In the last several years, every major food industry trade association has had major programs addressing scanning. The Food Marketing Institute (FMI) has had annual industry-wide conferences on scanning. The National-American Wholesale Grocers' Association (NAWGA) and the National Grocers' Association (NGA) have also had similar conferences. Thus, one would be hard pressed to find anyone who does not view scanning and, in particular, scan data as having enormous value. BUT, in the same vein of thought, one would in all probability not find any well organized complete scanning research program investigating the wealth of information contained in the scan data.

Prior to writing this paper, I had a computerized literature search performed to find any publications dealing with scan data research. In this search I excluded any Journal of Food Distribution Research publications since I was already aware of them. To my amazement, but in support of my earlier statement, only one article was found. The article was entitled, "New Gold Mines and Minefields in Market Research." The authors discussed the changes going on in marketing due to split-cable technology and scanner data and stressed the need to take full advantage of the new micro-level experiments possible with the new technology.

At this point in time, the first question that should come to the reader's mind is: WHY? The answer is not a simple one nor black and white as to why there is not a concerted effort on scan data research. The focus of the paper is on the status of scan data research. Before delving straight into this issue, a brief historical perspective on scanning is presented. Then, an overview of organized industry scanning research is given. This is followed by private marketing research firms and other research groups. Finally, the general subject of what has been the research as well as the problems and needs will then be presented along with the conclusions.
Commercial scanning started in 1974 when Marsh's supermarket in Troy, Ohio, installed one of the first scanners capable of reading the UPC (Universal Product Code) symbol. During this period of time, the discussion concerning scanning was centered around the feasibility issue. That is, what volume size retail food store could cost justify implementing scanning. The savings generated from scanning were delineated into two areas: hard savings and soft savings. Hard savings are those that result from improved speed and accuracy from the scanner's ability to identify a UPC symbol and to retrieve the product price from a computer file. Soft savings relate more to improved management information and control, and generally accrue over time as a result of using scanner-generated information that is processed through a computer.

The hard and soft benefit issues were addressed three years ago by Fletcher et al. (1984). Capps last year discussed the soft benefits issue. The industry consensus is that practically everyone accepts the feasibility of scanning for the retail food store except for some low sales volume retailers. On this latter point, Fletcher et al. (1986) published a report addressing the feasibility issue for the smaller retailers. They found, in general, that if the retailers used the newer scanning technology that scanning could be justified economically. BUT, in spite of the information available to retailers, one still does not see widespread scanning. In fact, in an Executive Report on scanning by Progressive Grocer, the editor states that the retailers are being divided into the "haves" and "have nots" as far as scanning competency is concerned. However, in the last several years the focus of trade association meetings, seminars and conferences has shifted to scanner data and the use of microcomputers.

Industry Scanning Research

Two major industry scanning research projects have been undertaken to address the issues of effectively utilizing scanner data. The first project was ScanLab™ (a trademark of General Foods Corporation) in 1981. The purpose of the project was to advance the level of understanding of the retailer in effectively using scanner data to improve merchandising decisions. The products that the project analyzed were detergent, dog food, shampoo and hair conditioner. They had matched store panels where one panel was the control group and the other was the test group which utilized the scan data information. The test stores which used the ScanLab information had a significant improvement. The ScanLab software is available through FMI but must be used on a mainframe or a minicomputer.

The second scanning project was the Pillsbury Scan Share Group which started in the fall of 1982. The objective of this group was to further the applications of scanner data for merchandising decisions. They delineated scanning data applications into three general categories: tracking, analysis and experimentation. This group also found significant benefits from utilizing scan data in merchandising decisions. However, three key lessons were learned if the benefits are to be achieved: 1) A close working relationship must exist between the user of the information and the data processor. 2) An investment must be made in software development and database development. 3) A company must prioritize its scan applications.

Other Research Groups

In recognizing the potential benefits from scanning, several private firms have entered the scanner data market. Some of the firms are Information Resources, Inc. Ad Tel (Test Marketing Group), NPD, A. C. Nielsen, Safeway's Scanner Marketing Research Services, ScanAmerica and SAMI. This list of firms is not all encompassing. The major focus of the private firms in utilizing scanner data has been to track advertising effects. Recently, they have been tying the split-cable technology with the scanner data. Besides these private firms, there are other groups investigating and utilizing scanner data. The individual retail food companies have been some of the leaders. They have done the work either by themselves, with a consultant or with a researcher from an academic institution. However, this type of work has been...
basically "proprietary." Besides the retail food companies, academic institutions have started research programs into the uses of scanner data. Some of the institutions are the University of Chicago, Cornell University, University of Southern California, University of Illinois, Texas A&M University and the University of Georgia.

Research Thrusts

Some of the research thrusts have already been addressed in the industry scanning research section. This research was of an individual store or chain need. Methods to utilize scanner data in order to address other needs of the retail food store operator are in the areas of shelf allocation, automatic reordering, price sensitivity, product mix, trade area analysis, etc. Once again, this work is generally of "proprietary" nature.

Another area of research thrust has been measuring the advertising effect on product movement. Information Resources, Inc. (IRI) has been one of the industry leaders in this area of marketing research. They have integrated the technologies of split-cable TV, UPC scanner and computers. Using these technologies, one can direct specific advertising toward a particular segment of the population and then measure product movement as well as obtain sociodemographic information on the purchasing households. The advertising work has also been proprietary.

One of the first scanner data research projects was done by Jourdan. Her work won her the FDRS Applebaum Award. Her study was to estimate own-price and cross-price elasticities of demand for a group of related beef cuts using scanner data. Her results were mixed since she did not have any household information in her model.

Two years ago MacNeary of IRI (1985b) presented a meat study using scanner data at FMI's 2nd Annual Industry Wide Scanning Conference. He divided IRI's scanner panel households into light and heavy beef buyers. Thus, he was able to examine purchasing dynamics. MacNeary did not find any differential impact on the two household groups from promotions. Thus, he concluded that it seems that one should not look to buyer-group segments as a focus of promotional programs but rather one should emphasize product areas.

With the increased use of scanner data for merchandising decisions, the accuracy of the data is paramount. If the data is erroneous, the results will be inaccurate which could lead to inappropriate decisions. Thus, Lesser and Smith did a study on the accuracy with which scanning data actually reflect disappearances. In their conclusions they found that, basically, scanning data provide a quite accurate measure of aggregate item sales over a multi-week period. However, for individual items for a particular store for a one-week period, they felt that one should exercise caution.

A recent major research thrust has been with the National Dairy Promotion Board. They have been charged to evaluate the effect of advertising on sales of dairy products. One component of their research program has been to utilize split-cable scanner data in evaluating advertising effect on sales of fluid milk, cheese and butter. A. D. Little was given the contract for the analysis. Initially, the researchers used analysis of variance procedures to compare the test and control household panels. Later on, they did an econometric analysis. However, in their report they state that the results are inconclusive. But, further work is being done using the split-cable scanner data. The Meat Board is also following a course similar to that of the Dairy Group.

Problems and Needs

One obvious criterion for performing scan data research is that the items that one desires to examine must have UPC symbols. Another criterion is that one must decide which items to examine first. In a store with a minimum of 10,000-15,000 UPC items, one would be immediately doomed to failure if some setting of priorities were not done. However, even with this many items UPC-coded, there are some critical products and categories that do not have UPC codes. For example, random weight items (e.g., cheese
and produce) do not have UPC codes. While meat has UPC codes, such codes are not in widespread use. Thus, many retailers have developed their own price look-up number for scanning these non-UPC-coded items since they constitute a significant proportion of the store sales.

As previously mentioned in the Lesser and Smith study, data accuracy is critical. If the data is not accurate, the research is meaningless. Along the same lines is the question of the accuracy of split-cable scanner data. Some items may not be purchased at a scanning store. For example, if a household purchases milk at a convenience store all milk purchases in that household may not be collected by scanning equipment. Also, if a consumer forgets the household ID card when shopping, none of those particular purchases will be recorded.

Even though there are some problems with split-cable scanner data, let's examine the flip side to this. What have been the data sets used to do consumer research? The major ones are the National Household Food Consumption Survey by the USDA, Consumer Expenditure Survey by the Bureau of Labor and MRCA survey data. All these databases are based on diary surveys which consumers fill out after making purchases. The question that comes to mind is: Just how accurate is this type of data? Based on a study by MacNeary (1985a), the answer is not reassuring. He found that household information based on recall was significantly different from the observed behavior. Furthermore, he found that one's memory seems to be associated with frequency of consumption rather than with frequency of purchasing.

Given these potential problems, are there any alternatives? One possible alternative is to use ScanAmerica's approach. A fully portable wand resembling a large laundry marker scans the UPC symbol of items purchased by a household independent of source of purchase. Each household has a wand. Periodically, the information stored in the wand is downloaded to a central computer. One obvious drawback to this approach is that price information is not being recorded. Another alternative is to have a procedure that could read scanner and non-scanner tape receipts.

Assuming that data problems could be corrected, financial considerations have to be taken into account. The cost of doing a comprehensive research project is prohibitive to a research institution without outside financial support. For example, the National Dairy Promotion Board spent over $1 million for one year's worth of split-cable scanner data of dairy purchases for only a couple of markets. Furthermore, the data was in raw form which must be transformed into a usable data set. This latter step takes an enormous amount of time, manpower and money.

Conclusions

Retail food store operators see the use of scanning data as a means to stay competitive and improve the bottom line. For example, the widespread interest and use of a DPP model illustrates the concern about competitiveness and the bottom line. Integrating scanner data with the DPP model will enhance an operator's position. In addition to the retail food store operator, commodity associations want to know how their product is performing in the market and whether advertising is helping them. The bottom line is that they want information which scanner data can provide. Food market projections done at the Georgia Experiment Station, University of Georgia, are sensitive to parameter estimates. Incorrect projections could lead to inaccurate policy implications. All of this points to the fact that the need for scanner data is tremendous. Researchers need the data available to them, but a concerted effort is required. This year a new regional research project (NEC-65) has been approved. Use of scanner data is a key component in the project. However, for this project to be effective, the cooperation of trade associations, industry and government is needed. In the meantime, one potential method of aiding scanner data research is for the USDA and the National Dairy Promotion Board to make the dairy scanning database available--within appropriate guidelines--to research institutions.
References


