PROCEDURE FOR ESTIMATING NATIONAL MARKET AND TOTAL SOCIOECONOMIC IMPACTS OF NEW SYSTEMS OF URBAN TRANSPORTATION

Eugene T. Canty and Thomas F. Golob, Transportation Research Department, General Motors Research Laboratories

This paper presents a general procedure for determining the potential national market and total socioeconomic and environmental impacts for an urban transportation system concept that can be considered for implementation in a large number of urban areas. The procedure involves the following closely interrelated steps: (a) statistical classification of all metropolitan areas into relatively homogeneous groups on the basis of their transportation requirements; (b) selection of the most representative area in each group; (c) performance of analytical case studies in each representative area in order to synthesize the optimal system design for that area and evaluate the impacts on user and nonuser population stratifications; (d) statistical analyses of the differences among areas within the same group; (e) performance of sensitivity analyses of each case study guided by these difference analyses; (f) extensions of the results of the case studies to the other areas in each group through the use of the sensitivity and difference analyses; and (g) aggregation of the market estimates for all metropolitan areas and of the total impacts for the country as a whole by user and nonuser population stratifications. Specific methods are given for many of the steps in the procedure, and guidelines are presented for some of the more traditional planning tasks such as case study analyses.

•IN THE study of new systems of public transportation, it is important to be able to estimate the potential range of application of the new system and the consequences of its implementation. These consequences include benefits to system users and other social, economic, and environmental impacts. When private funds are employed in system research and product development, the primary concern is with the size of the market for reasons directly related to the objective of maximizing return on investments, and the various system benefits and disbenefits serve as secondary objectives and as constraints. When public funds are so employed, the proper concern is with the magnitude and distribution of benefits and disbenefits. However, the market aspects must also be considered if the hope is to attract private investment capital into new system research and development.

The problem of estimating the total market and the total social-economic-environmental impacts for new urban transportation systems, even within the single country of the United States, is difficult because of the diverse transportation requirements and environments that characterize the hundreds of metropolitan areas throughout the country. It is not feasible to conduct a detailed design, analysis, and evaluation of a new system concept in every metropolitan area. Rather, it is desirable that a procedure exist whereby the total market and nationwide impact of a new system may be estimated based on limited case studies in some minimum number of selected metropolitan areas.

Sponsored by Committee on Information Systems and Data Requirements.

This paper outlines such a general procedure for determining the total market for a transportation system concept that can be considered for implementation in any of a large number of metropolitan areas and the total user and nonuser consequences of such implementation, based on case studies in a limited number of metropolitan areas and given a body of statistical data on all metropolitan areas. It is based on a method for the classification of metropolitan areas into homogeneous groups and the identification of the most representative areas within such groups and on methods for the design and conduct of case studies within representative areas and the extension of case study results to other metropolitan areas.

It is intended that the primary contribution of this paper should be relative to the overall tasks of planning case studies, selecting case study locales, and extrapolating results to other metropolitan areas, rather than to the more specific core tasks of design, analysis, and evaluation of a single system in a single metropolitan area. However, to ensure that the results of individual case studies are applicable to other metropolitan areas, it is desirable that a certain approach be taken to such design, analysis, and evaluation. Such an approach is therefore outlined.

SUMMARY

A generalized description of the procedure is as follows:

1. Stratify and cluster the total set of metropolitan areas into a number of distinct, relatively homogeneous groups on the basis of their transportation requirements;
2. Identify the most representative metropolitan areas in each group;
3. Perform an analytical case study of the new urban transportation system in each representative metropolitan area in which the optimum form of the new system and its likely social, economic, and environmental impacts on various user and non-user population stratifications (including the probability of its implementation during a specified time period) are determined;
4. Analyze the similarities and differences among the metropolitan areas constituting each group;
5. Utilize intragroup variances such as a guide to the performance of sensitivity analyses of the design and impacts of the new urban transportation system as a function of metropolitan area characteristics;
6. Extend the results of the individual case studies in the several representative areas to the remaining metropolitan areas, group by group, making use of the intra-group variance and the sensitivity analyses results; and
7. Aggregate the estimates for all metropolitan areas to determine the probable total market for the system (and for the subsystems and components of the system) and the probable total social, economic, and environmental impacts of system implementation for the country as a whole by user and nonuser population stratifications.

Although this general procedure is conceptually simple, its accomplishment is not a trivial matter. Each of the steps in the procedure must be performed in a manner consistent with the requirements of the remainder of the procedure. Thus, for example, the extension of case study results in a few areas to the remaining metropolitan areas forces one to be fairly rigorous about the concept of representativeness and thus about the factors entering into the stratification and clustering of metropolitan areas. Moreover, it requires one to plan the conduct of case studies and sensitivity analyses so that the data product is of a form that permits extrapolation to additional metropolitan areas.

DISCUSSION

Underlying Rationale

One of the cardinal assumptions in urban transportation planning is that one can relate transportation requirements to certain measurable characteristics of metropolitan areas, such as land use patterns and intensities and existing travel behavior.

If the vector C_i represents the set of n such characteristics $(C_{i,1}, C_{i,2}, \ldots, C_{i,n})$ for metropolitan area i, and R_i denotes the set of p transportation requirements for area i, then

$$R_i = R(C_i) \tag{1}$$

It is further assumed that the optimal system design configuration for that urban transportation system in a particular area is a function of the transportation requirements for the area and the base-line system specification.

$$S_i = S(R_i, S_o) \tag{2}$$

where S_i denotes the set of q system components constituting the optimal system design configuration for metropolitan area i, and S_o denotes the base-line system specification for the q components. Thus, the optimal design can be related to the metropolitan area characteristics, or

$$S_i = S'(C_i, S_o) \tag{3}$$

Through the optimal design of generic system S_o for all m metropolitan areas in the United States, the total national market for the xth system component can be specified as

$$S_{T,x} = \sum_{i=1}^{m} S_{i,x} \tag{4}$$

where it is understood that $S_{i,x}$ might be 0 for some x components and some i areas.

Similarly, the total of the yth set of impacts, $U_{T,y}$, attributable to the implementation of the urban transportation system in the m metropolitan areas can be specified as

$$U_{T,y} = \sum_{i=1}^{m} U_{i,y} \tag{5}$$

where U_i is the set of r impacts $(U_{i,1}, U_{i,2}, \ldots, U_{i,y}, \ldots, U_{i,r})$ in metropolitan area i.

However, the detailed optimal design task for each metropolitan area is a time- and resource-consuming process, and it is not feasible to perform such a task in each of the hundreds of metropolitan areas in the United States. Rather, it is desirable to conduct a detailed case study design of system S_o in one or more metropolitan areas and then to infer the relations of the optimal designs in the metropolitan areas not studied in detail from these case study designs. The design of S_o in area j (not studied in detail) may be stated as a function of design S_i (studied in detail) as follows:

$$S_j = S_i + \underset{i \to j}{\Delta} S \tag{6}$$

where

$$\underset{i \to j}{\Delta} S = f_{i,j} (C_i, C_j, S_o) \tag{7}$$

Here $f_{i,j}$ represents a function unique to observations i and j.

The problem encountered in extrapolating optimal designs for a large number of metropolitan areas from a small number of detailed studies, as specified in Eqs. 6 and 7, is twofold. First, the detailed C_i data are often not available for all metropolitan areas and, when available, are often not compatible. Second, the $f_{i,j}$ function relating the optimal system design changes between areas i and j is often unique to each i, j pair of metropolitan areas and consequently must be continually reevaluated. The procedure outlined in this paper attacks the problem on both fronts.

Let $C_i^\dagger$ represent that particular subset of s of the characteristics C_i of metropolitan area i, which is available and compatible for all m metropolitan areas. $C_i^\dagger$ statistics are what might be labeled as aggregate statistics, such as population demographics and route-miles of roadway by total functional classifications. It is hypothesized that the $\underset{i \to j}{\Delta} S$ relation between areas i and j can be expressed as a function of the C_i^* and C_j^* subsets of the C_i and C_j spaces, or

$$\underset{i \to j}{\Delta} S + g_{i,j} (C_i^*, C_j^*, S_o) \tag{8}$$

where $g_{i,j}$ is again some function dependent on the i,j pair.

The problem of the multiple g functions is approached by restricting ΔS calculations to homogeneous regions in a new orthogonal (or independent) space derived from the s dimensional C* space. This is accomplished by first applying the multivariate statistical technique of principal components analysis to the C* data. The result is the generation of an orthogonal space of t dimensions, denoted by F, in which most of the original C* information is preserved and where $t \leq s < n$. The $\underset{i \to j}{\Delta} S$ equation can then be written as

$$\underset{i \to j}{\Delta} S = h_{i,j} (F_i, F_j, S_o) \tag{9}$$

Next, the set of all m metropolitan areas is classified into μ homogeneous groups Q_k with respect to the locations of each area in F space. It is hypothesized that the $h_{i,j}$ function relating the v_k areas within the same Q_k group is continuous and differentiable in each of the t dimensions of F. Moreover, it is hypothesized that the t first derivatives of S are constant when the most representative area, i_k, say, in group Q_k, is related to any other area j_k in Q_k. Thus,

$$\underset{i_k \to j_k}{\Delta} S = (\partial S / \partial F_{i_k}) \cdot \underset{i_k \to j_k}{\Delta} F \tag{10}$$

In other words, since the new characteristics defining metropolitan areas i_k and j_k, F_{i_k}, and F_{j_k} are independent and since areas i_k and j_k are relatively close in space F (that is, relatively similar in terms of their urban characteristics C* and thus presumably similar in terms of transportation requirements R), the optimal design of system S_o in area j_k, S_{j_k} can be approximated from the optimal design in area i_k, S_{i_k} through a set of linear relations between S and F.

The total market $S_{T,x}$ for any subsystem or component x of system S_o may be estimated by extrapolating case results to similar metropolitan areas in the manner outlined above and by summing over all areas.

$$S_{T,x} = \sum_{k=1}^{\mu} \sum_{j_k=1}^{v_k} S_{j_k,x} \tag{11}$$

Similarly, the total impact U of social, economic or environmental condition (or all of these) on a particular type y or on a particular actor set y, that would result from the full implementation S_T of system S_o may be estimated by extrapolation of the corresponding results from the case studies to other areas within the homogeneous groups and by summation over all groups.

$$U_{T,y} = \sum_{k=1}^{\mu} \sum_{j_k=1}^{v_k} U_{j_k,y} \tag{12}$$

Stratification and Clustering of Areas

The method for classifying metropolitan areas into relatively homogeneous groups with respect to their transportation requirements is based on the staged implementation of complementary multivariate statistical analysis techniques. A detailed discussion of the general classification methodology is given by Golob et al. (13); a methodological summary with specific application to urban transportation problems is contained in this paper.

General references for the various multivariate statistical techniques employed in the stratification and clustering method (and also in the identification of representative areas and the intragroup variance analysis presented in later sections) are Anderson (2), Kendall (19), Morrison (22), and Cooley and Lohnes (12). References for specific applications of the techniques are given in the following discussions.

In the first stage of the classification method, those C_i characteristics of metropolitan areas that are related to the R_i transportation requirements for the areas are selected from the set of all available metropolitan area statistics, yielding the subset C_i^*. This selection process is conducted by transportation planners (preferably a multidisciplinary team) and is accomplished with regard to the scale aspect of the baseline system S_0 under study (e.g., arterial transportation requirements as opposed to major activity center distribution requirements).

Simple and canonical correlation analyses are conducted on the selected variables in the second stage of the method. In the simple correlation phase, the existence of a high degree of association between 2 variables is identified in order to eliminate extreme multicollinearity in the data and in order to eliminate variables with a large number of missing data observations. In the canonical correlation phase, the effects of variables judged as being marginally important in the data structure are explicitly identified. Linear combinations (components) of 1 variable subset are correlated with linear combinations of a second subset, where the second subset is made up of the first subset plus the marginal variables in question. The components for each subset are linearly independent of each other (orthogonal) and are chosen such that the correlation (called canonical correlation) between the components of the first subset and the corresponding components of the second subset are maximized. If each component of the first subset is significantly correlated with only 1 component of the second subset, and conversely, then the 2 component spaces are assumed to be essentially identical, and the marginal variables are assumed to have no significant effect on the data structure. The process is repeated for the various marginal variables. Correlation significance in both phases is determined through the use of statistical distribution tests applied to product-moment correlation coefficients.

The remaining variables are then factor analyzed by using a principal-components approach in the next stage of the classification procedure. Specific expositions on this multivariate technique are given by Harmon (16) and Hotelling (17); and Green and Tull (14) and Harder (15) discuss market research applications. The principal components analysis is used to reduce the dimensionality of the data in a manner such that a minimum of information is lost; describe the new orthogonal dimensions F (called factors or components) as linear combinations of the original variable dimensions; and estimate the measurements (often called factor scores) of the metropolitan areas on these orthogonal factors. The interpretation of the factors in terms of the original variables permits a description of the basic or underlying "dimensions" characterizing the metropolitan areas and is in itself useful in analyzing the similarities and differences among metropolitan areas with respect to their transportation requirements. Examples of (nontransportation specific) studies of metropolitan areas based on factor analyses can be found in Isard (18), Berry (3, 4, 5), Moser and Scott (23), and King (20).

The output of the stratification and clustering method—the classification of metropolitan areas into μ relatively homogeneous groups Q_k, where k = 1 to μ—is produced in the final cluster analysis stage of the method. The distribution of metropolitan areas in the orthogonal factor space, given by the F_i factor scores, served as the basis for the cluster analysis. The specific cluster analysis technique used is described by Rubin and Friedman (24) and involves a hill-climbing algorithmic search for the optimal partition of a data set, optimality being measured with respect to a selected criterion. The criterion chosen is the so-called Wilkes-Lambda criterion, defined as the logarithm of the ratio of the determinant of the total data scatter matrix to the determinant of the pooled data scatter within the individual groups. Heuristically defined data reassignments and program restarts from random initial partitions are employed in order to test for local maxima.

The cluster analysis stage of the method is reinitiated for each grouping of areas into a specified different number of groups. A plot of the monotonically increasing Wilkes-Lamba criterion function versus the number of groups for a particular application is used to identify the critical number of groups (if one exists) that best describes the natural clustering in the data. (The criterion rate of increase per cluster will decline after such a critical number is passed, as opposed to the rate of increase in the criterion function immediately prior to this critical number.) Of course, in any particular application of the method, the number of groups might be a specified constant if a given number of case studies are to be performed.

A number of previous studies have grouped spatial areas by using various techniques of numerical taxonomy. These techniques have been in general less flexible than the technique presented here, have not been staged within a comprehensive classification procedure, and have not been performed with respect to urban transportation requirements; the studies are those reported by Berry (5, 6, 7), King (20), and Stone (25). Similar studies specific to urban transportation are reported by Bottiny and Goley (8) and Zenk and Frost (26).

Identification of Representative Areas

The method for identifying the most representative areas within each homogeneous group Q_k is based on the application of 2 multivariate statistical techniques using inputs from the stratification and clustering process. A detailed discussion of the methodology underlying this method is given by Golob et al. (13), and references for the methodologies of the specific statistical techniques employed are given in the previous section.

The first technique used is a correlation analysis of areas. For each separate group of metropolitan areas, a correlation matrix containing the product-moment correlation coefficients between each area in the group and each of the other areas in the group on the basis of the measurements of these areas on the final set of C_j^* variables is generated. A count of the number of significant correlations for each area then gives an indication of the overall degree of association between the area and each of the remaining areas in its group, and the rank order of the areas in each group on the basis of this count is 1 input to the identification of representative areas. This analysis process is similar to that reported by Zenk and Frost (26).

The second technique is that of discriminant analysis. In this technique linear functions of the F_i factors that best differentiate the known groups of areas are calculated. These functions are then used to reclassify the areas into groups, and the areas that prove difficult to reclassify into their original groups are identified. Subsamples of areas are used in multiple calculation of the discriminant functions in order to compensate for the bias in discriminant classification.

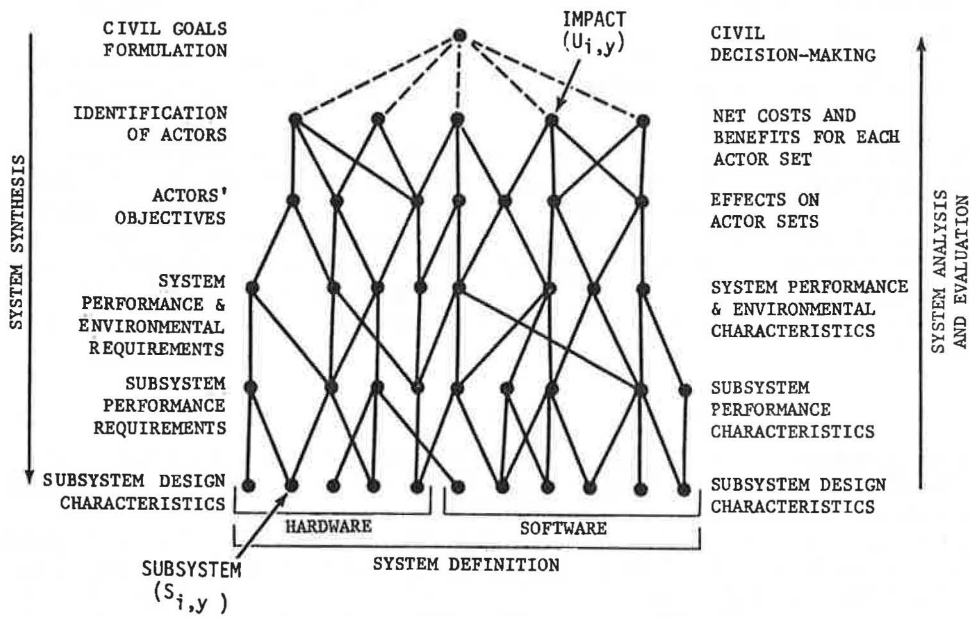
The identification of the most representative areas then involves the compilation of the results from the correlation analysis and discriminant analysis with an additional output from the cluster analysis stage of the stratification and clustering method: the matrix of generalized distances from each object to the center of each group. From these 3 inputs a rank ordering of areas in each group can be generated, partially through the subjective judgments of statistically trained analysts. The most representative areas then constitute the set of case study locales, unless considerations such as microdata availability and local planning cooperation dictate the use of the next most representative area for a particular group Q_k.

Case Studies: Design Synthesis and Analysis

Recognizing that the needs of several sets of actors are to be considered in the design of major civil systems, one should employ a design process equivalent to that shown conceptually in Figure 1. This iterative process of synthesis and analysis includes identification of the needs of multiple sets of actors, development of a system definition for that optimum design S_i in metropolitan area i, and an evaluation of the net costs and benefits U_y to each actor set y. [For discussion of this model concept and illustrative examples, see Canty (10), Alexander (1), and Liberakis (21).]

The design process shown in Figure 1 can be described as including the following activities:

Figure 1. Design process for civil systems.



1. Identification is made of the actors who would be served by or otherwise affected by the proposed new system, or would be influential in determining whether the system is to be implemented;
2. Objectives are defined for each actor set, relative to system effects;
3. The relative importance of each objective and system effect is determined for each actor set;
4. Relevance is established among subsystem design characteristics (both hardware and software), subsystem and system performance and environmental characteristics, and system effects;
5. Through the process of system synthesis, system performance and environmental requirements are translated into subsystem performance requirements and thence into subsystem design;
6. For each of a number of system configurations, through the process of system analysis, system performance and environmental characteristics are estimated;
7. System performance and environmental characteristics are evaluated in terms of their effects on the various actor sets;
8. The relative importance of each of the system effects to each actor set are considered in the evaluation of the net benefits or disbenefits to each;
9. A range of system design characteristics, both hardware and software, is considered so that one may accomplish a sensitivity analysis, tracing the impact, positive or negative, of design variations on the benefits or disbenefits to each actor set;
10. Depending on the degree of precision to which the process given above can be accomplished for any actor set, one may estimate the degree to which that system design is desirable, tolerable, or unacceptable to a particular actor set, and thus whether those actors can be expected to support, to be indifferent to, or to oppose implementation of the proposed system;
11. By variation of the system design, including both hardware and software subsystem characteristics, one may search for a preferred design—however, it is not likely that this preferred design will be optimal in the simplistic sense of maximizing total net benefits summed over all actors because the distribution of costs and benefits is also important;
12. An acceptable design is defined as one for which no net disbenefits (negative impacts exceeding positive benefits) are forecast for any actor set (preferably net

benefits are distributed among all actors in some relatively equitable manner that may require complex pricing policies, compensation formulas, assessment formulas, and transfer payments); and

13. The preferred design is the best of the acceptable designs, that is, the design that promises the greatest net benefits and also meets the criteria of an acceptable design.

One performs this process of design synthesis and analysis in a case study area i with a preconception of a base-line, or generic form, S_o of a particular transportation system concept and with an extensive body of data C_i on the case study area. Information on the people in the area, such as their value system, their transportation desires, and other transportation facilities, are included in C_i. The iterative process of design synthesis and analysis results in a preferred system configuration S_i and a set of social, economic, and environmental impacts (benefits and disbenefits) U_i. Per the adopted notation,

$$S_i = S'(C_i, S_o) = S^o(C_i) = S_{i,x};\ x = 1, 2, \ldots, q \tag{13}$$

The individual subsystems or components $S_{i,x}$ will vary in quantity or magnitude or both (including the allowance of 0 values where the system or a subsystem or component is not applicable to area i from those $S_{o,x}$ of the base-line design).

The process of analysis should also yield estimates of the forecast social, economic, and environmental impacts U_i of the proposed system on the metropolitan area. These estimates should be structured by type and by sets of actors as shown in Figure 1 to permit an analysis of the distribution of system benefits and disbenefits by type and by sets of actors (i.e., social, economic, ethnic, or civic groupings). Thus,

$$U_i = U_{i,y};\ y = 1, 2, \ldots, r \tag{14}$$

where $U_{i,y}$ is the net impact on actor set y in metropolitan area i. Also, note that $U_{i,y}$ is dependent on the characteristics C_i of metropolitan area i inasmuch as $U_{i,y}$ is a function of S_i, which in turn is a function of C_i, and inasmuch as the relative importance of the various social, economic, and environmental effects of S_i to an actor set y are dependent on the physical environment of area i, the affluence of actor group y in area i, and other factors included in the set of characteristics C_i.

It is desirable that the analysis include the development of probability statements relative to the likelihood of implementation of the preferred system and the impacts thereof in the case study areas in the planned time era. Let P_i denote the probability of implementation of the preferred system S_i in metropolitan area i. Also let $P_{i,x}$ and $P_{i,y}$ respectively denote the probability of implementation of the xth subsystem of S_i and the probability of occurrence of the yth impact, where these probabilities are not necessarily equal to P_i.

It is desirable also that the estimates of impact include a distribution overtime. It should be clear that the consideration of such probabilities and the estimation of such distributions are important elements in the system design and analysis process.

Intragroup Variance Analysis

A statistical analysis of the similarities and differences between metropolitan areas within the same homogeneous group Q_k is conducted in this step of the procedure. The primary multivariate statistical technique utilized is principal-components factor analysis, the methodology of which is discussed (and referenced) in a previous section. The objectives of the factor analysis are also similar to those specified in the earlier section except that the set of observations is the v_k metropolitan areas that are classified as members of group Q_k, and the process is repeated for each of the μ groups. These objectives are to reduce the dimensionality of the C* data in a manner such that a minimum of information is lost and to describe the new orthogonal factor dimensions as linear combinations of the original s variables constituting the C* space.

The factor spaces resulting from these analyses of the s C* variables are denoted by F^*_k $(f^*_{1,k}, f_{2,k}, \ldots, f_{w,k})$ for group Q_k $(w \leq s)$. These factors constituting the F* space may or may not be similar to the factors constituting the F space resulting from the factor analysis of the total metropolitan area population, depending on whether the areas exhibit similar or different distributions in C^* space when separated into groups or when pooled in 1 group. The independence of the factors in $F_k^\dagger$ space is important in the performance of the sensitivity analysis step of the procedure discussed in the next section, and, together with the reduction in dimensionality, aids in the interpretation of the differences among the areas classified as being homogeneous (relative to the areas in other groups).

Case Studies: Sensitivity Analyses

The case study process outlined in this paper includes 3 types of sensitivity analyses. The first of these, referred to in a preceding section, is employed in order to develop an "optimal" or preferred design and may be viewed as involving quantities of the form $\partial U_{i,y}/\partial S_{i,x}$ with C_i assumed to be constant.

The remaining sensitivity analyses have the purpose of developing estimates of the quantities $\partial S_{i_k,x}/\partial F_{i_k}^\dagger$ and $\partial U_{i_k,y}/\partial F_{i_k}^\dagger$ for use in the extrapolation of case study results S_{i_k} and U_{i_k} from area i_k in homogeneous set k of metropolitan areas to other areas j_k in k. (For simplicity, the subsubscript k will be omitted in the remainder of this section. However, it will be understood that the sensitivity analysis procedure outlined is repeated for each case study area, that is, for each of the μ values of k, assuming that there is 1 case study per homogeneous group.)

The quantities $\partial S_{i,x}/\partial F_i^\dagger$ and $\partial U_{i,y}/\partial F_i^\dagger$ are developed through a structured set of sensitivity analyses of the preferred system design and its impacts in metropolitan area i, where variations in the set of characteristics $C_i^\dagger$ are hypothesized. The analysis of intragroup variance outlined in the preceding section is utilized in structuring the sensitivity analyses. The relative loadings of the variables entering into C^* on the factors constituting the space F^* provide an indication of which of these elements of $C_i^\dagger$ should be varied as part of the sensitivity analysis, as well as a measure of their contribution to variations in $F_i^\dagger$.

Thus, those elements of C^* that are primarily responsible for the constitution of the factors in F^* are among the variables chosen. Let $\Delta C_{i,z}^\dagger$ denote some variation in the z th element of $C_i^\dagger$ such that

$$C_{i,z}^{\dagger\prime} = C_{i,z}^\dagger + \Delta C_{i,z}^\dagger$$

$$C_i^{\dagger\prime} = C_{i,1}^\dagger,\ C_{i,2}^\dagger,\ \ldots,\ (C_{i,z}^\dagger + \Delta C_{i,z}^\dagger),\ \ldots,\ C_{i,s}^\dagger \tag{15}$$

Corresponding changes are made in the larger body of C_i data to yield a new data set C_i' (for example, through a hypothetical change in land use density or per capita income values). The design process is iterated for C_i', yielding a new preferred system design S_i' and a new set U_i' of forecast impacts, such that

$$\Delta S_{i,x} = S_{i,x}' - S_{i,x} \tag{16}$$

$$\Delta U_{i,y} = U_{i,y}' - U_{i,y}$$

The change $C_{i,z}^\dagger$ is also mapped into space F^* by calculation of the changes in the factors of F^* in proportion to the loading of element $C_{i,z}^\dagger$ on those factors, producing a net change F^* such that $C_i^{\dagger\prime}$ is mapped onto $F_i^{\dagger\prime}$, where

$$F_i' = F_i^\dagger + \Delta F^* \tag{17}$$

This process is repeated for additional elements $C_{i,w}^\dagger$, resulting in alternative variations ΔF^* and alternative changes $\Delta S_{i,x}$ and $\Delta U_{i,y}$. The additional elements should be chosen in such combination as to sequentially produce variations in the w dimensions of ΔF^* or at least to ensure against multicollinearity in the variations in ΔF^*. From the set of approximately w variations in $C_i^\dagger$, one may estimate the partial derivatives.

$$\partial S_{i,x}/\partial F_i^\dagger \approx \Delta S_{i,x}/\Delta F_i^\dagger = (S_{i,x}' - S_{i,x})/(F_i^{\dagger\prime} - F_i^\dagger) \tag{18}$$

$$\partial U_{i,y}/\partial F_i^* \approx \Delta U_{i,y}/\Delta F_i^* = (U'_{i,x} - U_{i,y})/(F_i^{*\prime} - F_i^*) \tag{19}$$

Estimation of Total Market and National Impact

The total market for system S_o over m metropolitan areas is estimated by extrapolating the results of individual case studies in several metropolitan areas, which are each representative of a relatively homogeneous group of metropolitan areas, to the other metropolitan areas that constitute each such group, and by summing over the several groups. A similar process is involved in estimating the total national impact of such system implementation. It will be desirable to stratify the market for system S_o by subsystems and components and to stratify the national impact by social, economic, and environmental impacts on various sets of actors.

Let i_k be a representative metropolitan area in a homogeneous group k containing v_k members, and let a set of system design and analysis and sensitivity analysis exercises be accomplished for system concept S_o in area i. Let j be any of the $(v_k - 1)$ other members of group k. Because the points $C^*_{i_k}$ and $C^*_{j_k}$ are known in C* space and the corresponding points $F^*_{i_k}$ and $F^*_{j_k}$ in F* space, one may approximate the preferred value of system design S_{j_k} and system impacts U_{j_k} thusly:

$$S_{j_k} = S_{j_k,x};\ x = 1, 2, \ldots, q \tag{20}$$

$$S_{j_k,x} = S_{i_k,x} + (\partial S_{i_k,x}/\partial F_i^*) \times (F^*_{j_k} - F^*_{i_k}) \tag{21}$$

$$U_{j_k} = U_{j_k,y};\ y = 1, 2, \ldots, r \tag{22}$$

$$U_{j_k,y} = U_{i_k,y} + (\partial U_{i_k,y}/\partial F^*_{i_k}) \times (F^*_{j_k} - F^*_{i_k}) \tag{23}$$

The total market for system S_o may be expressed as follows:

$$S_T = S_{T,x};\ x = 1, 2, \ldots, q \tag{24}$$

$$S_{T,x} = \sum_{k=1}^{\mu} \sum_{j_k=1}^{v_k} P_{j_k,x}\ S_{j_k,x} \tag{25}$$

If separate probability statements are not developed for each j_k by extrapolation from the values for areas i_k, an alternative is to apply the same probability estimates to all members of the homogeneous group k. Thus,

$$S_{T,x} = \sum_{k=1}^{\mu} P_{i_k,x} \sum_{j_k=1}^{v_k} S_{j_k,x} \tag{26}$$

$$S_{T,x} = \sum_{k=1}^{\mu} \left\{ P_{i_k,x} \sum_{j_k=1}^{v_k} \left[S_{i_k,x} + (\partial S_{i_k,x}/\partial F_i^*) \times (F^*_{j_k} - F^*_{i_k}) \right] \right\} \tag{27}$$

Similarly,

$$U_T = U_{T,y};\ y = 1, 2, \ldots r \tag{28}$$

$$U_{T,y} = \sum_{k=1}^{\mu} \left\{ P_{i_k,y} \sum_{j_k=1}^{v_k} \left[U_{i_k,y} + (\partial U_{i_k,y}/\partial F^*_{i_k}) \times (F^*_{j_k} - F^*_{i_k}) \right] \right\} \tag{29}$$

APPLICATIONS

The procedure described here is being applied by the Transportation Research Department of the General Motors Research Laboratories in a study of a particular arterial transportation system concept—the Metro Guideway. The Metro Guideway concept is described by Canty (11) and consists of integrated facilities for dual-mode automobiles, personal rapid transit and group rapid transit, and automated freight movement. However, the procedure is not restricted in application either to the Metro Guideway system or even other transportation systems at the metropolitan scale. The general procedure should also be useful in the study of other urban systems such as education, housing, and medical care, and the study of transportation systems that have application at other levels of urban structure, e.g., major activity centers, central cities, and townships, as well as at the metropolitan scale. A general consideration of the applicability of various types of transportation systems at various levels of urban structure and scale is given by Canty (9).

REFERENCES

1. Alexander, C. A City Is Not a Tree. Architectural Forum, May 1965.
2. Anderson, T. W. An Introduction to Multivariate Statistical Analysis. John Wiley and Sons, New York, 1958.
3. Berry, B. J. L. An Inductive Approach to the Regionalization of Economic Development. In Essays on Geography and Economic Development (Ginsburg, N., ed.), Univ. of Chicago Press, 1960.
4. Berry, B. J. L. Cities as Systems Within Systems of Cities. Papers and Proc. of Regional Science Assn., Vol. 13, 1963.
5. Berry, B. J. L. Identification of Declining Regions: An Empirical Study of the Dimensions of Rural Poverty. In Areas of Economic Stress in Canada (Wood, W. D., and Thoman, R. S., eds.), Industrial Relations Centre, Queen's Univ., Ontario, 1965.
6. Berry, B. J. L. Essays on Commodity Flows and the Spatial Structure of the Indian Economy. Dept. of Geography, Univ. of Chicago, Res. Paper III, 1966.
7. Berry, B. J. L. Grouping and Regionalizing: An Approach to the Problem Using Multivariate Analysis. In Quantitative Geography, Part I: Economic and Cultural Topics (Garrison, W. L., and Marble, D. F., eds.), Dept. of Geography, Northwestern Univ., 1967.
8. Bottiny, W. H., and Goley, B. T. A Classification of Urban Areas for Transportation Analysis. Highway Research Record 194, 1967, pp. 32-61.
9. Canty, E. T. Transportation and Urban Scale. High Speed Ground Transportation Jour., Vol. 3, 1969.
10. Canty, E. T. The Demand-Responsive Jitney: A Socially-Oriented Transportation System Design Study. Presented at 13th FISITA Congress in Brussels, June 1970.
11. Canty, E. T. Metro Guideway: An Integrated Urban Transportation System. Highway Research Record 397, 1972, pp. 26-37.
12. Cooley, W. W., and Lohnes, P. R. Multivariate Procedures for the Behavioral Sciences, John Wiley and Sons, New York, 1962.
13. Golob, T. F., Canty, E. T., and Gustafson, R. L. Toward the Stratification and Clustering of Metropolitan Areas. Presented at National Meeting of Operations Research Society of America in Dallas, May 1971.
14. Green, P. E., and Tull, D. S. Research for Marketing Decisions, 2nd Ed., Prentice-Hall, Englewood Cliffs, N.J., 1970.
15. Harder, T. Introduction to Mathematical Models in Market and Opinion Research. D. Reidel, Holland, 1969.
16. Harmon, H. H. Modern Factor Analysis. Univ. of Chicago Press, 1960.
17. Hotelling, H. Analysis of a Complex of Statistical Variables Into Principal Components. Jour. of Educational Psychology, Vol. 24, 1933.
18. Isard, W. Methods of Regional Analysis: An Introduction to Regional Science. M.I.T. Press, Cambridge, 1960.

19. Kendall, M. G. A Course in Multivariate Analysis. Charles Griffin and Co., London, 1965.
20. King, L. J. Cross-Sectional Analysis of Canadian Urban Dimensions: 1951 and 1961. Canadian Geographer, Vol. 10, No. 4, 1966.
21. Liberakis, A. The Human Factor Introduced in Hospital Architecture by the Pattern Language. Ekistics, Vol. 186, May 1971.
22. Morrison, D. F. Multivariate Statistical Methods. McGraw-Hill, New York, 1967.
23. Moser, C. A., and Scott, W. British Towns: A Statistical Study of Their Social and Economic Differences. Oliver and Boyd, London, 1961.
24. Rubin, J., and Friedman, H. P. On Some Invariant Criteria for Grouping Data. Jour. of American Statistical Assn., 1967.
25. Stone, R. A Comparison of the Economic Structure of Regions, Based on the Concept of Distance. Jour. of Regional Science, Vol. 2, No. 2, 1960.
26. Zenk, G. K., and Frost, M. Factors Analysis for City Selection. In Systems Analysis of Urban Transportation, General Research Corp., Vol. 2, Santa Barbara, 1968.