THE ROLE OF SIPP IN MICROSIMULATION

Richard F. Wertheimer II, The Urban Institute

Adequate and reliable data are the key to the analysis of the distribution of benefits and costs of existing and proposed government programs. The Income Survey Development Program (ISDP) was established in 1975 to remedy the inadequacy of existing microdata surveys. The ISDP's main purpose is to develop a new survey, the Survey of Income and Program Participation (SIPP), to provide improved information on program eligibility, participation, and benefits, as well as to provide more comprehensive income and wealth data. For a detailed review of both SIPP and the ISDP, see Lininger (1980) and Ycas and Lininger (1980).

AN OVERVIEW OF CURRENTLY USED MICROSIMULATION MODELS

Since the mid-1960s there has been increasing acceptance of microanalytic simulation as the major tool for analyses of the government tax and transfer system.[1] The microsimulation method uses household survey data as an input to simulate the receipt of transfer income at the family level and then aggregates these benefits up to the program level. This methodology is used to evaluate existing tax and transfer programs and the effects of major and minor modifications to existing programs, and to develop and analyze new tax and transfer programs.

The following models are seven of those most commonly used:

o TRIM: The Transfer Income Model (TRIM) was developed at The Urban Institute as a replacement for RIM.[2] It was designed to answer questions relevant to the formulation of social welfare policies. The core of TRIM is a set of modules in which the rules and regulations of various tax and transfer programs are applied to the characteristics of individual households in a survey, and program outcomes are calculated by aggregating the individual incomes. TRIM is capable of projecting a sample to reflect a later year than the survey year. This "aging" procedure is accomplished by changing the sample weights to capture the changing age, race and sex distribution of the population, as well as the overall growth, and by applying multiplicative factors to income by type at the micro level. Since its inception, the TRIM model has been used continuously for policy analyses. The TRIM model is currently housed at The Urban Institute, HHS, and other government agencies.

o MATH: The Micro Analysis of Transfers to Households (MATH) model is an offshoot to TRIM.[3] (Mathematica Policy Research adopted a version of TRIM and incorporated some changes to it.) MATH maintains many of the same aging procedures as TRIM, some more complex aging procedures, and includes similar sets of rules to simulate tax and transfer programs, but also includes a simulation of a labor supply response to changes in a family's disposable income and effective wage rates resulting from proposed income maintenance programs. In addition, the MATH model can select public assistance recipients or participants from among the simulated eligibles. The MATH model has been used for many policy studies in the areas of energy, unemployment and welfare reform.

o KGB: The KGB model has recently been developed in the Office of Income Security Policy within the Department of Health and Human Services.[4] Its purpose is also to evaluate various welfare reform alternatives. It is a hybrid model in the sense that it incorporates features of the TRIM, MATH and DYNASIM models. The KGB model includes the TRIM input requirements and some of its modules which simulate cash transfer programs; the MATH model's labor supply response simulation module; and the DYNASIM unemployment experience modules. In addition, the model simulates behavioral responses to the availability of public employment programs. This model is relatively new, still evolving, and has had limited application to policy debate thus far.

o STATS: This microsimulation model was developed in the Office of Research and Statistics (ORS) within the Department of Health and Human Services.[5] This model, which has many features in common with TRIM, was developed concurrently and has been used by ORS to analyze welfare participation and social security payroll tax relief. Its aging procedure is similar to that of TRIM in that it adjusts population weights on the basis of age, race and sex to meet the Census Bureau projections. In addition, STATS makes some projection assumptions about the proportion of married males in the projected sample population. Also, like TRIM, economic "multipliers" are used to adjust nontransfer income sources, labor force participation and unemployment.

o DYNASIM/MASH: The development of the Dynamic Simulation of Income Model (DYNASIM), undertaken at The Urban Institute, incorporates a different approach to microsimulation than the TRIM, MATH, KGB or STATS models.[6] Rather than use reweighting procedures, DYNASIM simulates year-to-year changes in socioeconomic characteristics of sample members by applying estimated behavioral relationships to individual micro units. The model is capable of simulating the rules of the major transfer income programs. The primary use of this model has been in the area of retirement income transfers because this model is well-suited to long-term projections, and it is the only model which includes detailed simulations of private pensions and social security benefits.

o DYNASIM/MASS: This microsimulation model is an offshoot of DYNASIM/MASH and is identical in concept.[7] It was developed at Yale University and is currently also housed at The Urban Institute. This model was developed in response to the need for a model which was capable of simulating behavioral events, but was not as costly to use as the DYNASIM/MASH model. Thus, much of the rich detail of the DYNASIM/MASH model was omitted. For example, none of the transfer program rules is simulated, and some of the behavioral modules are simplistic. This model is relatively new and is still evolving.

o OTA's Personal Income Tax Model: The Office of Tax Analysis within the Office of the Secretary of the Treasury has developed a microanalytic model which simulates the fundamental interacting parameters of the U.S. tax structures.[8] This

model uses a unique data base which is constructed from a stratified random sample of individual income tax returns and demographic data obtained either from an exact match with social security administrative data or from a statistical match with either CPS or SIE survey data. The model has a projection capability which is similar to the reweighting procedures of the TRIM-like models.

Each of these models has general limitations and weaknesses. Some data limitations are universal in the sense that they are not specific to any model, but, instead, are limitations which are inherent to the data base. One example of this is the undercoverage problem in the Current Population Surveys.

In addition to undercoverage errors, the CPS suffers from underreporting and nonreporting of income. Similar problems exist on the Survey of Income and Education (SIE) file. Such underreporting and nonreporting of income create serious problems for the accurate simulations of tax and transfer programs. Other general data base limitations which affect all models include the following: (1) panel surveys such as the CPS and SIE collect income information for the year preceding the survey so that the survey reflects demographic characteristics of the household for the current year and income data for the previous year; (2) income generally includes only money income, excluding in-kind income or intra-family transfer, etc.; (3) income is generally reported in annual amounts only; (4) data on assets and health limitations are limited on many of the files; (5) expenditure data are almost nonexistent; and (6) with the exception of the 1970 Census Public Use Samples (CPUS) and the SIE, the data are unreliable at the subnational level.

In addition to these general data limitations, certain data constraints, specific to individual models, exist. These data limitations compromise the accuracy of the models' estimates since each introduction of synthetic data elements increases the probability of error.

In addition to suffering from data base limitations, each of these models could be improved in terms of its simulation routines. Often the lack of adequate data for estimating behavioral relationships or providing an adequate initial data base have discouraged modelers from undertaking these improvements.

For example:

- o Some models would be more realistic if a different accounting period were used.
- o All the models could benefit from more detailed and accurate income and assets data.
- o TRIM and MATH would be more useful if they could simulate unemployment insurance.
- o The models which simulate eligibility for transfer programs would benefit from inclusion of a labor supply response.
- o These same models would benefit from inclusion of behavioral program participation modules.

In short, there are many ways in which the microsimulation models could benefit from improved survey data. The next section explores the role which SIPP can play in the future in improving the capability and performance of these models.

THE ROLE OF SIPP IN MICROSIMULATION MODELS

SIPP will be yielding data better suited for microsimulation models than have been hitherto available. Moreover, there is the potential for making changes in survey design which will further increase their usefulness.

Unlike the CPS, which is designed principally for gathering employment data, SIPP is intended primarily to provide data on the income and wealth of the population. Thus, the items included in SIPP differ in a number of important ways from the items included in other, currently available data bases. These differences include:

(1) The reporting of income by detailed source. SIPP will provide detailed data on all sources and amounts of income. For example, the survey will include detailed data on wage rates and earnings, income by source, and a breakdown of other income by source (e.g., alimony, child support). In contrast, the March CPS currently provides data for only a limited number of sources, for example, lumping together most public assistance programs and including all types of pensions together. In simulations with the CPS, these aggregated amounts must be allocated to detailed sources before programs can be simulated.

(2) The inclusion of in-kind income. SIPP will include data on food stamp recipiency, on whether families live in public housing, and on medicare and medicaid coverage and use. These programs may represent significant sources of income to some recipients, but they are not included in the regular CPS surveys.

(3) The inclusion of monthly income detail. Although eligibility for most transfer programs is determined by monthly income, currently available data sources provide only annual income data. This data lack leads to problems with estimating eligibility with microsimulation models. However, SIPP will include detailed income data for each month of the survey year.

(4) The inclusion of detailed asset data. For many programs, eligiblity is determined by assets as well as income, but SIPP will be the first regular survey program to gather detailed asset data. These data should help increase the accuracy of simulations of program eligibility.

(5) The longitudinal nature of the data. SIPP will be the first data base to include longitudinal data on income for a large sample; and, because of this, it will allow new types of analysis using microsimulation models. Using the new data it will be possible to follow month-to-month changes in program eligibility, whereas with current data it is only possible to estimate eligibility at a single point in time.

(6) The following of splitoffs from families throughout the survey year. Current data sources present a picture of a family at only a single point in time, SIPP will present a longitudinal view of the family structure for the duration of the survey year. This type of data may increase the accuracy of the simulation of certain government programs. For example, estimates of exemptions for tax computations currently must rely on the structure of the family at a single point during the year, while with the new database, estimates may use the actual family structure throughout the year.

(7) The addition of other new data items. SIPP will include a wealth of data on personal background, education, and attitudes which may be of interest to analysts of tax and transfer programs.

In some sense, SIPP will be an embarrassment of riches. The microsimulation models as currently implemented are accustomed to a scarcity rather than an abundance of data--especially in the critical areas of income, assets, and labor force behavior.

Adapting the models to take advantage of this wealth of new data will be a large job. An example of this problem is the complexity of constructing for use in DYNASIM a variable measuring hours worked in a calendar year. In both the Census Public Use Sample (CPUS) and the Current Population Survey (CPS), this variable is constructed by multiplying hours worked last week by weeks worked last year. In the 1978 Research Panel,[9] data on both weeks and hours are available by job and by calendar quarter. In the 1979 Research Panel,[10] the situation is further complicated by the staggered interviewing which results in three different definitions of quarter (only one of which is a calendar quarter). No doubt, this is a far more accurate method of measuring total hours worked. But a significant programming effort will be required to extract the single number needed by DYNASIM from the dozens of variables supplied by the Research Panels.

Similarly, the presence of a variable indicating the parent of each child makes disentangling nuclear families from households more complicated but, almost certainly, more accurate.

TRIM, MATH, KGB, and STATS face similar difficulties in dealing with detailed income data that are available by person and by month rather than aggregated data that are available by family and by year. Routines must also be added to these models to use the detailed assets data. Assets portfolios and changes in these portfolios are available by person and by quarter rather than being nearly nonexistent.

At first, modelers will probably concentrate on adapting SIPP survey files to work in the models in their existing forms. This alone should improve the accuracy of the models. But over time, modelers will begin to expand the capabilities of the models to take advantage of the new data.

The first area likely to receive attention is the accounting period. Due to data limitations, most models simulate program eligibility on an annual basis even though the programs generally use a monthly accounting period. Each of the models could be adapted to run on the monthly (or quarterly) data available from SIPP.

A second important area which is like to receive attention is the simulation of program participation. Although there has been a fair amount of research done on this issue, there are still no generally accepted models to predict which families or persons among a set of eligibles will choose to participate in a transfer program. SIPP's detailed data on participation combined with its detailed data on the variables used to determine eligibility should provide the basis for estimating these models.

Other areas of somewhat lower priority include dynamic simulation of asset portfolios, simulation of taxes, and simulation of consumer expenditures. In each case, SIPP provides a combination of data items previously unavailable.

In short, SIPP can be expected to contribute to a significant advance in the state of microsimulation modeling.

FOOTNOTES

1. See Orcutt, Glazer, Harris, and Wertheimer (1980) for a history of the development of microsimulation models.
2. The TRIM model is discussed in Sulvetta (1976).
3. For a discussion of MATH, see Beebout (1977).
4. For a discussion of the KGB model, see Betson, Greenberg, and Kasten (1979).
5. Full documentation of this model is underway. Some examples of its use can be found in Projector and Murray (1978) and Bridges and Johnston (1976).
6. For a description of this model, see Orcutt, Caldwell, and Wertheimer (1976).
7. For a brief description, see Orcutt, Glazer, Harris, and Wertheimer (1979). DYNASIM/MASS also has been adapted by Hendrickson Corporation and renamed "MICROSIM." At present there is no documentation of MICROSIM.
8. For a description of this model, see Wyscarver (1978).
9. The Research Panels are pilot surveys which were used to test questionnaires, survey techniques, etc., prior to beginning SIPP.
10. Ibid.

BIBLIOGRAPHY

Beebout, Harold. "Microsimulation as a Policy Tool: The MATH model," Policy Analysis Series No. 14, Mathematica Policy Research, Washington, D.C., February 1977.

Betson, David, David Greenberg and Richard Kasten. "A Simulation of the Program for Better Jobs and Income: Technical Analysis Paper No. 17, ASPE, U.S. Department of Health, Education and Welfare, Washington, D.C., January 1979.

Bridges, Benjamin Jr. and Mary P. Johnston. "Estimation of Social Security Taxes on the March Current Population Survey," Studies in Income Distribution No. 4, Social Security Administration, Office of Research and Statistics, U.S. Department of Health, Education and Welfare, Washington, D.C., March 1976.

Chow, George S. and Richard C. Michel. "A Programmer's Primmer for TRIM: The Transfer Income Model," Urban Institute Working Paper 996-06, The Urban Institute, December 1977.

Doyle, Pat, David Edson, Norma Pappas, and Bill Boulding. "Creation of 1980 and 1984 Data Bases from the March 1978 Current Population Survey," Mathematica Policy Research, Washington, D.C., February 1980.

Doyle, Pat and Kevin Neyland (eds.) "The MATH Codebook," Mathematica Policy Research, Washington, D.C., September 1979.

Doyle, Pat and Kevin Neyland (eds.). "The MATH Technical Description," Two Volumes, Mathematica Policy Research, Washington, D.C., August 1979.

Greenberg, David H. "Participation in Guaranteed Employment Programs: An Exploratory Simulation," Public Service Employment, Supported Work, and Job Guarantees: Analytic Issues and Policy Implications, edited by John Palmer, The Brookings Institution, Washington, D.C., (forthcoming).

Institute for Research on Poverty. "A Microsimulation Model for Analyzing Alternative Welfare Reform Proposals: An Application to the Program for Better Jobs and Income," Microeconomic

Simulation Models for Public Policy Analysis, edited by Robert Haveman and Kevin Hollenbeck, Institute for Research on Poverty, Madison, Wisconsin, 1980 (forthcoming).

Kidd, Penny (ed.). "The MATH Users Guide," Two Volumes, Mathematica Policy Research, Washington, D.C., September 1979.

Lininger, Charles A. "The Survey of Income and Program Participation: Goals, Objectives, and Prospects," presented at the Survey Methods Section, American Statistical Association Meetings, Houston, 1980.

Orcutt, Guy, Amihai Glazer, Helena Jaramillo and Phil Nelson. "Microanalytic Simulation," Urban Institute Working Paper 1217-4, The Urban Institute, Washington, D.C., September 1978.

Orcutt, Guy, Amihai Glazer, Robert Harris, and Richard Wertheimer, "Microanalytic Modeling and the Analysis of Public Transfer Policies," in Robert Haveman and Kevin Hollenbeck (eds.), Microeconomic Simulation Models for Public Policy Analysis, New York: Academic Press, 1980 (forthcoming).

Orcutt, Guy, Steven Caldwell and Richard Wertheimer. Policy Exploration Through Microanalytic Simulation. The Urban Institute, Washington, D.C. 1976.

O'Reilly, Earmonn J. "Guide to the H.E.W. Welfare Simulation Model," Working Draft, U.S. Department of Health, Education and Welfare, Washington, D.C., December 1977.

Projector, Dorothy S. and Mary P. Johnston. "Income Distribution and Transfer Payments During the 1970s," unpublished Working Paper, Social Security Administration, U.S. Department of Health, Education and Welfare, Washington, D.C., November 1979.

Projector, Dorothy S. and Ellen G. Murray. "Eligibility for Welfare and Participation Rates, 1970," Studies in Income Distribution No. 7, Social Security Administration, Office of Research and Statistics, U.S. Department of Health, Education and Welfare, Washington, D.C., Number 78-11776, July 1978.

Sulvetta, Margaret B. "Analyst's Guide to TRIM--The Transfer Income Model," Urban Institute Paper, The Urban Institute, Washington, D.C., 1975.

The Urban Institute, "MASS Documentation," Urban Institute Working Paper, Urban Institute, Washington, D.C., no date.

Vroman, Wayne. "A Microsimulation Model of Unemployment Insurance," Urban Institute Working Paper 1280-2, The Urban Institute, Washington, D.C., January 1980.

Webb, Randall. "Toward a New Generation of TRIM: Efficient Data Structure and Supporting Features," Urban Institute Working Paper 1281-02, The Urban Institute, Washington, D.C., March 1979.

Webb, Randall, and George Chow. TRIM User's Guide. Urban Institute Working Paper, Urban Institute, Washington, D.C., September 1978.

Wyscarver, Roy A. "The Treasury Personal Income Tax Simulation Model," OTA Paper 32, U.S. Treasury Department, July 1978.

Ycas, Martynas, and Charles A. Lininger, "The Income Survey Development Program (ISDP): A Review," presented at the Survey Methods Section, American Statistical Association Meetings, Houston, 1980.