R/P_R\text{)}.$$

Thus, in our example where $P_R = .00903$ and $Q_R = .09157$:

$$\begin{aligned} Q_R/P_R &= 10.14064, \text{ so } n' = 10, \\ P_{n'+1} &= P_{11} = .00903 \times .14064 = .00127, \\ P_{n'} &= P_{10} = .00903 - .00127 = .00776. \end{aligned}$$

The cumulative number of purchases at $P_{n'+1}$ should equal M and it is worthwhile doing this calculation as an arithmetic check

$$\sum_{n=1}^{n'+1} nP_n = .28000 = M.$$

The complete results from this example are shown in Table C1.

C.3 The Dirichlet Parameter $\hat{S}$

The Dirichlet parameter $\hat{S}$ can be estimated using the *average* brand. But in practice we usually form a separate estimate $\hat{S}_j$ for each brand j, and then obtain the overall Dirichlet estimate as a *weighted average* across

Table C1. The Estimated Proportions Buying the Product-Field n Times

n	Proportion Buying P_n	Cumulative Proportion ΣP_n	Purchases nP_n	Cumulative Purchases ΣnP_n
0	.90000	.90000	.00000	.00000
1	.04532	.94532	.04532	.04532
2	.01972	.96504	.03944	.08476
3	.01111	.97615	.03333	.11809
4	.00697	.98312	.02788	.14597
5	.00464	.98776	.02320	.16917
6	.00321	.99097	.01926	.18843
		....		
10	.00776	.99873	.07760	.26603
11	.00127	1.00000	.01397	.28000

brands $j = 1, \ldots, g$ (or for a sub-set of brands g^*). This provides a diagnostic check on cases where the model fails to fit a particular brand, and these might be left out of the weighted average.

Each brand is analysed in turn. Here we describe the calculations for brand X in the example of Table A20 (or Table C7 later), the basic statistics for which are shown in Table C2. For simplicity, the subscript X will be dropped in this section.

Table C2. Input Statistics

	W	B as %	B	P_0	M
Product Field:	2.800	10.0	.100	.900	.280
	w	b as %	b	p_0	m
Brand X:	1.923	6.5	.065	.935	.125

The estimation of $\hat{S}$ for brand X has to be by iteration, as once again there is no explicit algebraic formula. We begin with a more or less arbitrary starting value, S', usually guessed at from previous experience of relevant $\hat{S}$ values, say $S' = 2$, or we start with $S' = 1$. The aim is to find an estimate such that the predicted number of non-buyers of the brand, p_0', is equal (or close to) the observed number p_0.

We estimate non-buyers of the brand separately for each frequency n of buying the product-field, by multiplying the probability P_n in Table C1 by a provisional estimate of the probability $p'_{(0|n)}$ of *not buying* brand X, as given by the model, conditional on having made n product purchases. These are summated up to n^*, plus the terms for n' and $n' + 1$, ie.

$$p'_0 = \Sigma \{P_n p'_{(0|n)}\}$$

where p' stands for provisional estimates, using the assumed Dirichlet value S'.

The procedure is to get a good estimate of p'_0 using the product-field proportions which we have found already. First p'_0 is obtained using P_0 and P_1 only, then using P_2 also, and so on through the whole distribution. In order to estimate the values of $p'_{(0|n)}$ for each iteration it is convenient to work with two terms, c' and d' (neither of these is meaningful but they help us to do our calculations).

These terms need to be adjusted for each P_n, using the observed m and M from Table C2 and the assumed working value of $S' = 2$. First, for P_0 and P_1

$$\begin{aligned} c' &= S' - ((m \times S')/M) \\ &= 2 - ((.125 \times 2)/.28) \\ &= 1.10714, \end{aligned}$$

and

$$\begin{aligned} d' &= c'/S' \\ &= 1.10714/2 \\ &= .55357. \end{aligned}$$

We now use the P_0 and P_1 to derive a first estimate of p'_0 for $S' = 2$

$$\begin{aligned} P'_0 &= P_0 + (P_1 \times d') \\ &= .9 + (.04532 \times .55357) \\ &= .92509. \end{aligned}$$

Next we use P_2 as well, still keeping $S' = 2$. The value of d' has to be revised and the sum for p'_0 is adjusted. Thus

$$\begin{aligned}
\text{New } d' &= \text{Old } d' \times (c' + (n-1))/(S' + (n-1)) \\
&= .55357\,(1.10714 + (2-1))/(2 + (2-1)) \\
&= .38882, \\
\text{New } p'_0 &= \text{Old } p'_0 + (P_2 \times d') \\
&= .92509 + (.01972 \times .38882) \\
&= .93276.
\end{aligned}$$

And for $n = 3$ we have

$$\begin{aligned}
\text{New } d' &= \text{Old } d' \times (c' + (n-1))/(S' + (n-1)) \\
&= .38882 \times (1.10714 + (3-1))/(2 + (3-1)) \\
&= .30203,
\end{aligned}$$

$$\begin{aligned}
\text{New } p'_0 &= \text{Old } p'_0 + (P_3 \times d') \\
&= .93276 + (.01111 \times .30203) \\
&= .93612.
\end{aligned}$$

And so on for all P_n up to n' and $n' + 1$. The final estimated value of p'_0 obtained in this cycle of calculations is .94093. This differs by .00593 from the observed value of $p_0 = .935$ for brand X in Table C2. To get closer a new iteration is made.

If the estimated p'_0 is greater than the observed p_0 (as here), we use a larger starting value S'', say twice the old value of S', ie. $S'' = 2 \times 2 = 4$. If p'_0 is less than p_0, we use a smaller S'', say half the previous one.

The first estimate of $p'_0{}'$ for P_0 and P_1 but now with $S'' = 4$ will in fact be the same as that for $S' = 2$, since the new d'' here is the same as before. The calculations are

$$\begin{aligned}
c'' &= S'' - ((m \times S'')/M) \\
&= 4 - ((.125 \times 4)/.28) \\
&= 2.21429,
\end{aligned}$$

and

$$\begin{aligned}
d'' &= c''/S'' \\
&= 2.21429/4 \\
&= .55357.
\end{aligned}$$

Using P_0 and P_1 to get the first estimate in the second iteration

$$\begin{aligned} P_0'' &= P_0 + (P_1 \times d'') \\ &= .9 + (.04532 \times .55357) \\ &= .92509, \end{aligned}$$

the same as p_0' for P_0 and P_1.

But the subsequent estimates of p_0'' (using P_2, P_3 etc) will be closer to the observed p_0. Thus, as before, if we now also use P_2 to upgrade our first value of p_0'' we obtain

$$\begin{aligned} \text{New } d'' &= \text{Old } d'' \times (c'' + (n-1))/(S'' + (n-1)) \\ &= .55357 \times (2.21429 + (2-1))/(4 + (2-1)) \\ &= .35587, \end{aligned}$$

$$\begin{aligned} \text{New } p_0'' &= \text{Old } p_0'' + (P_2 \times d'') \\ &= .92509 + (.01972 \times .35587) \\ &= .93211. \end{aligned}$$

All subsequent values of p_0'', taking account of $P_3, P_4, \ldots, P_{n'}, P_{n'+1}$, are calculated in the same way, revising d'' and p_0'', each time by a new d'' and a new p_0''. At the end of this iteration $p_0'' = .93812$, which is slightly closer to .935.

For the third and subsequent iterations, the new S value can be obtained from the two preceding ones (here $S' = 2$ and $S'' = 4$) by interpolation or extrapolation, depending on whether the latest estimated p_0 value is too high or too low. Thus, as we converge on p_0, the estimate of S' is improved:

S'	p_0'	$p_0' - p_0$
2	.94093	.00593
4	.93812	.00312
8	.93628	.00128
.	.	.
.	.	.
.	.	.

Further iterations are best done on a computer, where some iteration criterion like .0001, or a limit on the number of cycles, can be used as a cut-off point. In the present example, convergence to within .0005 is reached after 5 cycles, with an estimated S' for brand X of 12.592.

Estimates for brands Y and Z are found in exactly the same way. The final table of results, using the brand suffix j again is shown in Table C3.

Table C3. Estimates of the Dirichlet Parameter for each Brand

Brand	$\hat{S}_j$	m_j	m_j/M
X	12.592	.125	.4464
Y	27.291	.090	.3214
Z	36.241	.065	.2321

The three values of $\hat{S}_j$ here differ markedly – from 12 to 36 – whereas with a well-fitting model they should all be similar. The variation is because our small hypothetical buying pattern is unrealistic and does not follow a Dirichlet too well, but this does not detract from its illustrative value.

The overall Dirichlet parameter $\hat{S}$ is a weighted average of these $\hat{S}_j$, obtained by taking the market shares of m_j into account:

$$\hat{S} = \sum_{g^*} (\hat{S}_j m_j/M)/\sum_{g^*} (m_j/M).$$

In forming such an estimate one can drop the estimates for one or more brands if they look irregular, like the low 12.592 for brand X here perhaps, and estimate $\hat{S}$ from the $g^* = 2$ brands Y and Z and then check the fit of the model for all g brands, ie. including X here. More generally, we can fit the Dirichlet to any selection of g^* brands.

If we use the full set of g brands that make up the product-field (ie. $g^* = g$), the denominator in the above equation is 1 and $\hat{S}$ is merely the sum of the g $\hat{S}_j$s weighted by the market share of each brand. Numerically, the weighted estimate of $\hat{S}$ for our three brands is

$$\hat{S} = \frac{(12.592 \times .4464 + 27.291 \times .3214 + 36.241 \times .2321)}{(.4464 + .3214 + .2321)}$$

$$= 22.8062/1 = 22.8.$$

The values of the Dirichlet parameters $\hat{\alpha}_j$ (see Chapter 13, §13.3) are given by $\hat{\alpha}_j = \hat{S}(m_j/M)$ and here are

$$\begin{aligned}\hat{\alpha}_X &= 22.8062 \times .4464 = 10.1807\\ \hat{\alpha}_Y &= 22.8062 \times .3214 = 7.3299\\ \hat{\alpha}_Z &= 22.8062 \times .2321 = 5.2933\end{aligned}$$

which add up to $\hat{S} = 22.8$, as above.
We can also define the quantity $\hat{\beta}_j = \hat{S} - \hat{\alpha}_j$, which will be required in the next section.

Possible Short-Cuts. With suitable software the above calculations are less burdensome than they might seem. We have therefore not explored possible short-cuts very much so far. One set of possibilities is to curtail the calculations of the NBD in Section C.2 earlier and to iterate less often in estimating $\hat{S}$. Yardsticks for evaluating such short-cuts are how much the answers differ from the more accurate estimates, whether in terms of the $\hat{S}$ values or the $\hat{\alpha}_j$, or in terms of the estimates of behavioural statistics like b, w, w_p, b_s and the duplication patterns that are discussed in Sections C.5 to C.8.

Another possible short-cut where accuracy and adequacy needs to be explored is to estimate a single $\hat{S}$ for the average brand, using the above procedure. It should be particularly useful in cases where earlier analyses of the product-field have already shown the model to fit well; in these cases it is largely unnecessary to examine every brand.

C.4 The Dirichlet Proportions for Brand X

Having somewhat laboriously estimated the main Dirichlet parameter $\hat{S}$ we still cannot directly calculate any of the derived statistics like b and w for a given brand since there are no explicit algebraic formulae relating them to $\hat{S}$ (or to the other inputs to the model, ie. B, W and market shares). Instead, we now have to calculate a matrix of proportions for all those making n purchases of the product and r purchases of a brand, as set out schematically in Table C4. From these numerical values, the various statistics for brand X can then be tabulated at last, as will be discussed in C.5 to C.8.

In Table C4 the proportions buying in the second line are for the whole product-field, as presented in Section C.2 (the numerical values in our example being $P_0 = .90000$, $P_1 = .04532$, etc). The body of the table

Table C4. Matrix of Dirichlet Proportions for a Single Brand

		Purchases of the Product								
Numbers:		0	1	2	3		n^*	n'	$n'+1$	Total
Proportions:		P_0	P_1	P_2	P_3	...	P_{n^*}	$P_{n'}$	$P_{n'+1}$	1.0
	0	p_{00}	p_{01}	p_{02}	p_{03}	...	p_{0n^*}	$p_{0n'}$	$p_{0(n'+1)}$	p_0
	1		p_{11}	p_{12}	p_{13}	...	p_{1n^*}	$p_{1n'}$	$p_{1(n'+1)}$	p_1
	2			p_{22}	p_{23}	...	p_{2n^*}	$p_{2n'}$	$p_{2(n'+1)}$	p_2
Purchases	.				.	...	.	.	.	.
of the	.					...	.	.	.	.
Single Brand	.					...	.	.	.	.
	$r'-1$						$p_{(r'-1)n^*}$	$p_{(r'-1)n'}$	$p_{(r'-1)(n'+1)}$	$p_{r'-1}$
	r'							$p_{r'n'}$	$p_{r'(n'+1)}$	$p_{r'}$
	$r'+1$								$p_{(r'+1)(n'+1)}$	$p_{r'+1}$

Notes: $n = 0,1,2,\ldots,n'-1,\ n',\ n'+1$ for the Product Field
$r = 0,1,2,\ldots,r'-1,\ r',\ r'+1$ for the single Brand
All entries below the leading diagonal are identically equal to zero.

$\hat{\beta}$ = S. brand α

represents the proportion of the population making n purchases of the product and r purchases of brand X.

The row labelled "0" for purchases of X are the proportions not buying the brand for each group making $n = 0,1,2$, etc purchases of the product. We compute these values by the recurrence formula

$$p_{0n} = P_{0(n-1)} \times \frac{(\hat{\beta} + n - 1)}{(\hat{\alpha} + \hat{\beta} + n - 1)} \times \frac{Pn}{P_{(n-1)}} \quad \text{for } n = 1,\ldots,n'+1.$$

The starting value of p_{00} is equal to P_0 (ie. .90000). Putting into this expression our numerical values of $\hat{\alpha}$, $\hat{\beta}$, P_n and p_{00} from the last section we have:

$$p_{01} = .9 \times \frac{(12.6255 + 1 - 1)}{(22.8062 + 1 - 1)} \times \frac{.04532}{.9}$$

Proportion not buying brand

$$= .9 \times .55360 \times .05036$$

$$= .02509,$$

$$p_{02} = .02509 \times \frac{(12.6255 + 2 - 1)}{(22.8062 + 2 - 1)} \times \frac{.01972}{.04532}$$

$$= .02509 \times .57235 \times .43513$$

$$= .00625.$$

And so on to build up the whole row for non-buyers of the brand. For the next row, and the rest of the matrix, entries to the left of the leading diagonal are identically equal to zero. The remaining entries are found from the recurrence formula

$$p_{rn} = \frac{(n - r + 1)}{r} \times \frac{(\hat{\alpha} + r - 1)}{(\hat{\beta} + n - r)} \times p_{(r-1)n}, \quad \text{for } r = 1,\ldots,r'+1,$$

where the starting values of p_{1n} are obtained from the proportions of p_{0n} in the first row as:

$$p_{11} = \frac{1 - 1 + 1}{1} \times \frac{(10.1807 + 1 - 1)}{(12.6255 + 1 - 1)} \times .02509$$

$$= 1.0 \times .80636 \times .02509$$

$$= .02023,$$

$$p_{12} = \frac{2 - 1 + 1}{1} \times \frac{(10.1807 + 1 - 1)}{(12.6255 + 2 - 1)} \times .00625$$

$$= 2.0 \times .74718 \times .00625$$

$$= .00934.$$

And so on for all n. For the next row, $r = 2$, the leading diagonal is:

$$p_{22} = \frac{2 - 2 + 1}{2} \times \frac{(10.1807 + 2 - 1)}{(12.6255 + 2 - 2)} \times .00934$$

$$= .5 \times .88556 \times .00934$$

$$= .00414.$$

Again, values for all n are obtained from the recurrence formula. The complete matrix is shown in Table C5.

Some arithmetic checks should be made at this stage

$$\sum_{n=0}^{n'+1} p_n = 1.0,$$

$$\sum_{r=0}^{r'+1} p_r = 1.0,$$

$$\sum_{r=1}^{r'+1} rp_r = .125 = BW(\hat{\alpha}/\hat{S}),$$

$$BW(\hat{\alpha}/\hat{S}) = .1 \times 2.8 \times (10.1807/22.8062) = .125$$

and for each n,

$$\sum_{r=0}^{r'+1} p_{rn} = P_n.$$

Table C5. Dirichlet Proportions for Brand X

(In Base Period)

		Purchases of the Product									
Numbers		0	1	2	3	4	5	6	10	11	Total
Proportions		.90000	.04532	.01972	.01111	.00697	.00464	.00321	.00775	.00127	0.99999
	0	.90000	.02509	.00625	.00208	.00079	.00032	.00014	.00007	.00001	.93475
	1		.02023	.00934	.00433	.00205	.00100	.00050	.00031	.00004	.03780
	2			.00413	.00356	.00236	.00143	.00083	.00077	.00010	.01318
	3				.00114	.00140	.00119	.00087	.00127	.00017	.00604
Purchases	4					.00037	.00057	.00058	.00157	.00023	.00332
of	5						.00013	.00024	.00152	.00024	.00213
Brand *X*	6							.00005	.00116	.00021	.00142
	7								.00068	.00015	.00083
	8								.00030	.00008	.00038
	9								.00009	.00003	.00012
	10								.00001	.00001	.00002
	11									.00000	.00000

C.5 Single-Brand Measures of Buyer Behaviour

The standard measures of buyer behaviour can now be calculated from the estimates of $\hat{S}$, $\hat{\alpha}_j$ and the proportions in Table C5. They are of two types: (i) single-brand measures, which are illustrated in this section, and (ii) measures of multi-brand purchasing, which relate one brand to the whole product-field (Section C.6) or to individual competing brands (Section C.7).

The theoretical frequency distribution of purchases, for a single brand in the analysis period, is derived from the final column in Table C4, ie. p_0, $p_1,\ldots,p_{r'+1}$. In our example (Table C5) an estimated 93.5% of households do not buy brand X, 3.8% buy it once, 1.3% buy it twice, and so forth. From this we calculate that of those buying brand X in the base period, 58% are estimated to have bought the brand once, and that such once-only buyers contributed to 30% of sales.

The theoretical value of the brand's penetration is:

$$\begin{aligned}\hat{b} &= 1 - p_0\\ &= 1 - .93475\\ &= .06525 \text{ or } 6.5\%.\end{aligned}$$

The estimate of the average number of purchases per buyer, $\hat{w}$, comes either from observed values of B, W and the market share $(\hat{\alpha}/\hat{S})$, or alternatively from the numerical sum of rp_r in the Table C5:

$$\begin{aligned}\hat{w} &= BW(\hat{\alpha}/\hat{S})/\hat{b} = \sum_{r=1}^{r'+1} rp_r\,/\hat{b}\\ &= .125/.06525\\ &= 1.916 \text{ in either case.}\end{aligned}$$

These values are close to the "observed" $b = 6.5\%$ and $w = 1.923$ in Table C2.

Other buying statistics for a single brand can also be estimated, such as the incidence of period-by-period repeat-buying, including estimates conditional on being light, medium or heavy buyers in the first period (see Chapter 3). In practice we tend however to use NBD estimates here as they are simpler to obtain and more robust to any non-stationarities in other parts of the market.

C.6 Total Product Usage and Sole Buying

The Dirichlet model also enables us to estimate how brand X is bought in relation to other purchases of the product. One measure is w_p, the total product usage made by buyers of brand X:

$$\hat{w}_p = \{1(P_1 - p_{01}) + 2(P_2 - p_{02}) + \ldots + (n'+1)(P_{n'+1} - p_{0(n'+1)})\}/\hat{b}$$

$$= \{1(.04532-.02509) + 2(.01972-.00625) + 3(.01111-.00208) + \ldots\}/.06525$$

$$= \{.02023 + .02694 + .02709 + \ldots\}/.06525$$

$$= .2309/.06525$$

$$= 3.521\text{, or } 3.5 \text{ approximately.}$$

Another measure is the proportion of panel-members who only buy brand X and no other brand in the analysis period. The estimate of how many of these sole buyers there are is:

$$\hat{b}_s = p_{11} + p_{22} + \ldots + p_{(r'+1)(n'+1)}$$

$$= .02023 + .00413 + .00114 + \ldots$$

$$= .02606\text{, or } 2.6\%.$$

Expressed as a percentage of all buyers of the brand:

$$\hat{b}_s/\hat{b} = .02606/.06525$$

$$= .3994 \text{ or about } 40\%.$$

The average number of purchases made by sole buyers in the period is

$$\hat{w}_s = \{1 \times p_{11} + 2 \times p_{22} + \ldots + (n'+1) \times p_{(r'+1)(n'+1)}\}/\hat{b}_s$$

$$= \{.02023 + .00826 + .00342 + \ldots\}/.02606$$

$$= 1.3216\text{, or about } 1.3.$$

Theoretical results of all the measures so far for all three brands in the example are shown in Table C6.

Table C6. Theoretical Measures of Buyer Behaviour

Brand	Market Share $(m_j/\Sigma m)\%$	Penetration $\hat{b}_j\%$	Average Purchase Frequency $\hat{w}_j$	Total Usage $\hat{w}_{pj}$	Incidence of Sole Buyers $\hat{b}_{sj}\%$	Average Rate of Sole Buying $\hat{w}_{sj}$
X	45	6.5	1.9	3.5	40	1.3
Y	32	5.3	1.7	3.8	33	1.2
Z	23	4.3	1.5	4.1	28	1.2

C.7 Duplication of Purchase

The proportion of individuals who buy a pair of brands, *X* and *Y*, is estimated from the Dirichlet by calculating their separate and combined penetrations in the base-period, as noted in Chapter 13. This is done by forming a composite brand $(X + Y)$ – adding together their separate shares to give $\hat{\alpha}_X + \hat{\alpha}_Y = 10.1807 + 7.3299 = 17.5106$. This composite parameter is used to revise the matrix of Dirichlet proportions, following the procedure set out in Section C.4. From the new matrix a composite penetration figure, $\hat{b}_{(X+Y)}$, is obtained (.0880 in the present case). The proportion, $\hat{b}_{XY}$, buying both brands *X* and *Y* is then given by

$$\begin{aligned} \hat{b}_{XY} &= \hat{b}_X + \hat{b}_Y - \hat{b}_{(X+Y)} \\ &= .0653 + .0531 - .0880 \\ &= .0304. \end{aligned}$$

Conditional proportions can also be calculated:

$$\begin{aligned} \hat{b}_{X|Y} &= \hat{b}_{XY}/\hat{b}_Y = .0304/.0531 = .573 \text{ or } 57\%, \\ \hat{b}_{Y|X} &= \hat{b}_{YX}/\hat{b}_X = .0304/.0653 = .466 \text{ or } 47\%. \end{aligned}$$

In Chapter 10.5 it was seen how the average levels of duplication are systematically related to a brand's penetration. For this purpose an empirical coefficient D was derived. Here we find the theoretical Duplication Coefficient for the two brands X and Y as:

$$\begin{aligned}
\hat{D}_{XY} &= \hat{b}_{XY}/(\hat{b}_X \times \hat{b}_Y) \\
&= .0304/(.0653 \times .0531) \\
&= 8.767.
\end{aligned}$$

This is abnormally high compared with our experience of real-life cases, and is due once more to the artificiality of our small example.

So far we have estimates for the proportion of duplicate buyers and the Duplication Coefficient for a pair of brands. The calculations have in principle to be repeated for all other pairs of brands to obtain the full picture. The remaining answers in our three-brand example are:

$$\begin{aligned}
\hat{b}_{XZ} &= \hat{b}_X + \hat{b}_Z - \hat{b}_{(X+Z)} \\
&= .0653 + .0425 - .0826 \\
&= .0252, \\
\hat{b}_{YZ} &= \hat{b}_Y + \hat{b}_Z - \hat{b}_{(Y+Z)} \\
&= .0531 + .0425 - .0739 \\
&= .0217.
\end{aligned}$$

(Note that the table is symmetric, i.e. $\hat{b}_{XY} = \hat{b}_{YX}$, $\hat{b}_{ZX} = \hat{b}_{XZ}$, and $\hat{b}_{ZY} = \hat{b}_{YZ}$). Complete tables of (i) the proportions buying, and (ii) the conditional percentages are respectively:

Absolute Proportions	Brands		
	X	Y	Z
X	–	.0304	.0252
Y	.0304	–	.0217
Z	.0252	.0217	–

and

Conditional Proportions	% also buying		
	X	Y	Z
Buyers of			
X	–	47	39
Y	57	–	41
Z	59	51	–
Average	58	49	40

The theoretical entries down each column increase slightly with decreasing penetration (from X to Z). In principle this contradicts the Duplication of Purchase Law $b_{X|Y} = Db_X$ of Chapter 10 which expects them to be constant (with the same D for all pairs of brands). But it is in line with the patterns already observed empirically in 1972 and before (p.178). The Dirichlet is therefore more realistic than the Duplication of Purchase Law, but the difference is generally small.

In parallel with this, the pairwise duplication coefficients predicted by the Dirichlet are also not constant, although all of them are close to the average of about 9.2:

Duplication Coefficients	X	Y	Z	
X	–	8.8	9.1	
Y	8.8	–	9.6	
Z	9.1	9.6	–	
Average	9.0	9.2	9.4	= 9.2

Short-Cuts. Calculating all the pairwise duplications for a large number of brands requires extensive computing; for g brands there are $((g \times g) - g)/2$ pairs, which means 45 pairs for 10 brands, or 190 pairs for 20 brands. In practice, calculating the average D from the observed data as in Chapter 10 remains a useful simplification where these data exist.

Two theoretical short-cuts, which approximate the Dirichlet itself, could be:

(i) To calculate the g Ds for the *average* brand with each of the g specific brands, along the lines already set out.

(ii) To calculate the Ds for extreme pairs of brands – if L_1 *and* L_2 and S_1 and S_2 are the two largest and two smallest pairs, then calculate Ds for L_1L_2, L_1S_2, L_2S_1, and S_1S_2. This will cover the range of possible values. To simplify one can approximate to the required single value by a (possibly weighted) average.

In practice, it is easier to use a computer routine.

Other Multi-Brand Statistics. To illustrate briefly the wider range of statistics that can be estimated from the Dirichlet and which show how buyers combine their purchases, we note first the average number of brands bought per buyer of the product

$$\sum_{j=1}^{g} \hat{b}_j/B \quad = \quad .0653 + .0531 + .0425 \,/\, .100$$

$$= \quad 1.609.$$

Second, using the sole-buying figures in Table C5, the proportion of the total sample buying only one brand is

$$\sum_{j=1}^{g} \hat{b}_{sj} \quad = \quad .0261 + .0174 + .0120$$

$$= \quad .0555, \text{ or } 5.6\%.$$

Alternatively, we divide this sum by B to get the proportion of product buyers who buy only one brand, ie. $.0555/.100 = .555$, or about 55%.

C.8 Different Length Time Periods

Theoretical measures of buyer behaviour have so far been presented for some chosen base period. Now we shall illustrate the calculations for other time spans, of length T relative to the "unit" length of the chosen base period (where T can be greater or less than 1).

The proportions along the first row of Table C4 now depend on TA, TM and K (where A, M and K come from the existing product-field NBD in the base period). We predict B and W for length T using standard NBD theory

(Appendix A.4 and A.5). In our numerical example the unit length is 12 weeks, so for $T = 2$ (ie. 24 weeks) TB and TW are

$$\begin{aligned} TB &= 1-(1-TA)^{-K} \\ &= 1-(1-(2 \times 4.559))^{-.0614} \\ &= 1-.8675 \\ &= .1325, \\ TW &= TM/TB \\ &= .56\ /\ .1325 \\ &= 4.2264. \end{aligned}$$

This gives an estimate for P_0 of .8675 which is used in a complete re-run of the NBD model with the new parameters:

$$\begin{aligned} P_0 &= .8675, \\ 2A &= 2 \times 4.559 = 9.118, \\ 2M &= 2 \times .28 \quad = \quad .56, \\ k &= 2M/2A \quad = \quad .0614 \text{ (constant)}. \end{aligned}$$

These values are fitted into the earlier recursive formula

$$P_n = \left(\frac{2A}{1+2A}\right)\left(1-\frac{2A-2M}{2An}\right) P_{n-1},$$

so that

$$\begin{aligned} P_1 &= .901166\,(1-.93858)\,.86750 = .04802 \\ P_2 &= .901166\,(1-.46929)\,.04802 = .02296 \\ P_3 &= .901166\,(1-.31286)\,.02296 = .01422 \\ &\ \ \vdots \qquad\qquad\qquad\qquad\qquad\qquad \vdots \end{aligned}$$

With this new distribution of product purchases the cycle starts again from Section C.5. The distribution P_n, $n = 0,1,...,n'+1$, replaces line one in Table C4 and, along with the original estimates of $\hat{S}$ and $\hat{\alpha}_j$ (ie. 22.8062, and 10.1807, 7.3299 and 5.2933 respectively), the body of the matrix is revised. (The values of the brand choice parameters $\hat{S}$ and $\hat{\alpha}_j$ in the Dirichlet are unaffected by T).

The standard measures of buyer behaviour for brand X in the double period are

Measure	Single Period	Double Period
b_X	.0653	.0917
w_X	1.92	2.48
w_{PX}	3.54	4.91

For two periods of equal unit length several aspects of repeat-buying are now predictable, such as the incidence of repeat-buyers and new buyers, and their rates of buying. These measures are based on the penetration in a single period in relation to that in a double period. Thus, the proportion buying X in both periods is:

$$\begin{aligned} b_{X12} &= 2 \times b_{X1} - b_{X2} \\ &= 2 \times .0653 - .0917 \\ &= .0389. \end{aligned}$$

The proportion of new buyers over the same time span is:

$$\begin{aligned} \text{new } b_{X2} &= b_{X1} - b_{X12} \\ &= .0653 - .0389 \\ &= .0264. \end{aligned}$$

Finally, the incidence of repeat-buying is defined as:

$$\begin{aligned} b_{X12}/b_X &= .0389/.0653 \times 100 \\ &= 59.6\% \end{aligned}$$

which compares with 56.1% calculated from conventional NBD theory (using data from the first 12 weeks alone - Appendix A.8). Dirichlet and NBD period-to-period levels of repeat-buying are generally similar, and in this case the difference is just 3.5%.

At this stage the difference between NBD and Empirical versions of the Dirichlet model becomes important. The Empirical version cannot be used directly to make predictions over different time spans. Instead, a new product field distribution and new values of $\hat{S}$ and $\hat{\alpha}_j$ for T must be tabulated from raw data, if these are available. This means the base-period is re-defined at T and the cycle starts from Section C.4 again. Under stationary conditions K, $\hat{S}$ and $\hat{\alpha}_j$ ought to be constant - it is good to check this. Data requirements for the Empirical version are heavy and the model itself is inelegant; it is however useful to fit both versions and then compare the resultant statistics.

C.9 Multi-Brand Tabulations

We now consider the hand tabulation of observed data to obtain the multi-brand buying statistics discussed in Chapters 9 and 10 and in theoretical terms earlier in this Appendix. We use again the hypothetical 12-week example in Table A20 which is reproduced here in Table C7 with various summary statistics.

In this section we discuss tabulating the penetration B and average purchase frequency W for the product field, and in Section C.10, w_p, the total number of purchases of the product-field per buyer of a given brand. In Section C.11 and C.12 we consider tabulations of sole buyers and of duplicate buyers.

The first two summary columns in Table C7 show whether a panel-member bought the product (ie. any brand) at least once in the 12-week period and how often they bought it - eg. 11 times for panel-member 1. There are 20 panel-members (out of 200) who each made at least one purchase of some brand in the 12 weeks so that B = 20/200 = .1 or 10%. They bought 56 times, so that W = 56/20 = 2.8 and M = 56/200 = .28, as noted earlier.

The frequency distribution of product-field purchases, ie. one 11, one 6, one 5, three 4s, three 3s, two 2s and nine 1s (adding up to 56 as a check), can be derived from the Total column. This can be used as direct input for the Empirical-Dirichlet model, or to check the assumptions of an NBD.

As a numerical check it is worthwhile to count the number of purchases each week (at the foot of each column) and to check that the total (56)

Table C7. The Buying Pattern of Table A20 with Summary Statistics

	Purchases in Week:												The Product		Brands					
													At least		At least once			Totals		
Panel Member	1	2	3	4	5	6	7	8	9	10	11	12	once	Total	*X*	*Y*	*Z*	*X*	*Y*	*Z*
1	*X*	*X*	*Y*	*X*	*X*	*X*	*Y*	*X*	.	*Y*	*X*	*Y*	1	11	1	1	.	7	4	.
2	.	.	*X*	.	*Z*	*X*	.	.	.	*Z*	*X*	*X*	1	6	1	.	1	4	.	2
3	.	.	*Z*	.	.	*Z*	.	*X*	*X*	.	*Z*	.	1	5	1	.	1	2	.	3
4	.	.	.	.	*Y*	.	*Y*	*Y*	.	*X*	.	.	1	4	1	1	.	1	3	.
5	*Y*	*X*	.	.	*X*	.	.	.	.	.	.	*Y*	1	4	1	1	.	2	2	.
6	*X*	.	.	.	.	.	.	*Z*	*Z*	*Y*	.	.	1	4	1	1	1	1	1	2
7	.	*Z*	*X*	.	.	.	.	.	.	.	.	*X*	1	3	1	.	1	2	.	1
8	.	*Y*	.	*Y*	.	.	.	.	*X*	.	.	.	1	3	1	1	.	1	2	.
9	*Z*	.	.	.	.	.	*X*	.	.	.	*Y*	.	1	3	1	1	1	1	1	1
10	.	.	.	*X*	.	*Y*	.	.	.	.	.	.	1	2	1	1	.	1	1	.
11	.	.	.	*Y*	.	.	*X*	.	.	.	.	.	1	2	1	1	.	1	1	.
12	.	.	.	.	.	.	.	.	.	*X*	.	.	1	1	1	.	.	1	.	.
13	*Y*	.	.	.	.	.	.	.	.	.	.	.	1	1	.	1	.	.	1	.
14	.	.	.	.	.	.	.	*Y*	.	.	.	.	1	1	.	1	.	.	1	.
15	.	.	*Z*	.	.	.	.	.	.	.	.	.	1	1	.	.	1	.	.	1
16	.	.	.	.	.	.	.	.	*Y*	.	.	.	1	1	.	1	.	.	1	.
17	.	.	.	.	*Z*	.	.	.	.	.	.	.	1	1	.	.	1	.	.	1
18	.	.	.	.	.	.	.	.	.	.	*Z*	.	1	1	.	.	1	.	.	1
19	.	.	.	.	.	.	.	.	.	*Z*	.	.	1	1	.	.	1	.	.	1
20	.	.	.	.	.	*X*	.	.	.	.	.	.	1	1	1	.	.	1	.	.
21–200	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
The Product																				
Buying	5	4	5	4	5	5	4	5	4	6	5	4	20	.	13	11	9	-	-	-
Total Product	5	4	5	4	5	5	4	5	4	6	5	4	.	56	49	36	25	-	-	-

agrees with the sum of the row totals. No such cross-check is possible for the number of buyers each week.

C.10 Tabulating Product Rates of Buying

To get w_p for brand X, we have to count the total number of purchases of the product by the 13 panel-members who bought X at least once in the 12 weeks. This is tricky by hand, probably needing two fingers as we move down the column, one for the Total Product and another for buying X at least once. It is easy to make a mistake, and there is no check other than by repeating the count (preferably starting from the bottom the second or third time, to avoid possibly making the same mistakes as before). The final w_ps for the three brands are:

$$\text{Brand A} = 49/13 = 3.8,\ \text{B} = 35/11 = 3.2,\ \text{C} = 25/9 = 2.8.$$

C.11 Tabulating Sole Buyers

In Table C7 we see by visual inspection that panel-members 12 and 20 are 100%-loyal or "sole" buyers of X in the 12-week period, and similarly that there are sole buyers of Y (13, 14 and 16), and of Z. In the shorter period, weeks 1–4, panel-members 2 and 10 are sole buyers of X (but they buy other brands in subsequent weeks).

Tabulating sole buyers of a given brand, X say, by hand requires seeing whether a panel-member

(i) bought brand X at all, and
(ii) bought no other brand

in the chosen analysis period, and doing a gate-count accordingly. Numerically, some $b_s = 2/200 = .01$ (or 1%) of the sample only bought X in the 12 weeks, or $b_s/b = 2/13 = .15$ (or 15%) of X buyers were sole buyers.

If the number of purchases made by sole buyers is to be tabulated as well, a separate gate count of all numbers of purchase occasions should be kept, which for brand X happens to give 6 and hence $w_s = 6/6 = 1.0$ (w_s is usually low, but not as low as this). The results for all three brands are:

	b_s/b	w_s
X	15	1.0
Y	27	1.0
Z	44	1.0

The marked variation of b_s/b is an atypical feature of our artificial example.

Sole-buyers' rate of buying the product in an analysis period necessarily equals their rate of buying the brand. This provides a very handy computational short-cut for identifying sole buyers: they are simply any panel-member for whom $w = w_p$.

C.12 Tabulating Duplication Tables

Tabulating duplication tables by hand is a substantial task if the number of brands is at all large. For three brands, as in our example, separate counts are needed for just three combinations of duplicated buying, X and Y, Y and Z, Z and X. For 10 brands, $(10 \times 9)/2 = 45$ separate counts would be needed.

To obtain the number of duplicated buyers of brands X and Y from Table C7, we look down the columns for buying each brand at least once towards the right of the table and count that 8 panel-members were buying both brands, using gate counts:

X & Y	Y & Z	X & Z
8	2	5

Note how tallying for brands X and Z is more of a mental strain because they are "separated" in Table C7; this would be worse for more than three brands.

The tallies give a duplication table for b_{XY} of the symmetrical form

	Brands		
Absolute values	X	Y	Z
X	13	8	5
Y	8	11	2
Z	5	2	9

There is no indirect check on counting the duplications, so that all one can do is repeat the exercise.

To get figures like $b_{Y.X}$ ie. the percentage of buyers of brand X who also buy Y, we divide the $b_{XY} = 8$ by $b_X = 13$, to give $b_{Y.X} = 8/13 = .6154$, or 62%. The full duplication table is

Conditional proportions	Percentage also buying X	Y	Z
Buyers of			
X	(100)	62	38
Y	73	(100)	18
Z	56	22	(100)
Average	65	42	28

With real life examples the figures in each column (other than the 100%) are usually more uniform.

Average Purchases for Duplicate Buyers. To tabulate the average purchase rates of duplicate buyers, like $w_{X.Y}$ and $w_{Y.X}$, we run through the last columns of Table C7, doing separate counts of their purchases:

Buyers of X & Y Who Purchase: X	Y
7	4
1	3
2	2
1	1
1	2
1	1
1	1
1	1
15	15

This gives $w_{X.Y} = 15/8 = 1.9$, and $w_{Y.X} = 15/8 = 1.9$ as well. (The exact equality of the two averages is accidental, though the figures are usually fairly similar). Note that this type of tabulation is tricky to do correctly _ien working direct from the summary of Table C7, and so it may be quicker in the long run to write the figures out as above.

The complete results for the $w_{X.Y}$ (and w_X) rates of buying are given by the following non-symmetric table:

	Average Purchases of		
Buyers of	*X*	*Y*	*Z*
X	(1.92)	1.9	1.8
Y	1.9	(1.64)	1.5
Z	2.0	1.0	(1.44)
Average	2.0	1.5	1.7

The D - Coefficient. The Duplication Coefficient can be estimated as the ratio of the average of the six duplications for all brands to their average penetration, eg.

(73 + 56 + 62 + 22 + 38 + 18)/6 = 45 divided by (6.5 + 5.5 + 4.5)/3 = 5.5, so that D = 45/5.5 = 8.2

(which in our artificial example is abnormally high; usually it tends to lie in the range 1 to 2).

An alternative form of estimation is to take the average of all the ratios of $b_{Y.X}/b_Y$, calculated for all pairs of brands. Where exceptional duplications occur one can calculate a normative value of D for the more homogeneous subset of duplications.

BIBLIOGRAPHY

Aaker, D.A. and Morgan, S.M. (1971). "Modeling Store Choice Behavior". *J. Marketing Res., 8,* 38–42.

Aitchison, J. (1982) "The Statistical Analysis of Compositional Data". *J.R. Statist. Soc. B, 44,* 139–177.

Amstutz, A.E. (1967). *Computer Simulation of Competitive Market Response.* Cambridge: M.I.T. Press.

Anscombe, F.J. (1950). "Sampling Theory of the Negative Binomial and Logarithmic Distributions". *Biometrika, 37,* 358–382.

Arbour, A.G. and Kerrich, J.E.C. (1951). "Accident Statistics and the Concept of Accident Proneness". *Biometrics, 7,* 340–432.

Aske Research (1969 to 1981): *Dentifrice* (1969); *Ready-to-Eat Cereals* (1969); *The Structure of the Detergent Market* (1969); *Refrigerated Dough* (1969); *Cigarettes* (1971); *Petrol* (1972); *Results Concerning Brand-Choice* (1973); *Ready-to-Eat Cereals* (1974); *The Structure of the Toothpaste Market* (1975); *The Dirichlet Model* (1975); *Fitting the Dirichlet Model* (1976); *The Structure of the Biscuit Market* (1976); *Colour Cosmetics* (1981); London: Aske Research Ltd.

Aske Research (1970). *Loyalty Reports: A Coded Example.* London: Aske Research Ltd.

Bartko, J.J. (1961). "The Negative Binomial Distribution: A Review of Properties and Applications". *Virginia J. of Science, 2,* 18–37.

Barwise, T.P. (1984) "Mass Attitudes and Routine Choice Behaviour". Ph.D. Thesis London University.

Barwise, T.P. and Ehrenberg, A.S.C. (1979), "The Revenue Potential of Channel Four", *ADMAP,* Nov., 550–556.

Barwise, T.P. and Ehrenberg, A.S.C. (1983), "How Much Does UK Television Cost?". *International Journal of Advertising, 2,* 17–32.

Barwise, T.P. and Ehrenberg, A.S.C. (1985), Consumer Beliefs and Brand Usage, *J. Market Research Soc., 27,* 81–93.

Barwise, T.P. and Ehrenberg, A.S.C. (1987), "Consumer Beliefs and Awareness", *J. Market Research Soc., 29,* 88–93.

Barwise, T.P. and Ehrenberg, A.S.C. (1988), *Television and its Audience,* Beverly Hills and London: Sage Publications.

Bass, F.N., Jeuland, A. and Wright, G.P. (1976) "Equilibrium Stochastic Choice and Market Penetration Theories". *Management Sci., 22,* 1051–1063.

Bates, G.E. and Neyman, J. (1952). *Contributions to the Theory of Accident Proneness,* Parts I and II. Berkeley and Los Angeles: University of California Press.

Baum, J. and Dennis, K.E.R. (1961). "The Estimation of the Expected Brand Share of a New Product". Read at the 7th Esomar/Wapor Congress, Baden-Baden. Amsterdam: Esomar.

Bird, M. and Ehrenberg, A.S.C. (1966a). "Intentions-to-Buy and Claimed Brand Usage". *Operational Research Quarterly, 17,* 27–46 and *18,* 65–66.

Bird, M. and Ehrenberg, A.S.C. (1966b). "Non-Awareness and Non-Usage". *J. Advertising Research, 6,* 4–8.

Bird, M. and Ehrenberg, A.S.C. (1970). "Consumer Attitudes and Brand Usage". *J. Market Res. Soc. 12,* 233–247, *13,* 100–1 and 242–3, *14,* 57–8.

Bird, M., Channon, C. and Ehrenberg A.S.C. (1969). "Brand Image and Brand Usage". *J. Marketing Res. 7,* 307–314.

Boswell, M.T. and Patil, G.P. (1970). "Chance Mechanisms Generating the Negative Binomial Distribution". In *Random Counts in Scientific Work, Vol. 1* (ed. G.P. Patil). Pennsylvania State University Press.

Brace, D.A. (1959). Private Communication.

Brown, G.H. (1952, 1953). "Brand Loyalty – Fact or Fiction?" *Advertising Age, 23*, (9.8, 30.6, 14.8, 28.7, 4.8, 11.8, 1.8, 22.9, 6.10, 1.12) and *24* (26.1).

Butler, B.F. (1966) *Hendrodynamics: Fundamental Laws of Consumer Dynamics*. Croton-on-Hudson: The Hendry Corporation.

Cannon, T., Ehrenberg, A.S.C. and Goodhardt, G.J. (1970). "Regularities in Sole Buying". *Brit. J. Marketing, 4*, 80–86.

Castleberry, S. (1983) "Longitudinal field-experiment of repeat-purchase behaviour: *Effects of varying intrinsic quality, price, and brand availability*". Ph.D. Thesis, University of Alabama.

Castleberry, S., Ehrenberg, A.S.C. and Motes, W.M. (1987), "Extended Sales Tests of Product Quality", *J. Market Research Soc. 29*, 3–14.

Charlton, P. and Ehrenberg, A.S.C. (1976a), "Customers of the LEP". *Appl. Statist., 25*, 26–30.

Charlton, P. and Ehrenberg, A.S.C. (1976b), "An experiment in Brand Choice". *J. Marketing Res., 13*, 152–160.

Charlton, P., Ehrenberg, A.S.C. and Pymont, B. (1972), "Buyer Behaviour under Mini-test Conditions". *J. Market Res. Soc., 14*, 171–183.

Charnes, A., Cooper, W.W., De Voe, J.K. and Learner, D.B. (1971). "Models of Fact and Laws of Behaviour". *Management Science, 17*, 367–369.

Chatfield, C. (1969). "On Estimating the Parameters of the Logarithmic and Negative Binomial Distributions". *Biometrika, 56*, 411–414.

Chatfield, C. (1970). "Discrete Distributions in Market Research". In *Random Counts in Scientific Work*, Vol. 3 (ed. G.P. Patil). Pennsylvania State University Press.

Chatfield, C. (1975), "A Marketing Application of a Characterisation Theorem". In *Statistical Distributions in Scientific Work, Vol. 2*. (G.P. Patil, S. Kotz and J.K. Ord, eds). Dordrecht: D. Reidel.

Chatfield, C., Ehrenberg, A.S.C. and Goodhardt, G.J. (1966). "Progress on a Simplified Model of Stationary Purchasing Behaviour". *J.R. Statist. Soc. A., 129*, 317–367.

Chatfield, C. and Goodhardt, G.J. (1970). "The Beta-Binomial Model for Consumer Purchasing Behaviour". *Applied Statistics, 19*, 240–250.

Chatfield, C. and Goodhardt, G.J. (1973), "A Consumer Purchasing Model with Erlang Inter-purchase Times". *J. Amer. Statist. Ass., 68*, 828–835.

Chatfield, C. and Goodhardt, G.J. (1975), "Results Concerning Brand-choice". *J. Marketing Res., 12*, 110–113.

Coggill, C. and Simpson, S. (1984), *Analysis of Cambridge Division Sales*. Royston: PA Technology.

Collins, M.A. (1971). "Market Segmentation – The Realities of Buyer Behaviour". *J. Market Res. Soc. 13*, 146–157.

Connor, R.J. and Mosimann, J.E. (1969), "Concepts of Independence for Proportions with a Generalization of the Dirichlet Distribution". *J. Amer. Statist. Ass., 64*, 194–206.

Cramer, J.S. (1965). Private Communication.

Cresswell, W.L. and Frogatt, P. (1963). *The Causation of Bus-Driver Accidents – An Epidemiological Study*. London: Oxford University Press.

Darroch, J.N. and James, I.R. (1974), "*F*-independence and Null Correlation of Continuous, Bounded-sum, Positive Variables". *J.R. Statist. Soc. B, 36*, 467–483.

Darroch, J.N. and Ratcliffe, D. (1971), "A characterization of the Dirichlet distribution". *J. Amer. Statist. Assoc. 66*, 641–643.

Davis, E.J. (1964). "Test Marketing: An Examination of Sales Patterns Found during Forty-Four Recent Tests". In: *Research in Marketing*. London: The Market Research Society.

Day, G.A. (1969). Buyer Attitudes and Brand Choice Behavior. New York: The Free Press.

Dunn, R., Reader, S. and Wrigley, N. (1983), "An Investigation of the Assumptions of the NBD Model as Applied to Purchasing at Individual Stores." *Appl. Statist., 32,* 249–259.

Easton, G. (1979), "Stochastic Models of Industrial Buying Behaviour." *Omega, 8,* 63–69.

Ehrenberg, A.S.C. (1959a). "The Pattern of Consumer Purchases". *Applied Statistics, 8,* 26–41.

Ehrenberg, A.S.C. (1959b). "The Relative Merits of Independent Matched Samples and of the Panel Technique for Before-and-After Studies". *Commentary,* Number *1,* 1–7.

Ehrenberg, A.S.C. (1960). "A Study of Some Potential Biases in the Operation of a Consumer Panel". *Applied Statistics, 9,* 20–27.

Ehrenberg, A.S.C. (1962). "Verified Predictions of Consumer Purchasing Patterns". *Commentary,* Number *10,* 16–21.

Ehrenberg, A.S.C. (1963). "Bivariate Regression Analysis is Useless". *Applied Statistics, 12,* 161–179.

Ehrenberg, A.S.C. (1964). "Estimating the Proportion of Loyal Buyers.". *J. Marketing Res., 1,* 56–59.

Ehrenberg, A.S.C. (1965a). *Pack-Size Duplication of Purchase.* London: Aske Research Ltd.

Ehrenberg, A.S.C. (1965b). "Knowledge as our Discipline". *Commmentary, 4,* 211–235.

Ehrenberg, A.S.C. (1965c). "An Appraisal of Markov Brand-Switching Models". *J. Marketing Res. 2,* 347–362.

Ehrenberg, A.S.C. (1966a). "Ten Questions About Consumer Purchasing Behaviour and Some Answers". *Advert. Quart. 9,* 3–8.

Ehrenberg, A.S.C. (1966b). "Laws in Marketing – A Tailpiece". *Applied Statistics, 15,* 257–267, and in Ehrenberg and Pyatt (1971).

Ehrenberg, A.S.C. (1967a). "America and the Rest – Some Comparisons". *Commentary, 9,* 12–21.

Ehrenberg, A.S.C. (1967b). "Where Were You in the Revolution? – Marketing Research in the Future". *Admap, 3,* 6, 247–250.

Ehrenberg, A.S.C. (1967c). "The Neglected Use of Data". *J. Advert. Res., 7,* 1, 2–7.

Ehrenberg, A.S.C. (1967d). "On Not Understanding Media/Product Data". *Commentary, 9,* 203–212.

Ehrenberg, A.S.C. (1968a). "Models of Fact: Examples from Marketing". In: *Mathematical Model Building in Economics and Industry* (M.G. Kendall ed.). London: Charles Griffin, and in *Management Science B* (1970), *16,* 435–445.

Ehrenberg, A.S.C. (1968b). "The Elements of Lawlike Relationships". *J.R. Statist. Soc. A., 131,* 280–329.

Ehrenberg, A.S.C. (1968c). "The Factor Analytic Search for Programme Types". *J. Advert. Res., 8,* 1, 55–63.

Ehrenberg, A.S.C. (1968d). "The Time and Place for Readership Panels". *J. Advert. Res., 8,* 2, 19–22.

Ehrenberg, A.S.C. (1968e). "The Practical Meaning and Usefulness of the NBD/LSD Theory of Repeat-Buying". *Applied Statistics, 17,* 17–32.

Ehrenberg, A.S.C. (1968f). "On Clarifying M and M". *J. Marketing Res., 5,* 228–9.

Ehrenberg, A.S.C. (1968g). "Review of F.M. Nicosia's Consumer Decision Processes". *J. Marketing Res., 5,* 334.

Ehrenberg, A.S.C. (1968h). "Repeat-Buying of Textile Garments". *Op. Res. Quart. 14,* 421–432.

Ehrenberg, A.S.C. (1969a). "The Discovery and Use of Laws of Marketing". *J. Advert. Res., 9,* 2, 11–17.

Ehrenberg, A.S.C. (1969b). "Statisticians as Their Own Customers". *Conference of the Royal Statistical Society on "Customer Satisfaction",* April 1969, Sheffield.

Ehrenberg, A.S.C. (1969c). "Towards an Integrated Theory of Consumer Behaviour". *J. Market Res. Soc., 11,* 305–337, and in Ehrenberg and Pyatt (1971).

Ehrenberg, A.S.C. (1970a). "A Note on Never-Buyers". *J. Marketing Res., 7,* 536–8.

Ehrenberg, A.S.C. (1970b). "Predicting the Behaviour of New Brands". *Read at the Boston Conference of the American Marketing Association* and *J. Advert. Res., 11.*

Ehrenberg, A.S.C. (1971a). "Tempus Fidgets". *Management Science B, 17,* 369–370.

Ehrenberg, A.S.C. (1971b). "Prior Information and the Analysis of Sales". Read to a joint Meeting of the Market Research Society and the Royal Statistical Society, May 1971.

Ehrenberg, A.S.C. (1974). "Repetitive Advertising and the Consumer", *Journal of Advertising Research, 14,* 25–34.

Ehrenberg, A.S.C. (1975a) *Data Reduction.* London and New York: Wiley.

Ehrenberg, A.S.C. (1975b) "The Structure of an Industrial Market: the Case of Aviation Fuel", *Industrial Marketing Management, 4,* 275–285.

Ehrenberg, A.S.C. (1981) "Marketing experiments". In *Proceedings of the 24th Annual Conference,* London: The Market Research Society.

Ehrenberg, A.S.C. (1982) *A Primer in Data Reduction.* London and New York: Wiley.

Ehrenberg, A.S.C. (1986) "Pricing and Brand Differentiation", *Singapore Review of Marketing, 1,* 5–15.

Ehrenberg, A.S.C. (1987). "New Brands and the Existing Market", London Business School: CMaC Working Paper.

Ehrenberg, A.S.C. and Charlton, P. (1972). "An Analysis of Simulated Brand-Choice". *J. Advert. Res., 12,* 21–33.

Ehrenberg, A.S.C. and England, L.R. (1987), "Generalising a Pricing Effect", London Business School: CMaC Working Paper.

Ehrenberg, A.S.C. and Goodhardt, G.J. (1968a). "The Amount Bought by Above-Average Buyers". *J. Market Res. Soc., 10,* 157–171.

Ehrenberg, A.S.C. and Goodhardt, G.J. (1968b). "A Comparison of American and British Repeat-Buying Habits". *J. Marketing Res., 5,* 15–18.

Ehrenberg, A.S.C. and Goodhardt, C.J. (1968c). "Competitive Rates of Purchasing: Some Initial Results". *Admap, 9,* 157–168.

Ehrenberg, A.S.C. and Goodhardt, G.J. (1968d). "Repeat-Buying of a New Brand". *Brit. J. Marketing, 2,* 200–205.

Ehrenberg, A.S.C. and Goodhardt, G.J. (1968e). "The Incidence of Brand-Switching". *Nature, 220,* 5764, 304.

Ehrenberg, A.S.C. and Goodhardt, C.J. (1969a). "Loyalty Reports – A New Analysis Service". *Admap, 4,* 162–164.

Ehrenberg, A.S.C. and Goodhardt, G.J. (1969b). "The Evaluation of a Consumer Deal". *Admap, 5,* 388–93.

Ehrenberg, A.S.C. and Goodhardt, G.J. (1970a). "Pack-size Rates of Purchasing". *Applied Economics, 2,* 15–16.

Ehrenberg, A.S.C. and Goodhardt, G.J. (1970b). "A Model of Multi-Brand Buying". *J. Marketing Res., 7,* 77–84.

Ehrenberg, A.S.C. and Goodhardt, G.J. (1973) "Setting Budgets and Allocating Advertising Effects". In *Proceedings of the 19th Annual Conference.* New York: The Advertising Research Foundation.

Ehrenberg, A.S.C. and Goodhardt, G.J. (1974), "The Hendry Brand-switching Coefficient". *Admap, 10,* 232–238.

Ehrenberg, A.S.C. and Goodhardt, G.J. (1976), *Decision Models and Descriptive Models in Marketing.* Cambridge: Marketing Science Institute.

Ehrenberg, A.S.C. and Goodhardt, G.J. (1979a), *Understanding Buyer Behaviour.* New York: J. Walter Thompson and MRCA.

Ehrenberg, A.S.C. and Goodhardt, G.J. (1979b), "The Switching Constant". *Management Sci., 25,* 703–705.

Ehrenberg, A.S.C., Goodhardt, G.J. and Barwise, T.P. (1987) "Double Jeopardy Revisited", London Business School: CMaC Working Paper.

Ehrenberg, A.S.C. and Pyatt, G.F. (eds.) (1971). *Consumer Behaviour*. London: Penguin Books.

Ehrenberg, A.S.C. and Twyman, W.A. (1966). "On Measuring Television Audiences". *J. Roy. Statist. Soc. A., 130,* 1–59.

E.I.U. (1963). "A Special Report on Consumer Panels". In: *Retail Business, 61,* London: The Economist Intelligence Unit.

Engel, J.F., Kollat, D.T. and Blackwell, R.D. (1968). *Consumer Behavior*. New York: Holt, Rinehart and Winston.

Evans, D.A. (1953). "Experimental Evidence Concerning Contagious Distributions in Ecology". *Biometrika, 40,* 186–196.

Farley, J.U. and Ring, L.W. (1970)."An Empirical Test of the Howard-Sheth Model of Buyer Behavior". *J. Marketing Res., 7,* 427–438.

Feller, W. (1957). *An Introduction to Probability Theory and Its Applications (2nd ed.)* New York: Wiley.

Ferber, R. (1962). "Research on Household Behaviour". *Am. Econ. Review, 52,* 19–63, and in Ehrenberg and Pyatt (1971).

Ferber, R. (1963). "Observations on a Consumer Panel Operation". *J. of Marketing, 17,* 246–259.

Fisher, R.A., Corbett, R.S. and Williams, C.B. (1943). "The Relation Between the Number of Species and the Number of Individuals in a Random Sample of an Animal Population". *J. Animal Biology, 12,* 42–58.

Fourt, L.A. and Woodlock, J.W. (1960). "Early Prediction of Market Success for New Grocery Products". *J. of Marketing, 25,* 31–38.

Frank, R.E. (1962). "Brand Choice as a Probability Process". *J. Business, 35,* 43–56.

Frank, R.E. (1967). "Correlates of Buying Behaviour for Grocery Products". *Journal of Marketing, 31,* 48–53, and in Ehrenberg and Pyatt (1971).

Frisbie, G.A. (1980) "Ehrenberg's Negative Binomial Model Applied to Grocery Store Trips". *J. Marketing Res., 17,* 385–390.

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

Goodhardt, G.J. (1972). "Multi-Brand Buying and Preference Analysis". *J. Market Res. Soc. 14,* 46–53.

Goodhardt, G.J. and Chatfield, C. (1973), "Gamma-distribution in Consumer Purchasing", *Nature, 244,* No. 5414, 316.

Goodhardt, G.J. and Ehrenberg, A.S.C. (1967). "Conditional Trend Analysis: A Breakdown by Initial Purchasing Level". *J. Marketing Res., 4,* 155–161.

Goodhardt, G.J. Ehrenberg, A.S.C. and Chatfield, C. (1984) "The Dirichlet: A comprehensive Model of Buying Behaviour", *J.R. Statist. Soc.,* A, *147,* 621–655.

Goodhardt, G.J. Ehrenberg, A.S.C. and Collins, M.A. (1975, 1987). *The Television Audience; Patterns of Viewing*. Aldershot: Gower Press.

Grahn, G.L. (1969). "The Negative Binomial Distribution Model of Repeat-Purchase Loyalty: An Empirical Investigation". *J. Marketing Res., 6,* 72–78.

Greene, J. (1982), *Consumer Behavior Models for Non-Statisticians,* New York: Praeger.

Greenwood, M. and Yule, G.U. (1920). "An Enquiry into the Nature of the Frequency Distributions Representative of Multiple Happenings, with Particular Reference to the Occurrence of Multiple Attacks of Disease or Repeated Accidents". *J.R. Statist. Soc. A., 83,* 255–279.

Gurland, J. (1954). "Some Applications of the Negative Binomial and Other Contagious Distributions". *Amer. J. Public. Health, 49,* 1388–1399.

Gurland, J. (1957), "Some Interrelations among Compound and Generalised Distributions". *Biometrika, 44,* 265–268.

Haight, F.A. (1967). *Handbook of the Poisson Distribution.* New York: Wiley.

Herniter, J. (1969). *Probabilistic Market Models of Purchasing Timing and Brand Selection.* Cambridge: Marketing Science Institute.

Herniter, J. (1971), "A Probabilistic Market Model of Purchase Timing and Brand Selection". *Management Sci., 18,* 102–113.

Herniter, J. (1973), "An Entropy Model of Brand Purchase Behavior". *J. Advert. Res., 10,* 361–375.

Hinshelwood, C.N. (1967). "President's Anniversary Address". *Proc. Roy. Soc. B., 148,* 5–16.

Holloway, R.J., Mittelstaedt, R.A. and Venkatesan, M. (1971). *Consumer Behavior.* Boston: Houghton Mifflin.

Howard, J.A. and Sheth, J.N. (1969). *The Theory of Buyer Behaviour.* New York: John Wiley.

Hyett, G.P. (1958), "The Measurement of Readership". Seminar paper at the London School of Economics.

Irwin, J.O. (1964). "The Personal Factor in Accidents – A Review Article". *J.R. Statist. Soc. A., 127,* 438–451.

James, I.R. (1975), "Multivariate Distributions which have Beta Conditional Distributions". *J. Amer. Statist. Ass., 70,* 681–684.

Jephcott, J.St.G. (1972), "Consumer loyalty – a fresh look". In *Proceedings of the 15th Annual Conference.* London: The Market Research Society.

Jephcott, J.St. G. and Buck S.F. (1971). "Preference Analysis". *J. Market Res. Soc., 13,* 83–98.

Jeuland, A.P., Bass, F., and Wright, G.P. (1980), "A Multibrand Stochastic Model Compounding Heterogeneous Erlang Timing and Multinomial Choice Processes". *Operations Res., 28,* 255–277.

Johnson, N.L. and Kotz, S. (1969). *Discrete Distributions.* Boston: Houghton Mifflin.

Jones, J. M. and Zufryden, F.S. (1980), "Adding Explanatory Variables to a Consumer Purchase Behaviour Model – an Exploratory Study". *J. Marketing Res., 17*(3), 323–334.

Jones, J.P. (1986), *What's in a Name?* Lexington, Mass: Lexington Books.

Kalwani, M.U. and Morrison, D.G. (1977) "A Parsimonious Description of the Hendry System". *Management Sci., 27,* 467–477.

Kau, Ah Keng (1981) "Patterns of store choice". Ph.D. Thesis. University of London.

Kau, A.K. and Ehrenberg, A.S.C. (1984) "Patterns of store-choice". *J. Marketing Res., 21,* 399–409.

Kemp, C.D. (1970). "Accident Proneness and Discrete Distribution Theory". In *Random Counts in Scientific Work, Vol. 2.* (ed. G.P. Patil). Pennsylvania State University Press.

King, W.R. (1967). *Quantitative Analysis for Marketing Management.* New York: McGraw-Hill.

King, S.H.M. (1970). "What is a Brand?" *Advertising Quarterly, 24,* 6–14.

Kollat, D.J., Blackwell, R.D. and Engel, J.F. (1970). *Research in Consumer Behavior.* New York: Holt, Rinehart and Winston.

Kotler, P. (1967). *Marketing Management.* Englewood Cliffs: Prentice-Hall.

Kuehn, A.A. (1958). An Analysis of the Dynamics of Consumer Behaviour and Its Implications for Marketing Management. (Ph.D. thesis). Carnegie Institute of Technology: Pittsburgh.

Kuehn, A.A. (1962). "Consumer Brand Choice as a Learning Process". *J. Advertising Research, 2,* 10–17.

Kuehn, A.A. and Rohloff, A.C. (1965). "New Dimensions in Brand-Switching", in *New Directions in Marketing.* (F.E. Webster, ed.). Chigago: American Marketing Association.

Lawrence, R.J. (1966). "Models of Consumer Purchasing Behaviour". *Applied Statistics, 45,* 216–233.

Leckenby, J.D. and Kishi, S. (1984), "The Dirichlet Multinomial Distribution as a Magazine Exposure Model". *J. Marketing Res., 21,* 100–106.

Lilien, G. and Kotler, P. (1983), *Marketing Decision-Making: A Model-Building Approach.* New York: Harper & Row.

Lipstein, B. (1959). "The Dynamics of Brand Loyalty and Brand Switching". *Proceedings of the 1959 ARF Conference.* New York: Advertising Research Foundation.

Luce, R.D. (1959), *Individual Choice Behaviour: A Theoretical Analysis.* New York: Wiley.

Lundberg, O. (1940). *On Random Processes and Their Application to Sickness and Accident Statistics.* Upsalla: Almquist and Wiksell.

Maffei, R.B. (1960). "Brand Preference and Simple Markov Processes". *Operations Research, 8,* No. 2, 210–218.

Massy, W.F. and Montgomery, D.G. (1968). "Comments on Ehrenberg's Appraisal of Brand-Switching Models". *J. Marketing Research, 5,* 225–229.

Massy, W.F., Montgomery, D.B. and Morrison, D.G. (1970). *Stochastic Models of Buyer Behaviour.* Boston: M.I.T. Press.

McConnell, J.D. (1968). "The Development of Brand Loyalty: An Experimental Study". *J. Marketing Res. 5,* 13–19.

McPhee, W.N. (1963). *Formal Theories of Mass Behaviour.* Glencoe: Free Press.

Montgomery, D.B. and Ryan, A.B. (1974), "Stochastic Models of Consumer Choice Behavior". In *Consumer Behavior: Theoretical Foundations* (T.S. Robertson and J. Scott Ward, eds).

Montgomery, D.B. and Urban, B.G.L. (1969). *Management Science in Marketing.* Englewood Cliffs: Prentice Hall.

Montgomery, D.B. and Urban, B.G.L. (1970). *Application of Management Science in Marketing.* Englewood Cliffs: Prentice Hall.

Morrison, D.G. (1969). "Conditional Trend Analysis: A Model that Allows for Non-Users". *J. Marketing Res. 6,* 342–345.

Moser, C.A., and Kalton, G. (1971). *Survey Methods in Social Investigation.* London: Heinemann.

Mosimann, J.E. (1962), "On the Compound Multinomial Distribution, the Multivariate B-distribution, and Correlations among Proportions". *Biometrika, 49,* 65–82.

Mosimann, J.E. (1984) "Size and Shape Analysis". in *Encyclopaedia of the Statistical Sciences* (S. Kotz and N.L. Johnson, eds). New York: Wiley.

Motes, W., Castleberry, S.B. and Motes, S.G. (1984), "A Longitudinal Test of Price Effects on Brand Choice Behavior. *J. Bus. Res., 53,* 1–11.

Motes, W. and Woodside, A.G. (1984), "Field test of package advertising effects on brand choice behavior". *J. Advert. Res., 24,* 39–45.

Nelson, J.F. (1986). *GAPM.* Niverville, NY: Recurrent Statistics.

Nicosia, F. (1966). *Consumer Decision Processes.* Englewood Cliffs: Prentice Hall.

Parfitt, J.H. and Collins, B.J.K. (1968). "The Use of Consumer Panels for Brand-Share Prediction". *J. Marketing Res. 51,* 131–146.

Patil, G.P. (1962). "Some Methods of Estimation for the Logarithmic Series Distribution". *Biometrics, 18,* 68–75.

Patil, G.P. and Bildikar, S. (1964). "The Multivariate Logarithmic Distribution as a Probability Model in Population and Community Ecology and Some of its Statistical Properties". *Annual Meeting A.S.A.,* Chicago, December 1964.

Patil, G.P. and Joshi, S.W. (1968). *A Dictionary and Bibliography of Discrete Distributions.* Edinburgh: Oliver and Boyd.

Patil, G.P., Kamat, A.R. and Wani, J.K. (1964). *Certain Studies on the Structure of the Logarithmic Series Distribution and Related Tables.* Ohio: Aerospace Research Laboratories.

Paull, A.E. (1978), "A generalised Poisson Model for Consumer Purchase Panel Data Analysis". *J. Amer. Statist. Ass., 73,* 706–713.

Pearson, K. (1897), "Mathematical Contributions to the Theory of Evolution: on a Form of Spurious Correlation which may Arise when Indices are Used in the Measurements of Organs". *Proc. Roy. Soc., 60,* 489–498.

Powell, N. and Westwood, J. (1978), "Buyer-behaviour in management education". *Appl. Statist., 27,* 69–72.

Pyatt, G.F. (1969). "A Model of Brand-Loyalties". In: *Proceedings of the 1968 CEIR/SCICON Symposium on Model-Building in Business and Government.* (M.G. Kendall, ed.) London: Griffin.

Pymont, B. (1969). "The RBL Mini-Test Market". *Read at the 1969 Conference of the Marketing Research Society,* Brighton.

Pymont, B. (1970). "The Development and Application of a New Micro Market Testing Technique". Read at the 23rd ESOMAR Conference, Barcelona.

Quenouille, M.H. (1949). "A Relation between the Logarithmic Poisson and Negative Binomial Series". *Biometrics, 5,* 162–164.

Robertson, T.S. (1970). *Consumer Behavior.* Glenview: Scott, Foreman.

Rossiter, J.R. (1987). "Comments on Consumer Beliefs and Brand Usage and on Ehrenberg's ATR Model", *J. Market Research Soc. 29,* 83–88.

Rust, R.T. (1986). *Advertising Media Models,* Lexington, Mass: Lexington Books.

Sheth, J.N. (1967). "A Review of Buyer Behaviour". *Management Science B., 13,* 718–756.

Shoemaker, R.W., Staelin, R., Kadane, J.B. and Shoaf, F.R. (1977). "Relation of Brand-choice to Purchase Frequency". *J. Marketing Res., 14,* 458–468.

Sichel, H.S. (1982). "Repeat-buying and the Poisson-generalised Inverse Gaussian Distributions." *Appl. Statist., 31,* 193–204.

Shuchman, A. (1968). "Are There Laws of Consumer Behaviour?" *J. Advert. Research, 8,* 19–28.

Simon, L.S. and Freimer, M. (1970). *Analytical Marketing.* Harcourt, Brace and World.

S.R.S. (1965). *The S.R.S. Motorists Panel.* London: Sales Research Services.

Sudman, S. (1964). "On the Accuracy of Recording of Consumer Panels. Parts I and II". *J. Marketing Res. 1*(2), 14–20, and *1*(3), 69–83.

Tucker, W.T. (1964). "The Development of Brand Loyalty". *J. Marketing Res. 1,* 32–35.

Uncles, M.D. (1987). Using Consumer Panels to Model Travel Choices. In: *Contemporary Developments* in Quantitative Geography (ed. Hauer, J., Timmermans, H., and Wrigley, N.) Dordrecht: D. Reidel.

Uncles, M.D. and Ehrenberg, A.S.C. (1987a). "Patterns of Store Choice: New Evidence from the USA". In: *Store Choice, Store Location and Market Analysis* (ed. Wrigley N.) London: Routledge & Kegan Paul.

Uncles, M.D. and Ehrenberg, A.S.C. (1987b). "Understanding Buying Patterns for Brand Management", London Business School: CMaC Working Paper.

Walters, C.G. and Paul, G.W. (1970). *Consumer Behavior: An Integrated Framework.* Homewood: Irwin.

Wellan, D.M. and Ehrenberg, A.S.C. (1987a). "A Successful New Brand: Shield in the UK". London Business School: CMaC Working Paper.

Wellan, D.M. and Ehrenberg, A.S.C. (1987b). "Seasonal Segmentation: A Case Study", London Business School: CMaC Working Paper.

Wilks, S.S. (1962). *Mathematical Statistics.* New York: Wiley.

Williamson, E. and Bretherton, M.H. (1964). "Tables of the Logarithmic Series Distribution". *Am. Math. Statist., 15,* 284–297.

Wrigley, N. (1980). "An approach to the modelling of Shop-choice patterns". In *Geography and the Urban Environment. Vol III.* (D.T. Herbert and R.J. Johnston, eds). London and New York: Wiley.

Wrigley, N. and Dunn, R. (1984a&b) "Stochastic Panel-data Models of Urban Shopping Behaviour: 1. Purchasing at Individual Stores in a Single City, & 2. Multistore Purchasing Patterns and the Dirichlet Model". *Environment and Planning A, 16,* 629–650 & 759–778.

AUTHOR INDEX

SUBJECT INDEX*

*Very commonly occurring concepts are not indexed. Examples are Analysis period, Brand-loyalty, Buyer behaviour, Consumers, Generalization, Goods, Repeat-buying, Stationary situations. In appropriate cases, a reference to the basic definition is given (e.g. Penetration).