CENSUS DATA, LAND USE, AND TRANSPORTATION MODELING

R. D. Worrall
Peat, Marwick, Mitchell and Company, Washington, D.C.

The focus of this paper is on 4 specific topics: land use modeling, travel forecasting, impact analysis, and data maintenance and updating. The first 3 topics deal with selected applications of census materials to particular kinds of modeling activities. The fourth is concerned with the separate, but related, problem of the maintenance and updating of these materials over time.

LAND USE MODELING

All land use models have the common objective of estimating the most probable future distribution of urban activity, defined usually in terms of population and employment, in a form suitable for input to detailed transportation and land use planning studies. Nearly all of the models constructed to date can build, either directly or indirectly, on census inputs.

Consider for example, the EMPIRIC activities-allocation model, currently being applied by the Metropolitan Washington Council of Governments. The model is designed to perform 2 main functions:

1. To generate small-area (e.g., district-level) forecasts of the future distribution of population, employment, and land use for input to subsequent trip-generation and modal-split analysis, and
2. To provide a base for evaluating the impact of alternative planning-policy decisions on urban growth and development (e.g., the effect on growth of alternative transportation plans).

Figure 1 shows the role of the model within the overall framework of the comprehensive planning process. The model consists mathematically of a set of simultaneous, linear equations that relate predicted changes in the district-level distribution of population to similar changes in employment for the base year and the future distribution

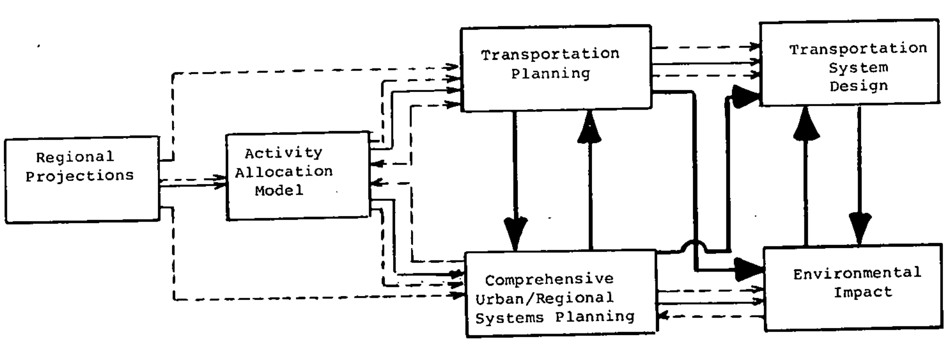

Figure 1. Role of an activity-allocation model in comprehensive regional and transportation planning.

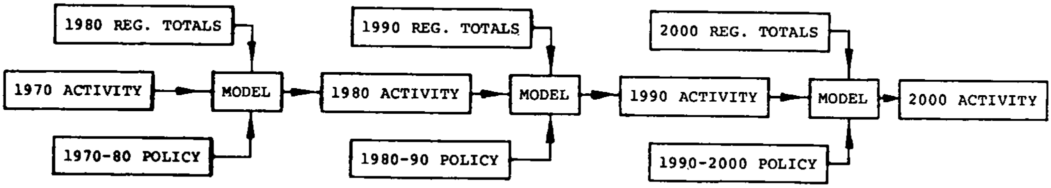

Figure 2. Forecast procedure.

and also relates these changes to alternative planning policy decisions. Activities are defined in the model in terms of classified, district-level population and employment counts. Alternative policies that may be included in the model include investments in future transit, highway, and public-utility systems; subregional growth and zoning controls; open-space policies; and location of major commercial or governmental developments.

The model is calibrated (Fig. 2) by using data collected for 2 separate points in time. It is then applied recursively to generate a succession of forecast distributions of population and employment for different points in time. Each of these forecasts is conditional on the pursuit of a particular set of planning policies and on the attainment of a given total level of regional population and employment. The model can be used both as a straightforward forecasting tool and also as a mechanism for evaluating the probable consequences (in terms of regional growth) of alternative planning policies (Fig. 3).

The model requires 2 separate sets of input data. For model calibration, information is required on the spatial distribution of land use, population, and employment, plus the values of each policy variable, for the start and end of the calibration period. For forecasting, equivalent data are required on each activity for the initial year of the forecast period, together with the predicted changes in each policy variable and the estimated future region-wide activity totals at the end of the forecast period.

In the case of the Washington model, this implies the following specific data inputs:

1. Activity data (in compatible form for 2 separate points in time, e.g., 1960 and 1970, and for each district) for population by age (4 to 6 classes), by household income (4 to 6 classes), and by household size (4 to 6 classes); for em-

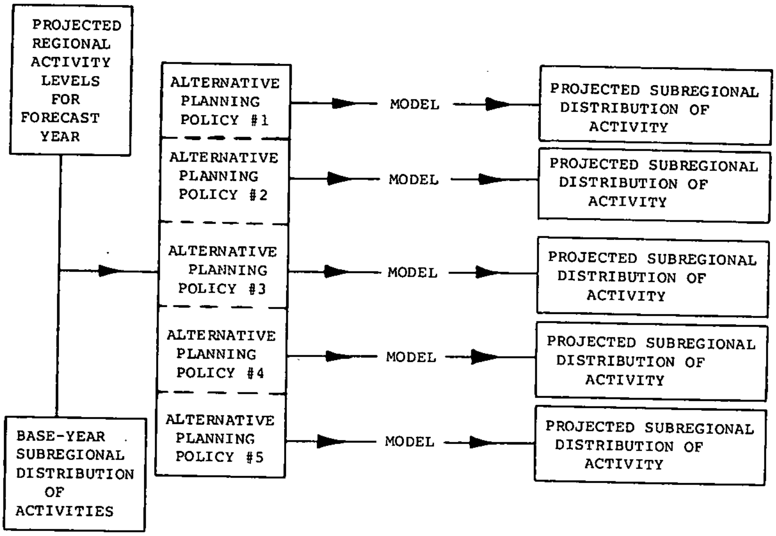

Figure 3. Policy evaluation of 5 alternatives in a single forecast period.

ployment by number of industrial, retail and consumer service, office-federal government, office-nonfederal government, and other employees; and for land use by total area in residential, industrial, consumer-service, office-federal government, office-nonfederal government, institutional, recreational, vacant and agricultural, and restricted, i.e., not available for development, uses (restricted use averages are also required for each forecast period);

2. Policy data for the 2 calibration years, 1960 and 1970, plus each forecast year, 1980, 1990, 2000, and so on, including district-level highway and transit travel-time matrices, district-level water-service measures, district-level sewerage-service measures, location and size of major commercial-governmental developments, zoning controls and open-space requirements, and acreages in restricted uses; and

3. Regional control totals for each activity, broken down as in item 1, together with equivalent region-wide forecasts for each forecast year.

Two, and possibly three, of these pieces of input data can be generated directly from census sources. The first and most obvious of these is the set of data relating to district-level population. Compatible breakdowns of population by age, household size, and household income are available as standard output from the census, for both 1960 and 1970, and for a variety of geographic units. For model-building purposes, tract-level tabulations will, in most cases, be satisfactory. Where district and tract boundaries do not coincide, the DIME system of the Bureau of the Census provides a convenient means for aggregating 1970 block-level statistics to the district level. This latter capability, however, applies only to 1970 data.

1970 employment data may similarly be derived directly from the 1970 census by making use of the proposed special tabulations of employment by place of work. Again, the DIME capability permits the aggregation of these data according to both study districts or census tracts. However, such employment information is obtainable only from the 1970 census and not from the 1960 or prior censuses. Small-area employment data for periods prior to 1970 must be obtained from other sources (e.g., first-work-trip data from an existing home interview study or data derived directly from state employment records).

The third potential application of census data to a model of the type discussed here is less clear. Transportation policies are expressed within the model in terms of changes in district-level highway and transit accessibilities. These in turn are derived from the calibration of a conventional, district-level gravity model. The detailed coding of home-to-work data in the 1970 census, coupled again with the special home-to-work tabulations proposed by the Federal Highway Administration, makes it possible to calibrate a gravity model based only on census inputs supplemented by network information. The resultant friction factors and travel-time matrices can then be used as input to the accessibility measures, which themselves are used as policy variables in the land use model.

EMPIRIC is, of course, only one of a variety of different land use models that might conceivably make use of census inputs. Its utilization of census data, as illustrated here, is broadly representative of the potential use that may be made of the data by other models. In each case, the major census products of interest are standard population and employment estimates. The model must have the capability of matching these data to a variety of different areal systems and separately determined planning-policy and land-area inputs. In this connection, the DIME capability incorporated in the 1970 census is of particular importance.

In the case of growth rather than cross-sectional models that require data for 2 separate points in time, there are obvious advantages in matching the time periods for model calibration and forecasting with those of the census. This necessarily implies (neglecting the possibility of a quinquennial census) a time interval for model building of 10 years. In instances where this is considered to be too large, it is clearly necessary to supplement census-based data with information from other sources. (This point will be taken up again later.)

TRAVEL FORECASTING

In the case of travel forecasting, a number of relatively straightforward applications of 1970 census products may be identified. These range from trip-generation and automobile-ownership analysis to students of work-trip modal split and trip-distribution.

Potential census inputs to trip generation analysis to a large extent parallel those outlined earlier for land use models. Again, census data provide a base for estimating and updating the values of district-level or zonal-level population and employment variables, which themselves serve as the independent variables in conventional trip-generation equations. In addition, the detailed journey-to-work data and work-place coding incorporated in the 1970 census also make it possible to perform a basic trip-generation analysis using only census data as input. This latter analysis, though restricted solely to work trips, may be broken down in some detail according to the characteristics of both the origin and the destination zone and also by mode.

It is also possible, again through special tabulations of the type proposed by the Federal Highway Administration, to use selected samples of census data for detailed household trip-making and trip-generation analysis. These, in turn, may be fed into a disaggregate trip-generation model. Similarly, the data may also be used as a basis for a modal-split (again, only for the work trip) and automobile-ownership analysis. A number of automobile-ownership models have already been successfully developed by using only census data as input.

In this context, however, the most important provision of the 1970 census materials is an extremely valuable basis for generating both a zonal-level and a district-level table of work-trip origins and destinations. This table, when supplemented by information on equivalent district-level or zonal-level travel times may be used directly as input to a trip-distribution analysis.

Considerable volumes of data have been assembled and considerable amounts of money have been spent during the past 2 decades on the collection of information for transportation planning. The utility of data from such large-scale transportation data-collection exercises is now seriously in question, partly because of their sheer expense and partly because of some doubt as to whether data in the quantities conventionally assembled by urban transportation planning studies are either necessary or useful. The journey-to-work data included in the census provide both a convenient and an inexpensive means for updating many existing data sets and can serve as an effective base for the design of future facilities. The fullest possible use should be made of this information before an agency engages in further original data collection.

TRANSPORTATION IMPACT ANALYSIS

Perhaps the most critical area of transportation analysis during the next decade is going to be that of identifying and evaluating the impact of alternative transportation plans and policies on the areas and populations that they are designed to serve. Census data can play an extremely important role in such analysis.

This may range from the simple characterization of the social and economic structure of the areas through which a new freeway or transit system is designed to pass to a detailed analysis of the incidence of transportation impacts on particular groups of the population or the identification of particular planning needs. An example of the latter might, for example, be the isolation of low-income, low-accessibility areas that require immediate transit system improvement. Alternatively, attention might be focused on the spatial distribution of job opportunities or the particular transportation needs of school children or elderly persons. All of these may build directly on census products.

In the case of impact analysis, census data provide at least an initial base for the evaluation of questions relating to household relocation, population relocation, accessibility to employment, incidence of impact on specific demographic and income groups, and potential effects of a new system on existing market areas and overall regional growth.

DATA MAINTENANCE AND UPDATING

My comments have focused on the application of census data to 3 particular areas of analysis. The fourth and final topic that I would like to consider briefly concerns not the use of census data but rather its maintenance and updating over time. Reference has already been made to some of the potential problems posed by the decennial nature of basic census materials. This in no sense negates the utility of the basic census products, nor does it ignore the utility of many of the regular intercensual activities of the Bureau of the Census. Three basic points should be made concerning the maintenance and updating of standard census data. These involve questions of update frequency, areal systems, and data sources.

In terms of transportation planning, it would appear that an average updating period of approximately 5 years is desirable for both population and employment data and also, if possible, for basic trip information. Although more frequent updating may appear to be desirable for particular instances, it is doubtful that the cost of this could be justified on a systematic basis. This in turn suggests a need for a regular program of data updating based to the extent possible on standard census methods and definitions.

Similarly, the DIME system and its related ACG and ADMATCH components must be continually updated over time if their utility is to be maintained. This requires an explicit allocation of resources by an agency employing the DIME system to its maintenance and updating, preferably on an annual basis.

All updated materials should be assembled at the same level and in the same format used to summarize the original census information. This implies, in most cases, either the block or census tract level, depending on the information in question and the geographic location.

The question of data sources for updating is perhaps the most critical single problem associated with the maintenance of a "current" data base. A variety of both secondary (i.e., existing) and primary (i.e., original surveys) sources may be considered.

Secondary sources that may be used for updating purposes include school enrollment, tax assessor and building department records, and, in some instances, state employment security information. The major problems likely to arise in this context are those of geographic specificity and potential incompatibility of definition. These secondary sources may, in many cases, be supplemented at a relatively small cost by small sample surveys designed to update those items of census data as a sampling frame and also by data from the regular data-collection activities of other local agencies. One interesting example is that of the Landmark Corporation in the tidewater area in Virginia. The corporation, primarily as a service to the local business community, conducts an annual sample survey of the basic demographic characteristics of the regional population and of market preferences and consumer behavior. The information is summarized and tabulated for a variety of geographic units including aggregations of census tracts. Such surveys, when complemented by the regular intercensual activities of the Bureau of the Census and the performance of selected special studies by individual agencies (for example, the collection of small samples of trip data by a transportation planning agency), provide a relatively solid base for updating base-year information and maintaining census-based data files in a reasonable current state.

SUMMARY

I have touched briefly on 4 particular topics concerning the potential applicability of census data to land use and transportation modeling. In no sense have I exhausted the subject, nor have I pursued any one of the topics as far as it deserves. My intent has been simply to illustrate the broad applicability of census data to the activities of the transportation and land use planner and to suggest some ways in which he may usefully take advantage of an extremely rich and valuable data source.