CHAPTER 5

SOME PRACTICAL APPLICATIONS

5.1. Theory and Applications

In this and the next chapter, we illustrate several practical uses of the NBD/LSD repeat-buying theory. They are drawn from the increasing variety of applications that have been developed in the last few years.

The first application deals with the simple question of whether repeat-buying patterns in one country are like those in another. The second application probes possible extensions of the results for frequently-bought goods to semi-durable goods such as clothing; included here are also major problems with the available data.

Various cases of below-normal levels of repeat-buying are examined in §§ 5.4 to 5.6. They relate to a new brand, to market segmentation, and to price-cutting.

These rather specific applications follow on from the more general application of the theory already illustrated in Chapter 3 of first of all understanding the general repeat-buying structure of any given product-field.

5.2. A Comparison of American and British Repeat-Buying Habits

The NBD/LSD theory was initially developed from British data. One task was therefore to see what repeat-buying was like elsewhere, and in particular perhaps in the U.S.

The existence of an empirically validated model was crucial to this purpose. Thus only after repeat-buying patterns in the U.K. had been found to be of the same general form for different brands, products, time-periods, and so on, was it plausible to suppose that repeat-buying elsewhere might take a sufficiently simple and general form to allow one to ask what "it" was like.

Furthermore, it was unnecessary to try and match specific American conditions to comparable British ones (e.g. types of product, frequency of purchase, length of analysis period, etc.). Instead, the easiest way to learn what American repeat-buying habits were like was to check whether they obeyed the British laws, i.e. the NBD/LSD theory as such. This simple use of a general model is one of the most basic forms of

practical pay-off to be gained from successful theoretical work. It is much more efficient than reference back to raw data or the use of controlled experimentation.

The American data analysed here come from George Brown's [1952/3] pioneering studies of brand-loyalty. They consist of certain small sub-samples from the *Chicago Tribune* Consumer Panel in 195 1 which Brown published in full, thus facilitating re-analysis. Four of the product-fields covered by Brown are reported on here, namely margarine, detergents, flour and regular coffee (selected as being least fragmented – bearing in mind Brown's small sample sizes – and non-seasonal). In these, 19 cases were identified (as listed in Table 5.1) which satisfied two criteria: first, the item in question was bought by at least ten households in each of two successive quarters of *195* 1 (10 was an arbitrary lower limit of sample size); second, the number of buyers and their average rate of buying were more or less steady or "stationary" over the two quarters.

To introduce the detailed findings, a typical example is given in Table 5.2. It is for Hills Bros. 1 lb. pack of regular coffee in Quarters I and II of 195 1 (taken from halfway down the list of 19 cases in Table **5.1).** In Brown's sample (exactly 100 buyers of coffee), there were just over 20 buyers of Hills Bros. 1 lb. pack in each of the two quarters. They bought nearly 60 packs, at an average frequency of 2.4 or 2.7 packs per buyer, as shown in Table 5.2*. The penetration and purchase frequency therefore showed little change from quarter to quarter.

The average quarterly buying rate of 2.55 packs per buyer for *both* quarters was used as input (averaged across the two quarters because of the small sample sizes). The NBD/LSD theory then leads to the following predictions about repeat-buying,from quarter to quarter.

The number of repeat-buyers: There should have been 14 households who bought Hills Bros. 1 lb. size both in Quarter I and in Quarter II.

The average rate of buying per repeat-buyer: The households who bought the 1 lb. size in both Quarters I and II should on average have bought about 3.1 packs per quarter.

The proportion of total sales accounted for by repeat-buyers: In each quarter, these repeat-buyers should have accounted for 81% of the total purchases of the Hills Bros. 1 lb. size.

The average rate of buying by "new" or "lapsed" buyers: The households who bought the Hills Bros. 1 lb. size *either* in Quarter II only *or*

* The analysis reported here was carried out in terms of "number of packs bought" rather than "purchase occasions" (see discussion in § 1.4 of Chapter 1 and § 4.1 of Chapter 4).

Table 5.1. Estimated and Observed Repeat-Buying Statistics for 19 Near-Stationary U.S. Cases in 1951

(Theoretical Norms "T" and Observed Values "0")

Brand	Size	Product	Quarters	The given quarterly data: Number of buyers 1st Q	2nd Q	Average packs bought per buyer 1st Q	2nd Q	Repeat-buying statistics: Number of repeat-buyers T	0	Average packs per repeat-buyer T	0	Percent sales by repeat-buyers T	0	Average packs per 1-Q-only buyer T	0
Am. Fam. Flakes	20 oz.	Soaps/Suds.	I & II	21	21	4.3	5.0	16	14	5.6	5.6	92	81**	1.4	2.6**
All Sweet	½ lb.	Margarine	III & IV	47	53	4.6	3.9	37	36	5.1	5.2	91	90	1.4	1.6
Am. Fam. Flakes	20 oz.	Soaps/Suds.	III & IV	21	18	3.8	4.6	15	14	5.1	5.2	91	91	1.4	1.4
Hills Bros.	2 lb.	Coffee	II & III	12	12	3.3	3.3	8	8	4.0	4.3	87	89	1.4	1.3
Hills Bros.	2 lb.	Coffee	III & IV	12	15	3.3	3.2	9	9	3.9	4.3	87	90	1.4	1.5
Nutley	½ lb.	Margarine	I & II	18	17	2.9	3.0	12	12	3.6	3.7	85	85	1.4	1.4
Tide	52 oz.	Soaps/Suds.	III & IV	13	11	3.0	2.9	8	7	3.6	3.7	85	77	1.4	1.6
Hills Bros.	1 lb.	Coffee	I & II	24	21	2.4	2.7	14	15	3.1	3.2	81	85	1.4	1.2
Chase & Sanborn	1 lb.	Coffee	III & IV	14	16	2.4	2.6	10	9	3.0	3.4	81	82	1.4	1.2
Parkay	½ lb.	Margarine	III & IV	26	30	2.5	2.3	17	17	2.9	3.0	78	76	1.4	1.5
Eight O'clock	1 lb.	Coffee	III & IV	17	18	2.1	2.6	11	10	2.8	3.2	77	78	1.4	1.2
Eight O'Clock	1 lb.	Coffee	II & III	18	17	2.4	2.1	10	9	3.1	2.8	76	70	1.4	1.4
Tide	52 oz.	Soaps	I & II	11	10	1.7	2.2	6	7	2.3	2.4	69	80	1.3	1.2
Pillsbury	5 lb.	Flour	III & IV	25	30	1.8	2.2	16	14	2.4	2.8	71	69	1.3	1.3
Pillsbury	5 lb.	Flour	II & III	26	25	1.7	1.8	13	15	2.1	2.1	66	68	1.3	1.4
Pillsbury	2 lb.	Flour	III & IV	14	12	1.9	1.7	7	7	2.1	1.8	64	59	1.3	1.4
Gold Medal	8 lb.	Flour	III & IV	18	18	1.4	1.4	7	10	1.7	1.6	49	62	1.2	1.2
Gold Medal	5 lb.	Flour	II & III	17	18	1.2	1.4	6	9	1.3	1.5	41	56	1.2	1.2
Average				20	21	2.7	2.8	13	13	3.3	3.4	77	78	1.4	1.4
Average discrepancy (1st Q–2nd Q)* or (T–O)*				2		0.3		1		0.2		5		0.2	

* Ignoring sign.

** See text.

Table 5.2. The Observed Quarterly Data for Hills Bros. Coffee 1 lb. Pack in Chicago. Quarters I and II, 1951

	Q.I	Q.II
Number of Buyers	24	21
Number of Packs bought	58	57
Packs per Buyer	2.4	2.7

Table 5.2a. The Estimated and Observed Repeat-Buying from Quarter I to II

(Theoretical Norms "T" and Observed Values "O")

	T	0
Number of Repeat-Buyers	14	15
Packs per Repeat-Buyer	3.1	3.2
Total Sales accounted for by Repeat-Buyers	81%	85%
Packs per "New" or "Lapsed" Buyer	1.4	1.2

in Quarter I only should on average have bought about 1.4 packs in the quarter in question.

These predicted figures are compared in Table 5.2a with the repeat-buying results tabulated directly from Brown's data for the two quarters. Clearly, the agreement is good. Thus the observed rate of buying per household who bought Hills Bros. 1 lb. size in both quarters was 3.2 packs, compared with the NBD/LSD estimate of about 3.1 packs. Again, there were 15 actual repeat-buyers in Brown's sample, compared with the estimated 14, and they accounted for 85% of the total purchases compared with the estimate of 8 1%. And so on.

The corresponding results for all of the 19 near-stationary cases are set out in Table 5.1. There is virtually no difference between the *average* observed and predicted values, and the discrepancies for the individual cases tend also to be small. For example, the number of repeat-buyers ranges from 36 to 7 sample households and is given by the theory to within about one household. The number of packs bought per repeat-buyer ranges from 5.6 down to 1.5 for the different cases and these figures are estimated to within an average of about 0.2 of a pack. And so on.

Some of the largest numerical departures from the theoretical results are the differences of 13 and 15 percentage points for the proportion of

total sales accounted for by repeat-buyers in the last two entries in Table 5.1 (Gold Medal Flour, the 5 lb. pack). These discrepancies are essentially caused by there having been three more repeat-buyers than expected (a 15% difference for a total sample of about 20 buyers) and are in fact not statistically significant.

A more interesting discrepancy occurs in Quarters I and II for the 20 ounce pack-sizes of American Family Flakes in the "Soaps and Suds-sers" field (the first item in Table 5.1). The observed percentage of sales accounted for by repeat-buyers is 11 points lower than would be normal, and the average rates of buying by "lapsed" and "new" buyers (2.9 and 2.4 – represented by their average of 2.6 in the table) are almost double the normal level of 1.4, a discrepancy which is well outside the limits of the other cases in Table 5.1 (or for stationary cases more generally).

Since the number of repeat-buyers and their average rate of buying are near the theoretical norms, it is the observed rate of buying by onequarter-only buyers that is abnormal. Further analysis showed additional peculiarities in the data. *All* the "lapsed" and "new" buyers bought more than one pack, and only 5 of the 14 repeat-buyers bought just a single pack in one quarter or the other – a very unusual distribution. There was in fact a marked tendency in Quarters I and II for buyers of this brand to have purchased an *even* number of packs.

In Quarters III and IV (the third entry in Table 5. 1), the average buying frequency of the "new" and "lapsed" buyers of American Family Flakes was however back to normal (i.e. 1.4), and there was also no longer any clustering of "even" purchases. This suggests that in Quarters I and II there may have been something like a manufacturer's offer of two 20-ounce packs physically "banded" together or sold at a reduced price. It would follow that this Quarters I and II discrepancy – about the only sizeable deviation from the theoretical norms that there was – was of a kind which would also have shown up in British data under similar marketing conditions*.

The conclusion from this study was therefore that American repeat-buying habits and British ones were the same. More recent U.S. data from the MRCA panel have widely confirmed this conclusion.

* If *purchase occasions* had been used as the analysis unit, the discrepancy for American Family Flakes in Quarters I and II would not have shown up in repeat-buying terms, but only as a special quirk in the average number of packs bought per purchase occasion. (We also note that American Family Flakes was generally marketed in a somewhat unusual way and was withdrawn from the market a few years later, but the Quarters III and IV results indicate that there was nothing *instrinsically* different about its repeat-buying patterns at the time.)

5.3. Repeat-Buying of Clothing

In the second case-history we move from "non-durable" branded food and other household products to clothing, i.e. "semi-durable" items most of which tend to be bought less frequently. We also move from full-scale panel data to an extremely limited set of data which had been obtained from single-interview sample surveys. This brings us up against major data problems. The study summarised here was an exploratory one*.

Prior to the study, Kemsley [1965] had already briefly reported that the frequency-distribution of purchases of clothing items in a given time-period tended to follow the Negative Binomial Distribution or NBD form. He had also mentioned certain measurement biases and other discrepancy problems in his data. The general tendency for the theoretical NBD-distribution to fit is confirmed here with data different from Kemsley's, but repeat-buying patterns in different time-periods and various discrepancy problems are pursued in more detail.

Most garments are bought relatively infrequently by any one person, at intervals of several weeks, months or even years. However, the main data available for analysis (a sample of 600 housewife interviews from the Continuing Clothing Survey carried out for ICI by Research Services Ltd) cover only very limited time-periods for any one informant's purchases: 2 weeks for stockings and 4 weeks for other garments. This restriction to a short time-span has limited the analysis that was possible and results will be discussed here only for the more frequently-bought items such as stockings, knitting-yarn, and socks, the patterns for the latter being similar to those for items such as panties and briefs and for other still less frequently-bought items.

A second difficulty with the data is that they suffer from measurement biases which occur in certain kinds of survey work: asking informants to recall their purchases over some previous period such as 2 or 4 weeks can lead to decreasing numbers of purchasing claims as the length of recall-period increases. The repeat-buying claims analysed here therefore represent a mixture of purchasing behaviour and measurement error which require separating out.

In discussing the results we consider firstly the extent to which people buy two or more items of the same kind in a single period such as a week. Secondly, we consider repeat-buying between two time-periods such as successive weeks – i.e. the proportion of buyers in one week

* Based on a report for ICI Fibres Ltd.

who buy also in the other week, and how much they buy. This work was undertaken before much was known about repeat-buying patterns for grocery goods in weekly time-periods (see § 4.9 in the previous chapter), but in any case, the question here was simply to establish what the empirical repeat-buying patterns were like. Major questions are of course whether there are any generalisable patterns in these results, and how they are to be interrelated or explained.

Stockings. Starting with the most frequently-bought garment, stockings, Table 5.3 shows the extent to which buyers buy more than 1 pair of stockings per week. Two observed frequency distributions are shown, for the week immediately prior to the interview, and for the week before that*.

In the week immediately prior to the interview, 54% of the 600 women in question had acquired no stockings, 30% had bought 1 pair, 12% 2 pairs, and so on. This pattern is almost exactly in line with the theoretical NBD which can be fitted using the percentage of non-buyers, 54%, and the average rate of buying per buyer, 1.44 pairs, as input. The good degree of fit is clear to the eye and is also summarised by the equality of the standard deviations of the observed and theoretical distributions. The fit for the "last-but-one-week's" data in Table 5.3 is equally good. The numerical differences between the two weeks' results – 46% claiming an average of 1.4 purchases in one week and 35% claiming 1.3 purchases for the earlier week – reflect the length-of-recall period measurement bias in the data**. The fact that the NBD fits despite this bias suggests that the bias is not all concentrated at some particular level of purchasing frequency.

Turning to the repeat-buying between the two weeks, the NBD formulae say that given the observed average rate of buying 1.4 pairs per buyer in the last week before the interview, 50% of these buyers should also have bought stockings in the preceding week and that such "repeat-buyers" should each on average have bought about 1.5 pairs of stockings per week. These theoretical predictions agree quite well with the repeat-buying pattern actually observed for buyers in the week prior to the interview, as is shown in Table 5.4 (57% buying 1.4 pairs on average).

* This was dictated by the form of questioning used in the surveys (essentially of the form "When did you last buy ...?" and "When before that?"), with interviewing carried out on Mondays and Tuesdays and purchase claims for the two preceding calendar weeks being analysed.

** The data are averaged over a number of different interviewing weeks and there was no marked trend in sales during this period.

Table 5.3. Weekly Purchasing Distributions of Stockings

(Observed Values "0" and Theoretical **Norms "T")**

Number of pairs bought	Week before interview: Last 0	Last T	Last-but-one 0	Last-but-one T
	%	%	%	%
0	54	(54)*	65	(65)*
1	30	31	26	26
2	12	11	7	7
3	2	3	1	2
4 +	2	1	1	1
Pairs per buyer	1 . 4	(1.4)*	1 . 3	(1.3)*
Standard deviation	0.9	0.9	0.8	0.8

* Fitted directly from the observed data.

However, when the number of repeat-buyers is expressed as a percentage of the (smaller) number of claimed buyers for the week before that, there is a large discrepancy with the theoretical figures. Thus 73% of the buyers in the last-week-but-one actually claimed also to have bought stockings in the following week, compared with the *theoretical* NBD estimate of only 43%*. The special interest of this observed excess of repeat-buyers is that it occurs only in the earlier week and therefore relates to the basic measurement bias in the data, such as the week-by-week drop of 11 percentage points (from 46% to 35%) in the number of buyers in Table 5.3. The repeat-buying results in Table 5.4 now indicate that this trend was not so much due to people who bought in both weeks but failed to claim the purchases they made in the earlier week, but to people who bought only in the *earlier* week but failed to report this. This possible explanation that *both* discrepancies – the excess of repeat-buyers and the week-by-week trend – are due to the same specific measurement error is something which could be tested experimentally in future work.

A second discrepancy from the general NBD pattern arises when we examine the frequency distribution of claimed purchases in the two weeks combined. This is shown in the "2-weeks" portion of Table 5.5.

* This estimate is based on the last-week-but-one's penetration of 35% and average purchase frequency of 1.3 (see Table 5.3); estimates based on the "last" week's figures would still be 50% repeat-buyers buying 1.5 times.

Table 5.4. Week-by-Week Repeat-Buying of Stockings

	Week before interview			
	Last		Last-but-one	
	0	T	0	T
No. of repeat-buyers / No. of buyers	57%*	50%	73%**	43%
Pairs per repeat-buyer	1.4	1.5	1.2	1.4

*% of buyers who had also bought in week ***before.***
% of buyers who had also bought in week ***after.*

Here the theoretical NBD gives a marked excess of purchases of 1 pair of stockings in the 2 weeks (26% versus the observed 17%) and a deficit of purchases of 2 pairs (14% versus 27%). This is not a question of the 2 pairs of stockings being bought on the same purchase occasion because there is no such "excess" of 2 pairs in the l-week distribution of Table 5.3. Nor can the known week-by-week measurement bias in the data be the whole explanation, since an "excess" of purchases of 2 pairs also occurred in some experimental data (again obtained by Research Services Ltd) covering a 6-week period but using a somewhat different measurement technique (l-week recall with a succession of six weekly interviews on the same informants), as is shown in the "6-weeks" portion of Table 5.5.

Having isolated this abnormally large group of people who regularly make one purchase in successive weeks, more experimental work would be required to explain it further*. Some questions are whether they tend to buy stockings of the same type each week (and whether the delay between the first and second purchase is due to uncertainty about the fit, or lack of spending money, or what), or whether they buy stockings of a different type each week (colour, make, price, etc.) and if so, why. Answers to such questions could affect practical merchandising decisions (such as 2-pair offers, and whether they should be of the same colour etc.).

Knitting-yarn. Next to stockings, hand-knitting-yarn is the most frequently bought textile item. The data cover the 4 weeks prior to each interview in the ICI Survey and are again subject to a length-of-recall-period bias, with fewer purchasing claims being made for earlier weeks than for ones closer to the interview (a trend from 6% buying to 13%).

* Subsequent work on full-scale panel data for grocery products has shown similar spreading out of purchases week by week, as already discussed in § 4.9.

Table 5.5. Purchasing Distributions of Stockings in Periods of 2 and 6 Weeks

(Observed Values "0" and Theoretical Norms **"T"**)

Number of pairs bought	Length of purchasing period			
	2 weeks**		6 weeks** *	
	0	T	0	T
	%	%	%	%
0	45	**(45)***	35	**(35)***
1	17	26	22	26
2	27	**14**	22	17
3	6	7	11	10
4	3	4	5	6
5	1	2	2	3
6	1	**1**	2	2
7 +	—	1	1	1
Average pairs per buyer	2 . 0	**(2.0)***	2 . 3	**(2.3)***
Standard deviation	1.4	1.5	1.8	1.8

***** Used in fitting.
****** Two-week recall.
******* Weekly recall in six successive interviews.

The most striking feature of the knitting-yarn data is that the incidence of repeat-buyers from one week to another is quite exceptionally high compared with the theoretical NBD result. This is illustrated in Table 5.6. Of the people claiming a purchase in the last week before the interview, 37% had also bought knitting-yarn in the preceding week, compared with a theoretical 25%. The results for the two earlier weeks are even more extreme still, in that for the last-but-third week before the interview, the percentage of the (low) number of buyers who also claimed to have bought in the following week is more than 40 percentage points higher than the theoretically expected NBD value (55% versus 12%).

A high incidence of week-by-week repeat-buying relative to the NBD norms might seem in line with what occurs for non-durable household products such as foods in short time-periods (see the "minimum time-period" discussion in Chapters 3 and 4), but it differs radically from what occurs for most other garments (see for example Table 5.9 for socks). The finding can hardly be explained by the measurement biases in the present data (as was the case for stockings in Table 5.4), since it occurs both forward and backward in time. Nor was it due to any sales trends. Instead, it seems to reflect something real for knitting-yarn.

The crucial point seems to be that we are dealing here with purchases of ***yarn,*** and not of completed garments. Thus there is known to be some tendency among consumers to buy the knitting-yarn needed for any garment in two or more instalments. This would explain exceptional high repeat-buying over the relatively short periods of time needed to complete a garment.

Table 5.6. Week-by-Week Repeat-Buying of Knitting Yam

(Observed values "0" and Theoretical Norms **"T")**

	Pairs of Weeks before interview				Pairs of Weeks before interview			
	Last		Last-but-one		Last-but-two		Last-but-three	
	0	T	0	T	0	T	0	T
No. of repeat-buyers / **No.** of buyers	**37%***	25%	**48%****	38%	**44%***	**16%**	**55%****	**12%**
Oz. per repeat-buyer	1.2	1.3	1.2	1.2	1.1	1.2	1.1	**1.1**

*% buyers also buying in the week ***before.***
% buyers also buying in the week ***after.*

Table 5.7. Quasi-Repeat-Buying of **Knitting-Yarn** in Periods of **2-weeks** Lengths

	Pairs of Weeks before interview			
	(Last + first)		(Middle two)	
	0	T	0	T
No. of repeat-buyers / No. of buyers	44%	43%	40%	44%
Oz. per repeat-buyer	1.6	1.6	1.8	1.6

A test of this hypothesis is whether repeat-buying covering time-periods longer than weekly ones comes nearer to the "normal" NBD level under stationary conditions. The scope for such a test is very limited with only 4 weeks' data available per informant, and with the data moreover being subject to a trend due to measurement bias. However, grouping the purchasing claims for the second and third weeks on the one hand, and the purchases claimed for the first and fourth weeks on the other hand, leads to two periods each of 2-week length which are of a fairly stationary form (i.e. with about 15% of the informants

buying in each such "period", at an average rate of roughly 1.5 oz. each) and for which the effect of any week-by-week instalment-buying is largely eliminated. This therefore provides a test, even if a rather artificial one. Table 5.7 shows that the "repeat-buying" patterns for these two 2-week periods in fact agree well with the appropriate theoretical NBD norms. This return to normalcy therefore suggests that for analysis-periods which do not relate to the instalment buying of knitting-yarn for any given garment, the "normal" NBD repeat-buying patterns might occur. But more direct study with better data is clearly needed.

Socks. The buying patterns for men's, boys' and girls' socks are similar, and we need only describe that for men. The pattern differs from the ones described above for stocking and knitting-yarn in two respects, but its major features are common to those of all *other* garments that have been examined (panties, briefs, shirts, blouses, etc.).

The observed weekly and fortnightly frequency distribution of claimed sock purchases are illustrated in Table 5.8. At first sight they seem in fair agreement with the theoretical patterns. However, there is one systematic discrepancy. Thus the observed distributions are in fact bimodal, there being a definite tendency for the number of purchases of 2 pairs of socks to exceed the number of single pairs being bought (e.g. 4% against 3% in the last week before interview). An "excess" of purchases of 2 pairs was, of course, also found for the two-week' and

Table 5.8. Purchasing Distribution of Men's Socks

(Observed Values "O" and Theoretical Norms "T")

Number of pairs bought	Weeks before interview: Last		Last two	
	O	T	O	T
	%	%	%	%
0	92	(92)*	83	(83)*
1	3	4	6	9
2	4	2	8	4
3	1	1	2	2
4 +	1	1	1	2
Pairs per buyer	1.9	(1.9)*	2.0	(2.0)*
Standard deviation	0.6	0.7	0.8	0.8

* Fitted directly from the observed data.

six-week purchasing of stockings (Table 5.5), although not in a single week (Table 5.3). But for socks, the excess *does* occur within a single week, i.e. generally on the same purchasing occasion. It might seem like commonsense that people will tend to buy 2 pairs of socks at a time, but in terms of quantitative detail, the question is whether it is commonsense (or even true) that *more* people should buy 2 pairs than buy 1 pair in a week. (This discrepancy appears to relate to the distinction between the *amounts* bought and the number of *purchasing occasions,* which has already been discussed in earlier chapters.)

Quantitatively more important than this discrepancy is the fact that the *week-by-week* repeat-buying pattern for socks which is shown in Table 5.9 is completely different from that for stockings, for knitting-yarn, and for non-durable household goods in general. Thus for other products which are bought at an average rate of about 2 units per buyer in one period, it has been found (as summarised by the theoretical NBD results) that a little over 55% of buyers in one time-period normally buy again in the adjacent period. For socks there is however an almost complete absence of repeat-buyers week-by-week (or fortnight-by-fortnight). Furthermore, the rare week-by-week repeat-buyers of socks that *do* occur buy at a rate of only 1 pair each, compared with the theoretically expected average rate of about 2.4. These findings are not unique to socks but occur also for garments such as panties and briefs, and for even less frequently-bought items such as knitwear, blouses, dresses, skirts, shirts and ties.

Table 5.9. Week-by-Week Repeat-Buying of Men's Socks

	Weeks before interview			
	Last		Las&but-one	
	0	T	0	T
No. of repeat-buyers / No. of buyers	4%	56%	4%	57%
Pairs per repeat-buyer	1.0	2.4	1.0	2.6

This radical departure from the NBD repeat-buying patterns does not seem to have anything to do with measurement biases. Instead, it fits in with commonsense experience that an item like socks is seldom bought in two successive periods as short as a week or so. The problem of what

repeat-buying of socks is like in *longer* time-periods, and whether or not it then fits in with the general NBD type of patterns, cannot be answered directly because appropiate data are not available.

Relating the low repeat-buying for socks to the general repeat-buying theory discussed in Part II, we note that one of the main empirical findings summarised in this theory is that under stationary equilibrium conditions, repeat-buying follows a single pattern for time-periods of various lengths such as months, quarters or half-years; only the numerical values of the parameters vary with the length of the period. These parameter values are themselves interrelated, which is what gives the NBD theory its power for descriptive, evaluative and forecasting purposes. All this has been found to be true for a wide range of household goods, but generally of course only for time-periods of *at least* a week or so, as discussed in § 4.9 of Chapter 4. There was a minimum time-period effect, but this took the form that the observed repeat-buying from one week to the next was *higher* than the NBD prediction. However, *day-by-day* repeat-buying patterns of most grocery goods, toiletries, etc. are generally quite different from week-by-week or month-by-month ones. Few people who buy detergents, corned beef, or toothpaste on one day do so again the next day. A certain "dead-period" generally intervenes before any repeat-purchasing occurs. Such day-by-day repeat-buying (or rather the lack of it) is a reflection of *short-term* shopping and usage habits and does not relate to the longer-term factors of brand-loyalty, brand-switching etc. which really underlie the structure of a market and the effectiveness of marketing policies.

The indications from the results here are that the same considerations apply to the garment market, except that the "dead-periods" are longer. The very low *week-by-week* repeat-buying found for most garments (socks, etc.) may be very much the same kind of short-term phenomenon as the low *day-by-day* repeat-buying for most non-durable household goods. In contrast, the weekly results for stockings in Table 5.4 – a product like non-durable household goods in average length of usage life – and the (rather artificial) fortnightly results for knitting-yarn in Table 5.7 indicate that it may be possible for repeat-buying levels of textile garments in suitably long time-periods to be of the general NBD kind. It is therefore possible that repeat-buying would once again fit in with the NBD pattern once some *minimum* length of period is exceeded, depending on the product. The further study of purchasing patterns in the textile market and the practical problem-solving application of the results largely depend on the availability of more suitable data.

5.4. Low Repeat-Buying: A New Brand R

The preceding case-history contained several examples where the incidence of repeat-buying was low compared with the theoretical NBD estimates, either because of faults in the data or because the time-periods analysed were short compared with the minimum inter-purchase time for the item in question. We now consider three other cases of low repeat-buying (taken from unpublished reports), and start with the case of a certain new Brand R: repeat-buying of R was low, particularly in non-consecutive time-periods.

For a new brand, the ultimate question is not so much whether people will try it at all or even whether they will buy it a second time. Thus with good product research, promotional support and retail distribution, a new brand should find no great difficulty in getting quite a high proportion of its *po ten tial* users to try it, and even to buy it a second or even a third time. Instead, the problem is whether a sufficient proportion of initial users will go on using the product more or less indefinitely.

In the present case-history, sales of a certain new Brand R had achieved a steady if unexciting level a year or so after its launch: about 4% of the population bought it each quarter, at an average frequency of about 1.4 purchases per buyer in the quarter.

The proportion of buyers of R in one quarter who bought R again in the next quarter was found to be 2 1%. This was relatively low, not only in absolute terms but also compared with a theoretical NBD/LSD norm in the high 30's (given 4% buying 1.4 times each in the first quarter). However, it was known that repeat-buying for new brands may not settle down to the "normal" level for an established brand for as long as a-year or two [e.g. Ehrenberg and Goodhardt 1968d, Ehrenberg 1970b], there still being a high in-and-out flow of first-time triers. Brand R might well have been suffering from this.

The question therefore was whether the 20% quarter-by-quarter repeat-buying did in fact represent a solid core of long-term buyers (with a high turnover of additional new triers), or whether the product was merely sufficiently attractive to be bought just two or three times. This was explored by carrying out quarterly repeat-buying analyses for the new brand in *non-consecutive* quarters (as has already been illustrated for *established* brands in Table 3.9 of Chapter 3).

In the present instance, the analysis showed that there was in fact a serious erosion of the repeat-buying of Brand R in the longer run: of the buyers in one quarter, only some 11% bought again two quarters

later, compared with 21% in the *next* quarter (and compared with a theoretical norm in the high 30's for established brands, whether in consecutive or non-consecutive quarters).

This form of analysis therefore showed that the new brand was failing to build up any sizeable following of buyers who would be loyal to it in the longer-term. The brand's steady sales level was due to its still continuing to attract new first-time buyers, but it was clearly working its way through the whole potential population in this respect, and sales would collapse before long. Management in fact withdrew the brand before this happened.

5.5. A Shortage of Repeat-Buyers or an Excess of Occasional Buyers?

Another case of low repeat-buying occurred in dealing with a certain food-product S. It was found that repeat-buying for virtually every brand was below the expected norm, by something like 7 percentage points on a quarter-byquarter basis (and more half-year by half-year). There was no obvious explanation such as some large sales trend (i.e. non-stationarity) or the like. Repeat-buying for the product-field as a whole was almost exactly on the norm, so that the low repeat-buying for each individual brand was not a case of buyers moving in or out of the market as a whole.

Most brands were sold in three or four different pack-sizes, ranging from "Small" through "Medium" and "Large" to "Giant". The Large and Giant sizes had mostly been launched only two or three years previously, and the marketing management of the client company felt that the relatively low repeat-buying of each brand might be due to an as yet irregular buying pattern for these larger sizes.

Repeat-buying was therefore studied for each individual pack-size of each brand. The results are illustrated in Table 5.10 and showed differences in the repeat-buying level of the pack-sizes, but in precisely the opposite direction to that hypothesised by marketing management: repeat-buying of the Large and Giant sizes was virtually normal (42% versus 38%) and it was that for the two *smaller* sizes that was markedly below the theoretical level (37% versus 52%). Table 5.10 also shows that the discrepancy lay only in the sheer *proportion* of repeat-buyers, and not in their frequency of buying or in that of the "new" buyers in the second quarter (where the observed and theoretical rates still agree closely).

Table 5.10. Quarter-by-Quarter Repeat-Buying for Individual Pack-Sizes of Product S

(Observed Values "0" and Theoretical Values "T")

	% Buyers in one Quarter who buy in next Quarter		Av. frequency of purchase per repeat-buyer		Av. frequency of purchase per "new" buyer	
	0	T	0	T	0	T
Small Size	34	**51**	2.2	2.1	1.3	1.3
Medium Size	40	**52**	2.2	2.1	1.2	1.3
AVERAGE	**37**	52	2.2	2.1	1.2	1.3
Large Size	44	42	1.7	1.6	1.2	1.2
Giant Size	39	33	1.5	1.5	1.2	1.2
AVERAGE	42	38	1.6	1.6	1.2	1.2

The question next arose as to whether the discrepancy for the two smaller pack-sizes signified a real short-fall in repeat-buying or instead perhaps an excess of occasional buyers. The point is that the preceding analysis is *relative*, in that "too many" once-only buyers in one period would show up as an apparent *short-fall* of repeat-buyers in the next.

Table 5.11 illustrates the more detailed analysis required here, namely examining repeat-buying amongst .previous light, medium and heavy buyers (a form of analysis already illustrated in Table 3.10 in Chapter 3). For the Small size (where in Table 5.10 there was a discrepancy of 17 percentage points in the observed and theoretical incidence of repeat-buyers), there was a marked short-fall of repeat-buyers amongst *light* buyers (i.e. ones who had made 1 or 2 purchases in the first quarter), whereas amongst *heavier* buyers (3 or more purchases in the first quarter) the incidence of repeat-buyers was on the mark. For the Medium size there-was similarly a short-fall only amongst the initial once-only buyers.

Table 5.11. Repeat-Buying of the Small Size by Light and Heavier Buyers

		Purchases of the Small Size in One Quarter				
		0	1	2	3	4+
		%	%	%	%	%
% Buying Small size	0	3	27	48	73	89
in the Next Quarter	T	3	39	62	76	88

The apparent short-fall of repeat-buyers in this product-field was therefore tracked down firstly to the two smaller pack-sizes, and secondly even there not to any real failure to attract repeat-buyers, but rather to an excess of occasional buyers. This appeared to be related to very marked price-cutting competition for the smaller sizes (especially amongst Own Label brands) and possibly to patchy retail availability, two factors which led to increased brand- or size-switching on an *occasional* basis and hence an abnormally high number of light buyers.

5.6. Low Repeat-Buying and High Brand-Switching

In a certain non-food product-field M, repeat-buying for each brand was generally at the normal level, as illustrated in Table 5.12 for the four leading brands M_1 to M_4. An exception occurred for two particular brands, called M_5 and M_6 here: quarter-by-quarter repeat-buying was at least 10 percentage points below the expected NBD/LSD norms.

These two brands had a certain product-characteristic in common. This appeared to be exceptionally relevant to consumers, as shown in their brand-switching behaviour. Thus in the kind of brand-duplication and switching analyses that are introduced in Part V (as illustrated for example by Table 9.7 in Chapter 9), there was a very clear clustering of brands M_5 and M_6 : switching between these two brands was at a substantially higher level than would be predicted from the general pattern of switching in this product-field.

Table 5.12. Quarterly Repeat-Buying in Product-Field M

(Observed Values "O" and Theoretical Norms "T")

	% Bought in next Quarter II	
Buyers in QI of:	O	T
Brand M_1 = 100%	70	63
Brand M_2 = 100%	60	58
Brand M_3 = 100%	**51**	**52**
Brand M_4 = 100%	58	**58**
AVERAGE	60	**58**
Brand M_5 = 100%	41	**54**
Brand M_6 = 100%	43	53
AVERAGE	42	53

Table 5.13. Quarter-by-Quarter Repeat-Buying and Brand-Switching between Brands M_5 and M_6

	In Next Quarter					
	% Buying the Same Brand		% Buying the Other Brand		Total: % Buying M_5 and/or M_6*	
Buyers in QI of:	0	T	0	T	0	T
Brand M_5 = 100%	41	**54**	16	**5**	**57**	**59**
Brand M_6 = 100%	43	53	11	4	54	57

* Double-counting 2 or 3% buying both.

This extra switching is shown in Table 5.13: the proportions of buyers of M_5 also buying M_6 or vice versa were almost 10 percentage points higher than the theoretical levels. This excess virtually made up for the short-fall of repeat-buying for each of the two brands, as is shown in the last column of Table 5.13 where the tendency for buyers of one of the two brands to buy *either* that brand or the other one was at virtually the predicted level (e.g. 57% versus 59% for M_5). There was therefore a definite tendency for consumers to treat brands M_5 and M_6 as substitutable for each other – it was more important to buy that particular *type* of brand (with the common product-characteristic) rather than a particular brand-name*.

The analysis therefore served to explain this particular short-fall in repeat-buying in "segmentation" terms, i.e. that there was a sub-group of brands which were similar in product-formulation and appeared to be treated as such by the consumer. Such special kinds of purchasing pattern however occur very rarely – repeat-buying of most brands follows the normal pattern, *without* having to take into account what other brands are or are not bought as well.

5.7. Summary

In this chapter, the NBD/LSD theory has been used to examine repeat-buying patterns under various previously unexplored conditions, such as for a different country, for a different type of product-class, and for data subject to major errors of measurement. This essentially involves comparisons of new kinds of data with previous results.

* This special degree of switching between the two brands may have been caused by uneven distribution and occasional lack of availability of one or the other brand, rather than by any positive breakdown of the "normal" intensity of repeat-buying for each brand. This possibility is so far unexplored in this particular market.

Such comparisons are facilitated by the use of a theory which effectively summarises previous empirical results and which therefore eliminates in these cases the need for controlled experimentation or for other matching procedures, or indeed the need for any direct recourse to the previous data in "raw" form.

The findings discussed show that repeat-buying in the U.S. takes the same form as in the U.K., and that the same repeat-buying patterns *may* also hold for semi-durable products like clothing in time-periods that are long enough to exclude the "dead-period" between one purchase and another (but further work on better data is needed here).

The analyses also show that usable information can sometimes be extracted from data which are subject to major measurement biases. Cases of abnormally low repeat-buying can also be tracked down to "real" factors such as the failure of a newly-launched brand to build up a repeat-buying franchise, an excess of occasional once-only buyers (due to price-cutting or patchy retail availability), or abnormally high brand-switching between brands which share a particular product-characteristic.

CHAPTER 6

FURTHER APPLICATIONS

6.1. Non-Stationary Situations and Technical Applications

Very little is as yet known about how or why consumers change their behaviour, and there is no valid theory of consumer dynamics. The theory of *stationary* buyer behaviour can however be of help in interpreting any specific dynamic, non-stationary situation. An extended example is given in § 6.2 below, namely the evaluation of the effect of direct marketing action.

Another application to a non-stationary situation is given in § 6.3, where we consider a new product launch. This differs from the case in § 5.4 of the last chapter partly in that it literally was a new product (and not merely a new brand in an existing product-class), partly in that the performance of the new product could only be assessed by using the NBD/LSD theory to expand data which were extremely limited in their scope, partly that the new product was in test market, and partly in that the new product was ultimately successful.

The remaining applications of the NBD/LSD theory discussed in this chapter centre on statistical problems. An evaluation of the effect of n-regular reporting by certain members of a consumer panel is given in $6.4, and some quick methods of assessing standard errors and the significance of extreme values (heavy buyers) are outlined in § 6.5.

6.2. Evaluating a Consumer Promotion

A major area of applying the theory of stationary repeat-buying is in cases where we know beforehand that the theory will not hold, i.e. cases where there *is* a trend in sales and we wish to examine the *nature* of the trend. Was the trend due to an unusual influx of "new" buyers, to holding on to an above-normal number of previous buyers, to an increase in *rates* of buying, or what?

A brief example of such an analysis was already given in § 2.4 of Chapter 2 (by way of evaluating the nature of a seasonal trend). We now give another case-history, where it is a matter of evaluating a deliberate marketing action, namely a consumer promotion.

The effects of many such marketing actions can now be evaluated in some detail. The basic question is not whether there was any positive effect at all – that is usually obvious from the level of total sales. Instead, the question is what *kind* of effect it was (e.g. extra penetration, or heavier purchasing from existing customers, or both).

The need is to compare what actually happened with what would have happened without the marketing action. One possible approach is through controlled experimentation. Thus the promotion (or other marketing treatment, such as a price-, product-, or packaging-change) can be run in certain regions of the country, and compared with "control" regions, the whole thing being subjected to the principles and procedures of statistical experimentation. But imposing effective experimental control in marketing is usually expensive, often difficult (e.g. due to competitors' activities or other variable factors), and sometimes altogether impossible (e.g. in assessing a seasonal trend). Furthermore, controlled experiments have to be planned in advance. An alternative approach is to use the theory of stationary buyer behaviour to provide the control data, i.e. to predict what would have happened without the trend.

The case described here is one where a controlled experiment could in fact have been mounted, if it had been planned in time. A special lesson to be drawn from the theory-based approach here is therefore that it can be initiated *after* the event (and tends also to be much cheaper than running controlled experiments).

A giant pack and seven marketing questions. The promotion or "deal" to be evaluated here was for the giant pack-size 'G' of a certain brand of a grocery-product*. The dealconsisted of a small pack of the brand being banded to the giant pack 'G'. These banded-packs were on offer during February of a certain year. Total sales of the giant pack that month went up by 25%. The problem was to understand *how* this sales increase had come about – more buyers or more per buyer? – and in particular to answer seven specific marketing questions:

(i) Had the deal increased the repeat-buying by previous buyers of G?

(ii) If so, had it appealed more to previous *heavy* buyers or to previous *light* buyers of G?

(iii) Had the banded-pack deal attracted more "new" buyers (i.e. people who had *not* been buying G recently) than would have occurred that month anyway?

* Based on a study for the J. Walter Thompson Company.

(iv) How many of the deal-packs were bought as substitutes for purchases of G that would have been made in any case?

(v) Was buying of the banded-pack concentrated among some "deal-prone" minority of consumers?

(vi) What was the *longer-term* effect of the deal on consumer loyalty to G?

(vii) What was the effect on *other* brands and pack-sizes (or from what other brands or pack-sizes were the extra sales gained)?

Answering such questions is becoming increasingly recognised as being of real practical (and theoretical) interest. Only by understanding what kinds of deals have what kinds of effects does *informed* guidance to marketing management become available.

Deal versus *no-deal.* Answering the above questions is basically very simple. It depends on comparing what *did* happen with what would have happened if there had been no banded-pack deal or other such cause of the sales trend. We therefore have to predict what the detailed patterns of buyer behaviour would have been if there had been no trend in sales.

We first note that in the four weeks in January just before the deal, 11% of all households bought the giant pack-size G, at an average frequency of 1.9 purchases. The NBD prediction from these two values is that in the absence of a sales trend from January to February, some 58% of the January buyers would have bought G again in the next four weeks (i.e. in February), at an average frequency of 2.3 purchases each. The corresponding prediction for previous *non-buyers* of G in January is that 5% of them would have bought it in February (i.e. been "new" buyers in the sense used here).

The actual figures observed in the February deal-period are that 14% of all households bought G, on average 1.9 times each. This compares with the pre-deal penetration of 11% in January, as shown in the right-hand column of Table 6.1. (In the absence of a trend, 11% of households would also have bought in February – again 1.9 times each – as denoted therefore by the theoretical norm "T".)

Table 6.1 also shows how the February penetration breaks down by previous buyers and non-buyers. Thus of the non-buyers of G in the previous month, 7% bought it during the deal month, But by no means all of these were attracted by the deal as such – there would have been 5% "new" buyers anyway and the *extra* new buyers amounted to only 2%. Given the large number of non-buyers in January (89% of the

Table 6.1. Observed Buying of G in February Compared with the Predicted 'No-Trend' Norms (Observed Values "0" and Theoretical Estimates "T")

	Buyers of G in January (11%)		Non-Buyers of G in January (89%)		AU Housholds (100%)	
	O	T	O	T	O	T
	%	%	%	%	%	%
Bought G in February	72	58	7	5	14	11
Purchases per Buyer	2.1	2.3	1.5	1.4	1.9	1.9

population), this incidence of *extra* new buyers in February however accounts for over half the 3% increase in the total penetration of G in February (14% versus 11%).

The remainder is accounted for by increased repeat-buying. Thus 72% of January buyers also bought G in February. The deal therefore attracted markedly more repeat-buyers than would have occurred without it (58%). That is the effect of the deal on ***repeat-buying.***

The evaluation is taken further in Table 6.2, to see whether the increased repeat-buyers was located more among the lighter or among the heavier buyers of G. The table shows that the proportion of people who had made *one* purchase of G in January and who then bought G again during the February deal-period was well above normal (63%

Table 6.2. Observed and Predicted Buying in February by Previous Light, Medium and Heavier Buyers

	Households who in January had bought G					
	Once		Twice		Three +	
	0	T	0	T	0	T
	%	%	%	%	%	%
Bought G in February	63	43	68	66	93	80
Purchases per Buyer	1.8	1.7	2.0	2.1	2.6	3.2

versus 43%). In contrast, the proportion of repeat-buyers among households who had previously made *two* purchases was almost exactly at the normal level (68% versus the normal 66%). For still heavier buyers – i.e. the buyers who had made three or more purchases in January – the differences from the predicted values (slightly more repeat-buyers buying slightly less frequently) were statistically insignificant on the small subsample of about 25 buyers in question*, and did not in fact contribute to the overall increase in the level of sales (i.e. 93% X 2.6 purchases observed versus 80% X 3.2 purchases predicted).

The results therefore add up to a rather consistent picture:

(a) Something like half the sales increase came from attracting *extra* buyers from amongst households who had not bought G in the previous month.

(b) The remaining extra sales came from an increase in the number of repeat-buyers amongst those households who had been *light* buyers of G in the previous month.

(c) Repeat-buying among previous *heavier* buyers of G remained virtually unaffected by the deal.

(d) The average number of purchases per buyer in each sub-group (and overall) was largely unaffected.

Substitution or Extra Sales? A further step in the analysis was to see where the deal-packs actually went, i.e. how the buying behaviour of individual households was affected by whether they bought a deal-pack or not. Direct tabulation showed that heavy buyers of G in January (i.e. households who then made two or more purchases) made a third of all the deal-pack purchases which were recorded in February. The previous analysis in Table 6.2 however showed that these heavier buyers were virtually unaffected by the deal. It follows that they bought these deal-packs merely as substitutes for regular packs which they would have bought anyway.

Similar calculations for the January once-only and non-buyers showed that almost a further third of the deal-packs (31%) were bought as substitutes, by either new buyers or by previous light buyers, as is shown in Table 6.3. Indeed, only just over a third of the deal-packs went directly to creating extra sales. This occurred by way of the

* The total sample was about 1000 households of whom 11%, or just over 100, bought G at all in January, on average about 1.9 times (almost 60 households buying once, 20 buying twice, and about 25 buying 3 or more times),

Table *6.3.* Purchases of the Banded-Pack in February either as a Substitute for the Regular Pack of G, or as an Extra

(Breakdown by Non-Buyers, Light and Heavier Buyers of G in January)

February Purchases of the Banded Pack:	Buyers of G in January:			
	None	Once	Twice +	Total
As substitute	21%	10%	33%	64%
As extra	21%	15%	0%	36%
Total Purchases of Banded Pack	42%	25%	33%	100%

increased number of "new" buyers and the increased repeat-buying among previous light buyers.

It follows from this high degree of substitute-buying that the deal-packs did not all go to some "deal-prone" minority. They were spread rather widely amongst households whose aggregate purchasing of G in the short-term was in many cases not affected.

The Longer-Term Effect for G. Next, we consider the longer-term effect of the deal. Here one needs to evaluate the loyalty patterns in subsequent time-periods.

The first period to consider was March, i.e. the four weeks immediately following the peak of the banded-pack deal in February. Overall, the incidence of buyers of G in March was 12% and had therefore returned almost to the pre-deal level of January (when 11% of housewives had bought G). In terms of its direct effect on sales, the deal was therefore short-lived (as is in fact general experience with promotions).

The remaining question was in what ways, if any, the underlying pattern of brand-loyalty to G in March was different from what it would have been without the deal. For example, were any of the additional buyers who had been attracted in the deal-period still being retained? Or had the increased buying in February perhaps been a matter of "stocking-up", leading now to ***reduced*** repeat-buying in the subsequent period?

Table 6.4 shows that by March, the households who had bought G in the January period were in fact unaffected one way or the other by the deal – 58% of them bought G again in March, at an average of 2.3 packs each, precisely as predicted under no-trend conditions. Only for

the previous non-buyers – the market segment which had been most affected by the deal in February itself – was there still some buoyancy (4% buying as against the expected 5%).

Table 6.4. Observed and Predicted Buying of G in March, for Buyers and Non-Buyers of G in January

	Buyers of G in January (11%)		Non-Buyers of G in January (89%)		All Households (100%)	
	O	T	0	T	0	T
	%	%	%	%	%	%
Bought G in March	58	58	6	5	12	11
Purchases per Buyer	2.3	2.3	1.6	1.4	2.0	1.9

The banded-pack deal had therefore little after-effect on the giant pack-size. As soon as the deal was over, repeat-buying behaviour as well as most of the in- and out-flow of "new" buyers of G, were virtually as if the deal had never occurred. Similar analyses in subsequent months also failed to pinpoint any more delayed effects for G.

The Effect on the Rest of the Market. A longer-term change in loyalty pattern was however demonstrated in the purchasing of the regular version of the *small* pack-size of the brand (i.e. the pack-size of which free samples were banded to the giant pack). In February, the time of the banded-pack deal for the giant pack, repeat-buying of the regular small pack became markedly depressed, from the normal level of 7 1% to only 43%. This loss of repeat-buyers lasted for all the subsequent periods which were covered by the data available in this case (four months from March to June). This detrimental and longer-term effect on the small pack-size illustrates the beginnings of an answer to the last two of the seven marketing questions posed initially, namely assessing *longer-term* effects and assessing the effect on *other* brands or pack-sizes*.

* At the time of this study, there was as yet no theory of brand-switching to provide norms for evaluating where these "lost" buyers of the small pack-size went, or where the extra buyers of the Giant size had come from. (Early stages of such a theory are now described in Part V of this book; a more developed version in Chapter 13; and a small-scale application has already been given in § 5.6 of Chapter 5).

"Once upon a time ...". To summarise, this case-history has shown how a specific marketing action can be evaluated without having to mount expensive controlled experimentation. All that is needed is empirically well-founded theory to tell us what would have happened in the ***absence*** of the marketing action. Such analyses not only isolate those aspects of buyer behaviour which were affected, but they also automatically show up those aspects which were ***not*** affected *. This makes for great simplicity in using the results of such evaluations.

Increased understanding of marketing phenomena is thus possible. However, *isolated* case-history evaluations of marketing action can provide no firm predictions for the future. Do banded-pack offers always affect new and light buyers, and not heavy buyers? Does the direct effect never last? Is the effect on the other pack-size always detrimental? More broadly, what kinds of consumer deals, and what kinds of marketing action generally, have what kinds of effect?

To provide a firmer basis for prediction, a variety of cases need to be evaluated for a wide range of different products and different marketing conditions. Generalisable laws governing their effectiveness can then be established (including the possibility that there is ***no*** generalisable effect). Much undigested data for such work already exists in market research files. The present example shows that some of the necessery analytic techniques exist as well.

The Effect of What? One basic query which remains is whether comparing the observed patterns in February with the repeat-buying predictions based on the January data really amounts to an evaluation of the consumer deal as such. In other words, were there not also ***other*** factors which could, in part or in whole, have caused the sales increase in February? There might have been deals for other brands, price-cuts, changes in advertising, a relaunch, and so on. Or the deal for G might for example have been superimposed on a seasonal trend [as in a case discussed elsewhere – Ehrenberg 1968e, § 9].

Sometimes such interacting activities and trends are difficult or even impossible to disentangle. But in this particular case-history it was clear that there was no seasonal trend, and no changes in ***other*** marketing

* This tendency to separate out "affected" and "unaffected" aspects of buyer behaviour is so far a very general finding. It is typified here by all the cases where the "no-trend" norms fit even under trend conditions, such as the incidence of repeat-buyers among the ***heavier*** buyers, and the average rates of buying per buyer in all the various sub-groups. (See for example also the case in § 2.4 in Chapter 2, the discrepant items in Chapter 3, and the various cases in Chapter 5.)

activities for G which had any net effect just then. Knowledge of general buyer behaviour in this market (i.e. the kinds of analyses illustrated in Chapters 3 and 9), together with cross-checks on the marketing activities and sales trends of the ***other*** brands and pack-sizes in the market, also indicated that the banded-pack deal for G was the only specific new marketing factor in February which could have had a differential effect on the sales of the giant pack-size G.

It was thus judged that the sales increase was due to the deal (and it only remained to understand in detail ***how*** the deal had worked). It would of course have been safer (but less useful) not to make such a judgment. However, irrespective of whether one exercises one's judgment in this way or not, the basic analysis will assess the ***combined*** effect of all the marketing factors that may have been in operation. Whatever the totality of causal factors at work, the analysis pinpoints just in what ways the sales trend which actually occurred in February differed from the usual no-trend situation. We are therefore left to use our general understanding of the market – our *judgment* – in interpreting the findings and in allocating causes to the observed effects. The crucial difference produced by the above analysis is that we really know what these effects were, so that our judgment has something hard to bite on.

6.3. **Expanding Limited Data: A New Product Q in Test Market**

The last case dealt with the interpretation of some given data about a trend. A different form of practical application on the stationary repeat-buying theory is in expanding limited data. This arises when fully continuous panel data are not available, for cost or other reasons.

Data may then be available from single interviews or from some kind of repeated interviewing of the same informants. In general, this will not provide as *reliable* or as *complete* a record of purchases. Assessing the internal reliability of the data and filling in the gaps are two practical uses of the theory which can be illustrated by an example of a new product launch*. The available data on consumer purchasing were not only very sparse, but in parts also unreliable.

The data consisted of how many packs of a certain new product Q informants claimed to have "ever bought", as measured from five separate samples of housewives interviewed at monthly intervals from about

* Based on a report for Cadbury Bros Ltd.

two months to six months after the launch of the product in test market *. Table 4.5 sets out the results. In the first survey (some 2 months after the launch), about 86% of the sample claimed not to have bought the product Q and 10% claimed to have bought it just once, 1.5% claimed to have bought two packs, and so on, making about 14% who claimed to have bought it at all. And so on for the remaining surveys.

Table 6.5. Claimed Purchases (Numbers "Ever Bought") of the New Product Q in 5 Successive Surveys after the Product-Launch

(Percent buying 0 or 1 times rounded to whole figures)

Survey Month	Number of Packs "Ever Bought"						
		0	1	2	3&4	5-7	8+
2	%	86	10	1.5	1.8	.5	.3
3	%	80	12	3.3	3.8	.7	.3
4	%	77	13	3.2	3.2	2.0	1.3
5	%	68	14	6.5	6.2	3.4	2.2
6	%	68	16	4.2	4.8	3.3	3.3

Aggregate data about the sales of Product Q were also available, from the manufacturers' shipment records and from measurements of sales at retail outlets (as carried out by the A.C. Nielsen Co.). These data showed a rather unusual pattern for a new product, namely more or less steady sales during the early months. (The more usual pattern for new products is a fairly steep rise and then a decline and subsequent levelling out [Davis 1964]. The reason for this steady level of sales of Q virtually from the start was probably that demand had outstripped production and supplies to retailers were therefore rationed,

Given this apparently "stationary" state of the sales data it was possible to try and interpret the purchasing data for Q in Table 6.5 by the direct application of the repeat-buying theory [rather than using this as a norm for interpreting trends or for setting "steady-state" targets, as discussed elsewhere – Ehrenberg 1970b].

* Product Q was a distinctive product (essentially a "new-new" product rather than a new brand in an existing market). Asking an "ever-bought" question was therefore feasible during the first year or so of Q's existence. The information gathered about the product in the surveys was essentially limited to this one question, being part of an all-purpose "omnibus" questionnaire.

One outcome of the resultant analysis was that much of the data appeared unbelievable. A general weakness of the data is illustrated in Table 6.6. This shows the observed data for Month 6 together with the theoretical NBD, fitted to the number of zero and 1 unit purchases*.

Table 6.6. The Number of Purchases of Q Claimed in the first 6 months, and the NBD fitted to the 0's and 1's

(Observed Values "0" and Theoretical NBD Norms "T"

	r	Number of Packs "Ever Bought"					
		0	1	2	3&4	5-7	8+
Month 6	O	67.8	16.5	4.2	4.8	3.3	3.3
	T	(67.8)*	(16.5)*	7.3	5.7	2.2	.6

* Used in fitting.

There is clearly a large excess of heavy purchasers (almost 7% claiming 5 or more purchases, against a theoretical estimate of only 3%). The data for the earlier surveys showed the same pattern. It appeared therefore that there was *over-claiming* of heavier purchases, as is not unusual in an interview situation, especially when using rather simple questions?. One way of dealing with the data was to reject the higher purchasing claims (2 or more) and work simply with the numbers of claimed non-buyers and claimed purchases of precisely one unit as input (since two observations are enough for fitting a two-parameter model such as the NBD). It seemed reasonable to suppose that informants could give *fairly* accurate answers to whether they had ever bought the product, and if so, whether they had only bought it once. And in the outcome, this approach gave workable and valid results, as will be seen below.

A second weakness of the observed data recorded in Table 6.5 appeared for Months 5 and 6. Thus the numbers of informants who claimed to have bought at least 2 packs of Q *decreased* by almost 3% from Month 5 to Month 6 (which is strictly impossible for having "ever-bought" something), while the total number of buyers did not change

* Fitting by mean value and number of zeros (as discussed in previous chapters) was not easily possible because the larger purchasing claims were grouped on the questionnaire as 3-4, 5-7, and 8+. It was therefore not possible to calculate an accurate mean value.

† The 3-4, 5-7, and 8+ groupings were already *pre-coded* on the questionnaire and this may have led to increased bias at the interview, especially when probing an initial "Don't Know" response.

(which is unlikely). Something seemed to be wrong and more detailed checks (along the lines of the more constructive analysis below) indicated that the Month 5 data were somehow aberrant.

The input information was therefore reduced to that shown in Table 6.7, i.e. cumulative figures for Months 2, 3, 4 and 6 of how many people had never bought Q and how many had bought it just once. The cumulative penetration was still rising, with 32% of the population having bought Q by Month 6 (100%-67.8%), but half of these buyers having done so only once.

Three major unknowns for these consumer purchasing data were then

(i) the average level of purchases each month,

(ii) the penetration (% buying) in each separate month,

(iii) the proportion of repeat-buyers from month to month,

this being the kind of information needed to make some more informed assessment of the product's acceptability and future potential.

The first figure to estimate was the mean value of the NBD fitted to the *cumulative* figure observed in each month. Estimates of these means are given in the first column of Table 6.8. Thus the average number of purchases of Q after 2 months was .21 packs per housewife in the population, and after 6 months about .68 packs.

These estimates suggest an average purchasing level per month of a little over .1 packs of Q – say .12 – per housewife (and about .1 1 in the first two months). Indeed, a monthly purchasing level of .12 packs provides a good fit to the estimated *cumulative* purchasing levels (as is shown by their agreement with the "smoothed" figures in the second column of figures) and also agrees with the lack of a trend in the monthly *sales* data measured at the retailer level, as also shown in Table 6.8. Furthermore, grossing-up the mean purchasing level of .12 packs of Q per housewife and the mean sales level of 5 (weight) units of Q per

Table *6.7.* The Observed Data on Q as Used for Analysis

Survey Month		No. of Packs Purchased	
		0	1
2	%	85.5	10.3
3	%	80.3	11.5
4	%	76.8	13.5
5	%	*	*
6	%	67.8	16.5

* Aberrant

retail outlet to the relevant housewife and retail-outlet populations, gave figures which closely agreed with each other. The monthly estimate of .12 purchases per housewife therefore appeared reliable.

This stationarity of the market also allowed one to form NBD estimates of the percentage of population buying Q in each separate (i.e. average) month for each survey period*. These monthly figures are given as the last column of Table 6.8 and also show no trend. Thus from each separate survey the estimate is that about 9% of the population bought Q each month.

Table 6.8. Sales and Penetration Estimates

Survey Month	Cumulative Purchases* NBD	Smoothed	Monthly Purchases*	Monthly Sales**	Buyers per month
2	.21	.21	.11 X 2	6 X 2	8.5%
3	.36	.33	.12	5	9.4%
4	.43	.45	.12	5	8.8%
5	***	.57	.12	4	***
6	.68	.69	.12	5	9.5%
Average	–	–	.12	5	9%

* Per housewife in the sample.
** Per retail outlet.
*** Aberrant.

All these various estimating procedures therefore lead to the single result that about 9% of the population bought in the typical month, and did so at an average level of .12/.09 = 1.3 packs per buyer. None of this – neither the lack of a month-by-month trend nor the quantitative level of the monthly figures – is self-evident from the cumulative data initially observed, as in Tables 6.5 or 6.7.

The next step is to form *repeat-buying* estimates. Given the monthly figures, the NBD estimate of the incidence of repeat-buyers from month to month is .2, i.e. that about 20% of the buyers in one month would buy again in the next. Numerically this repeat-buying level is very low, implying that the product's sales are largely relying on an inflow of new buyers each month. However, the observed cumulative penetrations did not increase all that rapidly month by month (see

* Thus given that 14.5% of the population bought Q at all up to the end of Month 2 (i.e. 100%-85.5% in Table 6.7), and given that the monthly purchasing level was .11 units per housewife in each of the two months (Table 6.8), the NBD estimate is that 8.5% of the population bought Q in each of the two months (i.e. 2.5% buying in ***both*** months).

Tables 6.5 and 6.7), and the average purchase frequency estimated for each month was very low, at 1.3 packs per buyer. The implication is more one of a long purchasing-cycle rather than one of a complete failure for repeat-buying as such to occur.

The final question was therefore what the buying pattern in *longer* time-periods would be. Table 6.9 shows NBD estimates of the percentage of informants who would buy the product at all in periods of up to 2 years (given the same conditions as in the first six months), and how many of them would buy once only. Thus after 2 years, half the population would have tried the product, and only 25% of these (13% out of 51%) would have bought once only.

Table 6.9. Purchasing Estimates for Periods of Various Lengths

		Periods of Length			
		1 month	6 months	1 year	2 years
% of Housewives who had:					
Ever Bought	%	9	32*	41	51
Bought 1 Unit Only	%	8	16*	14	13
Bought 1 Unit/Ever Bought	%	90	50*	33	25

* Directly Observed (Table 6.5).

This is the crucial result: given long enough, at least 75% of buyers would buy more than once! Consumer acceptability therefore exists (despite the numerically low repeat-buying level from month-to-month), as well as quite a high rate of initial trial. But the average rate of product-usage was seen to be low.

The two outstanding marketing problems were therefore what (if anything) to do firstly about this low rate of product usage, and secondly about the 50% of the population who would not yet have tried the product even two years or so after the launch.

As an addendum to this test-market evaluation based on very "thin" survey data, a check was possible with full consumer-panel data five years later, as shown in Table 6.10. In terms of purchasing of the product-class, the comparisons show a very fair degree of correspondence. There is no logical need for the two sets of data to agree (e.g. the market might have grown in the 5 years), and so the correspondence is of interest.

Table 6.10. Comparison of Launch Estimates of Q and Purchasing of the Product-Class 5 Years later

	Launch Estimates	5 years later
% of the population buying in 4 weeks	9	9
% of the population buying in 6 months	32	24
% of the population buying in 1 year	41	37
Purchases per buyer in 4 weeks	1.3	1.5

4.4. Statistical Errors in the Data

Errors in the data played a major role in the analyses of product Q in the last section and also in the analyses of clothing data in § 5.3 of Chapter 5. We now discuss briefly applications of the repeat-buying theory to three statistical types of potential errors in the data, namely

(i) the occasional 'outlying' value,

(ii) bias due to irregular reporting,

(iii) statistical sampling error.

Excessively heavy buyers. The fit of the NBD makes it possible to investigate the statistical status of exceptional observations or 'outliers', such as the occasional very heavy (or very frequent) purchaser of an item.

Purchasing data typically contain a very small proportion of rather heavy buyers (as was illustrated in Table 4.1 of Chapter 4). Occasionally there will be one or two purchasers in a given sample who seem to be "excessively" heavy*. For example, most buyers in a year may buy the item up to ten times, with a few making up to twenty or thirty purchases, and one buyer who made 65 purchases.

* A case occured in the late 1950's where a ***general excess*** of heavy buyers in a certain product-field might have explained some systematic over-reporting compared with sales data, and following a suggestion by Mr. D.A. Brown, this was investigated by fitting theoretical distributions to the data. However, the NBD gave a good fit to this data, even in the "tail" of the distribution, and this was found to generalise to other product-fields. The short-term conclusion was therefore negative, i.e. that this over-reporting was ***not*** due to an excess of heavy buyers – but on the more positive side, this was the start of work on the NBD model and repeat-buying generally (as mentioned in §4.2 of Chapter 4 and discussed further in § 11.4 of Chapter 11).

The question then is whether this very heavy purchaser can reasonably be regarded as "belonging" to the sample – the point being that he will affect many of the detailed analyses rather markedly, just because he is such a heavy buyer.

Since in general the NBD gives a good fit to observed frequency of distribution of purchase, analysis programmes of panel data can include an initial check of the fit with or without the heaviest purchasers, If the observed standard deviation significantly exceeds the theoretical one (the fitting being by mean and number of zeros), the implication is that at least the ***heaviest*** purchaser is statistically abnormal *. (It is then usually best to exclude such an "abnormal" value from detailed cross-analyses, as it can swing differences between sub-samples quite unduly, but at the same time to record it for inclusion in any grossing-up procedures.) Such a check is worth making, even though in practice abnormal cases have only rarely been found.

A particular case-history can perhaps illustrate this type of problem more generally. In the late '50s it was found in the Attwood Consumer Panel in Western Germany that one or two households in any one period would buy what seemed to be an abnormally large quantity of a certain class of detergents.

When Negative Binomial Distributions were fitted to the data, the theoretical standard deviations were significantly smaller than those actually observed, thus showing that these large purchases were in fact larger than the distribution of all the other purchases would appear to warrant, on the basis of the Negative Binomial Distribution.

This did not mean that these very large purchases were in any real sense "wrong", but it did show that they were statistically abnormal, since detergent purchases in other countries, and purchases in other product-fields in Western Germany, did not generally show this deviation from the Negative Binomial form. Such a conclusion does not in itself tell us what to do about the data, but it is useful to know that the readings ***are*** statistically abnormal, instead of merely having a subjective opinion to that effect.

Further analysis in this particular case showed that the major individual brands of detergents did not show such abnormalities at all, and that one was dealing with very heavy purchases of some relatively minor brand or brands. When the background of these purchases was

* This analysis can be more complicated if the data are subject to the "variance discrepancy"> when the observed standard deviation is generally *smaller* than the theoretical one (see § 7.8 of Chapter 7).

investigated, several special points were found. One was that certain of the lesser-known and relatively cheap brands were sold by door-to-door salesmen, and that these brands were then occasionally bought in very large quantities. One was therefore dealing with two types of purchase, namely those made from the ordinary retail outlets and those made at the door. It was reasonable that the NBD analysis should have shown an abnormality in the pattern of purchases before these two types of purchases were differentiated.

Irregular Reporting. A simple form of applying repeat-buying analyses lies in the study of potential biases due either to non-cooperation of some of the sample picked for a consumer panel, or to their *irregular* reporting even when they do co-operate. Thus in some consumer panel operations – especially those which are not run as well as they might be – appreciable proportions of panel-members may miss providing a proper return about their purchases in a given reporting period (usually a week) although they will do so again in the next period or so. Since analyses over longer periods such as 3, 6 or 12 months have to be based on continuous reporters, this means that they have to be based on samples which are not only relatively small but which may also be biased.

Table 6.11. A Comparison of the Market-Shares in a 4-week period of **Five Leading Brands F to** J, by Regularity of Reporting

Longer-term Reporters: About 500 panel-members who reported continuously for at least 16 weeks
Shorter-term Reporters: About 1500 panel-members who reported continuously for at least the 4-week analysis-period

4 Weeks	Total	Brand-Shares				
		F	G	H	I	J
Longer-term reporters	100%	20	18	10	9	8
Shorter-term reporters	100%	18	17	10	8	8

One question then is whether continuous reporters differ in their purchasing patterns from irregular ones. A deeper question is whether the periods in which the irregular reporters do *not* report are of a special kind (e.g. absences with no purchases being made, or very *heavy* purchasing periods, or whatever).

Investigations of such potential biases can start 'by comparing very simple measures, such as brand-shares. An example is given in Table 6.11, for five leading brands F to J in a certain product-field over a 4-week period.

The particular panel used to measure this field suffered from irregular reporting to quite an exceptional degree. Thus of a sample of about 1500 who reported for all 4 weeks in the 4-week period, only about 500 reported for all the following 12 weeks as well, with the remaining 1000 missing one or more weeks. However, Table 6.11 shows that there was no appreciable difference in the market-shares of the brands among the longer- and shorter-term reporters*.

Such comparisons can dig more deeply by utilising various indices of repeat-buying and of multi-brand buying (as discussed in Part V). Some comparisons of repeat-buying measures are given in Tables 6.12 to 6.14. Thus Table 6.12 shows little differentiation — by regularity of reporting — in terms of the frequency of buying the product in the 4 weeks analysed. (Cross-analysis of the breakdowns in the Tables 6.11 and 6.12, i.e. brand-shares amongst light, medium and heavy buyers of the product, also showed no difference.)

Table 6.12. Purchase Frequency of the Product-class, by Regularity of Reporting

4 Weeks	All Buyers	Purchases of the Product 1–4	5-8	9+	Av. number per Buyer
Longer-term reporters	100%	53	37	10	5.1
Shorter-term reporters	100%	50	39	11	5.0

A similar failure to identify any significant difference in the results occurs in Table 6.13 on examining the average frequencies with which the 4-week buyers of each brand buy either the product or the brand itself (the only sizeable exception being the 10% difference for total product purchases of buyers of Brand I).

The frequency distribution of purchases of each brand about these averages also shows no difference between the shorter- and longer-term reporters, as is summarised in Table 6.14: there are no differences in the proportions of buyers of a given brand who buy it once only, nor

* Any differences would be pinpointed more sharply by comparing the 500 longer-term reporters with the 1000 panel-members who were ***not*** regular over the total 16 weeks. However, for expository purposes, comparison of the longer-term reporters with the total sample of 1500 was more realistic, in terms of the practical use of the data.

are there any differences in such overall measures of the frequency distributions as their standard deviation.

The general finding so far is that whilst irregular reporting numerically decreases the sample size available for the analysis of longer-term buying patterns, it does not markedly affect the quality or validity of the resulting data.

Table 4.13. Average Purchase Frequencies of the Brand and of the Product, per Buyer of a Brand, Analysed by Regularity of Reporting

4 Weeks	Brands F	G	H	I	J	Average
Purchases of Product						
Longer-term reporters	5.7	5.7	5.9	6.1	5.6	5.8
Shorter-term reporters	6.0	5.8	6.0	6.7	6.0	6.1
Purchases of the Brand						
Longer-term reporters	3.3	2.8	3.0	2.9	2.4	2.9
Shorter-term reporters	3.1	2.8	2.8	2.8	2.7	2.8

Table 6.14. Two Measures of the Frequency Distribution of Purchases of Individual **Brands**

(The percentage of buyers buying once only, and the standard deviation)

4 Weeks	Brands F	G	H	I	J	Average
% Buying Once-only						
Longer-term reporters	9%	11%	6%	6%	7%	8%
Shorter-term reporters	9%	12%	8%	6%	7%	8%
Standard Deviation						
Longer-term reporters	1.9	1.7	1.5	1.4	1.1	1.5
Shorter-term reporters	1.9	1.7	1.5	1.4	1.3	1.6

Quick Estimates of Standard Errors. The sampling of informants on which consumer panels and surveys are based is another source of error in the data. 'The NBD model makes it possible to estimate quickly and easily various standard errors, instead of having to carry out extensive

tabulations of the raw data*. This is essentially a time- and labour-saving application (although it can play a deeper role in determining sample sizes when planning a new study).

First we consider the standard error of m, the average number of purchases per informant in a given period of time, i.e. the level of *sales*. For this we need the standard deviation or variance of the frequency distribution of purchases. In general, this would require working out the sums of squares, etc., from a specially tabulated frequency distribution. Given the fit of the NBD, the standard deviation can however also be estimated from the proportion of non-buyers p_0 and the mean m itself, two statistics which tend to be regularly provided in the relevant market research reports. The methods mentioned in Chapter 4 for estimating the parameters k or a can be used, the standard error of m for a sample of size n being then

$$\sqrt{\{m(1+m/k)/n\}} \text{ or } \sqrt{\{m(1+a)/n\}}$$

In certain product-classes the so-called "variance discrepancy" operates (see Chapters 4 and 7), so that the theoretical value of the standard deviation is an overestimate. But for many practical purposes all that is required is an indication of the order of size of the standard error. (In any case, the variance discrepancy phenomenon is itself regular and can be allowed for, e.g. by some correction-formula as in § 7.8.)

An even simpler result arises for the standard error of the difference in the mean purchasing rate m' in one period and the mean m'' in another (equal-length) period. The standard error of the difference $(m' - m'')$ can of course be computed directly from the pairs of readings for individual consumers, but can be estimated more quickly and easily form the NBD theory.

Thus if the means m' and m'' are based on *independent* samples of size n' and n'', it follows from the above results for a single time-period that a simple estimate of the standard error of $(m' - m'')$ can be calculated from the means and the proportions of non-buyers in each of the two periods, in the form of

$$\sqrt{\{m'(1+a')/n' + m''(1+a'')/n''\}}.$$

But if the same sample of consumers is used in both periods, as in work based on consumer panels, the two sets of readings will be correlated. In this case, to obtain an estimate of the standard error of the

* The discussion here treats the data as arising from simple random sampling. Sampling error estimates must also take into account the "design factor" of the sample (e.g. stratification, multi-stage sampling, etc.).

difference $(m' - \text{HZ}'')$, we take as the null-hypothesis that in the population as a whole the mean purchasing rates in the two periods are the same, i.e. that there has been no real trend. Then if r' and r'' are the purchases of any one consumer in the two periods, we require the variance of $(r' - r'')$ over all consumers in the population. This is equal to the expected value of the sum of $(r' - r'')^2$, since for all consumers the mean of $(r' - r'')$ has been assumed to be zero. Now in the stochastic NBD model of \$4.5, r' and r'' for any one consumer are taken to behave like independent readings from the same Poisson process with a mean μ for that consumer. The expected value of $(r' - r'')^2$ is therefore equal to the expected value of $(r' - \mu)^2 + (r'' - \mu)^2$, which is twice the mean μ of the Poisson (since the variance of the Poisson distribution is equal to its mean). It follows that taken over all consumers, the variance of $(r' - r'')$ is equal to twice m, the average rate of purchasing in the population. The sample estimate of this variance is therefore $(m' + m'')$. The standard error of the difference between the-means m' and m'' obtained from the sample is therefore:

$$\sqrt{(m' + m'')/n}, \text{ or } \sqrt{(2m/n)} \text{ for short,}$$

a very simple finding.

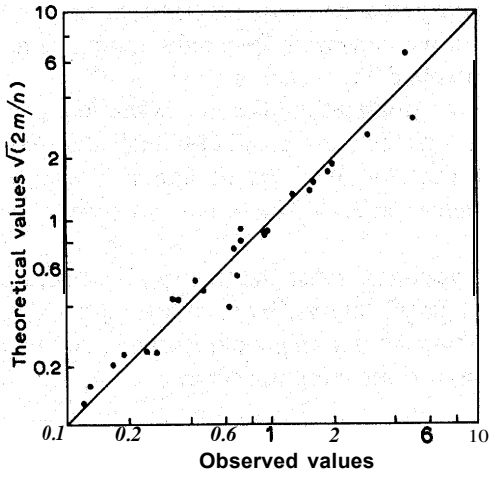

Fig. 4.1. Comparison of 'theoretical' and 'observed' values for the standard error of the difference between estimates of mean purchasing rate m in two periods of time.

The values given by this theoretical formula compare well with the values of the standard error computed directly from the individual differences $(r' - r'')$, as illustrated in Fig. 6.1. The formula holds even where the Negative Binomial distribution as such does not fit so well (i.e. for the "variance discrepancy" situation). This is not surprising, as the deriviation of the $\sqrt{(2m/n)}$ formula depends only on the Poisson part of the theoretical model, and not on the Gamma distribution.

The fuller mathematical development of the NBD theory is discussed in the following chapter. Deducing the theoretical formula $\sqrt{(2m/n)}$ and the comparison in Fig. 4.1 was however the first practical application of the underlying stochastic model for the NBD described in § 4.5 of Chapter 4 [cf. Ehrenberg 1959].

6.5. Summary

A common practical application of the theory of stationary repeat-buying is to evaluate the effects of marketing action under *non-stationary* conditions. The theory provides norms of what repeat-buying would have been like in the absence of the change in sales-level. In one case-history, the analysis showed that the extra sales due to a consumer promotion arose mainly from attracting extra buyers and partly from increased repeat-buying amongst light buyers, whilst heavy buyers of the item were unaffected by the promotion. Such uses of the theory to provide interpretative norms is generally much simpler and cheaper than running controlled experiments.

Another type of application lies in expanding limited data, as in interpreting data on the new product-launch discussed in § 6.3. The analysis showed that the new brand had achieved a high degree of consumer acceptance, although there was no directly observed data on this point.

More technical applications of the theory are illustrated by separating out "excessively heavy" buyers, by establishing the absence of bias due to irregular reporting of certain panel-members, and by providing quick ways of estimating various sampling errors.