

Copy of ...

PRELIMINARY DRAFT

A MICROSIMULATION MODEL
FOR THE EVALUATION OF GOVERNMENT
TRANSFER PROGRAMS

by

A.R. DOBELL and M.A. COHEN

JAMBIC 3
~~NOVEMBER~~ 1972

Conference paper for discussion purposes only; not for distribution or quotation. Any opinions expressed are the responsibility of the authors alone, and are not necessarily endorsed by any of their employers.

ABSTRACT

The Evaluation of Government Transfer Programs:
A Simulation Analysis

BY:

A.R. Dobell,
Director,
Quantitative Analysis Course
Room 307
2211 Riverside Drive,
Ottawa, Ontario,
K1H 7X5 Canada.
(613) - 996-7076

M.A. Cohen, Analyst,
Planning Branch,
Treasury Board Secretariat,
Room 433, Confederation Bldg.,
Ottawa, Ontario,
K1A 0R5 Canada.
(613) - 996-4270

This paper presents an approach to the evaluation of government transfer programs based on a simulation of the individual decision process. Individuals interact with the socio-economic environment by making demographic and career decisions - state trajectories and an associated net present value of financial flows for each individual result. These present values can be considered to be the output of a stochastic system and are conditioned on both the government transfer policy and the individual decision profile. The simulation program to generate a representative sample of state trajectories and criteria for policy evaluation based on the resulting present value distribution are discussed.

For consideration as part of the Session on Simulation (Methodology Section) or on Canadian Government Applications (Government Section).

CONTENTS

| | <u>Pages</u> |
|--|--------------|
| PREFACE | 1 - 5 |
| I. INTRODUCTION | 6 - 9 |
| II. THE INDIVIDUAL DECISION PROBLEM | 10 - 13 |
| III. THE RESULTING LONGITUDINAL RECORD | 14 - 19 |
| IV. THE SIMULATION MODEL | 19 - 26 |
| V. ADJUSTMENT MECHANISMS: THE FEEDBACK STRUCTURE | 27 - 29 |
| VI. TABULATION OF DEMOGRAPHIC DATA AND LINKS TO AGGREGATE PROJECTIONS | 30 - 31 |
| VII. CONCLUDING COMMENTS | 31 - 32 |
| APPENDIX | |
| A. OVERVIEW OF THE MODEL | 34 - 37 |

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, ^{which carried out} 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, ^{of the Canadian Government.*} ~~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.

The impetus for this work can be associated with an extensive ~~initial~~ project to examine costs and impacts of so-called contingent repayment plans for financial assistance to ^{post-secondary} 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, This paper represents only the conceptual framework for the model used in the policy evaluation and as such it does not deal with any policy conclusions.


* footnote on details of historical background of the project

acknowledgements
disclaimer of responsibility -

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,

Rod:

Perhaps the remaining material in the preface should be delegated to an Appendix of this paper. 

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 redistributive 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 at the present time; 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, 1973.

INTRODUCTION

This ^{Paper} ~~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 or other schemes for financial assistance to individuals. The approach follows the lead of Orcutt^[] in his pioneering work on microsimulation models, but with some difference in emphasis and therefore in analytical and computational structure. Specifically, the primary concern in the present work is with the impact of proposed programs on the distribution of lifetime costs and benefits over individuals and groups. This emphasis dictates an analytical structure which focuses on the life-cycle or financial life-history of individuals at the expense of detail on the composition of a whole population at any one time. For this reason, our computational work is organized differently from Orcutt's models (for example in the analysis with the Urban Institute)^[] and indeed the computational requirements in the present work are of a lower order of complexity.

The immediate stimulus for this work was the need to estimate the impact of policies respecting the provision of financial aid to students in post-secondary education. Existing work on cost/benefit analysis or rates of return to investment in education, and also our own earlier work on an aggregate model of contingent repayment schemes for financial assistance

[] (? Institute report to Ont. govt.)
or paper by G. Cook

to students (later repeated with minor modifications by Dresch and Goldberg^[] in the U.S.) is inadequate to meet this need, precisely because it fails to account for important redistributive impacts arising from changes in financing arrangements. The present model does provide a basis for estimating the impact effect of changes in the rules of operation of existing or proposed transfer programs. What it does not capture, except as discussed below, is the subsequent response of the system - and in particular the altered composition of the population - through various adjustment mechanisms. This limitation is, of course, most serious: the purpose of many programs is not simply to transfer incomes, but also to affect behaviour thereby. Unfortunately, the empirical understanding required for an adequate modelling of the overall system response is simply lacking.

The discussion is organized as follows. In the next section, the standard approach of microeconomic theory leads to description of the decision problem for an individual facing fixed income and cost data and attempting to develop decisions as to participation in various economic activities (as, for example, enrolment at post-secondary educational institutions) or other decisions on demographic matters. The result of the individual's conscious or unconscious decision process is a determinate life-history of economic activity, associated with an identified demographic record. Taken together, all such records provide us with the life-histories of a representative population, or a longitudinal sample of panel data.

When the individual's demographic history and record of economic activity are known, appropriate income and cost data can be estimated (indeed, the estimates of these presumably were the determinants of the individual decision), and thus other financial flows can be computed from known rules of operation for any selected government transfer programs. In Section III the individual record of economic activity, and the history of the associated financial flows, are described. *(with details delegated on appendix)* These records constitute the basic unit of analysis for later work.

In Section IV it is observed that in fact, despite substantial theoretical work in control-theoretic models of the individual's optimizing decision, the empirical knowledge necessary to generate the required population sample from this starting point is not available. Therefore we generate our population sample from a simple transition model which we expect will prove rich enough to enable us to construct a synthetic longitudinal sample agreeing well enough with the observed characteristics of the Canadian population in all crucial respects.

When one abandons the explicit representation of the individual decision, however, he also loses the adjustment mechanism by which the system regulates itself in response to changing balances or changing policies. Section V describes a crude procedure by which some estimates of the important adjustments in the underlying population and in the data of the model may be recaptured.

*perhaps
replace
with section
of macro-
problem*

Finally, Section VI outlines the links from outputs of this model to other aggregate projections or estimates. *A final* appendix provides a brief overview of the model structure.

Thus, to reiterate, the present work stands in the gap - which it cannot yet claim to bridge - between the theorist following microeconomic lines, studying the determinants of individual decisions on the assumption that relevant price or cost data are unaffected by these decisions, and the macro-theorist concerned to know how these data change so as to alter or shut off the flow of individuals in various directions in response to changing circumstances in the aggregate economy. Though it will not answer all the questions in this area, the model should enable us to synthesize much of the work at different levels of aggregation, and certainly should teach us something about the balance between allocative efficiency and equity or redistribution considerations in the evaluation of government transfer programs.

II. The Individual Decision Problem

In this section
A ~~An~~ analytic model ~~was~~ ^{is considered} designed to examine individual training and employment decisions. These decisions are reflected in an individual's participation in the ~~various~~ ^{variety of} ~~existing~~ 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 transfer policy set by the responsible government agency. One could treat such a policy as a set of control parameters ^(where the government is viewed as the controller) and thus the resulting individual decisions can be thought of as being elements of an optimal decision path arrived at by the individual in response to these parameters and the social and economic environments. Thus optimality is with respect to some as-yet-undetermined utility function and subject to the structural constraints of the social and economic system and the policy constraints set by government. This section will attempt to make this decision process more precise with the aid of some simple notation.

Let us begin with some government policy, represented by a vector of parameters θ from some underlying set* of policies Θ and a given vector of prices E which summarizes in part the social and economic background. Each individual will make a set of decisions, over time, $\delta(t; \theta, E)$ (from ^{the} ~~a~~ set of "available" (feasible) decisions ~~$A(\theta, E)$~~) which constitute a rational reaction to this policy. This decision path can be partitioned

* For example, the set of rules needed to specify a post-secondary education finance policy. Thus a range of possible advance terms and repayment terms is implied for the specific case of student assistance schemes.

$$S(t; \theta, E) = (D(t; \theta, E), K(t; \theta, E))$$

$K(t; \theta, E)$ is described by a Markov chain, conditioned on $D(t; \theta, E)$

$D(t; \theta, E)$ described by a general Stochastic process

Theoretical model of this type of a system is difficult 'a point process'

into a set of "demographic

decisions", ~~$D(t; \theta, E)$~~ , and a set of "activity decisions". ~~$K(t; \theta, E)$~~ .

The demographic decisions

~~$D(t; \theta, E)$~~ refers to the sequence of choices associated with

marriage, fertility, migration, etc., while ~~$K(t; \theta, E)$~~ refers

to ^{those} decisions concerning ^{ed with} educational and occupational activities.

Upon conditioning¹ 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 $S(t; \theta, E)$ whose i 'th component

represents the value in year t of the i 'th characteristic for

the individual. ~~To summarize then:~~

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

~~$S(t; \theta, E) = \delta(t; \theta, E) + \epsilon_t(\theta, E)$~~

~~$\epsilon_t(\theta, E)$ = a vector of random disturbance terms whose underlying probability density function is conditional on both the policy environment (θ) and the price environment (P)~~

perhaps should retain this after all

new paragraph → Given the state trajectory and the underlying structure of financial rewards and costs associated specifically with schooling and working, it is possible to derive, for the individual, a set of monetary flows over time. These flows are summarized in a matrix $SS(t; \theta, E)$ where:

1

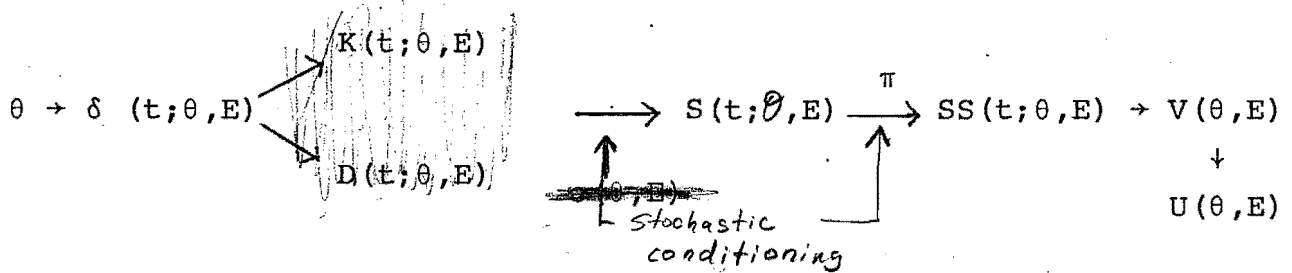
This conditioning process refers to the uncertain influence decisions can make on the resultant state histories. The effect of decisions is concentrated in their alteration of certain probabilities by, in effect, making the probability of a trajectory conditional on a decision and thus ultimately on a policy.

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

where π = 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, prices and the government policy. ¹

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:



Thus, in effect, we are saying that given policy θ ^{when} $U(\theta, E)$ ^{and a set of prices E} for each individual there is some utility $U(\theta)$ which results ^{from} through some rational decision process. Clearly rationality implies some concept of utility (or expected utility) maximization. Thus the individual decision problem becomes one of finding the optimal decision path $\delta^*(t; \theta, E)$. ¹

¹ This rule is actually stochastic as well but for the purposes of the ^{of session} argument which follows we shall treat it as being deterministic.

One could conceivably pose this problem as a mathematical program whereby $S^*(t; \theta, E)$ is

Find $S^*(t, \theta, E)$ derived ~~to~~ by maximizing $U(\theta, E)$

~~$S^*(t, \theta, E) \in M(\theta, E)$~~

~~$\max_{S^*(t, \theta, E) \in M(\theta, E)} U(\theta, E)$~~

subject to the social and economic structural constraints under some policy and price environment.

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, Sheshinski [], Ben-Porath [], or Zions and Southwick []. In fact, however, even such simplified treatments dealing with a single individual become quite complex at the level of rigorous theory and thus 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.

A complete listing of all characteristics and an example of a particular life history are to be found in Appendix I. The individual used will be referred to as E.C. in the remainder of this paper.

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 (Figure 3) 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.

The record illustrated in Figure 2 could represent the state trajectory of a typical college graduate with rather standard behaviour patterns. The educational/occupational history is displayed in the vector $K(T)$ and the spouse's activity is displayed in $KS(T)$. Region is indicated by $H(T)$.

Given ^{an} ~~the~~ individual record $S(t)$ ^{of this type,} ~~and initial~~ characteristics S_0 it becomes possible to impute appropriate cost, income, and tax ~~flows~~ ^{if we} for each year. That is, ~~knowing~~ ^{for example} the educational activity in which the individual is involved in one year, ^{then} ~~appropriate to that individual's~~ educational costs, tuition fees, foregone income, and taxes can be estimated. ^{Can be estimated.} ~~the age, educational attainment, and labour force status of the individual.~~ ^{post-educational} In later years, incomes and taxes ^(appropriate to the individual's labour force status) can again be estimated. Thus, associated with the $S(t)$ record 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 a background manual by Cohen [4] on ^a ~~the~~ deterministic model *which considered single life histories as ^{given} data input.*

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 ^{as outlined in} ~~the~~ *section II*. Ideally, one would prefer to analyze the various policy options θ by solving the optimization problem for each individual for every θ ^{in the policy set} ~~in the policy set~~. The techniques for such a calculation are clearly not available for any ~~but the most~~ **but the most** simplified of models and thus alternative methods ~~must be~~ **must be** considered.* The approach chosen in this study was that of Monte Carlo simulation. Consequently,

* See page 19a.

- * There is the point of view that states people are in fact not rational in this way, rather than being logical people are in fact systematic in their response to such stimuli. Thus the individual decision model has little predictive power and indeed observed behaviour forms a better basis for our purposes. It will be clear that observed behaviour at present forms the only basis (given the current state of data availability) for predicting the behavioural response to the type of policies under consideration and consequently the debate over rationality vs. systematic response is vacuous.

analysis of the policy space ~~is~~ must be preceded by a discussion of the simulation model used to generate a sample of state trajectories $S_i(t)$ ~~is~~.

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 hypothesis that observed behaviour of a population's education, occupation and demographic behaviour constitutes a basis for estimating a joint distribution (transition matrix) associated with $S_i(t)$ ~~is~~. Essentially, we are saying that known data can be used to identify a possible underlying transition structure consistent with observed behaviour for the system. In principle at least this point of view does not imply that future policy could not be 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 $S(t)$ for all classes of individuals.¹ 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 θ . Thus a rather detailed analysis of the policy space θ 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 ^{section} ~~chapter~~ will be devoted to a rather brief description of the simulation model as implemented. Details of program structure are dealt with in Appendix ~~II~~ and the problems of estimating the required joint probability distributions are considered in some depth in an ^{associated} ~~accompanying~~ background ^{paper} ~~piece~~ by Dobell, Cohen & Stone [].

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

¹ In reality each individual forms his own class but then we are back to the beginning with a stochastic description of ^{the original} optimization problem ~~§~~.

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, a state 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 conditioned on a set of individual characteristics. These characteristics could include age, sex, race, social class,

I am not sure if this type of description is helpful

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.¹ ~~Notationally can Markov~~

~~process has become:~~ If we group individuals in the population into classes reflecting the characteristics for which there is data, then ~~the transition matrix P~~ would be replaced by a family of matrices P^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.² 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:~~ *that each transition matrix P^i is somehow dependant on time and should be referred to as $P^i(t)$.*

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

¹ See page 23a.

1 # This of course is not to deny that the model is in principle restrictive but simply to suggest that it can be extended to accommodate existing data.

2 ~~the~~ A further issue related to time dependency in a Markov system has to do with "memory". That is, it is conceivable that the current state is not sufficient to specify the probability distribution for possible destination states in the next period but rather information on a number (if not all) previous states is required. The standard approach to dealing with this problem is to once again extend the state space so that state $S(t)$ could represent a particular history of previous states and the current state (in the old state space).

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 ~~W~~) on some specified trajectory.

a comprehensive
Thus ~~the~~ Markov Chain model of the education/occupation system ~~becomes~~ *would include a family of policy-time dependent transition matrices $P^i(t; \theta)$ and a family of net entrant ~~vectors~~ $N^i(t; \theta)$ vectors $N^i(t; \theta)$, each dependent on population group, time and policy.*

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 $P^i(t; \theta)$ on variables other than time t and policy θ , for example upon unemployment rates.

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.

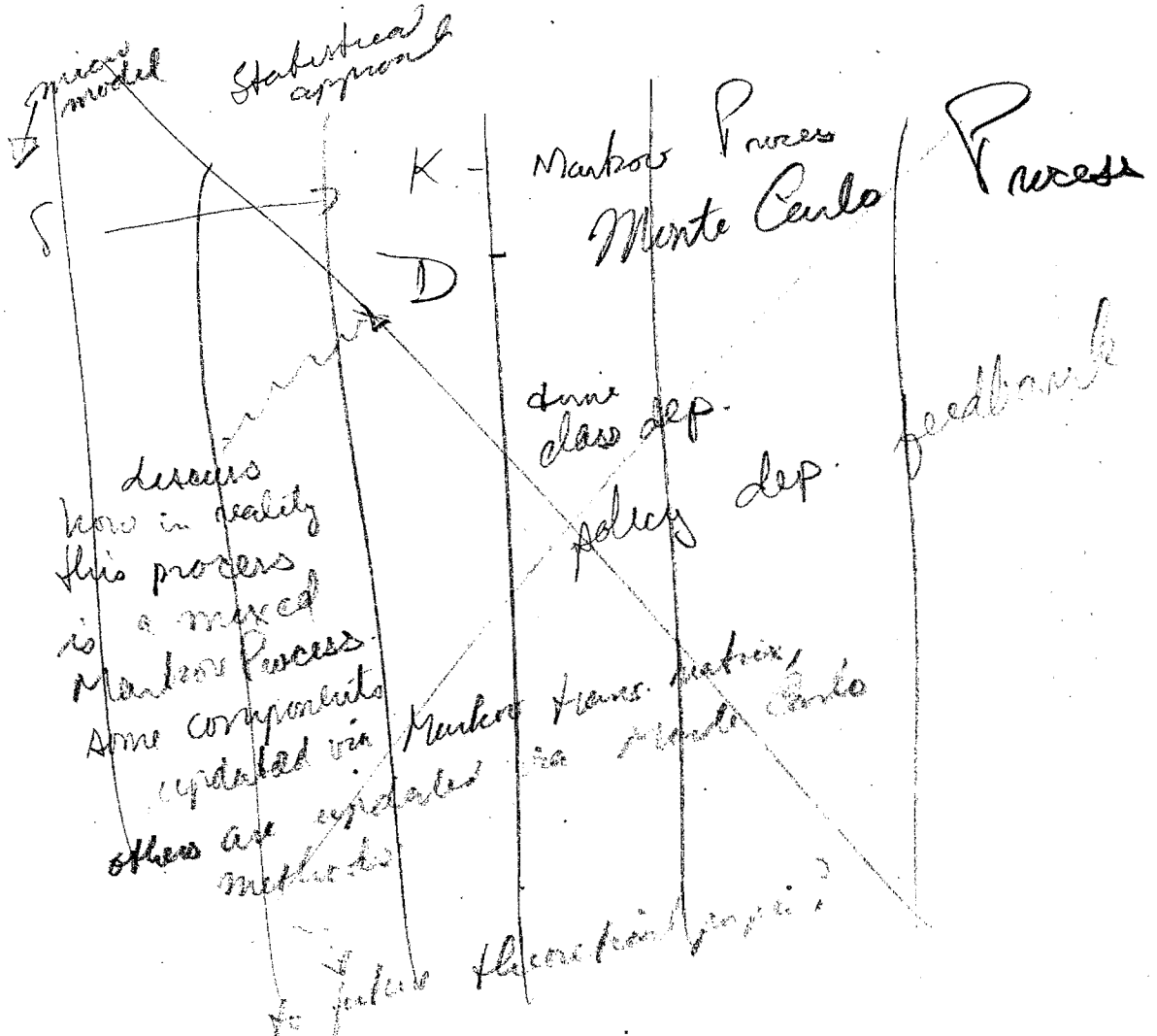
Once these observations are recorded, there remains no further need for detailed individual records; the financial history is aggregated with those of the previous individuals drawn from the same cohort, thus yielding, in the end, a summary financial history of the same form as the individual, but showing, for each year, the flows aggregated over all members of the particular cohort. These records, one for each cohort, are stored for later processing (^{see} Block 23 in Figure 1, ~~to be~~ described ~~later~~ in Appendix II).

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

* This leads to a considerable saving in ^{terms of} storage requirements and ^{the} complexity of the computer programs associated with the model.

such further processing takes the form simply of aggregating across all cohorts (Block 24) ^{in Appendix II} 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 computation 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 presently being undertaken 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.

Appendix I

The Individual Record:

The individual status record illustrated in Fig. 2 represents an example of the basic unit of analysis in the model. Each row of this record corresponds to a single census (age) and column corresponds to a single characteristic of the individual. A fixed set of characteristics which are recorded throughout the life span ~~is recorded~~ - such as sex or place of birth - is recorded by a single digit in record. That one digit has been generated for each sample individual. This record is operated on to produce frequency tables and summary statistics for the life record as present records of various interests. At any one time, the individual is in one of the states of the life record. After all, it is a family who lives in the world and not an individual. The life record is a record of the family's life.

The above table of responses
could represent the state trajectory of
a group of cells connected with
public behavioral behavioral
patterns of a subpopulation. The
trajectory is depicted as a line in the
state space. The arrows in
the diagram indicate the direction
of the trajectory. The trajectory
entirely inside the boundary of the
state space. (See Appendix A)
[It is in the paper [Figure 1
shown as a trajectory in the phase
space. States code.]

FIGURE 2

Individual History

• Constant Characteristics:

1 1 5 1958 2

State Trajectory:

| Age | Reg | ACT. | ASP. | * (1) | STATUS | | | | | | | | | | | | | | | | | | | |
|-----|-----|------|------|-------|--------|-----|-----|-----|-----|-----|-----|-----|------|------|------|------|------|------|------|------|------|---|----|----|
| | | | | | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) | (14) | (15) | (16) | (17) | (18) | | | |
| 14 | 6 | 1 | 0 | * | 0 | 0 | 2 | 2 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 15 | 6 | 2 | 0 | * | 0 | 0 | 2 | 2 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 16 | 6 | 3 | 0 | * | 0 | 0 | 2 | 2 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 17 | 6 | 4 | 0 | * | 0 | 0 | 2 | 2 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 6 |
| 18 | 6 | 5 | 0 | * | 0 | 0 | 3 | 3 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 19 | 6 | 9 | 0 | * | 0 | 1 | 5 | 4 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 20 | 6 | 10 | 0 | * | 0 | 2 | 5 | 4 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 18 | 0 | 0 | 0 | 19 | 15 |
| 21 | 6 | 11 | 0 | * | 0 | 3 | 5 | 4 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 18 | 0 | 0 | 0 | 19 | 15 |
| 22 | 6 | 12 | 0 | * | 0 | 4 | 7 | 5 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 18 | 0 | 0 | 0 | 19 | 15 |
| 23 | 6 | 20 | 20 | * | 23 | 4 | 7 | 5 | 5 | 5 | 5 | 4 | 0 | 1 | 0 | 0 | 0 | -1 | 18 | 22 | 0 | 0 | 19 | 15 |
| 27 | 6 | 20 | 22 | * | 23 | 4 | 7 | 5 | 5 | 5 | 5 | 4 | 1 | 1 | 0 | 1 | 0 | -1 | 18 | 22 | 0 | 0 | 19 | 15 |
| 28 | 6 | 20 | 22 | * | 23 | 4 | 7 | 5 | 5 | 5 | 5 | 4 | 2 | 1 | 0 | 2 | 0 | -1 | 18 | 22 | 0 | 0 | 19 | 15 |
| 39 | 6 | 21 | 22 | * | 23 | 4 | 7 | 5 | 5 | 5 | 5 | 4 | 2 | 1 | 3 | 2 | 0 | -1 | 18 | 22 | 0 | 0 | 19 | 15 |
| 40 | 6 | 20 | 22 | * | 23 | 4 | 7 | 5 | 5 | 5 | 5 | 4 | 2 | 1 | 0 | 2 | 0 | -1 | 18 | 22 | 0 | 0 | 19 | 15 |
| 46 | 6 | 20 | 22 | * | 23 | 4 | 7 | 5 | 5 | 5 | 5 | 4 | 2 | 1 | 0 | 1 | 0 | -1 | 18 | 22 | 0 | 0 | 19 | 15 |
| 47 | 6 | 20 | 22 | * | 23 | 4 | 7 | 5 | 5 | 5 | 5 | 4 | 2 | 1 | 0 | 0 | 0 | -1 | 18 | 22 | 0 | 0 | 19 | 15 |
| 56 | 6 | 22 | 22 | * | 23 | 4 | 7 | 5 | 5 | 5 | 5 | 4 | 2 | 1 | 0 | 0 | 0 | -1 | 18 | 22 | 0 | 0 | 19 | 15 |
| 65 | 6 | 22 | 22 | * | 23 | 4 | 7 | 5 | 5 | 5 | 5 | 4 | 2 | 1 | 0 | 0 | 0 | -1 | 18 | 22 | 0 | 0 | 19 | 15 |

Figure 3: STATUS CODES

1. K(T) - Activity Index:

- 1 - Grade 9
- 2 - Grade 10
- 3 - Grade 11
- 4 - Grade 12
- 5 - Grade 13
- 6 - C.A.A.T. 1
- 7 - C.A.A.T. 2
- 8 - C.A.A.T. 3
- 9 - Univ. 1
- 10 - Univ. 2
- 11 - Univ. 3
- 12 - Univ. 4
- 13 - Univ. 5
- 14 - Univ. 6
- 15 - Univ. 7
- 16 - Univ. 8
- 17 - Univ. 9
- 18 - Univ. 10
- 19 - Retraining
- 20 - Employment
- 21 - Unemployment
- 22 - Non-Labour-Force

2. KS(T) - Spouse Activity Index: Same Indices as K(T)

3. H(T) - Regional Index:

- 1 - NFLD.
- 2 - P.E.I.
- 3 - N.S.
- 4 - N.B.
- 5 - QUEBEC
- 6 - ONT.
- 7 - MAN.
- 8 - SASK.
- 9 - ALTA.
- 10 - B.C.
- 11 - CANADA

4. S(T) - State Vector

Constant Characteristics

- SO(1) - Sex 1 - Male
2 - Female
- SO(2) - Immigrant Status 1 - Domestic
2 - Immigrant
- SO(3) - Province (Region) of Birth 1.....11
- SO(4) - Cohort Identifier: - Year of Birth
- SO(5) - Sex of Spouse

Time Dependent Characteristics

- S(T,1) - Age at (Most Recent) Marriage
- S(T,2) - Number of Years of Post-Secondary School
- S(T,3) - Educational Achievement Category
 - 1 - Elementary
 - 2 - Some Secondary
 - 3 - Secondary Graduate
 - 4 - Some Post-Secondary/Non-University
 - 5 - Some Post-Secondary/University
 - 6 - Post-Sec. Graduate/Non-University
 - 7 - University First Degree (B.A., B.SC.)
 - 8 - University Second Degree (M.A., M.SC.)
 - 9 - University Third Degree (PH.D.)
- S(T,4) - Income Profile Category
 - 1 - Elementary
 - 2 - Some Secondary
 - 3 - Secondary Graduate
 - 4 - Some Post-Secondary and Post-Secondary/Non-University
 - 5 - University First Degree (B.A., B.SC.)
 - 6 - University Second Degree (M.A., M.SC.)
 - 7 - University Third Degree (PH.D.)

- S(T,5) - Decile Income Category 1.....10
- S(T,6) - Spouse Education Category
Same 9 Categories as S(T,3)
- S(T,7) - Spouse Decile Income Category 1.....10
- S(T,8) - Spouse Income Category
Same 7 Categories as S(T,4)
- S(T,9) - Number of Children Born
- S(T,10) - Marital Status
 - 0 - Single
 - 1 - Married
 - 2 - Widowed
 - 3 - Divorced
- S(T,11) - # of Months Unemployed in Year T
- S(T,12) - Dependent Status
- S(T,13) - # of Months Unemployed in Year T for Spouse
- S(T,14) - Age Difference of Spouse
- S(T,15) - Age at Graduation From Secondary School
- S(T,16) - Age at Graduation From Post-Secondary School
- S(T,17) - Age of Post-Secondary Graduation fro Spouse
- S(T,18) - Age at Entrance To Post-Secondary School

Child - Age Vector

CAG(I,J): Updated Age Age and Sex of i'th Child

*Appendix II
Model Summary
of the Program*

APPENDIX II

Model Summary

This appendix presents a summary of the computer program implementation of the analytical framework described in the paper. The program generates output which can assist in the analysis of the distributional impacts of government transfer programs. An important by-product of the analysis is the creation of a synthetic longitudinal sample of records of individual demographic and economic histories. As has been indicated in the paper the model includes a scheme for the creation of a sample of individual records and the use of this sample in the analysis of general transfer programs.

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 dealt with in a report [4] by 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 a 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 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.

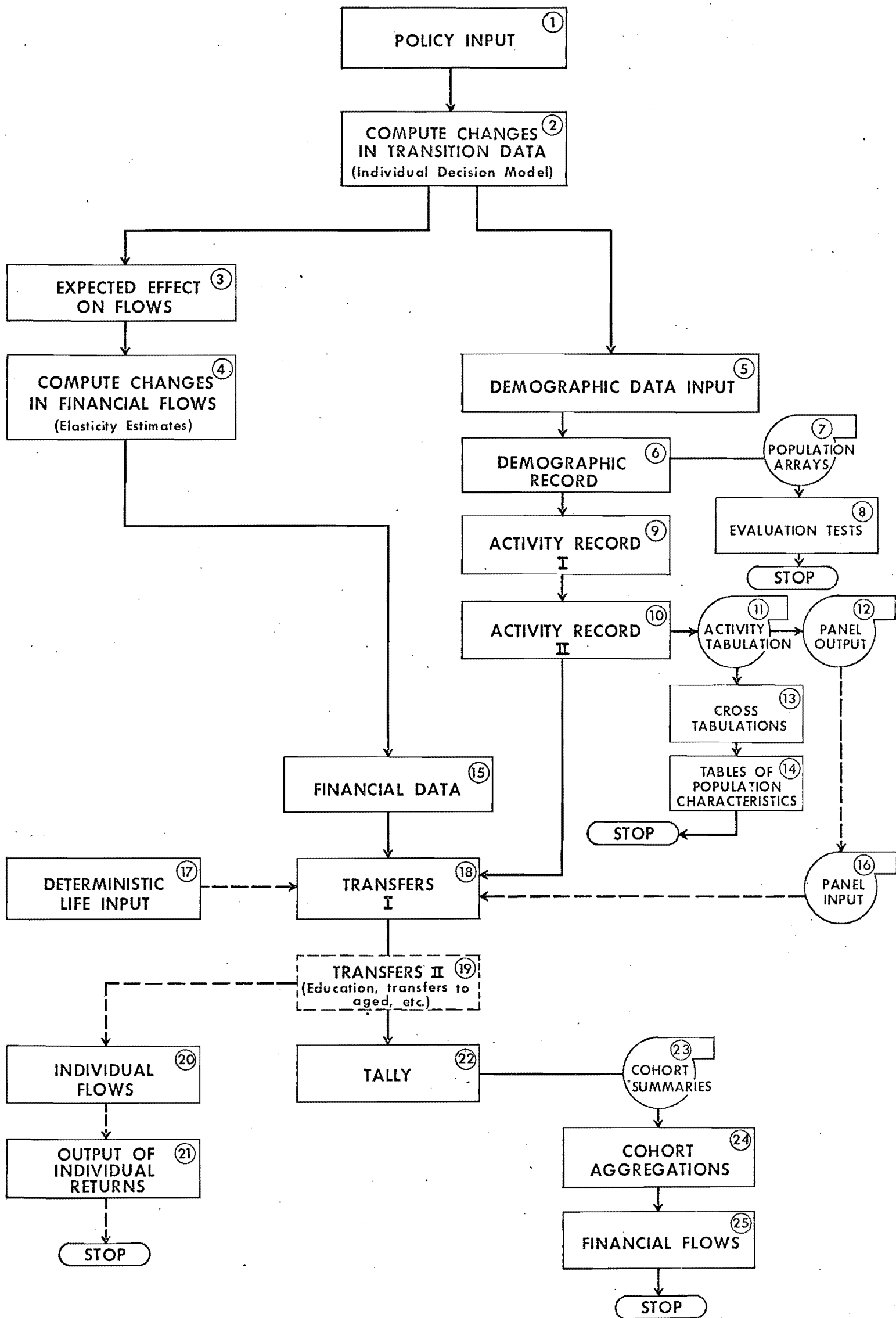
6. Comparison with aggregate analyses

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. Any application to policy analysis will require reconciliation of the 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 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.

FIGURE 1

SUMMARY FLOW CHART
MICRO-SIMULATION MODEL



- * 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, Forthcoming

- 6. Cook, G.C.A. Data Appendix, Planning Branch
 report, 1970.

* DRICUTT, G.H., and others. MICR ANALYSIS
OF SOCIOECONOMIC SYSTEMS: A SIMULATION
STUDY • NEW YORK • HARPER & ROW, 1961.

* DRESCH and GOLDBERG

INSTITUTE REPORT TO ONTARIO COMMITTEE

* SHERSHINSKI

ELIZABETH

* BEN-PORATH

* ZISLITZ ALB. S. W. L.

* SP. - ...

Bibliography.

Arnott R. Income data methodology,
unpublished report, 1993

1. Anderson, T.W. and L.A. Goodman, "Statistical Inference About Markov Chains", Annals of Mathematical Statistics, Vol. 28: 89-110 (1957).

Becker, G. Human Capital, National Bureau of Economic Research, New York, Columbia University Press, 1964.

Chen, M.A. MCSSAP User's Manual,
IOASEP paper

Cohen, M.A. Monte Carlo Simulation of Government Transfer Programs: The Deterministic Model, Quantitative Analysis Course Project Report, University of Toronto, 1992

Chen M.A and L. Cseh Micro-Simulation System Manual
Analysis, University of Toronto, Toronto 1993
Institute for Policy

Commission on Post-Secondary Education in Ontario
Draft Report, Toronto, Queen's Printer, 1990

Cook G. C.A. Data Appendix, Planning
Branch report, 1996

Dobell A.R. and M.A. Cohen, A Microsimulation Model for the Evaluation of Government Transfer Programs, Paper presented at the ORSA - TIMS - AIEE joint conference Atlantic City, November, 1972.

Dobell A.R., Cohen, M.A. and L.D. Stone. Synthetic Longitudinal Sampling and Demographic Record Linkage: Methodology and Validation, Institute for Policy Analysis, U. of Toronto, forthcoming working paper.

2. Dresch, S.P. and R.D. Goldberg, "Variable Term Loans for Higher Education - Analytics and Empirics", Annals of Economic and Social Measurement, Vol. 1, No. 1 (1972).

3. Durbin, E.P., Manpower Programs and Markov Chains, Santa Monica, California, Rand Corporation, October 1968. (Memorandum No. RM-5741-OEO).

4. Hartman, R.W., Credit for College; Public Policy for Student Loans, New York, McGraw-Hill, 1971.

Lacombe, B., Some Economic Aspects of Canadian Educational Systems, report prepared for Economic Council of Canada and the Quantitative Analysis Course, ~~RAM~~ Ottawa, 1973.

5. Lee T.C., G.G. Judge and A. Zellner Estimating the Parameters of the Markov Probability Model from Aggregate Time Series Data, Amsterdam North-Holland Publishing Company 1970.

5. Martens, A., M. Stelcner and T. Thonstad, Student Flows Through Primary and Secondary Schools in Quebec, Montreal, Institute for International Economics, 1970. (A Model for Planning Manpower and Education in Quebec, Progress Report 2.)

4. McReynolds, W.P., A Mathematical Model For An Educational System Occasional Papers 19 The Ontario Institute for Studies in Education Toronto 1971

6. Mood, A., More Scholars Per Dollar; Report Prepared for the Ford Foundation and Carnegie Commission on Higher Education, University of California, Irvine, Public Policy Research Organization, 1971.

7. Organization of Economic Co-operation and Development, Mathematical Models in Education Planning, Paris, 1967.

Petchenik, S.G., Financing Post-Secondary Education in Canada, _____, 1971.

8. Pike, R.M., Who Doesn't Get to University - and Why, Ottawa, Association of Universities and Colleges of Canada, 1970.

Stone, L.O. Regional Population Projections Methodology. IQASEP Paper _____

9. Thonstad, T., Education and Manpower: Theoretical Models and Empirical Applications, Toronto, University of Toronto Press, 1968.

Wessenthal, M. The Threat of Increasing
Numbers and Costs in Post-Secondary
Education, paper presented at Banff, Alberta,
Oct. 1970.

Wolfson, M. C. Compas User's Manual
IQ ASEP Technical Paper #