

*1st draft of the program  
document*

PRELIMINARY DRAFT

THE MICRO-SIMULATION PROJECT

by

A.R. DOBELL and M.A. COHEN

JUNE 1972

## CONTENTS

### PREFACE

- I. INTRODUCTION
- II. THE INDIVIDUAL DECISION PROBLEM
- III. THE RESULTING LONGITUDINAL RECORD
- IV. THE SIMULATION MODEL
- V. ADJUSTMENT MECHANISMS: THE FEEDBACK STRUCTURE
- VI. TABULATION OF DEMOGRAPHIC DATA AND LINKS TO AGGREGATE PROJECTIONS
- VII. CONCLUDING COMMENTS

### APPENDICES:

## PREFACE

This document describes an analytical scheme and a set of computer programs designed to permit analysis of income transfer plans or related financing policies. This structure, itself forming only part of a more extensive research program, was conceived in discussions with D.G. Hartle, initiated in several projects at the Institute for Policy Analysis of the University of Toronto, and continued as an activity of the Quantitative Analysis Course conducted by the Institute, in Ottawa, for the Public Service Commission. Work now continues at the Treasury Board Secretariat and the Department of Finance, with Dobell acting generally to coordinate and link the several components. In order to describe the nature and purpose of the present program, it is helpful first to indicate briefly how all this work fits together.

In an initial project to examine costs and impacts of so-called contingent repayment plans for financial assistance to students, the decision was taken to adopt an aggregate, or expected-value, approach, with more detailed analysis of redistribution among individuals to be deferred. This initial analysis, based upon broadly defined classes of individuals, and dealing only with education financing schemes, is described in the CORSAP manual written by Michael Wolfson [1]. In addition,

work on demographic models at approximately the same level of aggregation was undertaken at the Institute, as described in a methodological report by Leroy O. Stone [2]. A proposal by Dobell for integration of this demographic work with the CORSAP program and further computer programs describing other transfer schemes in order to assess regional impacts has not so far been implemented, but remains feasible and is being pursued.

In the meantime, related analytical work has been developed at the level of individual records, with a pilot project undertaken by Dobell and Cohen in the summer of 1970, as described in the MCSSAP manuals [3], together with related work undertaken by Professor G.C.A. Cook and outlined in her reports [6] to the Planning Branch of the Treasury Board Secretariat, forming the starting point. This work in turn was split into three parts for further development within the Quantitative Analysis Course over the past year. The first of these sub-projects was the construction of a computer program evaluating the impact of various education financing proposals upon an individual with a specified life-history; this program is described in Cohen's Quantitative Analysis Course project report [4]. The second part, the creation of a computer program capable of generating a representative sample of such life histories, was carried out through the Institute for Policy Analysis, and is described in a forthcoming report by Cohen,

Dobell, and Stone [5]. The third distinct activity, the creation of a new Monte Carlo simulation program for analysis of education financing schemes, integrating the two previous components, is outlined in the manual MCSSAP II cited earlier [3], which forms a companion to the present document, and a sequel to the MCSSAP manual. Finally, the overall logic and program structure is sketched in the present document.

Thus, considering only the analysis of proposals for education financing schemes, one may view the program structure as having three levels, within which full integration has not yet been achieved. At the first level is the deterministic model, which takes cost, income, and tax data as given, accepts the description of a single life history and the specification of a proposed policy, and computes the resulting transfers between the individual, the financing scheme, and some overall government budget. At the second level, the description of individual life histories is suppressed and the MCSSAP II program generates from estimated transition data a representative sample of such longitudinal records (or life-histories), computing summary descriptions of interpersonal transfers and cash flows to or from the financing fund. In principle, aggregation of this sample of individual records to the aggregate categories employed by the regional population projection model and by CORSAP would yield the population, enrolment, and employment projections

necessary for projections of cohort rates of return and aggregate cash requirements for the financing fund. Integration to this extent has not yet been attempted but will, in principle, provide a valuable check upon the consistency of the MCSSAP II results.

More general use of this program structure is also feasible, however. Considering the requirements for analysis of some unspecified transfer program, one sees that all the machinery is available in this general structure except the detailed description of the rules for operation of the specific scheme under study. Provided these require no more detailed information than the status codes contained in the existing demographic records, this detailed description can be expressed in a single subroutine inserted into the overall simulation structure. This structure can then be used to generate the same sort of summary information on redistributational effects and total cash flows as was developed for the transfer schemes dealing with education financing. In particular, standard flow of funds tables can be constructed to display intersectoral transactions.

Finally, this program structure can be employed independently of any analysis of transfer schemes, simply for the assembly and verification of demographic data. Since the simulation model generates a sample of individual records purporting to be representative of the current Canadian population,

it is crucial that the distribution of various characteristics within the artificially generated sample be checked for consistency with available data on the distribution of these characteristics within the population as a whole. The program therefore makes provision for output of sample observations suitable for cross-tabulation, and thus can, in principle, be employed for generating synthetically a body of longitudinal demographic data linking in a consistent way available cross-section data drawn from diverse sources.

In concluding this outline of where this work now stands, it must be emphasized that no validation of the demographic sample has yet been completed. Extensive work on this task is being undertaken over the summer of 1972; until it is finished, all of this model structure must be considered untested, and no guarantees or undertakings whatever can be made as to the accuracy of the data base or the estimates derived. While the authors are willing to cooperate in use of the program or in adaptation to other uses, no distribution of the program or results is anticipated before September, 1972.

## I. INTRODUCTION

This document presents an overview of work on an analytical framework and a set of computer programs designed to assist in analysis of distributional impacts of government transfer programs, with one important by-product being the creation of a synthetic longitudinal sample of records of individual demographic and economic histories. As indicated in the Preface, the work is closely related to several distinct projects described elsewhere, but the present document will be confined to an outline of the overall scheme for creation of a sample of individual records and use of this sample in analysis of general transfer programs. The particular application giving rise to the general program structure dealt with investment in education, and the remainder of this document will illustrate the general model in terms of this specific application. But, as will be clear from the discussion, many other transfer schemes might be studied with only the replacement of transfer and tally subroutines, and minimal changes elsewhere. The conceptual scheme remains unaltered.

The relationship between the various parts of the overall program structure can be illustrated in the aggregate flow diagram set out in Figure 1. In this diagram, using the identifying numbers in the upper corner of each program block, the various components of the model can be described as follows:

1. The complete model.

Consisting of the entire structure outlined in Figure 1, the complete model is described briefly in the present document.

2. The deterministic model.

Accepting pre-specified life-histories and computing the impact of various financing proposals upon them, the deterministic model consists simply of blocks 1, 15, and 17-21 in Figure 1. This model, useful for test purposes in the larger model as well as for creating particular detailed examples to accompany any general analysis, is the subject of a report [4] by Morris Cohen.

3. The demographic model.

If it is desired to create a representative sample of life histories such as described in the deterministic model, this task may be accomplished by Monte Carlo techniques. The program blocks to carry out the generation of this synthetic longitudinal sample are indicated in Figure 1 as blocks 5, 6, 9, and 10, with blocks 7, 8, and 11-14 providing for tabulated output permitting comparison with various sources of cross-section data. The creation and evaluation of this demographic sample is described in the Institute Working Paper by Dobell, Cohen, and Stone [5].

4. The simulation model MCSSAP II.

The simulation structure employed to provide estimates of distributional impacts of education financing schemes consists of blocks 1, 5, 6, 9, 10, 15, 18, 19, and 22 to 24. Various flow-of-funds or financial accounting statements can be created in block 25 if desired\*, and a tape containing a panel of longitudinal records can be substituted for blocks 5, 6, 9, and 10 in providing input to the model, as shown in block 16. Unless some

---

\* A forthcoming report by David Sewell of the Treasury Board Secretariat (Planning Branch) will deal in part with this issue.

feedback of policy onto demographic data is taken into account as described in item 5 below, input from such a tape is undoubtedly the most efficient procedure. Documentation of MCSSAP II is given in more detail in the manual by that same name [3].

5. Policy feedback.

The response of the education system and the individuals in it can in principle be taken into account in this model as indicated in blocks 1-4. Block 2 provides for estimates of the changes in individual decisions under altered financing policies, and blocks 3 and 4 provide for estimates of the impact of the resulting changes in student flows upon cost or income data. This structure is thus the only portion of the program which draws upon behavioural data or provides for any feedback or internal adjustment within the system. Unfortunately, while it is possible to take account of such responses in principle, the necessary data as to the impact of policy variables upon individual retention or participation decisions, or as to the elasticity of costs or incomes with respect to changing flows through the system, are entirely lacking. Only rather arbitrary sensitivity tests will be feasible in practical applications. Nevertheless, if such data do become available, through analysis of survey data or otherwise, they can be brought into the model framework in a consistent way.\*

6. Comparison with aggregate analyses.

From the population arrays tabulated in block 7 and the financial flows tabulated in block 25 can be derived aggregate series permitting comparison with other studies not attempting to deal with distributional concerns demanding detail at the level of individual records. Later reports will attempt to reconcile model output with results derived from such aggregate analyses.

It is perhaps worth pointing out that this facility makes the model structure useful as a

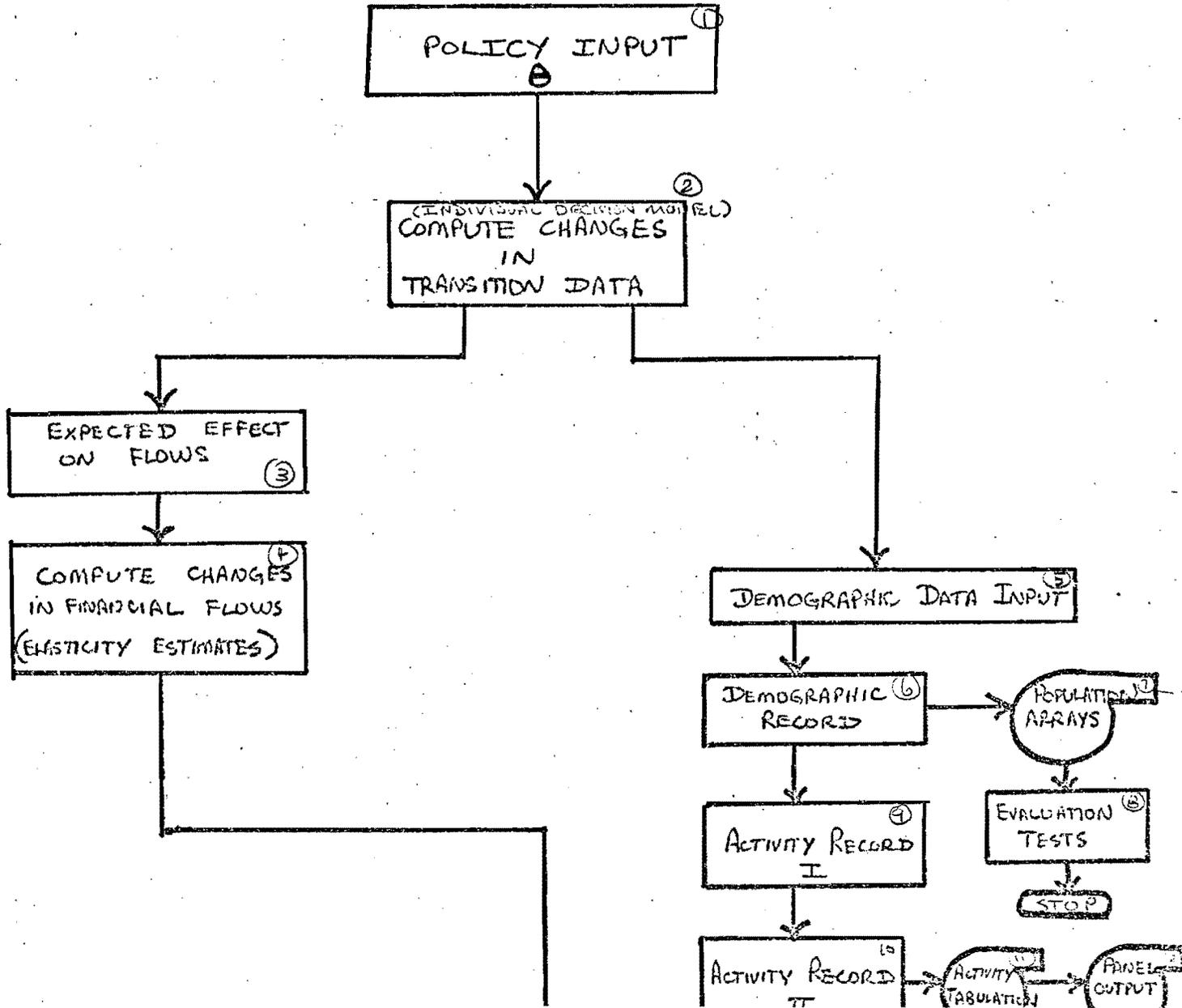
---

\* One may expect the work by Kathy Warren on accessibility to post-secondary education, and the work by Verne Henderson on the individual decision model for investment in education to yield data or model structures relevant to block 2.

framework for storing educational data in a consistent and meaningful way. This role as, in effect, a storehouse of educational data in a form permitting easy use and convenient updating was one of the objectives of the initial model work proposed two years ago, and has dictated some of the work on tabulation and output options which would otherwise be unnecessary.

The remainder of this report describes the analytical framework underlying the model structure. The first portion deals with a model of individual decisions concerning education. These decisions, in the aggregate, are then taken to be describable statistically by a transition process, rather than represented explicitly as decision functions in a dynamic programming framework. The aggregate model takes the data of the transition process as given, and proceeds to generate from those data a sample of individual lifetime records. These in turn are used as input to routines evaluating financial flows associated with particular histories, and tabulating the resulting distributions of benefits and costs. The remaining sections of this report will deal briefly with each of these steps in turn.

FIGURE 1  
SUMMARY FLOW CHART  
 MICROSIMULATION MODEL



## II. The Individual Decision Problem

An analytic model was designed to examine individual training and employment decisions. These decisions are reflected in an individual's participation in the various educational and occupational activities. In reality, this participation is conditioned by various stochastic elements not under the control of the individual. A further conditioning is provided for by the finance policies set by the responsible government agency. One could consider these policies as being control parameters and thus the resulting individual decisions can be thought of as being elements of an optimal decision set arrived at by the individual. Here optimality is with respect to some as-yet-undetermined utility function and subject to both the structural constraints of the system and the policy constraints set by government.

Thus given an input policy, represented by a vector of parameters  $\theta$ , the individual will make a set of decisions, over time,  $\delta(t;\theta)$ . This decision set can be partitioned into a set of "demographic decisions",  $D(t;\theta)$  and a set of "activity decisions",  $K(t;\theta)$ . The function  $D(t;\theta)$  refers to decisions associated with marriage, fertility, migration, etc., while  $K(t;\theta)$  refers to decisions concerning educational and occupational activities. Upon conditioning by stochastic elements, such as mortality, income mobility and success in school, a state history or trajectory for the individual results. This history is summarized at each time period  $t$  by a vector  $S(t;\theta)$  where:

$$S(t; \theta) = f(\delta(t; \theta), \lambda),$$

$$\delta(t; \theta) = (D(t; \theta), K(t; \theta))$$

$\lambda$  = a random vector drawn from some underlying population with a probability density function  $p(\lambda)$ .

$f$  = a mapping from decision space to state space.

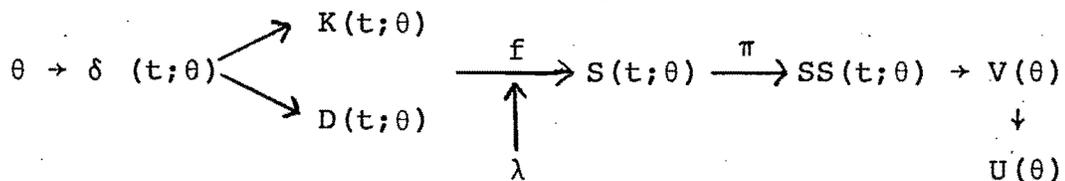
Given this state trajectory and the underlying structure of financial rewards and costs associated with schooling and working, a set of monetary flows over time is derived for the individual. These flows are summarized in a matrix  $SS(t; \theta)$  where:

$$SS(t; \theta) = \pi(S(t; \theta), \theta)$$

$\pi$  = a functional representation of the accounting and policy rules, income distributions and direct costs which combine to produce financial flows given an individual's state trajectory and the government policy.\*

Each row of  $SS$  represents the flows for a given time period  $t$ . One may also compute present values for the various cost and benefit streams (columns of  $SS$ ) and it is conceivable to relate the resulting vector of present values  $V(\theta)$ , to the individual's utility function  $U$ .

This process can be diagrammatically summarized as below:



\* This rule is actually stochastic as well but for the purposes of the argument which follows we shall treat it as being deterministic.

A number of simplified examples have dealt with versions of this individual decision problem in determining the optimal length of enrolment or extent of participation in educational activity as opposed to the labour force. See, for example, Sh<sup>e</sup>shinski [ ], Ben-Porath [ ], or Zions and Southwick [ ], as well as the work of Vern Henderson [ ], surveying this literature and related papers on rates of return in education. In fact, however, even such simplified treatments dealing with a single individual become quite complex at the level of rigorous theory; for the purpose of generating a whole sample of records reflecting such individual decisions, we adopt a statistical description derived from data on participation in education or labour force activities. The next section describes the resulting individual record.

### III. The Resulting Longitudinal Record

From the individual decision process as outlined above there results a record of participation in educational or labour force activities, and of family characteristics. This record is in the form of a matrix  $S$ , each row of which corresponds to one year (or age), and each column of which corresponds to one component of the so-called status vector  $S(t)$  describing the demographic state of the individual at any time. Figure 2 and the attached status codes display this record, together with the fixed vector  $S_0$  recording certain individual characteristics - such as sex or place of birth - which do not vary with time. One such record must be generated for each individual in a desired sample.

Figure 2 - status codes and sample of  $S$

Given the individual record  $S(t)$  - and initial characteristics  $S_0$  - it becomes possible to impute appropriate cost, income, and tax flows for each year. That is, knowing the educational activity in which the individual is involved in one year, educational costs, tuition fees, foregone income, and taxes can be estimated. Knowing the age, educational attainment, and labour force status of the individual in later years, incomes and taxes can again be estimated. Thus, associated with the record  $S(t)$  in any year is a further record  $SS(t)$

representing corresponding financial flows in the same year. The history S over all relevant years of the individual's life then yields the necessary record of demographic and labour force characteristics, while the history SS over the same period records relevant financial flows. These flows can be analyzed in various conventional ways to reveal realized rates of return, or the distribution of present values of benefits and costs among participating agents, as described in Cohen's manual on the deterministic model.

The next task is to generate a representative sample of such records, as described in the next section.

#### IV: The Simulation Model

Analysis of the social decision problem must be made in the context of the rational behaviour of individuals. Our particular view of rationality refers to the family of state trajectories arrived at by the population of individuals as they each solve their individual optimization problem (I). Ideally, one would prefer to analyze the various policy options  $\Phi$  by solving the optimization problem for each individual for every  $\theta \in \Phi$ . The techniques for such a calculation are clearly not available for any but the most simplified of models and thus alternative methods must be considered. The approach chosen in this study was that of Monte Carlo simulation. Consequently, analysis of the policy space  $\Phi$  must be preceded by a discussion of the simulation model used to generate a sample of state trajectories  $S_i(t; \theta)$ .

That is to say, the individual records described above, while generated in principle from each individual's rational decisions in pursuing his own goals, cannot be so treated for our analytical purposes. Instead, the sample of individual records is obtained by substituting for explicit optimization a statistical description of observed outcomes from past collections of individual decisions, and deriving individual records by simulating the resulting transition processes.

The simulation procedure is thus based on the following regularity hypothesis:

$H_1$ : Observed behaviour of a population's education, occupation and demographic behaviour constitutes a basis for computing the underlying joint distribution associated with  $S_i(t; \theta)$ . Essentially we are saying that known data are sufficient to

identify the underlying transition structure of the system. In principle at least this point of view does not restrict the consideration of future policy as an instrument for changing the underlying structure.

The requirements for the simulation model are thus nothing less than the joint probability distribution describing the set of all possible state trajectories  $S(t; \theta)$  for all classes of individuals.\* Upon aggregating over these trajectories, one can reconstruct the demographic, educational and occupational data describing the flow of individuals.

The aggregation of individual financial flows can also be achieved under the simulation procedure. The resulting flows from this computation could lead both to funding requirements and the distribution of financial costs and benefits under policy  $\theta$ . Thus a rather detailed analysis of the policy space  $\Phi$  is possible and one could consider the possibility of constructing a measure of social welfare by applying some crude social indicators to the simulation output. In this way, the second level of our policy problem, the social decision level, can be considered.

The remainder of this chapter will be devoted to a rather brief description of the simulation model as implemented. Details of program structure are dealt with in Appendix B and the problems of estimating the required joint probability distributions are considered in some depth in Appendix C.

The application of a simulation procedure leads in essence to a probabilistic description of the education/occupation

---

\* In reality each individual forms his own class but then we are back to the beginning with a stochastic description of

system. Thus the flow of students through training institutions and labour force activities can be viewed as being described as a stochastic process. In this study the particular case of a Markov chain model was constructed.

Consider the following:

$$X_t = X_{t-1} \cdot P + M_t$$

where  $X_t$  = a row vector in  $R^n$ ;  $n$  is the number of states in the system and  $X_t$  represents the distribution of the population over the  $n$  states at time  $t$ .

$P$  = an  $n \times n$  transition matrix.

$M_t$  = a row vector in  $R^n$  which represents the distribution over states of net entrants to the system at time  $t$ .

In general state,  $i$  represents an education/occupational activity. The population will thus distribute itself over all possible states in accordance with the structural constraints of the system as embodied in matrix  $P$ . Clearly it is too much to expect that a simple matrix will somehow capture the richness of experience implied by the micro-model of Chapter II. We will demonstrate that, in principle at least, the model can be modified to encompass most of the complexities previously discussed.

The first question we consider is homogeneity. We have previously referred to distributions associated with classes of individuals. This implies that transition through the system is actually conditional on a set of individual characteristics. These characteristics could include age, sex, race, social class,

parental income and so on. Any simplifications that one introduces in the way of aggregation over broad individual classes will therefore reflect data limitations and not a deficiency of model structure. Notationally our Markov process has become:

$$S_t^i = X_{t-1}^i \cdot P^i + M_t^i$$

where  $i$  refers to the  $i$ th class of individuals.

We are thus assuming that individuals do not change classes over time. This implies that our notion of class is associated with initial characteristics received either at birth or previous to entry into the system. If one wishes to relax this assumption, this "merely" implies an extension of the state space. Thus presence in grade 9 for a type  $j$  individual will correspond to a different state than that associated with presence in grade 9 for a type  $k$  individual and a single extended transition matrix will be sufficient.

A further complication is that of time dependence. One may postulate that the underlying structure of the system is changing and thus the transition matrix itself becomes a function of time. This implies the following notation:

$$X_t^i = X_{t-1}^i \cdot P(t) + M_t^i$$

A final complication has often been referred to as the "policy feedback" problem. If a government agency implements some policy  $\theta$  which, for example, affects the post-secondary education fee structure, then it is plausible to

believe that some of the transition probabilities will change. For example, lowering fees may increase the conditional probability of going from grade 13 to university. Alternatively, raising fees may lower that probability and increase the probability of dropping out of grade 12 to the labour force. In addition, it is conceivable that rigid control may be enacted to keep enrolments (net entrants  $M_t^i$ ) on some specified trajectory.

Thus the Markov Chain model of the education/occupation system becomes:

$$x_t^i = x_t^i \cdot P(t; \theta) + M_t^i(\theta)$$

It is also clear that both policy and the resulting state distribution may interact with the pricing mechanisms employed in allocating financial flows. Thus an external macro-model could interact with the system and could possibly affect the underlying structure. This would imply a dependence of the transition matrix  $P(t; \theta)$  on variables other than time  $t$  and policy  $\theta$ , for example upon unemployment rates. (For some evidence on the strength of such dependence, see Crean [ ]).

Inevitably, the model as implemented in this study was considerably less comprehensive. Complications of time dependence, policy feedback and reactions to pricing mechanisms were all laid aside. Consequently, a family of transition matrices conditional on a variety of demographic characteristics was used in the simulation program. This program effectively sampled the Markov chain once each year for each individual. In this way, the

required collection of state trajectories was generated. From these, a corresponding sample of financial flow statements, one for each individual, is also obtained.

For the deterministic (individual) model, it is this individual demographic record and associated financial history that serves as the basic unit of analysis, and even in the overall model some individual data are of interest for distributional questions. Accordingly, certain summary measures of the present value of individual benefits and costs are computed, and entered as individual observations into the tabulations for histograms to be printed at the end of the computation. We treat the individual records, in other words, as providing observations on a class of random variables, whose empirical distributions are tabulated as the computation proceeds.

Once these observations are recorded, there remains no further need for detailed individual records; the financial history is aggregated with those of the previous individuals drawn from the same cohort, thus yielding, in the end, a summary financial history of the same form as the individual, but showing, for each year, the flows aggregated over all members of the particular cohort. These records, one for each cohort, are stored for later processing (Block 23).

In this way, the cohort records can serve as input to report generators in any form desired. For the present model,

such further processing takes the form simply of aggregating across all cohorts (Block 24) to obtain total financial flows in each year of the simulation period, and organizing some of these into a sources-and-uses-of-funds statement for the hypothetical student financial assistance fund under study in the illustrative application. Other reports suitable for different transfer schemes can easily be designed and the necessary program blocks appended to the existing code following (or in place of) Block 25.

These observations conclude the description of the program structure as such; the remaining sections remark upon possible additional features or applications.

V. Adjustment Mechanisms: The Feedback Structure.

To the extent that the input data - transition probabilities, retention rates, and the like - remain fixed despite policy changes, the above program will generate essentially unchanged population samples (up to sampling fluctuations) in every run. It is for this reason that provision is made to record the demographic records on tape simply as a hypothetical longitudinal sample, which can be used as input to any of a number of programs computing the results of various transfer programs operating on the unchanged sample. To identify distinguishing features of different programs, indeed, it is a great convenience to be able to eliminate sampling fluctuations in this way. (This advantage of being able to replicate "random" sequences is frequently cited in the literature on pseudo-random number generators.) There are no more drawbacks in this procedure than there would be in working with any tape of individual records drawn from survey data or tax returns, for example.

But for estimating the consequences of policy changes, there are significant drawbacks to this procedure - it overlooks two key classes of adjustment mechanisms at work in any economic system. In the first place, even without any conscious policy changes, the system may operate to change the data relevant to individual decisions, and thus induce changes in behaviour. In

the education example, for instance, continuing flows of individuals through post-secondary institutions must increase the relative stock of skilled labour, and presumably bring about some erosion of relative incomes. This reduction in expected returns may be expected, of itself, to cause some individuals to reconsider decisions to continue educational activity. Thus the system generates the machinery to shut itself down where necessary, or expand flows where scarcities are signalled by high rewards. The lags are long and uncertain, of course, and the linkages sometimes very tenuous, but it would be gross error to ignore this machinery altogether.

More directly, policy shifts may operate directly to alter individual decisions, for example by offering financial assistance to those who might otherwise not continue their education. These impacts upon individual decisions will show up in our model as altered transition data, and hence as altered flows through various institutions.

Both types of alteration in the nature of the choice faced by the individual - because conditions have altered either through the self-adjusting mechanisms of the economy, or through discretionary policy changes - will be expected to affect the number of individuals choosing particular options, and hence the composition of the overall population. Thus the assumption of an unchanged underlying population sample becomes untenable in principle.

Unfortunately, there is almost no evidence to permit specification of alternatives. The model permits adjustment of retention rates and all other transition data or probabilities once the policy specification is complete; these adjustments represent an aggregation of the adjustments in individual behaviour predicted from the solution of the individual decision model described earlier. But in fact the elasticities of retention rates, participation decisions, marriage probabilities, or other individual educational, demographic, or labour force decisions are not known.

Similarly, the model permits changes in transition data to be used to estimate expected changes in total flows through institutions, and these in turn to determine changes in cost or income data. But, again, the elasticities of unit costs or expected incomes with respect to enrolments or manpower supplies are largely unknown as well.

Thus the model structure admits the possibility of an adjustment mechanism feeding back from policy shifts or changing circumstances to the demographic system or the actual sample generated, but no data are available to implement any such scheme. Should data become available, the program logic will be implemented; in the interim, some crude tests of sensitivity to changes in transition or activity data will be carried out.

## VI Tabulation of Demographic Data and Links to Aggregate Projections

As indicated above, the demographic computations employed in the overall simulation structure form a self-contained model which is of interest in its own right. This model generates demographic histories for a sample of individuals drawn according to specified sampling weights from prescribed cohorts. The resulting longitudinal records can be tabulated in three alternative modes.

- (a) Selected characteristics may be recorded from each individual history and written on an input file in the standard format required for input to programs for further statistical processing. In particular, this data array can be used as input to the SPSS (Statistical Package for Social Sciences) program for either statistical analysis or cross-tabulations.
- (b) The entire file of panel data can be written, in compressed form, to a magnetic tape for later input to simulation programs, or to specially-written programs for further statistical analysis.
- (c) The distribution of individual by age, activity, or income can be tabulated for specified years,

to provide output arrays in a form which can easily be checked against cross-section data sources or distributions obtained from aggregate projections.

## VII Concluding comments

Evidently a model of this kind is never finished, and indeed the present version is in no sense tested or validated even to a first level of accuracy. If evaluation tests through the summer (1972) proceed well, the ability of the model to generate a sample of records which will be "representative" in the sense of reproducing the distributions associated with given initial cross-section data and with aggregate projections will be verified. Similarly the ability of the model to assign financial flows to these individuals records in a manner consistent with available cross-section data and aggregate tabulations of flows-of-funds information, will be tested. These tests, together with final checks on program logic, will provide some assurance that the basic model structure is sound and that the demographic characteristics of the sample population are acceptable.

The immediate use for the model will then be in analysis of possible distributional impacts of alternative schemes for federal support of post-secondary students. For this purpose the aggregate outputs of the model in generating

projections of financial flows and crude enrolment levels will also be of interest.

More general applications are intended in study of social security programs and possible economic circumstances of the aged in the future, and in the overall balance of federal tax and transfer programs in affecting the personal distribution of income. This role of the model in integrating available data into a consistent overall framework for evaluating of the distributional impacts of government programs (along with some aggregate projections for checking against alternative sources of data) should make it a useful element in a kit of tools for longer-range planning. Further use of the model structure in such applications will be reported in the future.

- \* 1. Wolfson, M.C.      CORSAP User's Manual  
IQASEP Technical Paper #  
dated
- \* 2. Stone, Leroy O.    Regional Population Projections:  
Methodology, IQASEP paper,  
dated
- \* 3. Cohen, Morris A.    McSSAP User's Manual, IQASEP  
paper, dated  
  
Dobell, A.R. and Cohen, M.A.    McSSAP II, Q.A.C.  
report, May 1972.
- \* 4. Cohen, Morris A.    Monte Carlo Simulation of  
Government Programs: The  
Deterministic Model, Q.A.C.  
Report, April 1972.
- \* 5. Dobell, A.R., Cohen, M.A., and Stone, L.O.  
Synthetic Longitudinal Sampling  
and Demographic Record Linkage:  
Methodology and Validation,  
IQASEP Working Paper, Forth
- 6. Cook, G.C.A.      Data Appendix, Planning Branch  
report, 1970.