Volume Title: Consumer Sensitivity to Finance Rates: An Empirical and Analytical Investigation

Volume Author/Editor: F. Thomas Juster and Robert P. Shay

Volume Publisher: NBER

Volume ISBN: 0-87014-402-2

Volume URL: http://www.nber.org/books/just64-2

Publication Date: 1964

Chapter Title: Buying Intentions and Purchase Probability: II

Chapter Author: F. Thomas Juster, Robert P. Shay

Chapter URL: http://www.nber.org/chapters/c1286

Chapter pages in book:

CHAPTER 4

Buying Intentions and Purchase Probability: II

Introduction

THE analysis in Chapter 3 was concerned with the general relation between purchase probability and surveys of household buying intentions for particular durable commodities. In this chapter I examine some of the problems arising from the aggregation of purchase probabilities and their empirical counterpart, buying intentions.

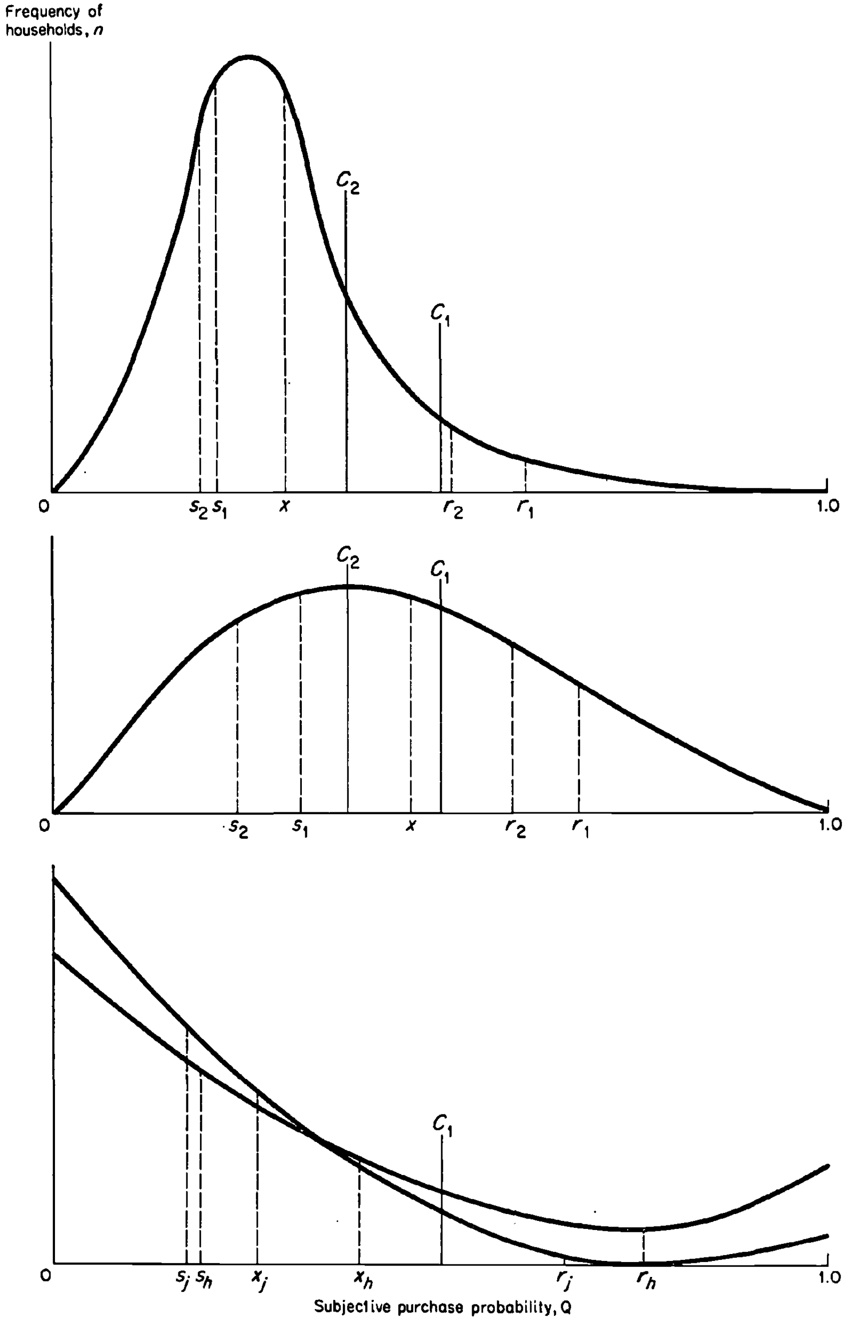
I start with two problems that are closely related to the analysis in the previous chapter. First, are the cut-off probabilities associated with alternative measures of buying intentions—the C_i—the same for all commodities? If not, are the differences systematically related to identifiable characteristics of the commodity? Second, are the C_i the same for all groups in the population, or do they differ with identifiable household characteristics such as income, life-cycle status, ownership of the commodity in question, and so forth.

As noted above, there are no direct observations on the cut-off probabilities associated with variant intentions questions. However, C_i is closely related to the empirically observable magnitude r_i. Given the probability function $n = f(Q)$, it has been shown that C_i must be somewhat below r'_i, the mean of the distribution function above C_i. I assume that the observed magnitude r_i—the fraction of intenders purchasing—is a reasonably good proxy for r'_i, although r_i is subject to a downward bias. If the C_i differ among commodities or population groups, it seems reasonable to suppose that the r'_i and r_i will differ also, since C_i must be strongly correlated with both.

For the moment, assume that *ex ante* r'_i is equal to *ex post* r_i. It follows that r_i must be above C_i; the difference between the two will depend on the shape of the distribution function and on the characteristics of the *i*th intentions question. If the distribution function has a rapidly changing slope in the region of the cut-off points, as in the top panel of Chart 11, C_1 and C_2 will be quite close to the corresponding means, r'_1 $(= r_1)$ and r'_2 $(= r_2)$. On the other hand, if the distribution function has a flatter slope in the region of the cut-off points, as in the middle panel of Chart 11, both C_1 and C_2 will be further below their corresponding means. Generally speaking, the more skewed the distribution function, the steeper the slope in the region of relatively low cut-off points and the flatter the slope in the region of relatively high cut-off points. The amount of skewness in the distribution function is, in turn, related to mean probability for the entire function; the smaller the mean the more skewed the distribution

CHART 11

Illustrative Relation Between Slope of Distribution Function and Cut-Off Probability



function and the smaller the difference between r_i and C_i for relatively low values of the latter.

On the other hand, the distribution functions are not necessarily bell-shaped with greater or smaller amounts of skew. A U-shaped function is not at all improbable. The largest single number of households may have probabilities equal to zero: many households may simply have no interest at all in purchasing particular items, or they may have just acquired a new one. At the high-probability end of the distribution, households with probabilities approximately equal to unity may exceed those with probabilities of 0.7 or 0.8. Suppose the commodity in question were automobiles, for example. A fairly large number of households buy a new car every year as a matter of course; purchase probability for these households, if the forecast period covers the customary time of purchase, is bound to be unity or close to it. If the distribution functions happen to be U-shaped, the mean probability above any cut-off point located around the trough will probably be higher the higher the mean in the distribution as a whole. This situation is illustrated in the lower panel of Chart II, where r_j represents mean probability for low-income households, r_h the corresponding mean for high-income households.

Differences in Cut-off Probability Among Commodities and Households

The question of whether or not C differs among commodities or among households, given the specifications of the buying-intentions questions, can now be examined. Table 19 summarizes the observed values of r, and the corresponding fraction of households reporting intentions to buy, for two of the alternative questions—those asking about "definite" and about "definite, probable, and possible" purchases within a year. The last column shows the estimated mean probability for the sample as a whole, i.e., the observed purchase rate.

The data provide strong evidence that C_i is not the same for all items. For example, r for those definitely intending to buy a car within a year—r_d—is 0.566. The corresponding cut-off probability, C_d, must be lower than 0.566. The function must be highly skewed (even more so than the function at the top of Chart 11), since the mean is 0.191; hence, the C_d corresponding to an r_d of 0.566 is unlikely to be lower than 0.40 or 0.45. Taking account of the fact that r_d is a downwardly biased estimate of r'_d, the true C_d is probably somewhat higher than 0.40. Yet the *mean* probabilities for households definitely intending to buy a television set (0.197), a garbage disposal unit (0.275), or a high-fidelity set (0.304) are all well below 0.40. The corresponding values of C_d must be even lower, although perhaps not so much lower as in the case of automobiles because these distributions are more highly skewed and the cut-off probabilities are relatively low. But C_d cannot possibly be the same for automobiles and for garbage disposal units.

Assuming that these differences among commodities in r_i represent real differences in C_i, it appears from Table 19 that, given the intentions question, C is relatively high for commodities owned by most households (cars, ranges, refrigerators), and relatively low for commodities owned by few

TABLE 19
ESTIMATED PROBABILITY THAT INTENDERS WILL PURCHASE SPECIFIED COMMODITY, AND PER CENT OF INTENDERS IN SAMPLE

	Estimated Purchase Probability Among:			*Per Cent of Sample Households with:*	
Commodity	Definite Intenders	Definite, Probable, or Possible Intenders	Entire Sample	Definite Intentions	Definite, Probable, or Possible Intentions
Automobile	.566	.349	.191	8.1	31.0
Furniture	.493	.330	.187	13.5	31.9
Carpets and rugs	.376	.201	.076	5.7	17.1
Washing machine	.409	.233	.075	3.3	11.9
High-fidelity equipment	.304	.144	.065	4.3	17.0
Refrigerator	.508	.249	.056	2.5	10.6
Range	.478	.257	.052	2.5	8.4
Clothes dryer	.284	.177	.046	2.8	11.4
Television set	.197	.118	.053	2.5	11.6
Air conditioner	.463	.230	.054	2.1	11.8
Food freezer	.317	.148	.027	1.6	8.3
Dishwasher	.392	.172	.026	1.9	7.5
Garbage disposal unit	.275	.149	.023	1.5	5.1

SOURCE: Table 2, Chapter 2.

households (food freezers, garbage disposal units, dishwashers). Also, as noted above, differences in r_i (hence C_i) among commodities are positively correlated with differences in x; that is, commodities purchased by relatively few households tend to have low cut-off probabilities, and vice versa for commodities with relatively high purchase rates. Since mean probability (= purchase rate) and degree of skewness are necessarily associated, the degree of skewness in the probability function is also correlated with both r_i and C_i.[1]

[1] There is no way (from the data) of isolating the basic cause of this relationship. The rank correlation between purchase rate (x) and mean probability for intenders (r) is somewhat higher than that between the fraction owning and mean probability; both correlations are significant at the 0.05 level. I would guess that the basic rela-

I turn now to an examination of the variation in C among households with different characteristics, given the commodity and the buying-intentions question. Here the results are quite different. There is no evidence that C_i, given the commodity, varies among households with different levels of family income, life-cycle status, or educational level. Since C is not observable, r_6 and $1 - r_6$—the fractions purchasing and not purchasing among those intending to buy within six months—are used as the test statistics. For each of thirteen durables, I tested the relation between r_6 and, in turn, income (eight classes), life-cycle status (five classes), and educational level (five classes). Of thirty-nine such tests, contrasting the observed values of r_6 with those predicted by the hypothesis of no association, only one yielded a value of chi-square large enough to be statistically significant at the 5 per cent level.[2] I conclude that r_6 (hence also C_6) is invariant with respect to family income, life-cycle status, or education.[3]

---

tionship is between ownership and C_i. If a commodity is already being used by most households, the probability function is mainly determined by replacement considerations. If the commodity is not owned, the prospect of acquiring something "new" may well encourage a high degree of whimsey and wishful thinking in the estimate of purchase probability and in the response to questions about intentions to buy. It is also possible that buying-intentions questions are subject to this bias while questions about purchase probability might not be, although again I have no way of demonstrating the point.

[2] The test statistics are actually $r_6 p_6 N$ and $(1 - r_6)p_6 N$, that is, the number of households reporting buying intentions and purchasing (not purchasing) in the first, second, etc., age, income, or education group. The one case with a significant value of chi-square was actually very close to but below the critical value of chi-square for the 0.05 level.

[3] Examination of the r_6 patterns within these classifications does not provide any clear evidence that C_6 might actually be somewhat higher when family income is higher, or when educational level is higher, etc. For the majority of commodities the correlation between r and family income is positive though necessarily nonsignificant, given the chi-square results. However, the probability functions will generally be less skewed for high-income groups than for low-income groups, since the mean probability (= purchase rate) in the sample as a whole is positively correlated with income. Given the (relatively low) level of the C_6, the less skewed the distribution function the further apart are C_6 and r_6, and the further C_6 is likely to be below any given r_6. Hence, the correlation between C_6 and family income might be even weaker than the small positive correlation observed between r_6 and family income. The same is true of the relation between r_6 and the other two variables. The relation here, if any, consists of a slight tendency for r_6 to be higher for groups with relatively high purchase rates—households with more income, more education, or with younger heads. But even this weak relationship may be too strong as a measure of the correlation between these characteristics and C_6.

If the distribution functions are U-shaped, the same conclusion will hold. In this case functions with relatively large mean values are likely to have more rapidly rising slopes above the minimum point relative to functions with smaller mean values. Thus, the mean probability in any segment above the minimum point is likely to be greater when the overall mean is relatively high (see Chart 11).

Differences in Purchase Rates Among Households

The fact that r_i (hence C_i) does not vary among households with different characteristics implies that purchase rates for nonintenders, s_i, are likely to vary among the same classes of households provided that x differs. In connection with a quite different problem, I showed in Chapter 3 that differences in x are generally associated with differences in s if r is the same for both groups unless very special conditions prevail.[4] Hence, if the means of these probability distributions above the cut-off point happen to be about the same for groups that differ with respect to mean probability in the distribution as a whole (x), the means below the cut-off point (s_6) will generally be correlated with x. The point is illustrated in Chart 12, which is similar to Chart 3 in Chapter 3.

The expected empirical results show up quite clearly when s_6 is computed for the thirteen commodities within the same income, life-cycle, and educational classes. In testing for independence between the observed values of s_6 (for the same income, life-cycle, and educational classes) and the mean for all classes combined, seven of thirty-nine independent tests showed chi-square values that would be observed only once in 100 trials and fifteen of thirty-nine showed values that would be observed only once in 20 trials. The variation in s_6, given the commodity, is highly correlated with the variation in x, as is bound to be true.[5]

The observed differences in the purchase rates of nonintenders within different income, age, etc., classes can be rationalized in several ways. It may be that the probability distributions are symmetrical above the cut-off point but not below. Many households purchase automatically whenever a specified contingency arises (a breakdown, a major repair bill, acquisition of knowledge about a particular product, etc.). Such households will be nonintenders unless the contingency is regarded as probable.

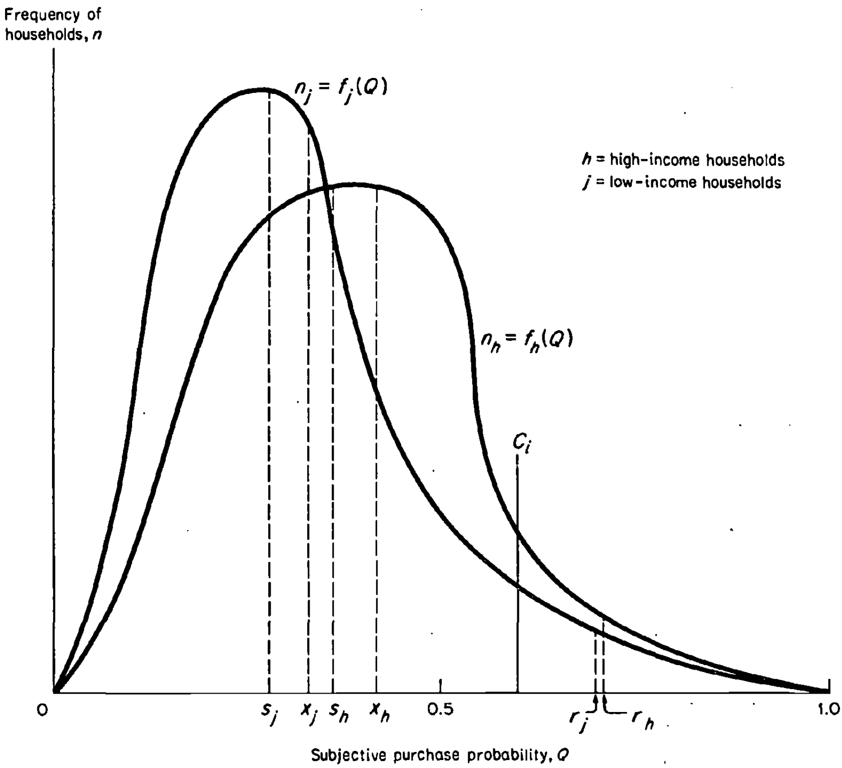
[4] See Chapter 3, p. 76.

[5] The same pattern in the r and s data can be observed in buying-intentions surveys taken from carefully drawn probability samples. A large number of such surveys, and a few reinterviews, have been conducted by the Survey Research Center at the University of Michigan. Lawrence C. Klein and John B. Lansing, in "Decisions to Purchase Consumer Durable Goods," *Journal of Marketing*, October 1955, present data (Table II) on the purchase rates of intenders and nonintenders, classified by income. Both the chi-square test and regression analysis indicate that the relation between s and income class is significant, but not the relation between r and income class. See also Peter De Janosi, "Factors Influencing the Demand for New Automobiles," *Journal of Marketing*, April 1959, Table 2A, where the same relation among r, s, and income is also apparent.

Because of financial constraints, the reaction of low-income households to such contingencies is less likely to involve a purchase than the reaction of high-income households to the identical contingency. In effect, the joint probability that any given contingency will occur *and* that the reaction will consist of a decision to purchase is considerably larger among high-

CHART 12

Illustrative Probability Distributions for High- and Low-Income Households



than among low-income families, even though the probability of occurrence for any given contingency is the same for both. But in this case, mean purchase probability will be larger among high-income nonintenders than among low-income nonintenders, provided these factors enter into the households' *ex-ante* estimate of purchase probability; observed purchase rates will differ in any event. I see no logical reason why a similar difference would necessarily be observed among high- and low-income intenders.

Another possibility is that *exposure* to contingencies may be different among these groups. If so, the (*ex post*) fraction purchasing may be quite different, while (*ex ante*) mean probability is the same. High-income households will generally have larger stocks of durables than others, and unanticipated replacement may be more frequent. On the other hand, high-income households are likely to have newer as well as larger stocks and that factor would tend to work in the opposite direction.

I did not perform extensive tests to determine the degree to which this set of relations would be altered if some other intentions question, with a different cut-off probability, had been used. A comparison of responses to the "within six months" and "within twelve months" intentions questions seems to show that r_{12} is not as unrelated to family income or life-cycle status as the above tests showed r_6 to be. This finding is quite consistent with the hypothesis that C_i does not vary among households. C_{12} must be lower than C_6, and the slope of the distribution function is apt to change more rapidly between these two cut-off points the more skewed the function. Since the probability functions for high-income groups have higher mean values than those for low-income groups, they are less skewed. Therefore, different means are more likely to be observed in the segment above C_{12} than in the one above C_6. The point is illustrated in Chart 11, where the difference in r_2 between the top and middle panels is greater than the difference in r_1. Other factors may also be present. For example, high-income households might have higher cut-off probabilities than low-income ones for intentions questions characterized by a high degree of uncertainty, or they may systematically underestimate their purchase probability over relatively long time horizons.

Two conclusions follow from these results. First, aggregation of buying intentions across different commodities will not produce a homogeneous total for the expected dollar value of purchases, since expected values are lower for some items than others, given the specifications of the intentions question and the price of the item. This is not likely to be a serious difficulty in practice, because the items that make up the large majority of intended purchases are relatively homogeneous with respect to the probabilities associated with any given intentions questions. The items for which probabilities seem to be much lower are generally those purchased by relatively few households; hence, they are comparatively unimportant in the aggregate of either expected or actual dollar values of purchases. Secondly, the distribution of the population by variables like income, demographic status, or education evidently makes little or no difference to the purchase probabilities associated with responses to questions about buying intentions. In effect, there is no convincing evidence that responses to (dichotomous) buying intentions questions should be weighted according to whether income is high or low, etc., and this is a great convenience in making practical use of such data.

Cut-off Probabilities and Purchase Rates Related to Durable Goods Stock

Up to this point I have analyzed variations in C_i among different commodities and differently situated households, finding significant variation in C_i among commodities that is apparently related to characteristics of the commodity itself, but no significant variation in C_i for a given commodity among households with different family income, life-cycle status, and educational level. Examination of the relation between C_i and some other household characteristics tends to weaken this tentative conclusion of invariance among households, and also points up some specific problems in aggregating probabilities or buying intentions.

Data on the durable goods stock position of each household are available from the CU surveys. In addition, respondents were asked whether any items in their durables stock "needed to be replaced." Answers to this question seemed to combine a judgment concerning age or condition of the stock with attitudes toward obsolescence.

The data on stock position and subjective replacement need cover nine commodities. These items divide naturally into three groups: (1) automobiles; (2) ranges, refrigerators, washing machines, and television sets; (3) clothes dryers, dishwashers, food freezers, and air conditioners. Automobiles are treated separately because of their relatively high unit cost, their extremely high ownership ratio, and the possibility of extensive multiple ownership. The second class consists of items that are owned by the bulk of households in the sample; the third, of those owned by relatively few households.[6]

Because buying intentions and purchases were comparatively infrequent for some of these commodities, especially those in group 3, I combined intentions and purchases for commodities in both the second and the third groups. Thus the basic data consist of average purchase rates for groups of commodities among households classified with respect to stock position, subjective replacement need, and buying intentions. Three such classes were formed among both intenders and nonintenders; those who owned the commodity but did not report that it needed replacement (S), those

[6] The ownership ratios range from roughly 74 to 79 per cent of all households for category 2, from 16 to 34 per cent in category 3; about 92 per cent of these households own at least one automobile.

who owned and reported a need for replacement (*R*), and those who did not own the commodity at the time of interview (*NS*).[7] Table 20 summarizes the estimated mean probabilities (= observed purchase rates) for intenders and nonintenders among households classified by ownership and by replacement need. The sample used for the analysis was asked about intentions to buy within six months.

TABLE 20
ESTIMATED MEAN PURCHASE PROBABILITIES (OBSERVED PURCHASE RATES) FOR THREE CLASSES OF COMMODITIES

Commodity Class	Sample Size (*N*)	Number of Intenders	*Observed Purchase Rates for:* Intenders (r_6)	Nonintenders (s_6)
1. *S*	2,583	138	.493	.142
R	759	225	.622	.242
NS	295	25	.640	.133
2. *S*	9,829	127	.331	.037
R	1,425	258	.264	.104
NS	3,294	205	.522	.056
3. *S*	3,290	53	.340	.048
R	136	19	.368	.068
NS	11,122	356	.272	.018

SOURCE: Basic data from Consumer Purchase Study, NBER.

NOTE: Commodity class 1 consists of automobiles; class 2: ranges, refrigerators, washing machines, and television sets; class 3: clothes dryers, dishwashers, food freezers, and air conditioners. *S* designates those who owned the items and did not indicate that they needed replacement; *R*, owners who stated that the specified items needed to be replaced; *NS*, nonowners. The number of nonintenders is the difference between the total number of observations (*N*) and the number of intenders. Sample size for each class is the number of households in the sample multiplied by the number of items in the class, i.e., *N* is four times as large in classes 2 and 3 as in class 1. The sample size distribution among households classified by *S*, *R*, and *NS* is essentially a distribution by commodity rather than by household; that is, the same household can be classified in *S* for ranges, *R* for refrigerators, and *NS* for television sets and washing machines.

The data examined in the first part of this chapter indicate that r_6, the estimated mean probability for intenders, does not vary systematically with family income, life-cycle status, or education. I inferred that the probability cut-off point, C_6, which is closely related to r_6, was also invariant with respect to these household characteristics. On the other hand, the data showed that s_6, the mean probability for nonintenders, was highly correlated with household characteristics such as income, being higher for those with relatively large incomes, etc. For the classifications summarized in Table 20, both r_6 and s_6 seem to vary with durable goods

[7] Obviously, the *NS* group could not have reported that replacement was necessary.

ownership and with the need for replacement; hence, C_6 cannot be the same for households with different S and R characteristics. The pattern of these relationships is quite interesting, and is puzzling in at least one respect.

AUTOMOBILES

For automobiles, mean probability among intenders (r_6) is significantly higher (0.05 level) for R households than for the S group; that is, households reporting that (one of) their automobile(s) needed to be replaced are more apt to buy than others. This is a sensible enough result except that both groups of households reported that they intended to buy, and there is no a priori reason why the condition of the automobile stock would not have been taken wholly into account when intentions were reported. It will be recalled that mean probability did not turn out to be significantly different for intenders in different income classes, but the replacement-need variable does not fit this pattern. Of course, C_6 for automobiles may really be the same for the S and R groups, with the difference in observed r_6 simply reflecting a different distribution of probabilities for these two groups above the identical cut-off point. This possibility cannot be tested with the data, although the difference in r_6 between these two groups is not so large that it can be disregarded.[8]

Mean purchase probability for automobiles is also closely related to stock and replacement needs among nonintenders, but this pattern is similar to that which had been observed previously for other household characteristics. The magnitude of s_6 is almost twice as large for R households as for either the S or NS ones. The sample sizes are all fairly larger here ($N > 250$), and the difference in s_6 between R and either of the other two classes is statistically significant at the 0.01 level. This difference probably reflects the influence of contingencies. Nonintenders indicating that an item needs to be replaced are essentially saying that they would purchase a new one if they could afford it because the old one is unsatisfactory. Thus, any favorable financial contingency is more apt to result in a purchase, and the unfavorable contingency of complete breakdown is also more probable.

HOUSEHOLD DURABLES

For the second and third commodity groups, most of the results are consistent with the proposition that r_6 is independent of the S and R character-

[8] The automobile purchase rate among NS intenders is also higher than among S intenders, and is about the same as for R. The difference in r_6 between S and NS intenders is not statistically significant, however, because the NS sample is so small.

istics, while s_6 is not. The r_6 values are not significantly different from each other for S and R households in either commodity group; the difference in s_6 is highly significant (0.01 level) for commodity group 2, not significant for 3.[9] In sum, contrasting the behavior of intenders and nonintenders among households that owned the item at the time of the survey, I find that intenders have essentially the same purchase rate, hence the same C, regardless of whether the item was thought to be in need of replacement. Nonintenders, on the other hand, are much more apt to purchase if they reported that the item needed to be replaced; this relation is especially strong in commodity group 2. All these results are wholly consistent with earlier findings about the behavior of intenders and nonintenders with different incomes, educational level, and life-cycle status.

However, the comparison of purchase rates for owners (S) and nonowners (NS) in commodity groups 2 and 3 shows some surprising results. Intenders who were nonowners had much *higher* mean probabilities than owners in commodity class 2; somewhat *lower* mean probabilities in commodity class 3. These differences are highly significant for class 2, not significant for 3. Differences in mean probabilities among nonintenders—s_6—follow a similar pattern, and these differences are highly significant in both commodity groups 2 and 3. To summarize: For commodity group 3, both intenders and nonintenders are *more* apt to purchase if they already own the item in question at the time of the survey; the differences are not significant for intenders, highly significant for nonintenders. For commodity group 2, in contrast, owners are significantly *less* likely to buy than nonowners among both intenders and nonintenders.

The upshot is that C_6 is probably not invariant among all classes of households. The difference in r_6 between owners and nonowners of ranges, refrigerators, etc., seems far too large to be due to either sampling error or differences in the shape of the distribution function above C_6. Hence, I conclude that C_6—and other C_i as well—must be substantially higher for nonowners than for owners of these commodities. Differences in r_6 among households are well within the limits of sampling variability for commodity class 3, partly because the sample sizes are relatively small. Finally, the data provide strong corroboratory evidence that purchase probability among nonintenders is significantly related to household characteristics associated with differences in purchase rates for the sample

[9] Very few householders who owned items in commodity group 3 reported that the items needed replacement, for the obvious reason that most of these commodities must have been acquired fairly recently. Hence, the sample size in the R group is quite small; and the difference between R and S, although in the expected direction, is not statistically significant at any reasonable level.

as a whole. Among nonintenders, those reporting that an item in their durables stock needs to be replaced are much more likely to buy than other households.[10] Purchase probabilities for owners and nonowners are also significantly different for nonintenders, but the differences go in opposite directions for commodity classes 2 and 3. The last observation is clearly in need of explanation.

DIFFERENCES BETWEEN MOVERS AND NONMOVERS

On a priori grounds, I would have thought that nonowners would be more apt to purchase than owners, other things equal, since the marginal returns from purchase of a newer model are normally less than the returns from a first purchase. This proposition is well documented by studies that show a negative association (in time series) between the size of the durables stock and the purchase rate.[11] Thus, the data for commodity group 2 seem sensible enough. But what about commodity group 3, where the correlation between ownership and purchase probability is significantly *positive* among nonintenders, positive but nonsignificant among intenders? One possible explanation is that owners of these commodities are more likely to buy than nonowners because they have higher incomes; another, that such households are more likely to move. Once having owned a dishwasher or clothes dryer the family is unlikely to do without one, and a move to a new house is frequently accompanied by wholesale renovation of the stock of kitchen and laundry equipment—the equipment previously owned being sold as part of the old house. Further, a prospective move of this kind may not be accompanied by "intentions" to buy new kitchen or laundry equipment. The decision to leave the old durables and buy new ones may come up for consideration only when a satisfactory new house has been located and a prospective buyer found for the old house. This interpretation is consistent with the data: only

[10] Another interesting piece of evidence that tends to corroborate these relationships, based on analysis of randomly selected population samples, is reported by Janet Fisher in "Consumer Durable Goods Expenditures, with Major Emphasis on the Role of Assets, Credit, and Intentions," *Journal of the American Statistical Association*, Sept. 1963, p. 653. In a regression of durables purchases during 1957 on selected independent variables, she reports that nonintenders who purchased durables in the preceding year are significantly more likely to buy than other nonintenders While intenders are significantly more likely to buy than nonintenders, it appears to make little difference to intenders' purchase rates whether they had or had not purchased durables in the preceding year.

[11] See, for example, Harold W. Watts and James Tobin, "Consumer Expenditures and the Capital Account," *Proceedings of the Conference on Consumption and Savings*, ed. Irwin Friend and Robert Jones, Philadelphia, University of Pennsylvania Press, 1960, Vol. II; and *The Demand for Durable Goods*, ed. Arnold C. Harberger, Chicago, University of Chicago Press, 1960.

about one out of five buying intentions in group 3 were reported by owners, but owners accounted for about half the purchases among nonintenders and almost half the total purchases.

The explanation via income differences seems to account for some of the observed positive correlation between s_6 and ownership, but not nearly enough. A test of the second explanation is provided by Table 21. Here

TABLE 21

ESTIMATED MEAN PURCHASE PROBABILITIES (OBSERVED PURCHASE RATES), BY HOME OWNERSHIP STATUS, FOR THREE CLASSES OF COMMODITIES

Commodity Class	Sample Size (N)	Number of Intenders	*Observed Purchase Rates for:* Intenders (r_6)	Nonintenders (s_6)
A. HOME OWNERS—NO PLANS TO MOVE, NO MOVE				
1. *S*	1,167	93	.462	.149
R	490	141	.660	.252
NS	168	11	.545	.127
2. *S*	7,145	65	.323	.028
R	976	169	.260	.099
NS	1,179	44	.432	.034
3. *S*	2,529	35	.314	.038
R	93	8	.250	.059
NS	6,678	176	.244	.018
B. NONOWNERS OF HOMES, AND HOME OWNERS WHO MOVED OR PLANNED TO MOVE				
1. *S*	916	45	.556	.129
R	269	84	.560	.222
NS	127	14	.714	.142
2. *S*	2,684	62	.339	.060
R	449	89	.270	.114
NS	2,115	161	.547	.069
3. *S*	761	18	.389	.083
R	43	11	.455	.094
NS	4,444	180	.300	.018

SOURCE: Basic data from Consumer Purchase Study, NBER.
NOTE: See Table 20 for definition of commodity classes and of *S*, *R*, *NS*, and *N*.

two classes of households are distinguished: first, those who owned their own homes at the survey date and neither reported intentions to buy a house nor purchased one; second, all other households, comprising those who did not own homes at the survey date, plus those who owned and either intended to buy or purchased another house. The same data as were shown in Table 20 are calculated separately for each of these groups of households in Table 21.

The r_6 column now seems to show significant differences among the *S*, *R*, and *NS* groups in only two places. In panel A, those intending to buy automobiles have a significantly higher (0.01 level) mean probability if they also said that their automobile(s) needed replacement; in panel B, nonowners intending to buy ranges, etc., have a significantly higher (0.01 level) mean purchase probability. The results here are much the same as in Table 20. Some of the differences among nonintenders (s_6) are substantially altered, however. Among panel A households, the difference in mean purchase probability between owners and nonowners of class 3 commodities is apparently rather small; but it is highly significant because of the very large sample size. On the other hand, the difference among movers (panel B) is much sharper than before, and is of course highly significant (0.01 level). Taking account of income differences, it seems quite possible that, among nonintenders for class 3 commodities, the propensity of households that already own these items to buy them relatively more frequently than nonowners is entirely due to the greater likelihood that the former will move and will purchase replacements in the process.

Summary

On the whole, the pattern of variation in mean purchase rates among households that differ as to ownership and replacement needs for durable goods is reasonably consistent with the findings and analysis in the first part of this chapter. For the most part, mean probability for intenders, hence also the cut-off probability, is invariant among households that differ with respect to ownership and replacement need. On the other hand, mean probability for nonintenders is highly correlated with *R*; nonintenders reporting that a commodity needs to be replaced are much more likely to buy than other nonintenders. This relation is especially strong for household durables typically owned by most families—ranges, refrigerators, washing machines, television sets and automobiles. For automobiles, the relationship between replacement need and purchase probability is quite strong for intenders as well. Finally, there are no significant differences among nonintenders in purchase probabilities for owners and nonowners, with one exception: owners of class 3 commodities have significantly *higher* purchase probabilities than nonowners. This relationship is mainly—perhaps entirely—due to the behavior of households that changed residence during the forecast period.

These results suggest a rather formidable aggregation problem. Take the question: How are stocks of durables related to purchases, net of other factors? My results show that ownership of class 2 durables is likely to be associated with a smaller probability of purchasing, holding replacement needs and intentions to buy constant. If replacement needs are not held constant, the relationship is likely to be reversed because only owners can be in need of replacement and replacement need is strongly correlated with purchase probability. On the other hand, ownership of class 3 durables is associated with a substantially *larger* probability of purchasing, holding replacement need and intentions to buy constant. Failure to account for replacement needs will make this relationship even stronger. Thus, the aggregate stock of durables may turn out to be completely unrelated to purchases, net of other factors, because the influence of ownership on purchase probability goes in diametrically opposite directions for the two important classes of consumer durables.[12]

Appendix: Predictions of Differences in Purchase Rates Among Households

The data developed in this chapter suggest another possible test of the proposition, discussed in Chapter 3, that mean *ex ante* purchase probability cannot be reliably estimated from the proportion of intenders—households with purchase probabilities above the cut-off probability associated with a given intentions question. In principle the appropriate test involves the time series correlation between the observed purchase rate (x), which is an unbiased estimate of mean *ex ante* probability (x'), and the observed proportion of intenders (p). Such correlations as have been estimated show that p explains a significant amount of the variation in x, but that substantial residual variation remains.[13] The residual variation is presumably due either to the presence of intervening events or to the fact that p is an imperfect (linear) predictor of x'.

It has been shown above that the observed purchase rates within various income and life-cycle classes differed substantially from each other, that the observed purchase rates for nonintenders were strongly correlated with x, but that the observed purchase rates for intenders (r) generally had little or no relation to x. As might be anticipated, these different

[12] If the stock were measured as aggregate value minus the amount in need of replacement, i.e., $S - R$, stock would show a significant negative relation to purchases because R has a significant positive relation. This procedure is not unreasonable, since it amounts to assigning a value of zero to those items in the stock that the household says are in need of replacement. Zero is too low a value, but if the choice is between zero and original cost the first is presumably a better estimate than the second.

[13] See Arthur Okun, "The Value of Anticipations Data in Forecasting National Product," *The Quality and Economic Significance of Anticipations Data*, Princeton for NBER, 1960 and Eva Mueller's "Comment," which follows the Okun article; and my *Consumer Expectations, Plans, and Purchases: A Progress Report*, Occasional Paper 70, New York, NBER, 1959. An exploratory investigation of time series data is also contained in Chapter 3, above.

income or life-cycle classes exhibit substantial variability in p. If it is true that p is a perfect predictor of x', the cross-section correlation between p and x for groups of households classified according to income or life-cycle status ought to be unity except for the differential influence of intervening events among the groups, sampling variability aside.[14] Put another way, if p is a perfect linear predictor of x in time series, net of intervening events, it ought also to be a perfect cross-section predictor—again, net of intervening events. And if this is the case, no other variable can have a significant relation to x in a cross section *unless* it represents an unforeseen event with a differential impact on behavior in the several groups.

To test this proposition, I regressed x on p for each of ten commodities, using values generated by a classification according to income level and life-cycle status (four income, three life-cycle classes). Regressions of x on p net of income were also computed for each commodity, as were a few regressions net of both income and life-cycle status. Use of an income and life-cycle classification is essentially a way of generating some variance in x among groups of households. Neither variable represents an unforeseen event; in the terminology of Chapter 7, below, both are initial-data variables. It follows that no net relation should be observed between the purchase rate and either income or life-cycle class if the correlation between p and x' is perfect except for unforeseen events; on the other hand, some net relation might be observed if p and x' have a less than perfect linear relation.[15]

The data are summarized in Table 22. Although differences in the proportion of households that report buying intentions are a reasonably good linear predictor of differences in the purchase rate, the correlation is by no means perfect. More important, for about half of the commodities examined, both family-income class and life-cycle status make a significant contribution to the explanation of differences in the purchase rate, net of differences in the proportion of intenders. There is no evi-

[14] In one respect this test may be better than the conceptually correct time series test. Some kinds of intervening events are specific to individual households or groups of households; others are general, in that they influence all households. Both these kinds of intervening events may have some influence on the time series correlation, but only the former (specific to groups) will influence the cross-section correlation.

[15] The simple correlation between purchase rate and either income or life-cycle class is determined by the choice of class intervals. Evidently, a classification into those with incomes below \$2,000 per year, between \$2,000 and \$25,000, and above \$25,000 will yield a higher correlation between the purchase rate and income than a classification into groups with more evenly spaced class intervals. But the association between purchase rate and income net of the proportion of intenders will not be sensitive to such arbitrary choices, sampling variability aside.

TABLE 22
CROSS-SECTION CORRELATIONS OF PURCHASES WITH BUYING INTENTIONS, LIFE-CYCLE STATUS, AND INCOME

	r^2_{12}	$R^2_{1.24}$	$r^2_{12.4}$	$r^2_{14.2}$
Automobile	.213	.646[a]	.161	.550[a]
Furniture	.656[a]	.867[a]	.640[a]	.614[a]
Washing machine	.219	.427	.287	.266
Refrigerator	.470[a]	.516[a]	.406[a]	.086
Air conditioner	.161	.777[a]	.044	.734[a]
Dishwasher	.532[a]	.543[a]	.174	.024
Range	.590[a]	.745[a]	.586[a]	.380[a]
Clothes dryer	.756[a]	.771[a]	.758[a]	.062
Garbage disposal unit	.357[a]	.395	.366[a]	.059
High-fidelity equipment	.067	.233	.006	.178
	$R^2_{1.234}$	$r^2_{12.34}$	$r^2_{13.24}$	$r^2_{14.23}$
Furniture	.894[a]	.039	.199	.641[a]
Refrigerator	.563	.370	.098	.004
Range	.865[a]	.657[a]	.472[a]	.381
Clothes dryer	.870[a]	.830[a]	.431[a]	.122

SOURCE: Basic data from Consumer Purchase Study, NBER.

NOTE: Subscript 1 denotes purchases; 2, buying intentions; 3, life-cycle status; and 4, family-income level.

[a] Significantly different from zero at 0.05 level.

dence that a systematic nonlinearity in the (p,x') relation accounts for these results. These findings support the intuitively plausible proposition that the linear correlation between p and x, though significantly higher than zero, is likely to be considerably less than the (unobservable) correlation between x' and x. In short, knowledge of p is by no means a perfect substitute for knowledge of x'.[16]

[16] I do not think it can usefully be argued that these results simply show that the relevant relations may be nonlinear. From a practical point of view, one wants to use survey data to aid in predicting purchases. If there is in fact a unique but nonlinear relation between p and x', it would take some time and additional data to establish the functional form of the relation. Even then, one could never be sure that the true relation had been found unless it were fairly obvious from the data—which it is not. And on a priori grounds, it does not seem plausible to me that the relation is unique but simply nonlinear.