

Retail merchandise buyers are shown to exhibit a nonregressive bias when making sales projections. A quantitative model based on the principle of statistical regression is found to outperform the judgmental sales predictions of experienced buyers. Implications for the appropriate roles of intuitive and model-based decision making in retail merchandise buying are discussed.

## Heuristics and Biases in the Intuitive Projection of Retail Sales

Over the years, marketing scholars have argued that managers' judgments are indispensable in marketing decisions (cf. Little 1970; Levitt 1978), spurring a growing body of research into marketing managers' judgmental processes (cf. Best 1974; Chakravarti, Mitchell, and Staelin 1979; Larréché and Moinpour 1983; Moriarty and Adams 1984). Though these studies have shed much light on marketing decision making, they have largely overlooked a great potential resource in the study of this phenomenon: the rich psychological literature on heuristics and biases in human judgment (for a review, see Einhorn and Hogarth 1981).

We apply findings from the human judgment literature to an important but little studied domain of marketing decision making, item sales projections of professional retail buyers. After describing this decision setting, we discuss its normative and behavioral implications. We then present empirical results on the heuristics used by, and biases of, actual retail buyers. Finally, we develop and test a model designed to improve on buyers' forecasts.

### THE DECISION SETTING

Retail merchandise buyers, according to one retail CEO, are "the most important factor in the success or failure of any retail venture" (Rachman 1979, p. 42). One of the most important and difficult tasks of these buyers is extrapolating the initial sales rates of new styles of merchandise (Seegal 1979). One reason these sales projec-

tions are so difficult is that they must be made quickly. In some cases a reorder placed beyond the second or third week an item is on the floor may arrive after the item is out of style or season (Seegal 1979). In addition, manufacturers, particularly in the apparel industry, commonly produce little if any merchandise in excess of their advance orders from retailers; therefore, retail buyers must place reorders very quickly if they want to receive additional stock (Goldberg 1979). Because of this urgency, buyers often must project sales on the basis of very limited data.

Consider, for example, a buyer who places an initial order on 10 styles of sport shirts, all from the same vendor, differing primarily in the print on the fabric. When ordering these items, the buyer may have no significant prior information indicating which style will sell best. Therefore, the buyer initially orders 24 units of each style. After one week on the salesfloor this group of styles has sold a total of 50 units. The buyer must place any reorder now or risk missing out on the manufacturer's leftover inventory. Therefore, the buyer must project each style's sales on the basis of one week of sales data.

One can view this buyer's forecast as having two components. One is the forecast of total sales for all 10 styles—essentially a projection of the "sales curve" for this group of items. This type of projection has been discussed elsewhere (cf. Carlson and Tully 1976; Vreeland 1963) and, though important, it is not the focus of our study.<sup>1</sup> Instead, we examine the second component of

---

\*Anthony D. Cox is Assistant Professor of Marketing, College of Business Administration, Georgia State University. John O. Summers is Professor of Marketing, Graduate School of Business, Indiana University.

---

<sup>1</sup>The sales curve for the *group* of items (i.e., *total* sales for this particular manufacturer's line of sport shirts as a function of time) will be determined by several factors such as the tendency for the basic fashion "look" these styles represent to diffuse, the weather, store promotions, etc.

the buyer's forecast: *What proportion of the group's sales total should be projected for each individual style?*

#### NORMATIVE PREDICTION OF AN ITEM'S PROPORTIONAL SALES RATE

Any single period's sales for a particular style reflect at least two factors, (1) the inherent salability of that product and (2) the chance events of that particular period. Therefore, if style A sold much better than the average style during the first week (e.g., 24% of the total for the 10 styles in the shipment), we can infer not only that A probably will continue to sell better than the average style in the shipment, but also that its first week's sales were likely to have benefited from chance events—for example, the mix of customers who happened to walk into the store that week. Therefore, when predicting the *future* sales of item A, the buyer *should* predict a rate of sales higher than that of the *average* item (i.e., 1/10 or 10% of the total), but should *not* predict a sales rate as high as A's *initial* rate of 24% because that rate is likely to have been partially due to "luck." In other words, the sales of the first week's best sellers are likely to move toward the mean (i.e., decrease as a proportion of the total for that line) during the second week. Similarly, the sales of the first week's slowest sellers also are likely to move toward the mean (i.e., increase) during the second week, because we can infer from the low initial sales rate that sales of these items were likely to have been hurt by the chance events of the first week.

The principle we have just described, of course, is called "statistical regression" or "regression toward the mean." It is a ubiquitous phenomenon explaining not only products' sales fluctuations, but a wide variety of phenomena including "[why] most outstanding fathers have somewhat disappointing sons . . . the ill adjusted tend to adjust, and the fortunate are eventually stricken by ill luck" (Kahneman and Tversky 1973, p. 249).

To illustrate the statistical regression phenomenon in actual retail sales data, we analyzed the first two weeks' sales data for six different groups of apparel styles (including children's apparel, women's blouses, and lingerie) obtained from two different department stores. Each group of styles was fairly homogeneous and each style within a group had been ordered in identical or nearly identical quantities. In every group, the majority of the styles' proportional sales regressed toward the mean between weeks 1 and 2. Of a total of 48 styles examined, 37 or 77% regressed toward the mean. This percentage is significantly greater than 50% ( $p < .001$ ). Given the pervasiveness of statistical regression effects in everyday life, this result should not be surprising. As stated by Kahneman and Tversky (1973, p. 249), "regression effects are all around us."

#### PSYCHOLOGICAL RESEARCH INTO JUDGMENTAL PREDICTION

Over the last 20 years, there has been a growing body of research on humans' statistical judgments (for a re-

view, see Einhorn and Hogarth 1981). This work has shown repeatedly that human judges often use heuristics that differ fundamentally from normative statistical models. For example, researchers have found such deviations in subjects' judgments about the likelihood of events (Tversky and Kahneman 1973), growth processes (Wagennar and Timmers 1978), and covariation (Jennings, Amabile, and Ross 1982).

Most relevant to our topic is the experimental work (cf. Bar-Hillel 1980; Kahneman and Tversky 1973, 1982; Tversky and Kahneman 1974) investigating intuitive predictions in probabilistic settings like that faced by the buyer in our example. These studies have found that human judges tend to adhere to the "representativeness" heuristic, predicting the outcome that is most representative of, or similar to, the diagnostic data. In our example, if one style accounts for 24% of total sales during its first week and there appears to be no causal rationale for a drop in its sales, the representativeness heuristic would dictate a prediction of 24% for that style's future sales rate. This heuristic differs from the "regression principle" in that (1) error in the diagnostic data (e.g., the inherent randomness in a week's sales) and (2) the mean (e.g., the sales of the average product) are both ignored. In a typical experiment by Kahneman and Tversky (1973, p. 250), graduate psychology students were asked to estimate a 95% confidence interval for the true IQ of a hypothetical person scoring 140 on an IQ test. As 140 is an extreme score, and as there is a considerable random element in a single IQ test score, we can infer not only that this individual is probably "smarter" than average, but also that this particular score probably benefited from chance factors—for example, how lucky the person was in his guesses. If one assumes a standard testing error of 10 points and a population standard deviation of 15 points, it can be shown that (though it is *possible* for the person's true IQ to be 140 or even higher) the best point estimate for this individual's IQ is about 128.<sup>2</sup> Despite the presumed statistical sophistication of the subjects, however, 68% centered their confidence intervals around 140, 10% around scores *above* 140, and only 22% made regressive predictions.

Kahneman and Tversky (1982, pp. 137–8) suggest that this pervasive difficulty in accepting the regression phenomenon may be due to its apparent violation of certain popular statistical intuitions—for example, that "any systematic effect must have a cause." Accepting some concrete *cause* for the criterion value to be less extreme than the diagnostic cue may be easier than accepting such an abstract concept as statistical regression.

<sup>2</sup>For a fuller understanding of this estimate, see Mood and Graybill (1963) and Nunnally (1978). Mood and Graybill show how a true score can be estimated on the basis of a sample score and the test-retest correlation. Nunnally shows how the test-retest correlation can be derived from the population standard deviation and the standard testing error.

### Research Question 1

The work of Tversky, Kahneman, and their colleagues suggests our first research question:

Do department store buyers follow the regression principle, the representativeness heuristic, or other processes when predicting products' proportional sales rates on the basis of their initial sales?

If buyers' forecasts *do* show a nonregressive bias, this fact would suggest that a regressive model might be able to improve on their forecasts. Previous research has shown the ability of simple quantitative forecasting models to outperform a variety of experienced judges (Hogarth and Makridakis 1981; Meehl 1954). The superiority of these models is due both to judges' biases such as the "regression fallacy" and to judges' inconsistency (Bowman 1963; Dawes and Corrigan 1974). To extend previous work on the "model versus man" issue, we develop a regressive sales projection model. If one assumes that item sales in weeks 1 and  $j$  are jointly distributed as roughly the bivariate normal, one can show that an item's proportional sales in period  $j$  (i.e., as a proportion of total sales for the line) should be predicted normatively by the following formula (see the Appendix for a detailed discussion).

$$(1) \quad \hat{P}_{ij} = 1/n + \rho_{1j}(P_{i1} - 1/n)$$

where:

$\hat{P}_{ij}$  is predicted proportional sales of item  $i$  in week  $j$  (i.e., as a proportion of total sales for the line),  
 $P_{i1}$  is actual proportional sales of item  $i$  in week 1 (i.e., as a proportion of total sales for the line),  
 $\rho_{1j}$  is the correlation of item sales between weeks 1 and  $j$ , and  
 $n$  is the number of items in the line (note that the mean proportion for the  $n$  items must always be  $1/n$ ).

Observe that sales for any given item  $i$  (as a proportion of total sales) will always be predicted to move toward the mean proportion for the line (i.e.,  $1/n$ ) except in the unlikely case that  $\rho_{1j} = 1$ , in which case  $\hat{P}_{ij} = P_{i1}$ .

### Research Question 2

The accuracy of this model in predicting the proportional sales rates of a line of fashion products is compared with that of a sample of retail buyers. To be consistent with Meehl's (1954) desiderata for model versus man comparisons, all calculations for the model are based on data that are also available to the buyers. Hence, our second research question is:

Will a regressive forecasting model outperform experienced retail buyers in predicting products' proportional sales when both are given the same data?

Two studies were conducted, one to address each research question. Study 1 examined the heuristics and biases in buyers' forecasts; study 2 compared the buyers' predictive accuracy with that of a forecasting model.

## STUDY 1. IDENTIFYING HEURISTICS AND BIASES IN BUYERS' FORECASTS

### Research Method

**Sample.** We believed it was imperative to use real managers as subjects (cf. Larréché and Moinpour 1983, p. 119). One criticism of much human judgment research is subjects' lack of experience with the particular phenomenon they are predicting. Such experience could conceivably affect the regressiveness of predictions, because experienced subjects are likely to have a clearer notion of which predicted outcomes are *plausible* and managers appear to adjust initial forecasts that seem implausible (Acito and Olshavsky 1983). In addition, anecdotal evidence suggests that experienced observers may sometimes *detect* regression toward the mean, whether they "understand" it or not (Kahneman and Tversky 1973). For these reasons, we decided to recruit actual retail buyers to project retail sales data.

Forty-one apparel and accessories buyers were recruited from five traditional department stores in the Midwest to participate in the research. Thirty-one of these subjects participated in study 1 and the remaining 10 participated in study 2. Ninety percent of the subjects had at least six months of buying experience. Prestudy interviews strongly indicated that apparel and accessories buyers, because of rapid merchandise obsolescence, often must project item sales on the basis of early returns, sometimes within a week. Furthermore, fashion apparel and accessories account for a major portion of department store sales.

**Data collection.** Ideally, we would have preferred to observe buyers projecting sales during the course of their daily work. However, as prestudy interviews indicated that buyers make projections at unpredictable times (because of frequent interruptions and schedule changes) and places (e.g., at home), much waiting would have been required to collect enough observations. Such difficulties have forced some researchers to spend as much as a year collecting field data on a single manager (e.g., Clarkson 1962). Because we believed field observation was infeasible, we gave the subjects a simulated sales forecasting task.

**Procedure and task.** One of the researchers met individually with 31 retail buyers at their respective firms. The buyers were shown the first week's retail sales and inventory data for each of 12 handbag styles and were asked to project each style's sales over the next three weeks. Other information also was provided, which preliminary interviews had revealed would normally be available to buyers making such forecasts: a profile of the store in which the line was being sold, a sample of the merchandise, the dollar sales for the entire handbag department during the initial week, and the planned department dollar sales for this and the subsequent three weeks.<sup>3</sup>

<sup>3</sup>To determine whether subjects were truly following the regression principle, the scenario also had to preclude realistically *other* causes

Several measures were taken to make this task as realistic as possible. First, before data collection, about 25 retail buyers were interviewed on sales projections and related topics. This information served as a "reality check" in the initial task design. Second, the task went through four iterations of pretesting and revision to make it more realistic. The pretest subjects included two current retail buyers, one former buyer, and one marketing professor who has conducted extensive field research on retail buying. The final pretest subject (a former department store buyer) stated that the task was "very plausible; this is basically the kind of data you work with." None of the subjects in the study expressed doubt that the scenario was real.

Two types of data were recorded, the subjects' numerical forecasts and their verbal protocols as they developed the forecasts. Concurrent protocols were collected to aid interpretation of subjects' forecasting processes. Researchers of intuitive prediction (e.g., Kahneman and Tversky 1982) have *speculated* about subjects' mental processes, but have not systematically marshalled process data to support their speculations. Verbal protocols also can help determine that subjects are responding to the scenario as presented, rather than (as Jenkins 1983 found) ignoring aspects that differ from their own work environments. At the start of each session a recorder was turned on; as expected (cf. Gordon 1980), subjects' consciousness of the machines appeared to decrease considerably after a few minutes, reducing the potential for reactive measurement.

### Analysis and Results

**Numerical analysis—regressiveness of buyers' forecasts.** To determine whether a subject's sales forecast for a given product was consistent with the regression principle, it was useful to state this forecast as

$$(2) \quad \hat{P}_{ij}^* = 1/n + k(P_{i1} - 1/n)$$

where:

$\hat{P}_{ij}^*$  is the proportion of total unit sales the subject *predicted* for item *i* in week *j*,

$P_{i1}$  is the *actual* proportion achieved by item *i* during the first week, and

$1/n$  is the average (across all items) proportional sales, simply equal to one divided by the number of products in the group

The variable *k* reflects the degree (if any) to which the subject predicted a product's proportional sales to regress toward the mean in week *j* relative to week 1. For example,  $k = .85$  would indicate that the subject pre-

dicted a product's sales to regress .15 (i.e.,  $1 - .85$ ) of the distance toward the mean. If  $k = 1.0$ , the equation reduces to  $\hat{P}_{ij}^* = P_{i1}$ , indicating the subject's prediction did not regress toward the mean at all. Because the values of all of the other variables in equation 2 are observed, a value of *k* can be calculated for each subject's individual predictions (e.g., sales for item 6 in week 3,  $\hat{P}_{63}^*$ ). The extent to which each subject regressed his or her predictions was summarized by averaging *k* across all of that subject's predictions.

The buyers regressed their forecasts only slightly, particularly for the week following the line's introduction. The mean *k*'s for weeks 2, 3, and 4 are, respectively, .947, .855, and .803. Thus, in week 2 subjects regressed their predictions an average of .053 (i.e.,  $1 - .947$ ) of the distance toward the mean, in week 3 about .145 toward the mean, and in week 4 about .197.

For comparison, we obtained some actual interperiod sales correlations for comparable fashion merchandise. They indicate the approximate degree to which subjects normatively *should* regress their predictions toward the mean according to the predictive model of equation 1. Actual interperiod correlations calculated from 11 sets of historical retail sales data (obtained from a large mid-western department store) produced mean correlations<sup>4</sup> of .57, .49, and .48 for weeks 2, 3, and 4, respectively. According to these data and equation 1, subjects *should* have regressed their predictions approximately .43 (i.e.,  $1 - .57$ ), .51, and .52 of the distance toward the mean in weeks 1, 2, and 3. *None* of the buyers' sales projections approach being as regressive as these historical correlations suggest their projections should be. For each week, buyers' mean degree of regressiveness is significantly ( $p < .05$ ) less than the amount dictated by these historical sales correlations.

**Process analysis—buyers' forecasting heuristics.** The verbal protocols revealed that most subjects used some form of the representativeness heuristic in making their forecasts. For example, more than a third of the subjects predicted a constant "sell-through" from week to week: if a style with a beginning inventory of 32 units sold eight units during its first week, the sell-through would be 25%; in the second week, a buyer using this approach would predict a 25% sell-through on the remaining 24 units, or sales of six units. Protocol segments from two typical subjects stated:

In the first style in the first week she had a 3% sell-through . . . therefore, I have no reason to believe she's going to sell more than 3% the following week.

If you're selling a certain percentage of your stock and [circumstances do not change] there's no reason you're gonna sell a greater percentage from week to week.

for a moderation of styles' initial sales rates. In particular, initially fast sellers might quickly run out of key sizes and colors, and their salesfloor visibility might decrease. The scenario precluded stockouts by depicting a product class (handbags) in which each style might come in a single size and color. The visibility issue was eliminated by stating that six units of each style were maintained continually on the salesfloor, replenished from backstock.

<sup>4</sup>These *r*'s were *not* averaged directly. Individual *r*'s were transformed by using Fisher's inverse hyperbolic *z* transformation, the mean of these *z*'s was calculated, then  $\bar{X}_z$  was back-transformed to obtain  $\bar{X}_r$ . For a discussion of the rationale for this procedure, see Morrison (1976).

The popularity of this approach partially explains why some buyers' predictions seemed to "regress" slightly toward the mean: those items that sell fastest will also have the fastest drop in inventory, so that if their sell-through remains constant, their unit sales will drop somewhat in relation to other items. However, the amount of this "regression" was always considerably less than that dictated by the statistical regression model. In addition, the *rationale* for this approach is essentially that of the representativeness heuristic. Retail buyers, always conscious of inventory turnover, often characterize the "hotness" of a new product in terms of its sell-through. For example, a "good" item may be one that sells 10% of its stock in a week. Therefore a constant sell-through would seem most representative of (or similar to) its initial sales rate. From the preceding quotations, these buyers appear to have felt they were merely projecting the status quo.

Only two of the subjects using the sell-through rule mentioned anything akin to the regression principle. Both suggested that extreme initial sales might result partially from peculiarities of the first week.

The ones that were 1's and 2's I planned up a little bit because . . . just because if it was a reasonable purchase, and you picked it the way you picked the others, that it probably was a fluke that it only sold one.

. . . the weak ones would probably get a little better; maybe they were mistreated.

About a quarter of the buyers simply predicted each style to have the same unit sales in the *second* week as in the first week—a clear use of the representativeness heuristic. A typical quote is:

. . . as long as it's not broken [i.e., no stockouts] I'd make it the same straight across.

Some buyers more or less arbitrarily added or subtracted units from a style's initial sales to predict its subsequent sales, essentially predicting future sales to be "representative" of initial sales plus a random component. For example, one subject stated:

Since we sold 5 here, we might sell 1 more or less the following week . . . one style's not going to sell the same every week.

By making the predicted sales patterns "realistically erratic" these buyers seemed to feel they had a better chance of predicting accurately.

Finally, a few subjects predicted that the fastest selling styles would accelerate in relation to other styles. One of these subjects said:

Generally, in my experience, if the first week is good, the second week is even better; I don't know whether it's word of mouth. . . . I don't see these fringe items as getting better . . .

In summary, though several subjects' forecasts regressed *slightly* toward the mean, only two appeared to use anything like the regression principle. Consistent with

previous research findings, most subjects made their forecasts by using some form of the representativeness heuristic. Given this underregressive bias in subjects' forecasts, one would expect a regressive model to improve on their forecasts. This issue is investigated next.

## STUDY 2. COMPARING THE PREDICTIVE ACCURACY OF THE BUYERS AND THE MODEL

### Research Method

One of the researchers met individually with 10 additional buyers. Each buyer was shown one week of *actual* sales and inventory data for a line of 12 blouse styles obtained from a midwestern department store and was asked to project each style's sales for each of the following three weeks. In this study, unlike study 1, we were able to obtain the products' *actual* sales for the three subsequent weeks and hence could assess predictive accuracy. Of course, neither the buyer nor the model had these subsequent data and none of the 10 subjects was selected from the firm that had provided the sales data for this task. These data did allow for occasional size or color stockouts, but this did not seem to be a severe problem. As in the first study, the buyers were given information about the store, a sample of the merchandise line, and the blouse department's weekly planned sales for the relevant test period. Finally, the buyers received weekly sales and inventory data from the *preceding* year for a *similar* line of blouses. These data allowed the observation of the interperiod correlation of sales across individual styles for comparable fashion items.

### Analysis

The performance of the regressive model (equation 1) was compared with that of the 10 buyers projecting blouse sales. The analysis focused on item sales as a proportion of total sales for each of the three weekly forecasts. The

mean absolute deviation  $\left( \text{MAD} = \left( \sum_{i=1}^n |P_{ij} - \hat{P}_{ij}| \right) / n \right)$  was used as the measure of forecasting accuracy.

Four methods of estimating  $p_{ij}$  were used to parameterize the model: (1) the interperiod sales correlations ( $r_{12} = .95$ ,  $r_{13} = .5$ ,  $r_{14} = .49$ ) calculated from the actual *previous year's* blouse sales data presented to the buyers, (2) the median of  $r_{12}$ ,  $r_{13}$ ,  $r_{14}$  (used for all forecasts), (3) the average interperiod item sales correlations<sup>5</sup> ( $r'_{12} = .57$ ,  $r'_{13} = .49$ ,  $r'_{14} = .48$ ) for 11 sets of sales data for comparable merchandise lines collected from the same department store, and (4) the median of  $r'_{12}$ ,  $r'_{13}$ ,  $r'_{14}$  (used for all forecasts).

### Results

Overall, the model outperformed the intuitive judgments of the retail buyers regardless of which procedure was used to parameterize the model (see Table 1). The four estimation procedures produced improvements in

<sup>5</sup>Average  $r$ 's calculated as discussed in footnote 4.

Table 1  
MEAN ABSOLUTE DEVIATION (MAD) FROM ACTUAL PROPORTIONAL SALES  
OF SUBJECTS' FORECASTS AND MODEL'S FORECASTS

	Week 2	Week 3	Week 4	Average across weeks
Mean MAD across subjects (standard deviation)	.0782 (.0074)	.0397 (.0042)	.0472 (.0170)	.0550 (.0061)
MAD for model (and % of S's "beaten" by model)				
Model estimation method 1 <sup>a</sup>	.0834 <sup>e</sup> (20%)	.0385 (60%)	.0285 <sup>f</sup> (100%)	.0501 <sup>e</sup> (70%)
Model estimation method 2 <sup>b</sup>	.0728 <sup>e</sup> (80%)	.0385 (60%)	.0284 <sup>f</sup> (100%)	.0466 <sup>f</sup> (100%)
Model estimation method 3 <sup>c</sup>	.0713 <sup>e</sup> (80%)	.0385 (60%)	.0289 <sup>f</sup> (100%)	.0462 <sup>f</sup> (100%)
Model estimation method 4 <sup>d</sup>	.0710 <sup>e</sup> (80%)	.0385 (60%)	.0285 <sup>f</sup> (100%)	.0460 <sup>f</sup> (100%)

<sup>a</sup>Weekly estimate of  $\rho_{ij}$  based on previous year's blouse data.

<sup>b</sup>Median of method 1 estimates of  $\rho_{ij}$ .

<sup>c</sup>Weekly estimates of  $\rho_{ij}$  based on 11 sets of sales data.

<sup>d</sup>Median of method 3 estimates of  $\rho_{ij}$ .

<sup>e</sup>The difference between subjects' mean MAD and model's MAD is statistically significant at  $p < .05$ .

<sup>f</sup> $p < .01$ .

MAD (averaged across weeks) ranging from 9 to 16% (.0550 for the buyers vs. .0501, .0466, .0462, and .0460 for the model). These improvements in MAD are all statistically significant at the .05 level or better. Furthermore, the last three model estimation methods produced lower average MAD's (across the three weeks) than *all* 10 retail buyers. The model's poorest performance was with  $r_{12} = .95$ , calculated from the previous year's blouse data. This finding is not surprising because .95 is a highly deviant interperiod correlation (only one other  $r_{ij}$  calculated from the 11 sets of sales data exceeded .65). Aside from this case, the model always produced weekly forecasts superior to those of the buyers as a group, with MAD not varying by more than .005 across the four estimation methods.

It should be noted that this phase of our research has several limitations. First, though drawn from several department stores, the sample for study 2 is small (10) and should be expanded in future research. Second, though the forecasting task was based on actual sales data and considered highly realistic by the subjects, it may differ in subtle but important ways from the tasks buyers face in their daily work. Finally, sales were forecast for a single line of merchandise. A more rigorous test would be to compare the buyers' and the model's forecasting performance for several different lines of merchandise. Given these limitations, our results should be viewed as suggestive rather than conclusive. What is suggested, however, is exciting.

#### DISCUSSION AND CONCLUSIONS

Our findings suggest that, when making sales projections based on initial sales rates, experienced retail buyers display similar biases and use the same types of heu-

ristics as found for naïve subjects in previous psychological research on human prediction. The buyers' past experience in performing similar item sales projections and their previous exposure to actual sales patterns were not sufficient to make them adequately regressive in their forecasts. As indicated by the verbal protocols, the vast majority of the subjects employed some form of the representativeness heuristic in developing their sales projections.

The findings also provide support for Hogarth and Makridakis' (1981, p. 129) observation that "... quantitative, and particularly simple models can outperform humans in a wide range of situations." The regressive model proved superior to the buyers' judgmental forecasts in predicting products' proportional sales rates. Furthermore, the same basic modeling approach (i.e., utilizing the principle of statistical regression) is applicable to an important class of forecasting problems that has largely been overlooked in the normative forecasting literature. Quantitative forecasting methods traditionally have required several periods of data (cf. Wheelwright and Makridakis 1977). However, the model discussed here requires only one period of sales data to forecast the proportional sales rates within a group of products. This model is most appropriate for a group of products having similar expected life cycles and seasonality and for which the forecaster has equal prior sales expectations (e.g., a given manufacturer's new assortment of short-sleeved knit shirts). The model appears particularly useful for products (e.g., fashion goods) whose sales must be forecast on the basis of very early sales data and for which conventional sales extrapolation methods (e.g., exponential smoothing) therefore may not be appropriate.

As noted before, the model forecasts the sales of individual styles as a *proportion* of total sales for the entire product line. Forecasts of styles' sales *in units* could be obtained by combining these proportional projections with a forecast of the entire line's sales in units. This total line forecast could be generated either by the buyer or by another model (e.g., Carlson and Tully 1976; Vreeland 1963).

Our study also suggests the usefulness of applying formal decision tools to retail buying—an area of marketing traditionally dominated by “seat of the pants” decision making. At least one researcher (Hertz 1975) has noted the great unrealized potential for the use of decision models and decision aids by retail buyers. We have developed one such decision aid, which uses readily available data and whose calculations can be made on a hand-held calculator. Future research should be directed at identifying other aspects of buyer decision making that could be aided by simple quantitative models.

In advocating the use of models in retail buying, we do not suggest the elimination of human decision making. Several authors have suggested complementary roles for models and men in forecasting (e.g., Bowman 1963; Meehl 1954; Moriarty 1985). The challenge of future research, drawing from all these points of view, will be to identify the appropriate role of quantitative models in retail buying, as well as in other facets of marketing decision making.

#### APPENDIX

Assume a population of diagnostic cues,  $X_1$  (a set of products' first week sales), and a corresponding set of probabilistic events to be predicted,  $X_j$  (these products' sales in week  $j$ ), are *roughly* distributed as bivariate normal. The formula for the expected value of  $X_j$  given  $X_1$  ( $E(X_j|X_1)$ ) is simple, fairly robust to normality violations, and can be found in most introductory mathematical statistics books (cf. Freund 1971, p. 374).

$$(A1) \quad E(X_j|X_1) = \mu_j + \rho_{1j}(\sigma_j/\sigma_1)(X_1 - \mu_1)$$

where:

- $\mu_j$  = the mean sales across these fashion items in week  $j$ ,
- $\sigma_j$  = the standard deviation across all items in week  $j$ , and
- $\rho_{1j}$  = correlation between  $X_1$  and  $X_j$  (i.e., items' sales in weeks 1 and  $j$ )

For example, if item  $i$  had sales of  $X_{i1}$  units in period 1, one should predict sales of  $\hat{X}_{ij} = \mu_j + \rho_{1j}(\sigma_j/\sigma_1)(X_{i1} - \mu_1)$  units for item  $i$  in period  $j$ .

Dividing both sides of equation 1 by total sales in period  $j$  ( $T_j$ ) produces a model for predicting item sales *as a proportion of total sales* ( $P_{ij}$ ).

$$(A2) \quad \hat{P}_{ij} = \hat{X}_{ij}/T_j = (1/T_j)(\mu_j + \rho_{1j}(\sigma_j/\sigma_1)(X_{i1} - \mu_1))$$

To operationalize this model, values are needed for all variables on the right side of the equation. Estimating  $\sigma_j$

is particularly difficult. Fortunately, however, the authors' analyses of several sets of item sales data for fashion products like those used in the study (e.g., styles of women's blouses) demonstrated that the coefficient of variation ( $\sigma/\mu$ ) for any particular grouping of fashion items varied by less than 10% from week to week. This finding suggests that  $\sigma_j/\sigma_1$  in equation 2 can be replaced by  $\mu_j/\mu_1$  (because  $\sigma_1/\mu_1 \cong \sigma_j/\mu_j$  implies  $\sigma_j/\sigma_1 \cong \mu_j/\mu_1$ ).

$$(A3) \quad \hat{P}_{ij} = (1/T_j)(\mu_j + \rho_{1j}(\mu_j/\mu_1)(X_{i1} - \mu_1))$$

Substituting  $(1/n)T_j$  for  $\mu_j$ ,  $(1/n)T_1$  for  $\mu_1$ , and simplifying the equation gives

$$(A4) \quad \hat{P}_{ij} = 1/n + \rho_{1j}(P_{i1} - 1/n)$$

which is the equation 1 presented in the text.

#### REFERENCES

- Acito, Frank and Richard Olshavsky (1983), “Model to Mimic Managers' Mental Forecasting Process Possible,” *Journal of Business Forecasting*, 2 (Fall), 11–15.
- Bar-Hillel, Maya (1980), “The Base Rate Fallacy in Probability Judgments,” *Acta Psychologica*, 44 (May), 211–22.
- Best, Roger (1974), “An Experiment in Delphi Estimation in Marketing Decision Making,” *Journal of Marketing Research*, 11 (November), 447–68.
- Bowman, E. (1963), “Consistency and Optimality in Management Decision Making,” *Management Science*, 9 (January), 310–21.
- Carlson, Phillip and John Tully (1976), “Fashion Life,” *Southern Apparel Magazine*, 37 (September), 95–6.
- Chakravarti, Dipankar, Andrew Mitchell, and Richard Staelin (1979), “Judgment-Based Marketing Decision Models: An Experimental Investigation of the Decision Calculus Approach,” *Management Science*, 25 (March), 251–63.
- Clarkson, Geoffrey (1962), *Portfolio Selection: A Simulation of Trust Investment*. Englewood Cliffs, NJ: Prentice-Hall, Inc.
- Dawes, Robyn and Bernard Corrigan (1974), “Linear Models in Decision Making,” *Psychological Bulletin*, 81 (February), 95–106.
- Einhorn, Hillel and Robin Hogarth (1981), “Behavioral Decision Theory: Processes of Judgment and Choice,” *Annual Review of Psychology*, 32, 53–88.
- Freund, John E. (1971), *Mathematical Statistics*. Englewood Cliffs, NJ: Prentice-Hall, Inc.
- Goldberg, Joel (1979), “Understanding the Computer as a Merchandising Aid,” in *The Buyer's Manual*, P. Cash, ed. New York: National Retail Merchants' Association, 203–44.
- Gorden, Raymond (1980), *Interviewing: Strategy, Techniques, and Tactics*. Homewood, IL: Dorsey Press.
- Hertz, Paul (1975), “Quantitative Management Science and Merchandising Decision Making in Department Stores: An In-Depth Study of Applications, Implementation, and the Future,” unpublished Ph.D. dissertation, New York University.
- Hogarth, Robin M. and Spyros Makridakis (1981), “Forecasting and Planning: An Evaluation,” *Management Science*, 27 (February), 115–38.
- Jenkins, A. Milton (1983), *MIS-Design Variables and Deci-*

- sion Making Performance*. Ann Arbor, MI: V.M.I. Research Press.
- Jennings, Dennis, Teresa Amabile, and Lee Ross (1982), "Informal Covariation Assessment," in *Judgment Under Uncertainty: Heuristics and Biases*, D. Kahneman, P. Slovic, and A. Tversky, eds. New York: Cambridge University Press, 211-30.
- Kahneman, Daniel and Amos Tversky (1973), "On the Psychology of Prediction," *Psychological Review*, 80 (July), 237-51.
- and ——— (1982), "On the Study of Statistical Intuitions," *Cognition*, 11 (March), 123-41.
- Larréché, Jean-Claude and Reza Moinpour (1983), "Managerial Judgment in Marketing: The Concept of Expertise," *Journal of Marketing Research*, 20 (May), 110-21.
- Levitt, Theodore (1978), "A Heretical View of Management 'Science'," *Fortune*, 98 (December 18), 50-2.
- Little, John (1970), "Models and Managers: The Concept of a Decision Calculus," *Management Science*, 16 (April), B-466-85.
- Meehl, Paul (1954), *Clinical Versus Statistical Prediction*. Minneapolis: University of Minnesota Press.
- Mood, Alexander and Franklin Graybill (1963), *Introduction to the Theory of Statistics*. New York: McGraw-Hill Book Company.
- Moriarty, Mark (1985), "Design Features of Forecasting Systems Involving Management Judgments," *Journal of Marketing Research*, 22 (November), 353-64.
- and Arthur J. Adams (1984), "Management Judgment Forecasts, Composite Forecasting Models, and Conditional Efficiency," *Journal of Marketing Research*, 21 (August), 239-50.
- Morrison, Donald (1976), *Multivariate Statistical Methods*. New York: McGraw-Hill Book Company.
- Nunnally, Jum (1978), *Psychometric Theory*. New York: McGraw-Hill Book Company.
- Rachman, David (1979), "The Organizational Structure of a Retail Firm: The Buyer's Role," in *The Buyer's Manual*, P. Cash, ed. New York: National Retail Merchants' Association, 29-44.
- Seegal, Herbert (1979), "Keys to Profit," in *The Buyer's Manual*, P. Cash, ed. New York: National Retail Merchants' Association, 357-76.
- Tversky, Amos and Daniel Kahneman (1973), "Availability: A Heuristic for Judging Frequency and Probability," *Cognitive Psychology*, 5 (September), 207-32.
- and ——— (1974), "Judgment Under Uncertainty: Heuristics and Biases," *Science*, 185 (September), 1124-31.
- Vreeland, Carl (1963), "The Jantzen Method of Short-Range Forecasting," *Journal of Marketing*, 27 (April), 66-70.
- Wagennar, William and H. Timmers (1978), "Intuitive Prediction of Growth," in *Environmental Assessment of Socio-Economic Systems*, D. Burkhardt and W. Ittelson, eds. New York: Plenum Press, 103-22.
- Wheelwright, Steven and Spyros Makridakis (1977), *Forecasting Methods for Management*. New York: John Wiley & Sons, Inc.

## ADVERTISERS' INDEX

The Ehrhart-Babic Group .....	F-1
NameLab, Inc. ....	p. 270
SAMI/Burke, Inc. ....	Cover 4