MODELING INDIVIDUALS' BEHAVIOR:
EVALUATION OF A POLICYMAKER'S TOOL

Alan L. Gustman

Modeling Individuals' Behavior
Evaluation of a Policymaker's Tool

ABSTRACT

With a continuous decline in the cost of manipulating data and a
continuous increase in the richness of data banks, policymakers have
increasing opportunities to build and apply so-called micro-simulation
models--models that attempt to simulate the behavior of the individuals in
a large population under a specified program. The efforts of the
Department of Labor to use a model in evaluating proposed changes in the
unemployment insurance system point up both the power and the weaknesses
of such models. Any user who applies these models without attempting to
understand which of their strengths and weaknesses are most important for
analyzing the problem at hand is asking for trouble. Easy to use or not,
these models are not user friendly.

Alan L. Gustman
Department of Economics
Dartmouth College
Hanover, NH 03755

With the entry of the era of the computer, policy analysts confront increasing possibilities of simulating through the use of large-scale models the likely effects of some existing government program or of weighing the consequences of some proposed changes in a program. Building a model that can simulate such effects is a problem for specialists; but those who make and evaluate policies increasingly will be relying on the projections provided by such models as a basis for their work. Accordingly, there is a growing need for all policy analysts to have a sense of both the strengths and the limitations of the models.

Nature of the Models

Among the various tools for simulation, the so-called micro-simulation model has been used to an increasing extent by U.S. government agencies and by various states. So far the principal applications of such models have been in assessing the effects of federal programs that entail payments to individuals, including food stamps, housing support, and aid to families with dependent children.

Micro-simulation models are built up out of two kinds of data.

To begin with, the U.S. government periodically conducts censuses and large scale surveys that cover a sufficient number of individuals in the United States so that reliable information can be gained about differences among people who live in different states. Censuses and large surveys of this sort provide some characteristics for each such individual, such as place of residence, age, family status and so on. They also report whether the individual is working or not, income from work, and useful but incomplete information on income from some government programs.

Data such as these are not sufficient by themselves to simulate effects

of any existing or prospective program. Data of the type collected in a general census must be linked to other sources of information. One purpose of the linkage is to enlarge and enrich the description of the individuals and their economic activities provided by the census and large survey data. Another purpose is to permit analysis for some period other than that of the census, and thus to specify the behavioral responses of the individuals concerned. If income taxes are the focus of the study and if the object is to simulate the behavior of all taxpayers, it is obviously necessary to know much more than a census of population will reveal: for example, detailed information on the sources of income of each individual, the propensity of each individual to cheat, and so on. Since estimates of that kind are not available in the record related to each individual, they must be obtained either by matching records from the census with those from other sources, or by enlarging each such record by some approximating process on the basis of information gleaned from other sources. These other sources of information can come from anywhere: from studies based on smaller samples of individuals, from the experience of operating officials or their management information systems, or even from pure guesswork.

Enlarging individual records by an approximating process is a tricky operation. The first step is to create categories in the population that need to be separately tracked. The appropriate categories depend on the subject being studied: a food stamp program will require one set of categories, an income tax program another set. If income taxes are the subject of the study, for instance, and if outside evidence suggests that the incidence of cheating is higher among independent businessmen than among wage earners whose earnings are reported by their employers, distinguishing

between the two groups in the general census data will be important. Each group so identified might be assigned a different proportion of cheaters, according to the outside evidence; and the specific designation of those assumed to be cheaters could then be determined by random choice within each group. In that way, the two sources of data will be linked and the model will be ready to spell out some of the aggregate implications of the income tax program to be tested.

A micro-simulation model called TRIM (acronym for Transfer Income Model) has provided the core for various micro-simulation models that have been put to a number of different uses in the U.S. government and elsewhere. Although TRIM has been widely used, however, there has been little critical evaluation of the suitability of such models either for predicting the impact of existing programs under changing economic circumstances, or for predicting the effects of changes in the programs.[1] The opportunity to undertake such an evaluation arose in connection with a micro-simulation model that was used to analyze the federal government's unemployment insurance program - a model we can call the Unemployment Insurance (U.I.) model.[2] - This model, although used independently of TRIM, both draws on information from the TRIM model and is meant to serve as an addition to TRIM which would enhance the power of TRIM in analyzing the phenomenon of unemployment. To get some sense of the strengths and weaknesses of models such as these, it is helpful first of all to take a closer look at how the model actually works.
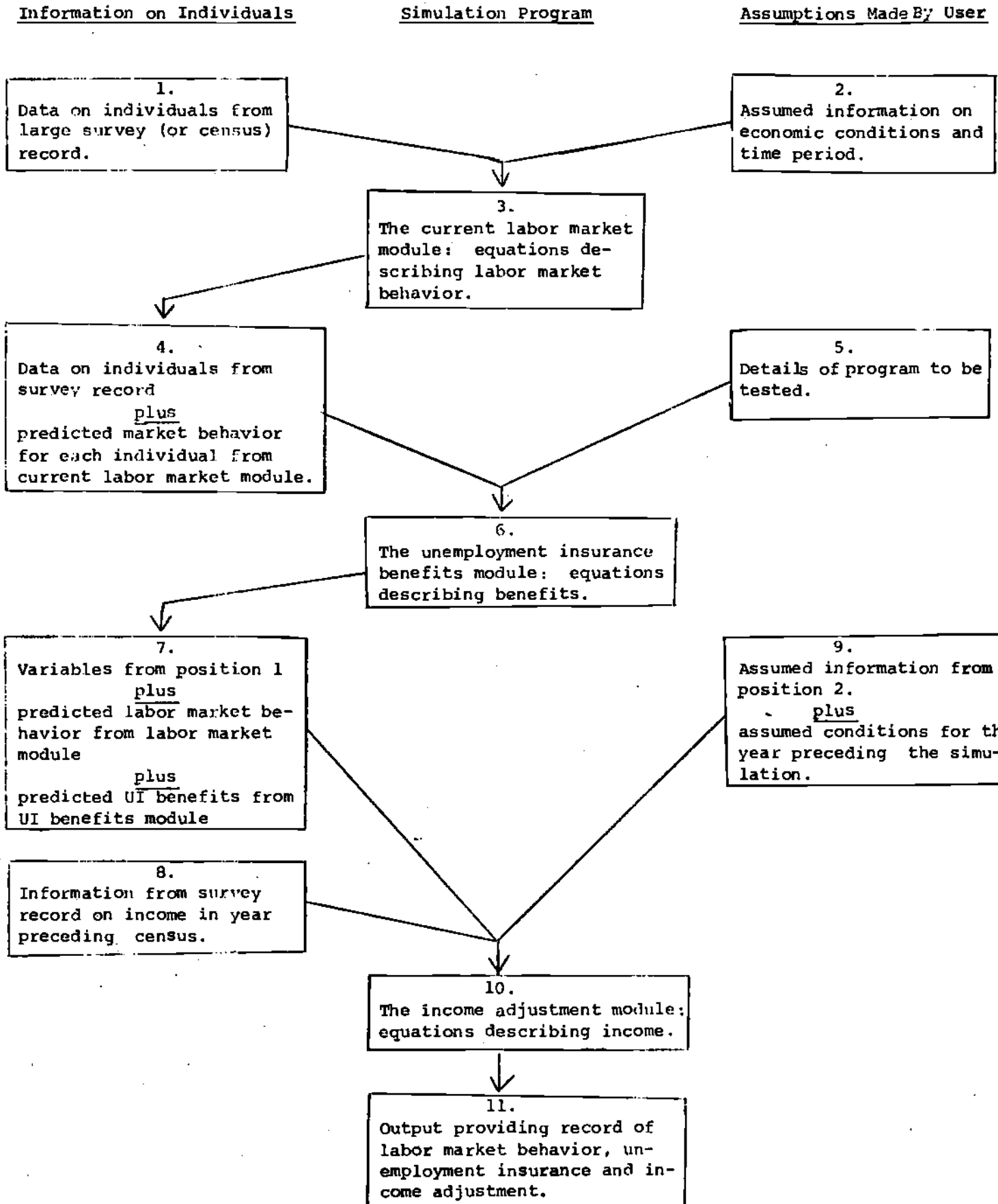
How the Model Simulates

The model's structure. The model may be viewed as a set of simulation programs which takes as inputs two types of information: data on individuals from a large survey and assumptions made about economic and other conditions during the period for which the simulation is being run. These inputs are processed through three modules, each of which consists of a set of equations, producing as a product the model's predictions.

The process begins with relevant information that is available from the survey record for each individual. To date only one such survey has been used as the data input for the UI model – the same survey that was used to estimate some of the model's equations. As is the case with the TRIM model in general, with some modification of the program, this basic input into the simulation model may be drawn from the census or a large survey. Included in each survey record are such variables as age, race, sex, marital and family status, industry of employment, occupation, and state of residence. In Figure 1, that source of data is indicated in position 1.

Note that at this stage the file can contain no information on the employment status of the individuals it covers for the period of the simulation. The task of the simulation model is to create estimates of such information, estimates that are consistent with the assumptions about the year of the simulation and about economic conditions for the simulation period that have been made in position 2. More specifically, this is the

Figure 1

Schematic Diagram Indicating How the Unemployment
Insurance (UI) Model Simulates

Information on Individuals	Simulation Program	Assumptions Made By User

1.
Data on individuals from large survey (or census) record.

2.
Assumed information on economic conditions and time period.

3.
The current labor market module: equations describing labor market behavior.

4.
Data on individuals from survey record
plus
predicted market behavior for each individual from current labor market module.

5.
Details of program to be tested.

6.
The unemployment insurance benefits module: equations describing benefits.

7.
Variables from position 1
plus
predicted labor market behavior from labor market module
plus
predicted UI benefits from UI benefits module

9.
Assumed information from position 2.
plus
assumed conditions for th year preceding the simu-lation.

8.
Information from survey record on income in year preceding census.

10.
The income adjustment module: equations describing income.

11.
Output providing record of labor market behavior, un-employment insurance and in-come adjustment.

job of the labor market module, which occupies position 3. That module is comprised of a set of equations which have been estimated from a number of sources of data. For example, time series data from other sources may provide the rate of participation in the labor force of people in a given demographic category in different years. Such data may be used to construct a regression that specifies for each such demographic group what proportion is in the labor force at each point in the business cycle. These regressions can then be used to predict the probability that individuals within the given demographic group will be in the labor force during a particular period. Of course, the modeler has no way of knowing which specific individuals in the demographic group covered by his census (or other large survey) records will actually be in the labor force. But that specific designation can be assigned on a random basis to the appropriate number of individuals who are in that demographic group.

Using that approach, each individual in the population is described according to the principal indicators of the individual's status in the labor market. In addition to indicating whether the individual is in the labor force, the record will indicate whether the individual is employed or unemployed; if unemployed, whether a new job seeker or someone who has lost or quit a job; and if unemployed, for how long. The records so generated occupy position 4 in the figure. At this point, however, the process of providing a full record for the experience of each individual has only begun. A module in the model, dubbed the current labor market status module, has been completed; but two more stages in the process still lie ahead.

As we shall presently see, the fact that the UI model performs in a series of separate stages has profound implications. In the language of the statistician the model is recursive and therefore to be distinguished from a model that is represented by a system of simultaneous equations. A simultaneous system generates all its solutions in a single step, producing values that are consistent with all the equations. But a recursive model operates sequentially; each stage builds on what has been learned or estimated in early stages, but without modifying the solutions provided by the earlier stages.

With the completion of the current labor market module, the model enters on the second stage of an estimating process. That stage, indicated in position 5 on Figure 1, calls for feeding in the details of the policy whose consequences are to be estimated by the simulation process. For example, if the purpose of the simulation is to analyze the effects of a new policy proposal, the information required would be the proposed changes in program rules. The program specifications, together with the output from the current labor market module, become inputs into a set of equations that predict unemployment benefits; that set of equations is dubbed the unemployment insurance benefits module, representing the second major building block of the model, and is located in position 6.

A third module, the income adjustment module, consists of a set of equations that estimate the income of individuals during the period of the simulation, and according to categories that are compatible with the output of the other two modules. Once the income of the individuals has been introduced in the model, it is possible to relate the unemployment benefits

of individuals to the income of individuals in a variety of ways relevant to policy. For instance, the impact of the program on people in different income categories can be compared; such benefits can be compared to prior earnings, or to unearned income or to income that will be received from other transfer programs.

Among the various estimates that the model has produced are the number of persons that are expected to receive unemployment insurance, the number of weeks unemployment insurance benefits are received by unemployed persons, and the fraction of lost incomes which are replaced by unemployment insurance. But the equations of the model—or of any of the modules in the model—can generate a wide range of information. Separate values can be calculated for any group in the population, provided of course that the group has been distinguished in the equations: whites can be distinguished from blacks, women from men, northern states from southern states, and so on. The weights indicated from the population survey would, of course, have to be appropriately adjusted so that each group of the population was weighted in accordance with its importance during the year of the simulation rather than the year of the survey: but with that adjustment made, values can then be aggregated to generate information for each distinguished group. Moreover, the outcomes can be distributed among individuals in the population, so that when a proposal to alter a program is under consideration, the characteristics of the winners can be distinguished from those of the losers.

The Equations

Choosing an approach. When devising equations that seek to predict the performance of individuals in a market such as the labor market, the

modeler can choose from several quite different approaches. One approach
simply uses a record of behavior in the past, without specifying what
produced the behavior. Thus, the employment performance of 19-year-old
black males can be traced through the ups and downs of the 1960s and
1970s, and the assumption can be made that similar economic conditions in
the 1980s will produce a similar record of employment for the group. In
essence, this was the approach most commonly used to develop the equations
for the unemployment insurance model.

The problems associated with such an approach are well known. Because
the factors that actually determine employment performance have not been
specified, the relationships observed in the base period may be altogether
misleading; the observed performance during that period may have been due
to an unspecified factor that is no longer present in the 1980s. For
example, if the various entitlement programs such as Aid to Families with
Dependent Children were a factor in determining the employment performance
of young women during the base period, a change in the program in the 1980s
could greatly impair the usefulness of the base period data for projection
purposes. Analogously, if labor market opportunities were an important
determinant for labor force participation by minorities or young adults, and
these opportunities were more favorable during the base period than during
the period of projection—due to the presense, for instance, of large scale
public service jobs and training programs during that period—then any
equation which ignored the role of such opportunities may not be useful
for simulating behavior during the period of projection.

A quite different way of attempting to project the behavior of various
groups would be to try explicitly to identify the causal factors that pro-

duce differences in the decisions they make, and thus their performance over time, and then to express these behavioral relationships in the equations. In the case of labor market behavior, a number of well developed hypotheses exist that serve to explain the likelihood of unemployment, the reasons for unemployment and the duration of unemployment for various groups in the population. Consider the decisions made by workers, such as the decision to quit a job. Young workers learn about jobs by trying them. Some find a job offering acceptable employment conditions early in their search. Others may continue to change jobs until they marry or reach their middle or late twenties, experiencing intermittent spells of unemployment as a result. Thus we find that within the group of younger workers--and indeed for all workers--the longer the worker has held a job, the less likely he is to quit. The reasons why other factors in the labor market, such as unions, reduce quits have been extensively analyzed, as have the conditions which influence the decision of an unemployed worker whether to accept an offer for a job.[3]

Factors affecting the decisions of firms, and thus the availability of employment to workers, may be a little more obvious. When there is a regular and well expected cycle in demand, layoffs are well anticipated, and date of recall may be made known to the worker at the time of layoff. Longer periods of slack demand may result in permanent layoffs for some, especially for those who have not been with the firm long enough to have received training that is specialized and of greatest use within the firm. The equations in the model that purport to predict employment behavior, however, make none of these distinctions; among those with the same demographic characteristics who worked full time last year, they do not distinguish the individuals most likely to experience a layoff nor the individuals whose

layoffs are likely to be permanent.

The modelers' problems are exacerbated--and the risk of error enhanced-- by the fact that there are neither historical records nor well developed hypotheses that would account for some aspects of behavior in the labor market. For example, the previously mentioned layoff decision is not fully understood. Some firms that experience a downturn in demand for their pro- ducts fire some employees while retaining the rest on a full-time basis, whereas others cut the hours of work for most workers and retain practically all of their labor force. Some firms cut wages rather than employment to cushion the downturn, while others will not reduce wages. When the reasons for these differences are not well understood, the modeler has little choice but to rely on hunch. Some of the equations, such as the one predicting weeks of employment, are perforce of this type, being based on simple mechanistic assumptions.

The quality of the data. Somewhat apart from the question of the nature of the equations being used is the quality of the data on which the equations are based. In the development of the equations for the labor market module, no single data source could be used that would permit the modelers to estimate all of the various relations that they wished to express. For instance, data that described the differences in labor market status among individuals were not broken down into all the categories that interested the modelers, even for a single base year.

Time series that consistently described the employment behavior of constituent groups were even less adequately disaggregated. For a model

that is expected to project the incidence of unemployment through the business cycle, inadequacies of that sort become particularly threatening.

In any event, the data on labor market behavior by relevant groups is developed by splicing various sets of statistics developed for somewhat different primary purposes. One such set represents a survey at a point in time (1976), the Survey of Income and Education, which is essentially an expanded version of the monthly Current Population Survey, asking more questions and surveying a larger group. The Survey of Income and Education provides more reliable information about activities in state of residence than can be obtained from the monthly surveys of the population. The other sources are a diverse collection of surveys which provide various time series of the labor market behavior by month or by year. In order to generate an estimate for the whole population from the smaller samples covered in the time series surveys, measures from these surveys of the labor market behavior of different ages and sex are attributed to those of the same age and sex who were sampled in the Survey of Income and Education. But neither the Survey of Income and Education nor the time series data present all the variables that are of interest to labor market analysts.

The character of the equations. On the basis of such incomplete data, the model proceeds in highly eclectic fashion to develop equations for unemployment behavior in the business cycle. Some of the equations are naive; for instance, some assume that a given type of unemployment will always bear a fixed proportional relationship to another figure, irrespective of the phase in the cycle. For example, unemployment rates of nonwhites are assumed to be related in the same way to the unemployment

rates of whites no matter what their industry or occupation. Some relations

take what statisticians would describe as a reduced form--they merge

together the effects of direct program influences on a particular variable

with the responses of individuals and firms to the rules of the program.

For example, the equations which predict duration of unemployment and

relate unemployment duration to state of residence, among other variables,

may be viewed in this way. Strictly speaking, no equation is truly

structural--each equation corresponding to a separate aspect of behavior.

Those equations which might in some sense be structural, such as the

equation analyzing labor force participation, omit too many variables to

allow them to be clearly classified.

Some of the implications of this eclectic approach can

be seen by examining more closely the elements that went into the making of the

last equation referred to above--the one that expresses the probability that the

individual will be in the labor force during the period to be simulated.

This equation does not incorporate factors reflecting opportunities in

the labor market. For instance, the participation of the individual in

the labor force is not directly related either to the level of wages being

offered or to the availability of jobs. Nor is the influence of such

factors picked up indirectly, through other measures that might reflect the

demand for labor or the level of wages offered, such as changes in the

level of output over time. Instead, labor force participation is projected

essentially on the basis of three factors: income of the family, personal

and family characteristics, and the labor force participation of the in-

dividual in the prior year. These relationships, in turn, are devised in

part from other models, such as the Urban Institute's DYNASIM model. Once the relationships have been defined, they remain unchanged in the unemployment insurance model through the business cycle. The danger of using unchanged relationships of this kind in all phases of the cycle is apparent. For example, although an individual's being in the labor force may be related to whether the individual was in the labor force in the preceding year, that relationship would surely be different at different stages in the cycle.

Another key equation seeks to measure the probability of unemployment for each of various groups. Such groups in the labor force, distinguished by age and sex, are covered in the model. In this instance, each equation is derived from annual data on unemployment over a period of eleven years. The probability of unemployment for any of these groups is estimated from equations that rely on three factors for their predictions: a time trend, observed in the 11-year base period; the age and sex of the individual; and the rate of unemployment that is projected for white males of a given age, for each age in the 35- to 54-year range. The unemployment rate for these white males embodies some of the market opportunities as they vary over the business cycle that were not considered in the analysis of labor force participation.

The estimates of the probability of unemployment that emerge from this procedure, however, are not sufficiently disaggregated for the subsequent steps in the model, when the effects of various unemployment insurance systems are to be considered. The provisions of unemployment insurance vary by individual states, each of which lays down different requirements for eligibility with respect to previous wages and weeks of employment. The relation of the level and duration of benefits to the individual's work

history also varies among states. For later steps, therefore, it is necessary to determine unemployment probabilities by state of residence, and to have the closest possible relation between current unemployment and past employment history. Further disaggregation of the unemployment prediction is achieved by adjusting the probabilities of unemployment with so-called adjustment factors derived from the average relationships for state, occupation and industry that appeared to exist among these categories in a base period, 1967 to 1977. Those relationships, in turn, are gleaned from an analysis of surveys which identify the state of residence of the unemployed.

Estimates based on techniques such as these raise familiar questions. For one thing, the regression equations are obviously vulnerable. With only eleven annual observations, how stable should we assume these estimates to be? In so brief a time period, how can we distinguish the effects of an enduring trend from the effects of a cyclical movement or of a random disturbance? Inasmuch as any effect ascribed to "trend" is simply an expression of some otherwise unidentified factor, how secure is the assumption that the "trend" will continue in the future?

The use of simple adjustment factors to estimate the unemployment experience for those living in different states, or working in different industries and different occupations, also entails high risks. Presumably, holding age and sex constant, those in different states or jobs will experience somewhat different patterns of unemployment through the business cycle. But because the adjustment factors are unchanged over the cycle, those possibilities cannot be taken into account.

The vulnerability of the various equations in the model can be illustrated from other elements of the model as well. Within the current labor market module, for instance, the equations that estimate the duration of unemployment suffer from essentially the same weaknesses that are present in the equations that estimate the probability of unemployment. And added questions are introduced by the equations that are contained in the unemployment insurance benefit module. That module, in fact, illustrates why simulation models often compel modelers to take great leaps across uncharted territory in order to complete the simulation process.

The logic of a simulation model often requires the modeler to fill in with data for which no empirical basis exists. For instance, in order to receive unemployment benefits, an individual must make an application. But no hard data appear to exist on which to base an estimate of the proportions of unemployed that will in fact apply. In developing the model, analysts assumed—not unreasonably—that the probability would vary with age, sex, student status, and various other characteristics, including duration of unemployment. But the probabilities that were then assigned to each of these categories appear to have been nothing more than the guesses of the modelers.[4]

The income adjustment module also offers numerous illustrations of the demands that a large-scale model imposes on the modelers. Individuals that are laid off, for instance, not only lose their current wages but also may lose pension rights; just what pension rights they lose, however, depends on the terms of individual pension plans, a fact that prevents the modeler from dealing with such problems in any but a rough-and-ready fashion.

Of course, model-builders are as conscious of the problems described above as those who profess to criticize their handiwork. Accordingly, in the process of structuring the equations in a model such as the unemployment model, analysts are constantly on the alert for opportunities to align their equations with reality and experience. For instance, strenuous efforts are made to pick up changes in the composition of the population over the years, including changes in age, sex, and marital status.[5] Moreover, whenever there is an opportunity to compare the estimates produced by the equations with real data, such comparisons are made; and if the estimates are materially different from reality, the estimating equations are adjusted to try to bring about a closer fit.

Some of these efforts move a step in the right direction. But others are extremely dangerous, such as adjusting the equations to remove any difference between a predicted and actual value without understanding where the difference comes from and whether it will persist. Given the inherent character of the model and the data it uses, none of these adjustments increases the reliability of the model by very much, and some of the adjustments may undermine its usefulness.

The Recursive Structure

It will be remembered that the unemployment insurance model is recursive in structure; that is, the model produces results sequentially without allowing for the possibility that the result generated at later stages in the model's operations might be helpful in reestimating the values generated at earlier stages. The reasons for employing a recursive structure in this instance are fairly clear. As we saw in our description

of the model, its ultimate results emerge from a series of modules, each
of which stands on its own statistical feet. Linking those modules in a
way that could simultaneously produce solutions for all of them would be
a very difficult statistical task. Yet failing to link the modules means
that some major relationships are disregarded in the estimates that the
model produces.

Consider, for example, the equations explaining the probabilities of
labor force participation and unemployment, equations which are found
within the current labor market module. The question whether the
unemployment was temporary or permanent, whether it arose from layoffs,
quits, or new entrants, does not get introduced until a later stage in the
model. Because the model is recursive, those questions are not allowed to
influence the previously developed estimates of labor force participation
and unemployment. Yet there is no doubt that the nature of the existing
unemployment can, for example, affect the rate of labor force participation.

Another illustration of the risks imposed by the recursive structure
of the model is highlighted by the changes that have recently been adopted
in the unemployment insurance programs. As a result of such changes in
federal law, unemployment insurance benefits are now taxed in all but
lower income families, the maximum-benefit period has been reduced, and
eligibility requirements have been tightened--all these measures having
been taken to reduce the costs of the insurance program. These changes
will affect the equations of the unemployment insurance benefit module;
but they are also likely to have an important impact on the values generated
in the current labor market module, such as duration of unemployment, reason

for unemployment, incidence of unemployment, and probability of labor force participation. Given the way the model has been arranged, however, and given the structure of each of its equations, the changes in the unemployment insurance will not affect the values generated by the current labor market module.

Omissions of this type have two important consequences--they can cause errors in prediction even if policy remains unchanged, and they may cause serious side effects of any new policy to be missed. In the past, for instance, unemployment insurance programs have been made more generous during recessions, a fact that has increased the likelihood or the duration of unemployment. Equations in the current labor market module, however, may attribute labor market outcomes to the cyclical movements in economic activity when in fact they are due in part to the increase in unemployment benefits. Accordingly, the predictions of some equations of the current labor market module may not be correct for any future recessions in which unemployment benefits were not increased.

More generally, there is considerable evidence that the nature of a state's unemployment insurance system affects the level of unemployment in the state as well as the relative incidence of different kinds of unemployment. Ample evidence exists, for instance, that where employer payroll taxes are not raised sufficiently to cover the full cost of benefits for any additional workers the firm decides to lay off, temporary layoffs are more frequent.[6] In the recursive model, however, the differences in unemployment among states are estimated without direct reference to the unemployment insurance system because the insurance characteristics come later in the recursive sequence. To be sure, the current labor market module would be

inadequate in its present form for reflecting different behaviors at the firm level; major changes in that module would be required in order to pick up such effects. Nevertheless, the risk of the recursive structure is clear: if the analyst accepts the model's logic and assumes that the unemployment insurance system does not affect the patterns of unemployment, the analyst is likely to be misled when modeling the effects of any proposed change in the system.

None of these problems is easily handled, a fact of which analysts are acutely conscious. Apart from the fact that each module stands on its own feet, there is the equally formidable fact that the feedback effects that have to be stipulated are disconcertingly complex. This is illustrated in a recent analysis of the effects that insurance benefit levels and benefit periods have upon the duration of unemployment of insured workers. It turns out that different individuals have different preferences for leisure relative to income, a fact that impedes easy generalization regarding their responses.[7]

All told, therefore, the recursive feature of the model poses considerable added risks. That fact comes as no surprise to those who fashion such models; their choice of the recursive approach is imposed on them by cost and convenience, not chosen by them as a preferred technique. But policymakers who are the users of the output of such models must always be aware of the implications of the recursive feature.

Implications for Model Users

The purpose of exploring this model is not to recommend to the policymaker whether or not to use such models. They may be better than

an educated guess. They may even be better than the predictions that would
be obtained from simpler models, such as a model that focussed only on a
few key aspects of labor market behavior. But then again, they may not.

Given the state of this art, these models will be subject to consider-
able error. While some of the equations seem perfectly reasonable, others
are incomplete or incorrectly specified. Give the structure of the models,
the errors in the various equations may be expected to interact in ways
that make it very difficult to determine their overall effect on the model's
predictions. Particular errors affect the suitability of a model for some
uses more than others. Reduced form equations mix together the effects
of current policy with relevant behavioral responses, undermining the model's
suitability for predicting the effects of major changes in policies. The
same is true for the effects of aligning the model, adjusting the equations
so that their predictions come closer to currently observed outcomes.
Incompletely specified equations may predict effects that are not there
and will ignore some effects that are. Instability also appears likely
to characterize the time series equations that are used to simulate the
effects of policies in future periods, mixing together cyclical influences
and secular trends.

Where these models seem least prone to error is in projecting the
effects of current programs a few years into the future, under the assumptions
of little change in economic conditions or in program rules. For such a
projection, the biggest source of change is changes in population. Yet
even these projections will be subject to error if the factors that are
omitted from the equations vary in relation to the factors there are included

in ways that are different from the past. Correspondingly, the fact that a model generates predictions which turn out to be accurate in the short term under conditions of unchanging policy is no test of the suitability of the model for projecting the effects of major changes in policy; nor has such a model been tested for its ability to project effects over a long horizon.

However these models are to be applied, considerably more information should be developed about their potential errors in use: what types of error are involved, how wide their effects are likely to be, and how the various types of errors may interact. Such an effort would seem to be worthwhile even if the funds required have to be diverted from further work on refining and extending the model. Policymakers who use these models urgently need to know how sensitive the results of the model may be to differences in economic assumptions, inadequacies in the model's structure, and errors in the underlying data.

Notes

1. For an exception, see General Accounting Office, "An Evaluation of the Use of Transfer Income Model—TRIM—to Analyze Welfare Programs," (Washington, D.C.: November, 1977).

2. More precisely, the authors call the model "A Simulation Model of Unemployment Insurance." The model was constructed for the Department of Labor by researchers at the Urban Institute under the direction of Dr. Wayne Vroman. The explanations of the model are in a series of memos. (Vroman, Wayne, "A Simulation Model of Unemployment Insurance," Urban Institute Working paper number 1280-7, August 1980 and Vroman, Wayne, "The Urban Institute Simulation Model of Unemployment Insurance Detailed Project Memos," Urban Institute Working paper No. 1280-4 August 1980). The model has been used for policy analysis by the Labor Department's Office of the Assistant Secretary for Policy, Evaluation and Research, as well as by the Unemployment Insurance Service and the National Commission on Unemployment Insurance

3. Behavioral theories which explain unemployment are reviewed in Gustman, Alan L., "Analyzing the Relation between Unemployment Insurance and Unemployment," in R. G. Ehrenberg, editor, Research in Labor Economics, (5) (1982): 69-114.

4. The memo dated May 24, 1979, contained in Vroman, "...Detailed Project Memos," op. cit., includes the following quotation on p. 15: "It should be stressed that the Table 6 probabilities are initial estimates based on apriori considerations, and not on any regularly published program data."

5. See, for instance, Sulvetta, Margaret B., An Analyst's Guide to TRIM - The Transfer Income Model, (Washington, D.C.: The Urban Institute, June 1976).

6. Feldstein, Martin S., "The Effect of Unemployment Insurance on Temporary Layoff Unemployment," American Economic Review, (68)5 (December 1978): 834-846; Feldstein, Martin S., "Temporary Layoffs in the Theory of Unemployment," Journal of Political Economy, (84)5 (October 1976): 937-957; Baily, Martin Neil, "On the Theory of Layoffs and Unemployment," Econometrica, (45)5 (July 1977): 1043-1063; Topel, Robert, "Unemployment Insurance Financing and Unemployment: Empirical Investigation of Adverse Incentives," Report to the Office of the Assistant Secretary for Policy, Evaluation and Research, U.S. Department of Labor, November, 1982.

7. Moffitt, Robert and Nicholson, Walter, "The Effect of Unemployment Insurance on Unemployment: The Case of Federal Supplemental Benefits," Review of Economics and Statistics, (64)4 (February 1982): 1-11.