

Some analytical approaches to the branch location problem

By BENGT ANDERSSON¹), Dr. JOHN U. FARLEY²) and
CHRISTER GUSTAFSSON³)

This paper outlines four examples of analysis involving branch locations. The structure of the models and the choice of variables differ over markets like the three studied here, but the analytical procedures are very similar. Each makes use, in one way or another, of a computer – as a data bank for easy storage and easy inquiry, and as a computational aid in terms of regression analysis for forecasting and in terms of simulating the market for the distribution of possible rates of return for a given investment. Each technique can easily be made a routine management aid. In general, the analytical methods are capable of handling more variables simultaneously and of considering more alternatives than management – two capabilities of the computer which will more and more shape analysis of problems such as those discussed here. The method also relieve management of routine drudgery and provide time for more productive purposes.

Locating new outlets to serve retail customers is a repetitive problem faced by a wide variety of business-banks, chain retailers of many types of merchandise, gasoline refiners, newspaper and magazine distributors, and travel agencies for example. A variety of general discussion about this problem are available (4, 19).

Managers responsible for locating branch outlets describe the decision-making process approximately as follows: First, some measures of the potential market served by the proposed site are collected. The market is defined in a way which is relevant for the particular product or outlet-

1) Civilekonom, assistant, Handelshögskolan, Gothenburg.

2) Associate Professor, Graduate School of Industrial Administration, Carnegie-Mellon University and visiting Lektor, Handelshögskolan, Gothenburg.

3) Fil. kand., assistant, University of Gothenburg.

The authors are deeply indebted to Svenska Sparbanksföreningen for financial support in the preparation of this article, and to Director J. O. Berg for his personal support and encouragement.

Some analytical approaches to the branch location problem

By BENGT ANDERSSON¹), Dr. JOHN U. FARLEY²) and
CHRISTER GUSTAFSSON³)

This paper outlines four examples of analysis involving branch locations. The structure of the models and the choice of variables differ over markets like the three studied here, but the analytical procedures are very similar. Each makes use, in one way or another, of a computer – as a data bank for easy storage and easy inquiry, and as a computational aid in terms of regression analysis for forecasting and in terms of simulating the market for the distribution of possible rates of return for a given investment. Each technique can easily be made a routine management aid. In general, the analytical methods are capable of handling more variables simultaneously and of considering more alternatives than management – two capabilities of the computer which will more and more shape analysis of problems such as those discussed here. The method also relieve management of routine drudgery and provide time for more productive purposes.

Locating new outlets to serve retail customers is a repetitive problem faced by a wide variety of business-banks, chain retailers of many types of merchandise, gasoline refiners, newspaper and magazine distributors, and travel agencies for example. A variety of general discussion about this problem are available (4, 19).

Managers responsible for locating branch outlets describe the decision-making process approximately as follows: First, some measures of the potential market served by the proposed site are collected. The market is defined in a way which is relevant for the particular product or outlet-

1) Civilekonom, assistant, Handelshögskolan, Gothenburg.

2) Associate Professor, Graduate School of Industrial Administration, Carnegie-Mellon University and visiting Lektor, Handelshögskolan, Gothenburg.

3) Fil. kand., assistant, University of Gothenburg.

The authors are deeply indebted to Svenska Sparbanksföreningen for financial support in the preparation of this article, and to Director J. O. Berg for his personal support and encouragement.

number of people living near or passing by the site in a day, for example. These gross measures are adjusted, often qualitatively, to eliminate those segments which are not relevant for the particular product or outlet, although there is often great uncertainty about relevant segmentation criteria. Second, other businesses in the general vicinity are considered; some are more competitive with and some complimentary to the proposed branch. Sometimes the effects of the new location on the firm's proposed branches, either existing or planned, are considered. Finally, management sifts the evidence, feigns the alternative arguments pro and con, and decides on a certain location or set of locations.

Recently, the branch location has been approached more analytically, and this paper discusses four examples of more rigorous and systematic attacks on the problem. Systematic data gathering and analysis is, of course, a trademark of the computer era, and branch location lends itself well to certain types of analysis which are quite easy for the computer. The detailed results from these four studies (which deal with banking, newspaper sales and gasoline retailing) are proprietary, but general discussion of methodology and of results is possible.

Systematic Search and Evaluation

The accelerating complexity of the branch location problem motivated development of more orderly procedures for evaluating alternative sites (20, 21). An example of such a procedure, developed for banking (23), involves the following steps:

1. A list is prepared of all variables considered by management as important influences on the success of a location. This is a kind of "brain-storming", and a list of the type shown in Table 1 results. The list will vary over industries and may even differ substantially over decisions for a given firm – in the case of downtown versus suburban locations, for example. It is important to note that data on most of these factors are available from public sources, so the data collection task is not nearly as formidable as it might appear.
2. Variables with identical values for all alternative locations are eliminated.
3. The remaining variables are evaluated for all alternatives.

4. The variables are ranked by importance and some weights are assigned to each. Validation of this weighting scheme, a difficult but vital part of the work, may be done subjectively by management or by statistical techniques described later.
5. Weighted scores for each alternatives are computed using the values developed in 3/ and 4/. These scores become part of the basis for management's decision. Sensitivity analysis can be done by varying the weights or evaluations of some variables to examine how closely the results depend on certain variables or on certain aspects of the weighting scheme.

The procedure may also depend on the mix of business a branch realizes, so different classes of revenues or of branches must often be analysed separately. For instance, the lending and the deposit functions of a new branch are important for profits (16), but the lending patterns are affected less by location decisions. The same is true for accounts of large firms, but other types of deposits are intimately related to branch location (1).

While the procedure does not require a computer, automating the analysis can have several benefits. First, the values of the measures in Table 1 may be the starting point for a useful data base for a management information system. For example, graphic display of characteristics of established versus proposed branches may be very useful in the last stages of discussions about a location decision. Second, the sensitivity analysis is much less tedious and is likely to be done much more thoroughly if a computer is used. Third, the computer makes available a number of analytical techniques which are logical extensions of the procedure described here.

Statistical Procedures for Forecasting

The technique described in the last section is a very useful starting point, but there are two apparent problems. First, the procedure ranks alternatives, but it gives no indication whether any or all of the locations is desirable *per se* – that is, it produces no forecast. Second, a somewhat less subjective weighting procedure is desirable. Multiple regression (15) analysis provides a key for the solution of these problems.

To review briefly, regression techniques can be used under rather general circumstances to solve jointly for a set of weights ($b_0, b_1, b_2, \text{etc.}$) in the following type of linear equation:

$$Y = b_0 + b_1X_1 + b_2X_2 + \dots + b_nX_n + e \quad 1/$$

Table 1.

**A TYPICAL CHECKLIST FOR A BRANCH LOCATION
PROBLEM**

1. *Common economic factors*
 - a/ The basic economy of the location area
 - b/ Similar measures for neighboring areas

2. *Traffic and transportation*
 - a/ Major traffic arteries in the area
 - b/ Traffic volume at the proposed location
 - c/ Condition of the roads
 - d/ Public transportation in the area

3. *Population*
 - a/ Current population in the area
 - b/ Characteristics of the population
 - b1/ Income
 - b2/ Age
 - b3/ Sex
 - b4/ Occupation etc.
 - c/ Projected population growth

4. *Residential*
 - a/ Dwelling units in the area now
 - b/ Rate of home construction
 - c/ Characteristics of existing houses
 - c1/ Age
 - c2/ Type of construction etc.

In this case, Y is the dependent variable that is to be predicted – normally sales volume in a location problem. The X 's are a set of n independent variables upon which the forecast depends – variables such as those in Table 1. The e is the forecast error. The weights are calculated with a procedure which minimizes certain forecast errors for a set of observations of the data. The general forecasting procedure for branch location is:

1. Values of independent and dependent variables are measured for a set of outlets already in existence. The success of the new outlets is predicted on the basis of what factors affect the success of those already in operation.
2. The data are used to solve for weights as in equation 1/. The magnitude of the computations is such that a computer is required for all but the smallest problems. However, computer programming is unnecessary, since regression is a standard statistical procedure and efficient programs are available in the library of virtually every computer center.
3. The set or weights produced in step 2/ are used to forecast sales of proposed new outlets with values of independent variables measured for these new outlets.

Such a procedure has been applied to AB Pressbyrå's problem of locating kiosks which sell newspapers, magazines, confectionary, tobacco and a variety of other products (5). Pressbyrå has 1200 kiosks spread over Sweden. Competition is fierce in sales of the items in the kiosk assortment, and newer kiosks are bigger and more expensive, often occupying permanent locations in new buildings. Because of the extensive system and because of significant shifts in traffic and population patterns, location decisions occur frequently. The decision is thus repetitive and a formal model would be a great aid.

Development of a forecasting model started with a sample of kiosks with similar mixes of product sales, so that sales volume could be used as a proxy variable for gross profit. Because it is unlikely that sales of all outlets respond in the same way to a set of independent variables, separate analyses were planned for the following groups: outlets mainly serving neighborhood customers, outlets at main traffic points, and outlets mainly serving particular institutions.

A multiple regression which predicted sales of the sample of kiosks with a high degree of statistical significance included the following classes of independent variables: number of people living in the neighborhood, traffic volume past the site (autos, pedestrians, trams and busses), measures of competitive businesses within 300 meters of the outlet, and certain aspects of the kiosk's appearance.

The set of independent variables is much smaller than that listed in Table 1, since many of the measures in that list are redundant and others have no significant effect on sales. Cutting down the list on independent variables is legitimate, since the goal of the analysis is forecasting and not analysis of the structure of the equation. Notice also that all the independent variables are easily measurable before the kiosk is built, so a model can easily be used to screen large numbers of alternatives at virtually no cost above that of data collection.

Combining Demand and Cost Analysis

Regression is useful for forecasting and it can also be integrated into a break-even analysis procedure. Such joint analysis of both revenue potentials and costs has been applied to selection of sites for gasoline stations (9). The role of the firm's real estate experts, who has historically handled all phases of the analysis, remains very important, but their work is complemented by other statistical analysis.

Cost analysis

Cost analysis of this operation was straight forward. Standard plans are available for the few models of stations that are built, and historical records provide good estimates of construction costs. The real estate experts found promising sites and estimated costs of acquiring and preparing the land. Standard accounting figures provided estimates of operating costs at various volumes. The latter, combined with properly discounted values of capital outlay, provide a clear picture of out-of-pocket costs involved with a given site.

Revenue estimates

Revenue estimates were based on regression much like the one described in the previous section. Historical sales data were analyzed for a set of stations serving arterial highways and for another group serving neighborhood market clientele. (Consumers' purchasing patterns differ markedly between the two groups, so separate analysis was necessary).

The dependent variable in each regression was total volume of gasoline

sales, which was then broken down into different octane grades on the basis of experience of stations with similar volumes. Sales of oil, tires and batteries, and other services were similarly estimated on the basis of volume of gasoline sold. This approximation was adequate for purposes of the present analysis, but validation is necessary in all specific applications.

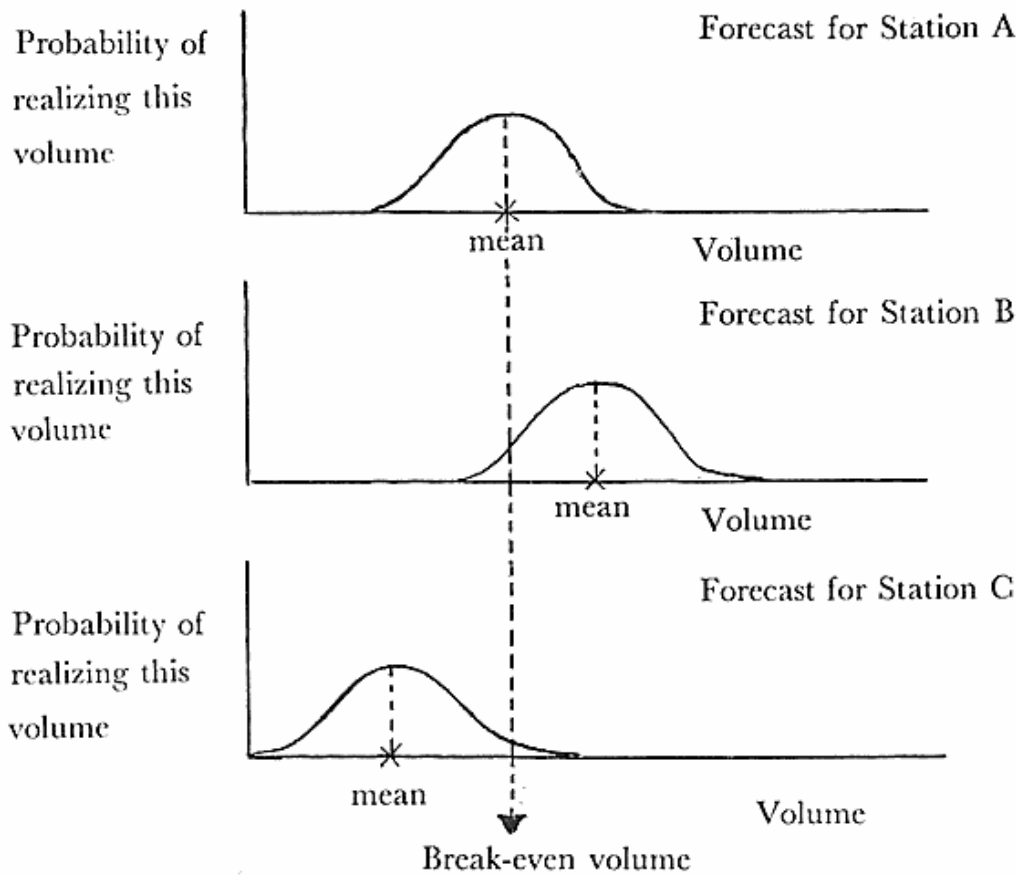
The independent variables for the regressions involving neighborhood stations were: total auto traffic along streets with legal access to the proposed station, neighborhood income, neighborhood population, neighborhood socioeconomic class, the diameter of the largest sign planned for the station, the number of pumps planned, the area of the approved lot, the number of competitive stations within a radius of a kilometer, and the number of hours per week the station planned to operate. The independent variables for the regression involving the arterial stations were the same, except the three describing neighborhood population, income and social class were eliminated because these stations get little business from the local neighborhood. For both regressions, the coefficients had signs as expected, and the goodness of fit were also statistically significant, although the models did not fit as well as those described in the last section.

The regressions involve mixtures of variables including market potential, the physical capacity and drawing power of the proposed station, and certain competitive conditions. The formulations assume that the quality of management – unknown in the planning phase but by far the most important factor in the station's success – will be about average. This is a rather crude assumption but it is workable, since the test of the model is its ability to forecast. The final revenue forecast is the volume forecast by the regression multiplied by an estimated gross profit per gallon of sales.

Combining cost and revenue analysis

Since the forecast is produced by a statistical technique, it is an average and there is an error variance associated with it. Combining cost and revenue analysis allows use of this variance in evaluation of both the regression model's performance and the forecasting ability of the real estate experts. For example, if the forecast just equals the volume required to break even on the investment, as for station A in figure 1, the proposed station on average will neither make nor lose money. Station B will make money on average, and station C will lose money. Station B is normally preferable under these circumstances.

Further analysis of the forecasting behavior of the real estate experts uncovered two characteristics which caused slight variation in the analysis:



DISTRIBUTIONS OF FORECAST ERRORS ABOUT THEIR MEAN VALUES
Figure 1

1. As a set of stations in a given neighborhood becomes large, there is a tendency, in addition to competitors cutting down sales volume as indicated in the regression, for prices and hence profit margins to be depressed. This fact, which in effect raises the break-even volume in such neighborhoods, was incorporated into the analytical procedures. Previously, evaluation of forecast errors considered only errors in predicting physical volumes and disregarded errors in price forecasts, which were based on standard or "normal" prices.
2. True human forecasters tended to screen out stations which would just break even and those which would on average lose money. However, one basis for evaluating their work was the number of stations operating well above their forecasts. As a result, the experts would often forecast volume for an attractive site not at

expected volume but at or near break-even volume. As a result, the stations actually built were profitable and produces volumes substantially over forecast, on average. Applying similar criteria to the regression model produced results almost identical to those forecast by the experts for a set of stations for which both sets of forecast were available. The problem of systematic underforecasting can and probably should be resolved by preparing cost and demand analyses separately, and by avoiding criteria for evaluating forecasts which treat positive and negative errors so asymmetrically.

As in the example of kiosk location, the regression involve only variables which can be measured or forecast before the station is built, so the entire procedure can easily be formalized as a routine decision aid for preliminary evaluation of potential sites. In addition to performing the regressions, the computer can easily be programmed to perform the entire cost-revenue analysis and print out a report in a format convenient for management's use.

Risk Analysis Simulation

The methodology described thus far is oriented toward producing a point estimate forecast of what, on average, the sales volume of a proposed outlet will be and whether, on average, the proposed location can be expected to surpass a break-even volume. The point estimates themselves are not exact, of course, so each has a variance or a spread about it in the form of a probability distribution of possible outcomes other than that point estimate. While the forecast outcome can be expected on average, and is usually the most likely event as well, there is always some likelihood that the actual volume achieved will differ considerably from the forecast. In addition, some of the independent variables and cost estimates may be viewed as having distributions rather than fixed point values.

Risk analysis is a branch of operations research which evaluates the managerial significance of the uncertainty inherent in the values of these variables and also of complex-relationships among them (13). Underlying risk analysis is a computer simulation which uses probability distributions to replace point estimates in those cases where there is uncertainty. Such probability distributions may govern errors at many points in the analysis.

Risk analysis has been applied to the problem of locating branch banks by Carter and Cohen (6). The basis flow chart logic for their procedure is shown in Figure 2. Probability distributions are assessed for such aspects

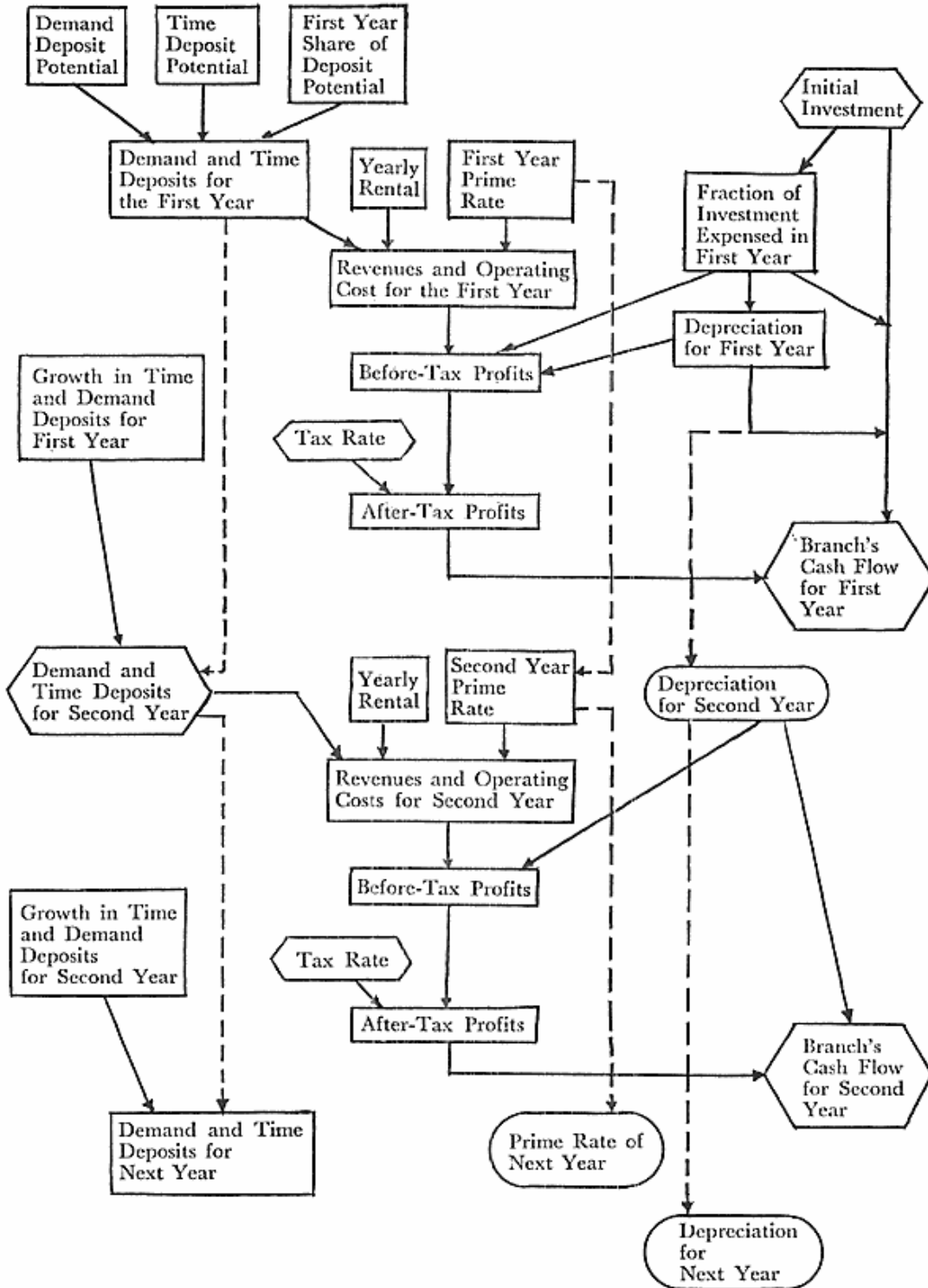
of the analysis as the initial depreciation policies. For each run of the simulation, values of these variables are drawn at random from their distributions and the analysis of the branch carried through as in Figure 2 as if the drawn values were the actual realized values. The final output of each run through the routine is an estimated rate of return for the branch; the final output from a large number of runs of the simulation, each involving different values of the input variables, is a distribution of the probabilities of different rates of return being realized. This latter distribution allows management to look not only at a point estimate of average rate of return (as with the techniques described in the last two sections) but also at measures related to riskness. Management might, for example, be willing to accept a somewhat lower expected rate of return in the variance about this rate were smaller than one associated with another alternative.

Availability of fast computers with large memories permit simulations of this type to be done efficiently and quickly, so risk analysis can be made part of the routine analysis of any branch decision that might face the firm. It is also quite feasible to incorporate regression forecasting methods into the analysis of deposit potential and cost variation; the distributions of the errors about the estimated regression lines can then be used as probability distributions in the simulations.

Some Important Research Questions

These procedures provide management with useful bases for analysis, but the branch location problem is by no means solved. First, none of the procedures provides much help in the original task of seeking out potential locations. Some heuristic routines for searching over census data for unserved market segments or neighborhoods might be very useful for developing an efficient system (17). Second, the regression models are a useful basis for forecasting, but their structure has not been validated and it is thus impossible to evaluate numerically the exact influence on one variable on the volume of a branch. Third, people often spread business over several types of related and often competing institutions – in the case of financial business, one customer may use a commercial bank, a savings bank and a postal account, for example. The motivation for using each service is often quite different, and the tendency of customers to shift this mix as a result of a branch opening is a vital but unstudied question. Such analysis might be cast in much the same framework as studies of consumers' tendencies to switch from one brand of a product to another – an area where considerable research has been done (10). Fourth, one of the techniques described here really deals adequately with interdependencies among different

Figure 2. Flow Chart of Yearly Branch Operations



outlets of a firm. In particular, while some of the volume of a new brand may result from expansion of industry sales or may be drawn from competitors, a substantial part of the new volume may simply be shifted from other branches of the same firm. Procedures have been suggested for dealing with this problem (12, 7) but the measurement problems are subtle and considerably more work on the behavior of the individual customer is needed. Finally, the difficult problems of defining a relevant market area (8, 14, 18) and relevant market segments (11) for individual branches also require further work.

Results of research on these problems which can be incorporated into anyone of the formal analyses will help clear up areas of managerially significant uncertainty in branch location.

Summary Discussion

This paper has outlined four examples of analysis involving branch locations. The structure of the models and the choice of variables differ over markets like the three studied here, but the analytical procedures are very similar. Each makes use, in one way or another, of a computer – as a data bank for easy storage of data and easy inquiry, and as a computational aid in terms of regression analysis for forecasting and in terms of simulating the market for the distribution of possible rates of return for a given investment. Each technique can easily be made a routine management aid. In general, the analytical methods are capable of handling more variables simultaneously and of considering more alternatives than management – two capabilities of the computer which will more and more shape analysis of problems such as those discussed here. The methods also relieve management of routine drudgery and provide time for more productive purposes.

REFERENCES

1. Alkjær, E.: *»Befolkningen og Bankerne«*, København 1962.
2. Andersson, B., *»Lokalisering som Konkurrensmedel«*, Handelshögskolan i Göteborg, 1967.
3. Andersson, B., *»Lokalisering som Konkurrensmedel«*, rapport från Konferensen M-68 i Malmö, 1968.
4. Appelbaum, W., *»Store Location and Development Studies«*, Worcester, Massa., 1961.
5. Bengtsson, B., Bergman, B., Gustafsson, C. and Testén, C., *»Lokalisering för Mångföretag inom Detaljhandeln«*, Handelshögskolan i Göteborg, 1968.
6. Carter, E. E. and Cohen, K. J., *»The Use of Simulation in the Selecting Branch Banks«*, *Industrial Management Review*, vol. 8. no. 2, spring 1967, pp. 55–69.

outlets of a firm. In particular, while some of the volume of a new brand may result from expansion of industry sales or may be drawn from competitors, a substantial part of the new volume may simply be shifted from other branches of the same firm. Procedures have been suggested for dealing with this problem (12, 7) but the measurement problems are subtle and considerably more work on the behavior of the individual customer is needed. Finally, the difficult problems of defining a relevant market area (8, 14, 18) and relevant market segments (11) for individual branches also require further work.

Results of research on these problems which can be incorporated into anyone of the formal analyses will help clear up areas of managerially significant uncertainty in branch location.

Summary Discussion

This paper has outlined four examples of analysis involving branch locations. The structure of the models and the choice of variables differ over markets like the three studied here, but the analytical procedures are very similar. Each makes use, in one way or another, of a computer – as a data bank for easy storage of data and easy inquiry, and as a computational aid in terms of regression analysis for forecasting and in terms of simulating the market for the distribution of possible rates of return for a given investment. Each technique can easily be made a routine management aid. In general, the analytical methods are capable of handling more variables simultaneously and of considering more alternatives than management – two capabilities of the computer which will more and more shape analysis of problems such as those discussed here. The methods also relieve management of routine drudgery and provide time for more productive purposes.

REFERENCES

1. Alkjær, E.: *»Befolkningen og Bankerne«*, København 1962.
2. Andersson, B., *»Lokalisering som Konkurrensmedel«*, Handelshögskolan i Göteborg, 1967.
3. Andersson, B., *»Lokalisering som Konkurrensmedel«*, rapport från Konferensen M-68 i Malmö, 1968.
4. Appelbaum, W., *»Store Location and Development Studies«*, Worcester, Massa., 1961.
5. Bengtsson, B., Bergman, B., Gustafsson, C. and Testén, C., *»Lokalisering för Mångföretag inom Detaljhandeln«*, Handelshögskolan i Göteborg, 1968.
6. Carter, E. E. and Cohen, K. J., *»The Use of Simulation in the Selecting Branch Banks«*, *Industrial Management Review*, vol. 8. no. 2, spring 1967, pp. 55–69.

7. Cohen, K. J. and Elton, E. J., »*Intertemporal Portfolio, Analysis Based on Simulation of Joint Returns*«, Management Science, Forthcoming.
8. Engström, C. G., »*Utvecklingslinjer inom Varudistributionen*«, Finlands Samhälls-ekonomi, pp. 239-251.
9. Farley, J. U., »*Model versus Expert in Predicting Service Station Volume*«, Working Paper, Graduate School of Industrial Administration, Carnegie-Mellon University, 1963.
10. Farley, J. U. and Kuehn, A. A., »*Stochastic Models of Brand Switching*«, in J. Schwartz, ed., Science in Marketing, New York, 1966, 446-64.
11. Frank, R. E., »*Is Brand Loyalty a Usefull Basis for Market Segmentation?*« Journal of Advertising Research, Vol. 1, no. 2, June, 1967, pp. 27-33.
12. Hartung, P. H. and Fisher, J. L., »*Brand Switching and Mathematical Programming in Market Expansion*«, Management Science, August, 1955, pp. B 231-243.
13. Hertz, D. B., »*Risk Analysis in Capital Investment*«, Harvard Business Review, January-February, 1964, pp. 95-106.
14. Huff, D. L., »*Defining and Estimating Trading Area*«, vol. 28. Journal of Marketing (July, 1964), 34-38.
15. Johnston, J., »*Econometric Methods*«, New York, McGraw Hill Book Co., 1963, Chs., 2-6.
16. Katona, G., »*Business looks at Banks*«, University of Michigan Press, 1957..
17. Kuehn, A. A., and Hamburger, »*A Heuristic Program for Locating Warehouses*«, Management Science, Vol. IX.
18. Littlefield, J. E., »*On the Art of New Branch Locating*«, Working paper, Graduate School of Business, University of North Carolina, 1968.
19. Mertes, J. E., »*Creative Site Evalution for the Small Retailer*«, Norman, Okla., U. of Oklahoma Press, 1962.
20. Moore, P. W., »*An Outline of 12 Ways to Determine Branch Locations*«, Banking, September, 1963, p. 48.
21. Moore, P. W., »*Nine Axioms of Branch Location*«, Banking, February, 1964, p. 61.