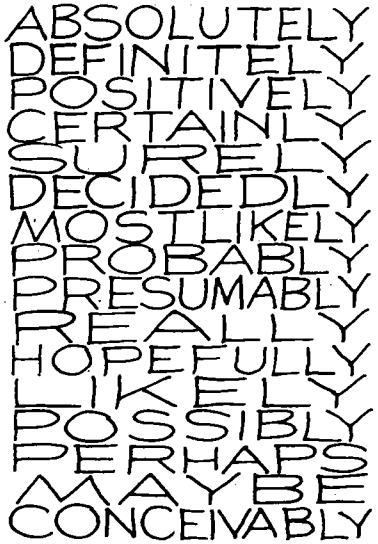

THE USE OF SUBJECTIVE PROBABILITY METHODS IN ESTIMATING DEMAND

Hanns Schwarz *Vice President, Daniel Yankelovich, Inc.*

THE CONTRIBUTIONS of statistics to sampling and analysis of data in marketing research are widespread and liberally documented. For example, we generate facts about the behavior and attitudes of very large populations from studies of relatively small (and inexpensively gathered) samples of respondents. The methods for selecting such small groups to represent large populations come from statistical sampling theory. Such sample surveys are now common practice and indispensable in research. Methods of analyzing data emanating from surveys (as well as from other sources) are also research tools provided by statistics.

Moreover, new statistical applications are being developed continually. Some of these relate neither to survey sampling nor to analysis of data, but to the basic questioning process itself—the essential core of marketing research. In 1971, these applications are not yet in widespread use nor are they well

documented. This paper describes one technique which has been used successfully in recent marketing research: the use of subjective probability methods in predicting demand.

THE NEED TO ESTIMATE DEMAND BY SURVEY

One important function of commercial research is to make estimates of consumer demand for a product or service. On occasion, such estimates are required for existing products, although forecasting methods based on past sales data are then usually entirely adequate. When new or modified products or services are involved, however, other techniques are called for. In some cases, test markets are used and sales results in these few markets become the basis for national estimates. Experimental variations of market testing using simulated retail outlets are used in instances in which information is required prior to a commitment to large-scale production, in which a product is not far enough along in production to supply one or more entire test markets, in which secrecy is considered desirable, in which there is not enough time for a test market, or in which the test market method has not proved to be particularly useful in the past.

In many instances neither a test market nor an experimental simulation can be used to predict demand, notably in the case of large, durable products such as cars or major appliances, new services not as yet adapted to the consumer market, or new products only in the concept stage that can neither be sold nor shown except, perhaps, in the form of pictures. In such instances, a research survey of the likelihood of purchase is needed.

THE TROUBLE WITH ESTIMATING DEMAND BY SURVEY

Studies of the likelihood of purchase are probably as old as research itself, and over the years, they have been tried with many variations and with various degrees of technical expertise. But the issue almost always boils down to a question such as "How likely are you to buy_____?" All too often, results of such studies have been impressively wrong, usually ending in estimates of buying intention substantially higher than subsequent levels of actual purchases. In certain cases—notably in attempts to assess interest in a new type of product that does not yet exist—there are instances in which 10% or less of those who said they would buy actually did so when the product reached the market. Of course, the same problem occurs in the case of existing products, too. Research literature abounds with examples of gross discrepancies between stated purchase intentions and subsequent purchasing behavior. A recent example from our own work concerns automobile buying. Of a total of 72 persons who reported that they planned to buy a car within

the next six months, only 33 (46%) actually had done so when they were reinterviewed after six months.

The reasons for such overstatement by survey respondents have been explored and discussed many times. The desire of survey respondents to be agreeable—to say "Yes"—particularly because it costs them nothing to do so, has often been cited. Moreover, there is an ego-enhancing aspect to reporting that the purchase of a new car, for example, is imminent. It has been generally agreed that this overstatement phenomenon is not due to conscious falsification but to a failure to consider seriously all of the factors which will play a part when the time of actual purchase comes. In the absence of such serious consideration, it is easiest to say "Yes."

The technique discussed in this paper is based on the experience that most survey respondents are not capable of giving considered answers to direct questions on purchase likelihood. The technique is based on a thorough study and understanding of the particular purchasing process involved, the identification of key factors which are likely to affect this purchase and a quantification of how and to what degree these factors are operating for each survey respondent to affect his likelihood of actual purchase. The method may be termed "a systematic, subjective computation of purchase probability," and it entails the use of subjective probabilities that differ somewhat from classical statistical probabilities. The latter derive either from the full understanding of the mechanics of a process (e.g., dice shooting, coin tossing) or from extensive empirical observation so that they are precise or very nearly precise probabilities. Subjective probabilities on the other hand are not nearly so precise. They are essentially educated guesses derived from judgment, experience, and so on, and generally are used under conditions of uncertainty which very frequently characterize business decisions.

THE USE OF PROBABILITIES IN ESTIMATING DEMAND

In order to demonstrate most clearly how the procedure works, a hypothetical study involving an estimate of consumer demand for a new, atomic-powered cabin cruiser is used for illustrative purposes. While this is a greatly simplified example, it will serve to show how this statistical technique could be used in a variety of subject areas. The essential details of this hypothetical study are that it consists of personal interviews with 200 men who report that they expect to buy a cabin cruiser within the next year. This sample of men is selected through telephone screening to be representative of those "in the market" for a cabin cruiser. These men are shown pictures of the new cruiser as well as a detailed written description and are questioned as to their likelihood of buying it. In addition, answers are solicited to questions relating to the various factors that will be used later to assess overall purchase probability.

TABLE 1. Likelihood of Purchasing Cabin Cruiser as Expressed by Respondents

	NUMBER	PERCENT
Total Respondents	**200**	**100**
Definitely intend to buy	10	5
More likely than not to buy	70	35
Not likely to buy	120	60

The results of the key purchase likelihood question are given in Table 1. Forty percent of the respondents have expressed at least a moderate level of interest in buying the cabin cruiser. Yet it is unreasonable to assume that all of these, in fact, will purchase one. In other words, not every person who has reacted positively has a probability of 1.0 of purchasing it. At this point, the process of making a more realistic assessment of the purchase probability of each respondent starts.

To begin with, the assumption is made that those who have expressed a definite purchase intention are more likely to become actual purchasers than those who have merely admitted to the likelihood of purchase. Therefore, the latter group will be "weighted down" to reflect this lower purchase probability. There is no sure way to determine precisely what weight to apply in the absence of any hard data comparing expressed intention and subsequent purchase of cabin cruisers. Such data have been compiled for other products at other times, however, and can at least serve as a rough guide. For example, in the automobile study cited earlier, the probability of purchase for those who claimed to be planning a purchase was 0.46. Another guide comes from the experience of the researcher and his associates whom he may consult in order to arrive at a weighting scheme based on a "jury of informed opinion." In this hypothetical instance, the consensus is that respondents who have said they are "likely to buy" will receive a weight of 0.4, so that at this intermediate stage, the estimate of demand would be

10 "definitely intend to buy" × weight 1.0 = 10
plus (70 "likely to buy" × weight) 0.4 = 28
divided by 200 = 19% purchasing.

The remainder of the process entails similar weighting of respondents based on various purchase-related factors, thereby further modifying purchase probability. We could argue that further modification may be unnecessary as respondents have considered these purchase factors in answering to the buying intention question. However, there is considerable opinion (Juster 1966) to the effect that responses of this type are, on the whole, not considered ones and that modifying factors do need to be taken into account.

TABLE 2. Weighting the Likelihood that Respondents Will Buy Cabin Cruisers by Time of Intended Purchase

EXPECT TO BUY	WEIGHT
Within 3 months	1.0
In 3 to 6 months	0.8
In 6 to 12 months	0.6

There is likely to be a considerable number of such factors, but for the sake of simplicity, only four will be considered here. Two types of factors exist, the first bearing upon the likelihood of respondent's purchase of *any* cruiser at all within the next year despite his declaration that he will (because this was a requirement for eligibility in this study) and the second bearing upon the likelihood that he will select the particular cruiser being studied.

Of those factors bearing on the purchase of any cruiser, the first concerns just how soon the purchase is expected to be made—within three months, in three to six months or in six to twelve months. Experience indicates that the closer the intended purchase, the more likely that the purchase, in fact, will be made; in other words, those who say they expect to buy within three months are more likely to make actual purchases than those who expect to buy in three to six months, who, in turn, are more likely to become actual purchasers than the six-to-twelve month group. After another round of discussions among several knowledgeable researchers, the weighting scheme shown in Table 2 is developed. In the automotive study cited previously, the data clearly showed that the discrepancy between stated intention and actual purchase was substantially less when the intention was to buy within three months of the time of interview than when the projected purchase data was four or more months away.

The second purchase-related factor deals with the attitudes of the wives of married respondents toward the impending purchase. In this hypothetical instance, the three types of responses coded in the study are shown in Table 3, along with the weights that again are decided by consensus among a group of researchers.

TABLE 3. Weighting the Likelihood that Respondents Will Buy Cabin Cruisers by Wives' Attitudes

RESPONSE	WEIGHT
Wife has no voice in decision	1.0
Wife has a voice and wants to buy a cruiser	1.0
Wife has a voice and prefers alternative purchase	0.6

TABLE 4. Weighting the Likelihood that Respondents Will Buy Cabin Cruisers by Concern over Atomic Power

RESPONSE	WEIGHT
Unconcerned over use of atomic power	1.0
Evidences concern over atomic power	0.4

The third purchase factor bears on the likelihood of purchasing the particular cruiser being studied and deals with the degree of apprehension on the respondent's part about the fact that this is an atomic powered cruiser and might therefore be dangerous in the event of improper shielding of the power source. The weights agreed upon after discussion are shown in Table 4.

The final purchase factor that also bears on the likelihood of selection of the particular cruiser studied concerns the degree to which it fits in with individual price requirements. Those who are planning to spend either more or less than the expected retail price of the cabin cruiser are weighted down as in Table 5.

Using the various weights which have been described, each respondent's probability of purchase is assessed. To illustrate the process:

(1) Respondent A, after seeing a picture of the cabin cruiser and reading a description, reports that he "definitely intends to buy": weight = 1.0.
(2) He claims he will make the purchase within three to six months: weight = 0.8.
(3) He is married but reports that his wife will have no voice in the purchase decision: weight = 1.0.
(4) He reports no apprehension over the use of atomic energy as the power source for this cruiser: weight = 1.0.
(5) The amount he is planning to spend in his upcoming purchase is somewhat less than the expected price of the cruiser: weight = 0.3.

TABLE 5. Weighting the Likelihood that Respondents Will Buy Cabin Cruisers by Consistency with Price Requirements

RESPONSE	WEIGHT
Price in line with plans	1.0
Price higher than planned	0.3
Price lower than planned	0.8

These weights provide the basis for estimating overall purchase probability. Each weight itself is a limited sort of purchase probability—what is called a *marginal probability*. To illustrate how several marginal probabilities are combined to form an overall probability, consider the following illustration.

> The probability of obtaining "heads" in a simple coin toss is 0.5, a marginal probability. The probability of tossing two consecutive heads is 0.25, the product of the marginal probabilities for tossing "heads" on the first toss and for tossing "heads" on the second toss $(0.5) \times (0.5) = 0.25$.

Thus, respondent A's overall purchase probability is simply the product of the various weights (or marginal probabilities): $(1.0) \times (0.8) \times (1.0) \times (1.0) \times (0.3) = 0.24$. The sum of these purchase probabilities across all respondents represents the estimate of demand for the new cabin cruiser. In all, 80 respondents had originally reported that they "definitely intend to buy" or are "likely to buy." If one were to take these responses at face value, he might conclude that as many as 40% (80 of 200) of those in the market may buy this cruiser, though one would be hard pressed to find a modern researcher who would be willing to make so liberal an interpretation of these data. Implicit in this type of estimate is that each respondent who reacts in a positive way is treated as having a purchase probability of 1.0. A more stringent estimating method reduces this purchase probability to 0.4 for those who are only "likely to buy," but leaves the probability at 1.0 for those who report that they will "definitely buy." This, in fact, is how the intermediate estimate of 19% cited earlier was obtained.

Using the system of weighting for various purchase factors brings the estimate lower still; for no one receives a purchase probability of 1.0 without specifically "earning it." In other words, even respondents who will "definitely buy" must give all the "right answers" with respect to the four purchase factors in this illustration before they receive a purchase probability of 1.0. Thus, it would not be at all unlikely that the sum of probabilities for the 80 respondents who reacted positively to the cruiser might be in the area of 5.0 leading to an estimate of demand of 2.5% (5/200), a far cry from 40% and 19%.

In using this subjective probability technique, we must observe one caution. We must overcome the temptation to include an overlong list of factors because if one keeps on multiplying probabilities for too long, the results will come dangerously close to 0. Thus, the factors chosen should be only those that are important purchase influences. In any given case, there are many marginal factors which have only small effects. The analyst must exercise good judgment in limiting the list to the key factors.

Estimates developed through this technique have been accurate. In the study of automobile-buying intentions, for example, the data were highly predictive of actual consumer demand six months later. Although the rule of confidentiality of client data precludes disclosure of specific figures, predicted total sales of the make of automobile as well as various component models were very close to subsequent actual demand.

A PRACTICAL RATIONALE

The technique that has been described here is very unlike the exacting and precise statistical procedures used in sampling and data analysis, described elsewhere in this volume. The development of the weighting schemes appears to be a rather subjective process that is seemingly defenseless against the question: "Why did you choose to use a weight of 0.8 and how can you prove it is better than a weight of 0.6?" It is *not* possible to prove in advance that 0.8 is a better weight than 0.6. However, this lack of certainty should not be permitted to create a "hang-up" to eliminate the use of weights because the means of arriving at them is less than precise. It seems quite clear that the use of weights that are arrived at intelligently is better than having to assume that the person who says "Yes, I will buy," in fact, will do so. In other words, it makes sense to assume that, all other things being equal, the man who is concerned over radiation is less likely to buy an atomic-powered cruiser than one who is not. The method described here makes use of this knowledge on the theory that it would be less accurate to ignore it.

In other respects, too, the method described here is less precise than classical statistical theory would dictate; for example, the assumption of the independence of the various weights is made. Despite some of these loose ends in the theoretical framework of the technique, however, we see only one practical limitation on its usefulness: the skill of the practitioner. Successful application depends on the ability to understand the essential workings of the particular purchasing process being studied and the pinpointing of those key factors that will affect purchasing behavior. As we gain more experience, it is not unlikely that some standardization of weights according to product class may be achieved (e.g., for cars, major appliances, minor appliances) thereby increasing the general usefulness of the method.

On a pragmatic basis, the technique has proved to be a valuable one. It gives good, realistic results and appears to offer a viable solution to the problem of determining purchase likelihood via questionnaire.

REFERENCE

F. Thomas Juster. 1966. *Consumer Buying Intentions and Purchase Probability.* National Bureau of Economic Research.

We can now combine the distributions for males and females. The resulting distributions are shown in Figure 4. We note that the distribution for product C is shifted to the right, the less acceptable scores, as compared to the H results. There is considerable overlapping of the distributions representing the variation, so even on a pictorial basis we have problems interpreting the relative acceptability of the test products. We must resort to the average scores, which for the 50 individuals who tasted the same product are the following:

C = 0.22 Liquid H = 1.18 Solid H = 1.06.

Our yardstick for judging differences decreases, because these new average scores are based on 50 people. Our new yardstick, taking a 1-in-20 chance of being wrong, is 0.45.

It is apparent the difference between the acceptability of the two H products is not large enough to say one is more acceptable than the other. There is very little risk, however, in concluding that both H products are more acceptable than C, the competitive product.

Thus, through testing and with the application of statistical methodology these new products were shown to be palatable and to live up to the concept of a high-protein food. Two of the early criteria in the long process of introducing a new food product have been met.