MANAGEMENT SCIENCE
Vol. 13, No. 10, June, 1967
Printed in U.S.A.

SALES FORECASTING WITH THE AID OF A HUMAN BEHAVIOR SIMULATOR*†

GEORGE SCHUSSEL

Northrop Corporation

This article describes a recent application of simulation to a forecasting problem. The simulation was of the human behavior characteristics of a diverse group of retail dealers. The purpose of the simulation was to propose and test a new method of forecasting a manufacturer's sales to his retail dealers. A simulation model of dealer behavior was constructed and proved very effective in helping the manufacturer to generate good sales forecasts. Unsophisticated retailers were found to have a sufficiently systematic set of procedures to permit simulation of these procedures by a computer model and it was concluded that effective simulation of the human decision processes of a large non-homogeneous group of businessmen is possible.

The purpose of this article is to report the results of research that was done using simulation of human behavior characteristics as an aid in forecasting. The objective of the research was to test the feasibility of simulating the decision behavior of a large, non-homogeneous sample of people and to test the usefulness to a firm of simulating the external environment in which the firm operated. It was felt that the only valid way to test these points was to construct the simulation in such a manner that it would be used for forecasting. The heart of the research work consisted of a simulation model that was constructed of the film reordering techniques of thirty-three photographic dealers. The simulation was constructed to aid a manufacturer in determining the orders for film that would be placed at his warehouse by these dealers. Due to historical experience, the manufacturer had a good idea of what the retail sales of his various film types were, however, he had difficulty in forecasting his own sales because of the perversity in the way that the dealers reordered film. Because the film product of the manufacturer was very perishable, forecasting sales became an extremely important input to the production scheduling decision.

Even though the manufacturer felt that he could forecast retail sales fairly accurately, his past attempts at converting the forecasts of retail sales into the forecasts of company sales had proved very inaccurate. From this problem came the idea of interviewing dealers and constructing a simulation model of their behavior which would take a forecast of retail sales as input and convert this forecast into a forecast of orders placed at the warehouse.

The basic information for the simulation of dealer behavior was derived from

* Received July 1966; revised November 1966.

† This article is based on a doctoral thesis submitted at the Harvard Graduate School of Business Administration. In the interests of brevity, much of the detail of the thesis was omitted in this report. For this detail, the interested reader is referred to D.B.A. thesis, *Forecasting in the Photographic Industry: Testing a Simulation Model,* at Baker Library, Harvard Business School, Boston, 1966.

detailed field interviewing of the thirty-three dealers in the study. Retail sales forecasts were obtained from company executives.[1]

In order to test the efficacy of the simulation model and to establish a bound on the accuracy of the model, it was necessary to devise a testing procedure for the model. This test consisted of determining the actual retail sales for the group of stores over a fifteen-week period and using these actual sales as input for the simulation model. The output of orders from these dealers was available for the fifteen-week period and with the true sales as input, the basic accuracy of the simulation model was testable. Weekly inventory counting at each store in the study, plus the record of shipments made from the manufacturer were sufficient to determine the actual weekly retail sales over the period of the study.

Several different methods of generating sales forecasts were used in conjunction with the simulation to forecast orders for the fifteen-week period. Also, in addition to the forecasts, several other order forecasts were obtained by more conventional means. These other forecasts also served to test the usefulness of the simulation model as a forecast aid. From the results of the study, it was concluded that simulation of the human decision processes of a large non-homogeneous group is possible. The simulation model performed well when actual retail sales were used as input. It was also concluded that using the simulation in conjunction with forecasted retail sales was the most accurate of the various methods which were examined of forecasting orders to the manufacturer. Three executives had been asked simultaneously to forecast retail sales of, and direct orders from, two dealer samples which were made up from the thirty-three dealers in the study. One sample consisted of nine large volume dealers and the other was made up of the remaining twenty-four smaller volume dealers. In every forecasted case (six) the combination of forecasted sales and simulation of dealer behavior proved more accurate than the direct forecast of orders, even though the executives included in the study were less familiar with forecasting retail sales than orders. General conclusions about the usefulness of this type of procedure for forecasting are made at the end of this article.

The fact that unsophisticated retailers seem to have a sufficiently systematic set of procedures to permit simulation by a computer model is encouraging because it may permit the use of the sales forecasting technique presented in this article by other firms which market products through retail dealers. The only new technique that this article presents involves the use of the simulation as the transfer function between a forecast of retail sales and a forecast of wholesale orders. The necessity of a sales forecast was not alleviated by the use of the simulation model. The use of this technique, however, may permit some firms to achieve more accurate and more easily performed forecasts than the more inaccurate forecasts which are now made.

[1] To eliminate confusion in the remainder of this article, "retail demand" will mean the amount of film sales requested by customers of the dealers; "sales" will refer to the actual sales made by the dealers. The quantity "sales" is always less than or equal to "retail demand" because of film stock-out conditions. The word "orders" refers to sales of film by the manufacturer to the dealer.

The Polaroid Corporation of Cambridge, Massachusetts cooperated in making the research reported here possible. During the time that the research was being done, Polaroid manufactured five types of film which so dominated Polaroid's retail consumer film sales that they could have been considered the entire Polaroid film line. The five types of film had different sales backgrounds in the sense that they had been in the market widely varying lengths of time. At the beginning of the study, film type 1 had been available in the market place for eleven years, while types 4 and 5 had only been sold one and one-half years. Type 2 had been in the market place seven years, while type 3 had been available for two years. Film types 1 through 3 were used by an older type of Polaroid camera which was no longer made, and therefore the sales of these film types were decreasing over time. Film types 4 and 5 were used by the standard production Polaroid cameras and just three months prior to the inception of the study, two new cameras had been brought out which used these film types. The sales of film types 4 and 5 were therefore growing relatively rapidly. Because of the substantial differences of time in the market and the fact that they were in different stages of their life cycles, the five film types helped to generalize the study with respect to product history. This generality is important because it permits more confidence in the success that might be expected by adopting this article's methodology to firms who wish to forecast sales for products other than film.

Methodologies

Four methods were used in the study to forecast dealer orders to the manufacturer. One of these methods was a direct forecast of dealer orders without any intermediate analytical steps, while the other three methods required intermediate analysis before arriving at a completed forecast for orders. These forecast methodologies were essentially parallel approaches to the same problem. The one important difference among the methods is that a direct approach to forecasting orders does not require any other forecast as an input. The two-step forecasts involving the simulation model require the input of a retail sales forecast. The simulation model then acts as a transfer function on this sales forecast.

First Method—Regression on Past Orders

The first approach to forecasting dealer orders was a conventional statistical approach that used the regression technique for forecasting. The input data for this method consisted of the past total monthly quantities of each type of film ordered by the dealers in each sample for two and one-half years. (The manufacturer sold five types of film to the dealers.) These order figures were seasonally adjusted on a monthly basis. The method used was the standard "ratio to moving average" technique [1]. Least squares regression lines with time as the independent variable and orders dependent, were then obtained. A total of ten regression equations were determined, one for each film type in each sample. These regression equations were then used for forecasting with new values for the time variable. An example is:

$$\text{Orders} = 1458.118 - 9.2274 \text{ time},$$

which was the regression equation for film type 1 in the nine-dealer sample. Time is equal to 1 for January, 1963, and is incremented by 1 for every month thereafter, e.g., time is 30 for June 1965 and therefore "orders" has a predicted value of 1181. These resulting figures were then multiplied by the monthly seasonal factors to obtain the forecasted orders. These forecasts were single points, not distributions as some of the other forecasts were.

Second Method—Simulation on Regression Sales Forecast

The regression method outlined above was also used to forecast historical retail sales figures. These forecasts of retail sales figures were then input as point forecasts for the simulation model.

As mentioned before, the simulation technique for order forecasting required retail sales forecasts for input, and the regression sales forecast was one method used for generating these sales forecasts.

Third Method—Executive Forecast of Orders

As an alternative to the regression forecast of retail sales, three executives at Polaroid were asked to forecast retail sales and wholesale orders from the two samples of dealers for the fifteen-week study period. The retail sales forecast was used as part of the input for method 4, and the executives' wholesale order forecasts constituted the "third method." These forecasts were obtained by showing the three executives relevant data on the past sales and orders of the study dealers and asking them to use this information, in addition to their own opinions of future trends, to determine the sales and order forecasts.

Fourth Method—Simulation on Executive Sales Forecasts

The effectiveness of the simulation model was to be primarily tested by comparing the accuracy of the forecasted order output from the simulation model to the forecast of orders made by the company executives. This fourth method used the executive retail sales forecasts as input for the simulation model. The output from this method could be compared with the executives' order forecasts since both the executives and simulation model had the executives' sales forecasts for input. This fourth forecasting model, like the second, used the simulation model.

After the executives had had an opportunity to study historical sales and order data on the two dealer samples, they were individually asked questions about the future sales and orders of the two dealer samples. When they were questioned about sales, five-point probability distributions of expected sales for the first five-week period in the study were obtained. Questions concerning the rest of the study period (the remaining ten weeks) were then asked, conditioned on hypothetical outcomes for this first period. The second period forecasts were therefore conditional on the results of the first period and were handled thusly in the simulation model. The cumulative probability distributions for the first period (five weeks) and the distributions for the second period (ten weeks) were later used in conjunction with a random number generator to generate demand for the simulation.

The output order forecasts from the simulation were probabilistic in nature and were described by a mean and standard deviation. The distributional nature of the forecast output by method 4 was due to two factors. First of all, the input sales forecasts were distributional; and second, because of the use of Monte Carlo techniques, the simulation model itself was inherently probabilistic.

Simulation Test—Simulation of Orders Using Actual Sales

A fifth type of "order forecast" was used to test the accuracy of the simulation model. This test consisted of putting known dealer sales for the fifteen-week period into the simulation model and comparing the model's output of orders with the actual orders that were received by Polaroid.

The actual sales of the dealers were obtained by taking weekly inventories at each store for a period of fifteen weeks (16 inventory takings). This information, plus the record of dealer shipments from the Polaroid warehouse, was adequate for determining the weekly sales of the dealers. Other factors, such as loans of film between dealers, were covered in questions at the end of each inventory taking.

The simulation model, acting as a transfer function, then converted these sales figures into order figures to Polaroid. The inherent accuracy limits on the simulation model were then evaluated by comparing these "forecasted orders" with the actual ones that were received.

Interview Results

Inventories and Ordering

The purpose of carrying an inventory is to have a readily available stock of goods to sell. Although all merchants carry inventories of the products they sell and regularly order to build up their inventories, many different types of systems are used by retailers to maintain their inventories.

The first requirement of keeping an inventory is to have some method of knowing when the inventory has fallen to a point where a reorder is necessary. Among the dealers in the study, this was done in one of two ways. The first involved a periodic inventory taking at intervals such as one week or a month. The stock in every item being reviewed by the store was counted and those that were below a desired level were reordered. The other basic type of system can be called the constant reviewing system. Every time that an item was sold, the person doing the selling noticed how much of that item was left in inventory. When this amount reached a specified low point, it was then reordered. In practice, this type of review was usually just a quick visual survey when the item was removed from the shelf.

The advantage of the periodic review is that in practice it is usually a more reliable method. The only person who needs to be aware of the reorder point is the person who reviews the counting of the items. Also, because most of the counting is done at one time, the ordering can be done at one time. The disadvantage of this periodic counting of inventory is that unless a unit control system is used, it requires the investment of time and effort by someone to go around and do the counting. The constant review system requires noticing how much of an item is left after each sale and this usually can be accomplished without making a large additional investment of time or effort. Although, in theory, the constant review system can supply better information than the periodic inventory, in practice this was not true for the dealers in the study. This was because no stores were found where everyone knew the reorder points of all the types of film carried and conscientiously wrote down the items that were below their reorder points.

Instead of ordering only when an item was below its reorder point, certain stores ordered up to desired inventory levels all of the products supplied by a manufacturer when they ordered anything from him. In this case, the necessity for ordering one item served as the "trigger" for other items to be ordered.

After the decision to place an order with a manufacturer was made, the typical dealer next decided the quantity to order in each item. Often the amount ordered varied seasonally, being greater when retail sales were larger. Here again, the seasonality varied from dealer to dealer. Some dealers had what could be called a "desired inventory level" which they ordered up to. Typically, this was an estimated month's sales of an item. Therefore, in busy seasons, the desired inventory level was higher.

As opposed to the desired level concept, some dealers liked to reorder in constant quantities. For example, because it is factory packaged in cartons of fifty, Polaroid film is conveniently ordered in multiples of fifty. Constant quantity reordering required more orders in a busy season but did not require explicit estimates of sales or desired levels to be made.

The day of the month also had an effect on some dealers. By ordering as far away from the payment due day as possible, the dealer could effectively have the manufacturer finance part of his inventory. In theory, this should be much more important for ordering from a manufacturer whose products represent a large volume for a dealer. This was borne out in the interviewing. Dealers who purchased and sold relatively large quantities of Polaroid products were more apt to take advantage of the billing dates. Four of the nine dealers in the large volume sample and four of the twenty-four in the small volume sample paid attention to the day of the month before ordering.

Four camera stores and five discount stores were in the sample of nine large dealers. The twenty-four dealer sample was comprised of eleven camera stores, five department stores, two discount stores, two drug stores, two jewelry stores, one Army-Navy store, and one book store. The ordering and inventory systems varied considerably among the diverse store types and also among different stores within any particular type. It proved impossible to make generalizations about ordering systems used by the various store types, with the following two exceptions:

1. Department store camera departments used the most explicit planning in ordering. Most department stores kept records of previous year's sales and used these in addition to present trends for determining order quantities.

2. Drug stores used the constant review trigger level type of system. It would be almost prohibitive for a drug store to count periodically all of the stock of the many thousands of items it carries.

Retailers also made decisions about how often to place orders. The dealers who ordered more frequently required more of a time investment in counting, making up and submitting their orders. Also, since more shipments were required for the same volume of film, the average shipping cost, as a percentage of total cost of goods, was higher. However, the dealers who maintained a small inventory with frequent reorders lost no more sales because of stock-outs than those who maintained a large inventory with infrequent reorders. The added cost of more frequent orders was balanced off by being able to maintain a lower inventory level. Because of faster inventory turnovers, the probability of having out-of-date film was less for those dealers who frequently reordered.

Dealer Differences

The dealers in this study were placed into two separate samples so that the orders placed by the larger dealers would not swamp the ordering of the small dealers. The nine dealers in the large volume sample averaged slightly over 10,000 rolls and packs of Polaroid film sold in 1964; the largest dealer sold 20,000 and the smallest dealer sold 5,000. The twenty-four dealers in the small volume sample average 1,500 units in sales in 1964; the largest sold slightly under 4,000 rolls and packs and the smallest sold 340. With the exception of the store size in terms of sales volume, the most important distinguishing characteristics of the dealers are listed below.

1. *Having a periodic review, and how often they had it.* Three dealers counted their stock twice a month, four counted it once a month, and twenty-one of the thirty-three sample dealers did not use a periodic review.

2. *Ordering constant amounts of film versus having a desired inventory level; and the values of these amounts.* Twelve dealers conceived of the ordering process as one of bringing stock up to a certain desired level; the other twenty-one dealers thought of it more in terms of ordering fixed amounts. The actual values varied from as few as ten units to as many as 400.

3. *Average delay from the time the decision to place an order is made to the delivery of this order.* For twenty-one of the dealers, the delay was one week; for the other twelve, it was two weeks. This delay included the store delays for processing outgoing orders and the incoming shipments.

4. *Types of film carried.* Twenty-five of the dealers carried all types of film in the study. Five did not carry Type 3; two did not carry Type 1; and one dealer did not carry either Type 1 or Type 3.

5. *Regular and emergency order trigger levels on film.* These figures varied substantially with the size and type of dealer, ranging from zero to 200 units as the trigger level.

6. *Batching orders, i.e., whether a dealer reviews and possibly adds to his orders items which are not below their trigger level.* The idea of batching orders arises because it is very simple to add other items to an order which is being placed with a manufacturer. The order is considered a batched one only if it contains some items

which would not have been ordered had not an order been placed with the manufacturer at this time. By this definition, fourteen dealers batched their orders; while nineteen did not.

7. Ordering so as to take advantage of the manufacturer's billing dates, thereby gaining an extra two weeks to a month of financing on film inventories.[2] Eight dealers tried to take advantage of the billing dates; twenty-five paid no attention to billing dates as far as Polaroid was concerned.

8. Percentage of total sales made to industrial accounts. This percentage varied widely among the different film types handled by dealers. Some dealers had no industrial sales of any types, while others sold as much as 80% of their Type 1 film to industrial users.

There were many other distinguishing characteristics among stores in the study (store type, sales volume, number of employees, etc.). These factors are related to the ones mentioned above. For example, a store with a large sales volume will tend to have high trigger levels and reorder amounts. However, in the researcher's opinion, these other distinguishing characteristics did not present any meaningful differences with respect to ordering patterns that could not be handled by quantitative descriptions of the above points. In fact, very few consistent clues to ordering behavior were discovered by examination of these more visible characteristics.

The information that was derived from the interviews was used to first conceptualize and then construct the model representation of the manner in which retailers order film. This model explicitly covered all of the above mentioned differences among the dealers and constructed a logical framework for these differences. The model was first quantified into flow charts and then programmed into source code so that it could be simulated. This simulation was a single, conceptual representation of the ordering behavior of all of the interviewed dealers. The individual dealer differences were handled in the model by tests at appropriate points which defined a unique path through the model.

Simulation Model

Programming

The simulation model was programmed in FORTRAN and run on the IBM 7094 at the Harvard Computing Center. The total programming to accomplish the simulation was divided into two jobs. The first program (the demand program) took the basic input data and generated specific consumer demand for each film type for each dealer in each period. The second program, which was was the heart of the simulation, took the demand and simulated the dealer decision behavior.

[2] Polaroid's payment policy was that payment for any order placed between the 25th of one month and the 10th of the consecutive month is not due until the 10th of the following month. For example, payment for an order placed on May 28 is not due until July 10. Any order placed between the 11th and 24th of the month has payment due the 10th of the following month.

Demand Generator

In addition to some secondary chores, the primary purpose of the first program (demand generator) was to take the general sales forecasts made by the executives and break down these forecasts into specific demands for each dealer, for all five film types and for each week. The basic inputs for this program were:

1. An expectation of the way that retail sales would vary over the fifteen-week period.
2. Factors that assigned what percentages of the entire forecast were to be given to each dealer and each film type. The record of past orders to Polaroid for 1964 was the criterion used to determine these percentages.
3. The total forecast distribution of sales for each film type in the period under consideration.

The program first generated a figure for the total demand by using the cumulative probability distribution curve for forecasted sales in conjunction with the random number supplied by the random number generator. The actual conversion from random number to sales figure was done by means of a cumulative probability curve. The theory for this is rather simple and is explained in [2].

Main Program

The dealer logic chart Fig. 1 is a representation of the part of the simulation model that emulated the dealer's logic in making his reorder decision. This dealer logic was part of the main program. The main program took the output of the first program, the individual dealer demand, and generated the final order pattern to Polaroid.

Much of the programming in the main program was devoted to bookkeeping and other secondary chores. The dealer logic itself, however, is rather important and is partially described below.

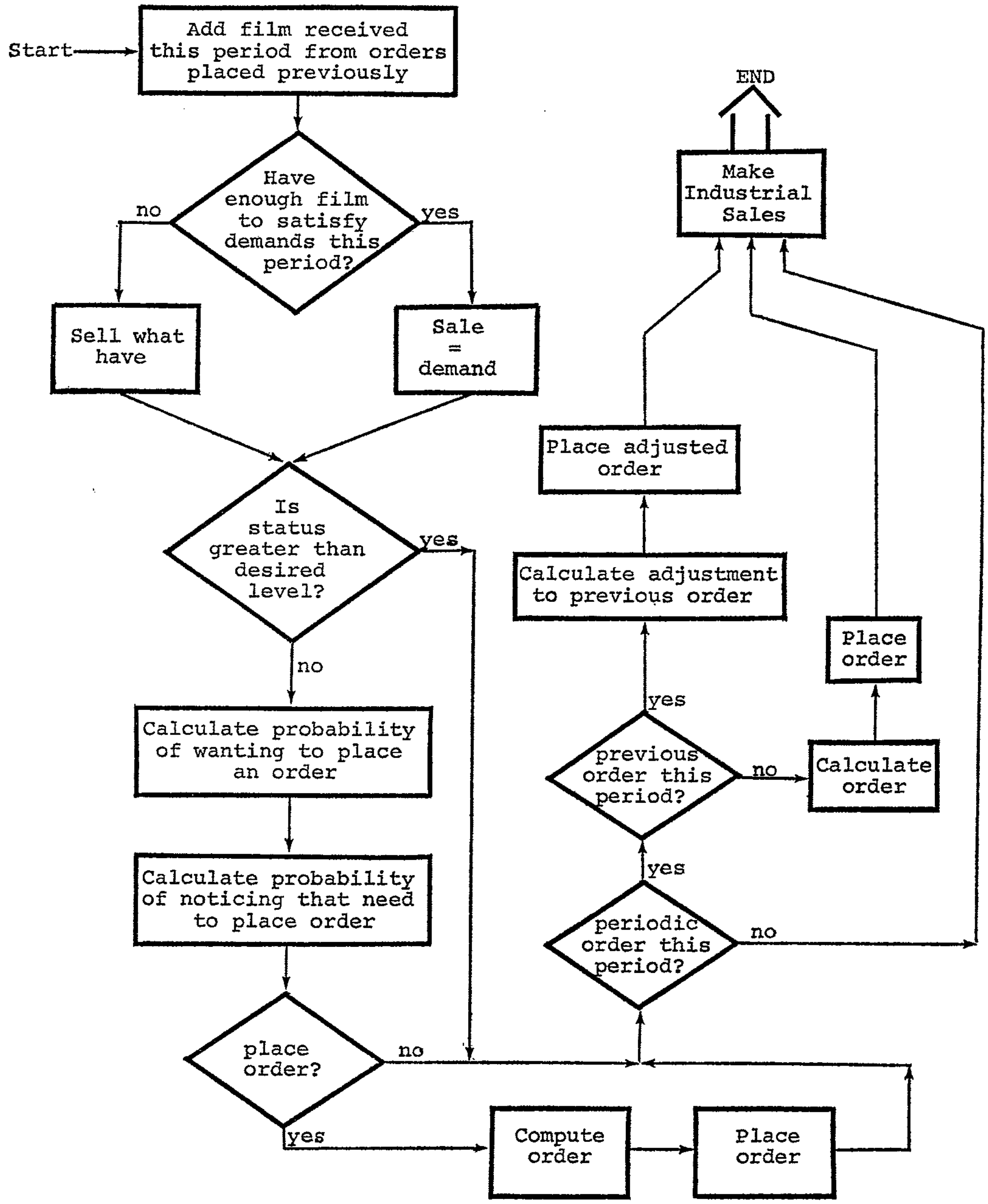
Dealer Logic

The first thing that the program did for each cycle through the model was to add the film orders received by the dealer during that period because of orders placed during previous periods.

The retail demand was next presented and if the dealer had enough film to cover all of his demands, sales equal to the demand were made. If not enough film was available, sales were made up to the level of film in stock.

The trigger level order routine was next entered. All dealers, whether or not they had a trigger level type of reorder system, had a low point trigger which could cause them to order stock. If orders were never placed other than at periodic intervals, then this trigger was considered to be negative.

With "status" defined as the amount of film on hand plus that on order, the program next tested whether the status was greater than the desired inventory level of the dealer. It it was not, a probabilistic order factor was calculated. This factor was an interpretation of the trigger level questions that were asked of the dealers. The interpretation was that this factor was the percentage "chance" that the dealer would want to place an order if he noticed that a film type was below



GENERAL FLOW CHART OF DEALER DECISION LOGIC

FIG. 1

its trigger level. This percentage "chance" was calculated by linearly interpolating or extrapolating from two known points. These two points were two levels of film where it was assumed the correct corresponding "chance" was known and could be derived from the following conceptualization of the dealers' decision process. The amount of film that was the answer to the trigger level question on the research questionnaire was assumed to be the point at which there was a 50% chance that the dealer would want to place an order. At zero units of film, it was assumed that there was a 100% chance that the dealer would want to place an order. Any amount of film in stock plus that on order corresponded to a "chance" that could be calculated by linearly interpolating or extrapolating from these two points. For example, if the trigger level was twenty rolls on hand and on order, then the factor was .75 for ten rolls, .5 for twenty rolls and .25 for thirty rolls.

Initially, the researcher did not have the above model in mind when he asked the dealers the trigger level questions. However, after a few interviews, the researcher noticed that about one-half of the inventory levels of all combinations of dealers and film types were below the amounts the dealer had given as trigger levels. Observation of the inventory levels of the rest of the dealers led to the above model.

Next, a second probabilistic factor was calculated independently of the first. This second factor was time related and could be interpreted as the probability that the dealer would notice that his film stock has reached a reorder point. Dealers tended to order more before a holiday and this factor reflected this point. In addition, some dealers paid attention to the payment due date so that they could get added financing on their inventories. More than one dealer was encountered who would not place an order from around the 20th to the 24th of the month.

Therefore, there were two time probabilistic factors, one for the dealer who watched the cutoff date and one for the dealer who did not. The proper factor was selected and multiplied by the first probabilistic factor. The rationale for this was that the probability that an order would be placed was equal to the product of the probability that the dealer would notice he needed to order times the probability that he would want to order if he noticed he needed to. A random number was generated and if this number was less than the product of the probability factors, the order was placed.

The actual order placed next depended on whether the dealer was a "desired inventory level" or a "constant order" dealer. The difference was noticed in the interviews, where it was ascertained that some dealers conceived of the ordering process as a bringing of stock up to a predetermined point, and others simply put in an order of a constant given amount (such as fifty) of a type when they were below trigger.

Before the trigger level routine was left, a test was made to determine whether this dealer batched his orders. A dealer who batched his orders was simply one who reviewed all of his Polaroid film stock for ordering when he had to reorder one type. If the dealer did not batch his orders, the trigger level order routine was cycled for all of the film types. However, if he did batch orders and an order for one type of film had been placed, then the trigger level routine was left and the batch routine was entered.

This batch routine used secondary trigger levels and secondary desired inventory levels. If an order had already been placed for one type of film, the likelihood was increased that the dealer would include others that were below their desired levels although still above the primary trigger levels. Therefore the secondary trigger levels were higher than the primary trigger levels. When this type of secondary order was placed, however, the amount ordered was usually less than would have been the case if the film had been ordered because of being below the primary trigger. Accordingly, the secondary desired levels were always lower than the primary desired levels. Perhaps this section might be made clearer by a numerical example. Assume the primary and secondary trigger and desired levels to be as follows:

TABLE I

	Film Types				
	1	2	3	4	5
Primary Trigger Level	10	15	10	20	15
Secondary Trigger Level	20	20	20	30	25
Secondary Desired Level	40	60	25	70	60
Primary Desired Level	50	75	30	90	75

If the inventory levels of the dealer were 5, 18, 12, 85, and 31 and the level of the first film type succeeded in triggering an order, then the desired order would be for 45, 42, 13, 0, and 0 units of film. Because Polaroid film can only be ordered in multiples of ten the actual order would be rounded to 50, 40, 10, 0, and 0 units of film. If, because of the random nature of the triggering device, the first film type had not succeeded in triggering the order, then no order would have been placed because no other film type is below its primary trigger. Therefore, for batched stores where a primary order had been placed, the batch routine added an order for any film whose status was below the secondary trigger level.

If no primary order was placed, or one was placed but the store was one that did not batch order, the trigger level routine was exited from. The next section of program took care of finishing out the primary trigger order for the dealer. It could have been possible that a level was lower than the trigger and yet, because of the random number that was generated, no order for that particular film type had been placed. If an order had been placed in the primary trigger routine, it is only reasonable to expect that any other film that was below its primary trigger would have been added, even if the dealer didn't batch orders. A subsequent section in the program took care of this.

The next section of the program was the periodic review section. The logic in placing the trigger level orders first was that in the stores using the periodic order concept the trigger level was an emergency order level which took precedence over regular periodic orders. Once the program had progressed to the periodic review section, a test was made to see if a periodic review was scheduled by the dealer during the current period. If the dealer did not have a review scheduled for the current period or if this dealer did not use the periodic review concept, the whole section was skipped.

If the store did have a periodic order review during this period, the next test was for whether an order had already been placed because of a trigger. If no order had been placed, the periodic review was then effected. The order was calculated according to the constant order or the desired level policy and it was placed. If an order had been placed, the next test was for whether all of the film types had been ordered. If they had been, the program proceeded to the next section. If some film types had not been ordered during this period, the program entered a section which calculated a supplemental order section, if necessary, to the one that had already been placed.

The supplemental section was necessary because if a dealer regularly ordered only at periodic intervals and a partial order had already been calculated for him in this period, the model had interpreted this partial order as an "emergency" order to fill in the dealer's stock until the next regular reordering period. Since the model now became aware that this was the regular ordering period, the amounts formulated as an "emergency" order for this period were converted into the regular amounts that would have been ordered in a regular periodic order. The supplemental section made this conversion. Since all of this happened in one time period, the "emergency" order and supplemental order showed up as one order.

The program took care of a dealer's industrial sales after the periodic review section was over. These industrial sales were different from ordinary sales in that they were usually made to a small number of customers who called relatively infrequently and ordered relatively large amounts of film at one time.

The rest of the program consisted of certain bookkeeping operations, output, and the statistical section which computed the basic statistics on sales and orders that were of interest.

The main part of the program was cycled through four different cycle indexes: the number of dealers, the number of periods (15), the number of executives (3), and the number of simulation cycles. Because of their interactions, the five film types were taken care of interdependently on each cycle. The other four factors, however, operated independently and therefore could be handled by cycling. From the innermost to the outermost, the central part of the program was cycled for all dealers, then all periods, then the executives, and finally the simulation cycles (which, along with random numbers, introduced the distribution aspect to the outcomes).

The final output from the simulation was a single page giving the mean and standard deviation of the total orders placed with Polaroid for each of the three five-week periods in the study. In addition, the value of the total sales experienced by the dealers at retail was given. There were many other things besides sales and orders that could be calculated from this program: average inventory levels, fluctuations in inventory, lost sales due to stockouts, etc. However, these were not of immediate interest to the object of the research and therefore the only printouts concerned the above points.

Results

Some of the forecast methods used in the research required forecast distributions on sales as part of their input. These methods and the other methods that used the simulation model as a transfer function generated output forecasts of orders that were distributions instead of point forecasts. These distributions were the result of the probabilistic nature of the reordering mechanism embodied in the simulation model and the distributional input for those forecasts that had such input. The results presented below do not discuss the distributional nature of these forecasts; instead, they treat the mean of these forecasts as certainly equivalents. This is a simplification that is only strictly valid for problems with quadratic losses. The reader who is interested in an analysis of the meaning and importance of the output distributions is referred to the thesis footnoted by the title.

Measurement Statistics

Each forecast method generated two cases of fifteen forecasts: one forecast for each of three five-week periods, five types of film, and two samples. In order to determine which of the forecast methods is "best" we must derive a summary statistic that reflects, somehow, the accuracy of the forecast. This statistic can and should be derived from the economic environment in which the forecasts are used. The effectiveness of the forecast can be judged only in terms of the losses that will be incurred because of errors in the forecasts. Therefore, any statistic that is used devoid of derivation from a cost structure will not necessarily guarantee the correct choice of which is the "best" forecast method. This important point is often overlooked because of the ease of using well-known classical statisitics. However, one major advantage of the simulation methodology is that it permits us to model specific real world cost structures and develop specialized statistics where necessary.

The derivation of two summary statistics is presented below. The assumptions which are made do not have to be met exactly for these statistics to be applicable. If one forecast method is best as judged by each of several different statistics that were derived from different assumptions, then we can feel fairly confident about its superiority under general circumstances.

If one of the forecast methods generated forecasts that were always somewhere between every other method's forecasts and the true value, we could definitely say that this method was superior under all reasonable evaluation statistics. Unfortunately, this is not the case for any of the methods used.

For the derivation, it is assumed that the costs associated with any forecasted outcome, corresponding action, and actual outcome can be estimated. It is also assumed that we have a method which tells us what the optimal action is for each forecast and that the forecast errors are not compensating. With this information, we can simply itemize the cost associated with every forecast plan and total them. The method giving the lowest cost is "best." For example, assume actions have been taken on the second method's forecast for film type 1 in period I. Assume Polaroid sales were forecasted at 620 while in actuality they were 750. The difference between the money cost to the company resulting from incorrect action based on the forecast of 620, as opposed to the correct action based on a forecast of 750 can be called $x_{1,\mathrm{I}}$. The summation of the x's over all three periods and five film types would serve as a summary statistic that satisfies the condition of reflecting the associated cost structure. This statistic, or any linear transformation of it, is the proper statistic to be used for ranking the methods.

Absolute Deviation

It is interesting to look at the assumptions embodied by the use of classical statistics, such as absolute deviation and squared deviation. This is now done and the forecasts by the various methods are ranked according to these measures. Assume the following conditions:

1) Costs associated with the action corresponding to a forecast that is too large are equal to the costs associated with a forecast that is too small by the same amount.
2) The cost of any particular forecast error is the same for all types of film.
3) Errors in different periods or film types cannot cancel each other out. They are not compensating in the sense that over-production in one period or type cannot be used to counteract under-production in another period or type.
4) Losses are strictly proportional to the magnitude of the forecast error. This condition is commonly known as "linear losses."

With these assumptions, the proper ranking statistic is absolute deviation summed over the fifteen forecasts corresponding to each film type and period. The first condition implies symmetry and permits the absolute value of over and under deviations to be merely added. The second condition is necessary to permit additivity across the five film types. The third condition eliminates interdependencies among the fifteen figures so that minimization of the pieces will provide minimization of the whole. Looked at in another way, the third condition is also necessary for additivity. The fourth condition states that these deviations or a common linear transformation of them are the figures that should be added.

Squared Deviation

The use of the sum of squared deviations as a ranking statistic results from the same set of assumptions listed above except for number 4. This assumption is replaced by a statement of the form that losses are proportional to the magnitude of the square of the forecast error. As the error becomes larger the losses become larger more rapidly than linearly proportional losses.

Ranking of the outputs of the forecast methods according to sum of absolute deviations and sum of squared deviations are presented in Tables II and III. Any inference made from this ranking is strictly valid only for the assumptions outlined; however, informally we can infer that the methods with the smaller figures are generally superior. It should also be noted that no use was made of the fact that some methods generated distributions instead of point forecasts; only the means of these forecasts were used.

TABLE II

Ranking of Forecast Methods

Forecast Method	Sum of Absolute Deviations
Sample of 24 Stores	
1) Simulation on actual sales	1997
2) Simulation on First's forecast of sales	2293
3) Simulation of Second's forecast of sales	2620
4) Simulation on Third's forecast of sales	2803
5) Regression forecast of orders	2940
6) Third's forecast of orders	3010
7) First's forecast of orders	3340
8) Second's forecast of orders	3560
9) Simulation on regression forecast of sales	4652
Sample of 9 Stores	
1) Simulation on actual sales	1836
2) Simulation on Second's forecast of sales	5008
3) Simulation on regression forecast of sales	5414
4) Second's forecast of orders	5870
5) Simulation on Third's forecast of sales	5899
6) Regression forecast of orders	6100
7) Simulation on First's forecast of sales	6176
8) Third's forecast of orders	7020
9) First's forecast of orders	7860

TABLE III

Ranking of Forecast Methods

Forecast Method	Sum of Squared Deviations
Sample of 24 Stores	
1) Simulation on actual sales	513,157
2) Simulation on First's forecast of sales	558,061
3) Simulation on Second's forecast of sales	713,810
4) Regression forecast of orders	800,600
5) Simulation on Third's forecast of sales	827,319
6) Third's forecast of orders	1,075,300
7) First's forecast of orders	1,256,200
8) Second's forecast of orders	1,476,800
9) Simulation on regression forecast of sales	2,501,398
Sample of 9 Stores	
1) Simulation on actual sales	342,158
2) Simulation on Second's forecast of sales	3,068,170
3) Simulation on First's forecast of sales	3,719,958
4) Second's forecast of orders	3,787,100
5) Regression forecast of orders	3,995,400
6) Simulation on regression forecast of sales	4,338,180
7) Simulation on Third's forecast of sales	4,727,055
8) Third's forecast of orders	4,919,800
9) First's forecast of orders	8,407,600

It is not the purpose of this article to explore the problems associated with measurement statistics. However, it should be mentioned that as we move away from the above mentioned assumptions, the analysis of certain points can become difficult. If assumption 3 is reversed so that errors are compensating, then it is possible that a method which produced very poor forecasts for specific films and and specific periods could, by chance, produce an overall excellent series of forecasts. This, however, is a statistical problem and not a problem associated with the forecast methods.

Method Evaluations

Certain facts stand out from a study of the rankings. The first is that the simulation model was able to generate the best order output when actual sales were used as "forecasted sales." This was to be expected, and had the simulation not done this, the validity of the model would have been in doubt. The effectiveness of this test of the simulation is especially noticeable in the nine dealer sample.

If the manufacturer is interested in forecasting his own orders from retailers, he would also be interested in the accuracy of the order forecasts from this model. A statistic reflecting the basic accuracy obtainable from the simulation is the average error percentage of the order output for each sample when actual sales were used as input. The mean error percentage of the simulation model's output in this case was 14.1% for the twenty-four dealer sample and 5.5% for the nine dealer sample. If the two samples are combined on a base weighted according to total orders, the average output error is 8.1%. Total orders in all film types amounted to 14,140 units from the twenty-four dealer sample and 32,060 units from the nine dealer sample.

A final and important point about the rankings in Tables II and III is that in all four cases, two samples and two statistics, every executive's sales forecast operated on by the simulation was superior to the straight forecast of orders made by the same executive. This is probably the most important single outcome of thestudy and is discussed further in the following section.

Conclusions

Most of the ordering systems used by the retailers in this study were informal variants of trigger level and constant review systems. A hope that there would be visible store characteristics, which would aid in the determination of the type of ordering system used, was not borne out. Few conclusions could be drawn about the type of reorder system that a store used from characteristics such as store type, location, and sales volume. Similar stores (e.g., camera stores in the same general location with approximately the same sales volume in Polaroid film and other items) were found that used different inventory systems, kept different types of records, and reordered at different intervals. As mentioned before, drug stores primarily used trigger level systems in ordering film, while department stores primarily used periodic reviews, along with records of previous sales of the items, as a basis for ordering.

A basic reason for this lack of uniformity in the ordering habits of similar stores is that in most cases, both periodic review and trigger level ordering systems with different reorder frequencies can accomplish essentially the same purpose for a retailer. Moreover, the amount of inefficiency that is introduced into the overall management of a small business by having an inventory that is not exactly optimal is not great enough to significantly influence the retailer's success. The result is a wide diversity of systems among comparable stores. It is therefore the opinion of this researcher that studies which attempt to correlate product ordering with some visible store characteristics either will not be successful or will be successful only because they propose a clever or novel way of looking at the problem.

In most cases, it would probably not be valid to try to force another product into the specific reordering model for Polaroid film used in this thesis. The most outstanding characteristic of simulation as a tool of analysis is that it is the most specific, non-generalizable form of analysis available. For example, if one tried to use the simulation to help in forecasting orders of Kodak film it could be possible that the interactions in the ordering process among the large number of different films types made by Kodak would be such that this model would not work very well. The theoretical interactions could be of the same form, but the large number of different types and the way that errors propagate might make useless for Kodak the simulation constructed for this thesis. Therefore, it is possible that the specific model derived for this thesis cannot be generalized even to other film producers.

Even though the simulation could do a good job of emulating the dealer ordering mechanism, it would be of little value if there was no way of forecasting sales. The fact that the model can create an order output that is close to the actual (when the actual sales are used) is primarily interesting in the sense that a bound is determined on how good one can expect the results from the model to be. More interesting is how well the simulation model performs on available forecasts. On this point, the model used in this research did well, even though it had no influence on the sales forecasts that were the basic inputs for the simulation runs.

The fact that all of the executives' sales forecasts used in combination with the computer simulation were superior to the direct order forecasts is especially impressive when it is realized that these executives were more used to forecasting orders than sales. The reason that the simulation yielded such improvements was that it reflected the cyclical nature of the order. If more effort could have been put into the forecast, better results probably would have been forthcoming. Since the simulation was only a transfer function, it could not produce a good order forecast from a poor sales forecast.

Considering all of the results that were derived from the research on both samples of dealers, it appears that a behavioral simulation model of the type constructed can be useful in the analysis and prediction of retailer behavior. More generally, we can state that a behavioral simulation model can be useful for analysis or forecasting in marketing problems where one tries to simulate the external environment of the firm. Whether the simulation approach is the correct one in terms of cost justification depends on the specific problem area.

Unsophisticated retailers seem to have a sufficiently systematic set of procedures to permit simulation of these procedures by a computer model. This is not such a surprising conclusion, since we would expect successful dealers to have a rationale for their actions. Those that do not have rational (not necessarily sophisticated) patterns of business behavior probably have gone out of business. For many products, these procedures may be determined by interviewing retailers. There is nothing particularly unusual about film that would lead us to believe that we can successfully model the reordering of film and not of other products.

There were many particular characteristics, both of the market place and of the particular product line, that led to the model developed in this research. However, these were not requisite characteristics for developing a simulation model of the reordering process. For example, if film were heavily promoted, a section in the simulation model could have taken this into account. By definition, if one wishes to simulate a process, he must model the special characteristics of that process. Just because another process does not possess those characteristics does not mean that it can not be modeled likewise. It does mean that the model constructed for the first case probably will not fit the second.

A simulation model can be useful for understanding a process quite apart from the prediction of that process. In the process of constructing the model, the researcher developed an understanding of the way in which dealers operated in reordering film. In an industrial situation, such an understanding could be brought to bear on problems that the simulation model did not handle explicitly.

A simulation model of the type that was constructed for this research could also be useful in experimentation leading to a better understanding of the market environment. One variable could be varied while the rest are held constant, and variations in the output could then be compared with the changes in the input variable, so that a better understanding of the environment can be obtained. For example, a step function of sales could be arbitrarily introduced, so that the resulting ordering pattern could be studied. Sensitivity tests would also be possible. If a change in corporate policy affecting dealer ordering was being contemplated, this change could be programmed into the model and the effect on the forecasting of orders generated by the model could be studied. This type of information would be valuable in corporate decisions.

Because the primary purpose of this research was to examine the ordering process which acts as a transfer function between retail sales and wholesale orders, little attention was paid to factors not immediately relevant to this transfer function. An example of this was the type of regression analysis that was used—a strictly standard regression. Little thought was given to novel ways of forecasting sales because this could validly constitute another topic of research. It is obvious, however, that any improvements that can be made in retail sales forecasting would tie in directly with the results of this study. In fact, the results of the research suggest that the transfer function between the retail sales and wholesale orders is tractable and that further work should be done in the area of forecasting retail sales.

References

1. MILLS, FREDERICK C., *Statistical Methods*, 3rd ed. Henry Holt and Co., New York, 1955, pp. 360–65.
2. SCHLAIFER, R., *Probability and Statistics for Business Decisions*, McGraw-Hill, New York, 1959, pp. 323–325.