Working Paper No. 257.

Household Dynamics and Socio-Economic Forecasting: A Micro-Simulation Approach.

M.Clarke, P.Keys and H.C.W.L.Williams.

Paper prepared for the European Congress of the Regional Science Association, London, August 1979.

School of Geography,
University of Leeds,
Leeds LS2 9JT.

August 1979.

Household Dynamics and Economic Forecasting:

A Micro-simulation Approach.

Contents.

1. Introduction.

"In its broadest sense aggregation is a process whereby a part of the information available for the solution of a problem is sacrificed for the purpose of making the problem more easily manageable".

H.A. John Green (1964).

One of the more significant developments in analytic planning methods over the last decade has been the refinement and application of models specified and estimated at the level of the individual decision maker. This is perhaps best illustrated in the case of travel demand analysis in which there has been a dramatic shift from the conventional aggregate approach, involving models specified in terms of trip groups, to one embracing probabilistic choice concepts in which theoretical statements are married to a representation of behaviour at the micro-level. This transition has been prompted not only by the inefficiencies, expense and insensitivity of the modelling frameworks developed through the 1960s, but also by the new planning contexts which have emerged in the seventies, involving shorter time horizons, a wider range of policies, and evaluation frameworks which required a greater consideration of distributional issues.

While the behavioural paradigm was already fashionable in human geography in the late 1960s, it was noteably Hagerstrand,(1970) and Morgan Thomas (1969) who urged the reconciliation of aggregate macro-level descriptions of socio-economic phenomena and their spatial manifestations with disaggregate micro-level concepts. A discussion of the development of these ideas is given by Leigh and North (1978). The formal aspects of the aggregation process have, of course, long been of interest to economists (see, for example, Green, 1964), but it was not until the advent of high speed electronic computers that the representational and computational aspects involved in aggregation have been studied in depth. They are now in fact the subject of detailed investigations in travel demand analysis.

Orcutt and his collegues (1961, 1976) recognized that for many phenomena of interest in socio-economic studies, several attributes of the behavioural units (individuals, households, etc.) had to be specified if a significant amount of the variability between these units was to be captured. To prevent information loss on aggregation it was thus required to manipulate a representation whose data states were of high dimension. To avoid a sparse matrix problem, Orcutt et al. (1961) (see also Wilson, 1974 and Wilson and Pownall, 1976) noted the possibility of the efficient computational manipulation of information, through processing samples of behavioural units multiply classified by the various attributes of interest. Dynamical processes, such as births and deaths, which involve transitions between states, could then be effected by Monte Carlo simulation, and high dimensioned coupled differential equations governing the system dynamics could be solved. Orcutt's work in the Urban Institute has culminated in the development of a sophisticated demographic model, DYNASIM, which is more accurately described as a model of household dynamics implemented at the micro-level.

The processing of samples with attribute lists appears to be ideally suited to the study of spatial systems in which the household activity pattern demands a large number of attributes for its description. The recent cross-sectional location and mode choice studies of Wilson and Pownall (1976), Kreibich (1979) and Bonsall (1979) are excellent examples of this approach, in which the choices, constraints and decision processes can be examined at the individual level and any necessary aggregation process performed without loss of essential information.

While there have been many simulation studies of specific firms and industries following the suggestions and analyses of Shubik (1960), Cyert and March (1963) and others, the sheer complexity and assortment of individual firms and lack of suitable data, have prevented similar progress on the numerical aggregation process being made in the production sector. Forecasting models of that sector and its interaction with the consumption sector are typically based on the industrial groupings in conventional input-output and econometric models. Orcutt et al. (1976) have, however, specifically considered the embedding of a micro-model of household dynamics within an environment created by a macro-economic model. This interface of macro- and micro- concepts will be an important ingredient of the regional application described in the latter half of this paper.

In the paper we consider some of the representational and methodological issues involved in the formation of dynamic models for interacting labour and housing systems. Of central interest will be the numerical approach to the aggregation process itself, and in particular we aim to marry the dynamical considerations of Orcutt et al. (1961, 1976) and Kain and Apgar (1977), with the study of micro-level interdependency by Wilson and Pownall (1976).

In Section Two we consider the representation of household dynamics and response, and the design of models for the labour and housing systems. Transitions in these systems are then considered in a <u>macro</u> stock-flow representation, and the methods of forming allocation models are presented in Section Three. An illustration of the interface of macro- and micro-economic systems is then presented in Sections Four and Five in a regional application. Here the implications for individual households in Yorkshire and Humberside of various projections of the demand for labour in the early 1980s are presented.

2. Household Dynamics and Interdependency.

Any quantitative understanding of the dynamics, response and regulation of a socio-economic system must be expressed within a pre-specified frame of reference. The process of system description through state definition, often tailored to process information collected at one or more cross sections, not only determines what can be addressed by a model but also reveals what the modeller (as an observer of the system) considers important in a particular problem context.

An understanding of the role of space, and in particular in what circumstances space may be regarded as a passive backcloth against which demographic, social and economic processes are 'mapped', or is an active ingredient of change, must ultimately be traced to the actions, both present and past, of individuals, households, firms and government. Indeed, the qualitative and quantitative explanation for activity and consumption patterns at the level of individual spatial actors are the hallmarks of the behavioural perspective. The construction of behavioural macro-models involves aggregation over the actions of these micro-units.

The analysis of the heterogeneity exhibited in the characteristics of households is an important methodological approach to the understanding of the various processes governing their evolution and response to change. The traditional and periodic travel, expenditure and household surveys allow a cross-sectional micro-level description of the household to be drawn up in which are recorded the following vectors:

$\underline{d}$: the demographic characteristics

$\underline{m}$: sources and measures of income

$\underline{e}$: expenditure (consumption bundle)

$\underline{A}$: an activity pattern.

The characteristics of each individual I_i^j belonging to an M^j membered household H_j, of an N membered population, may be summarized in the following notation

$$H_j\{\ldots\ \{I_i^j(\underline{d}_i;\ \underline{m}_i;\ \underline{e}_i;\ \underline{A}_i)\}\ldots.;\ (\underline{d}^j,\ \underline{m}^j,\ \underline{e}^j,\ \underline{A}^j)\}.$$
$$j = 1,\ 2,\ \ldots.\ N.$$

Various summary household characteristics may also be recorded in $\underline{d}^j, \underline{m}^j, \underline{e}^j, \underline{A}^j$.

This information, which is inevitably grouped and cross-classified in official publications, is the essential raw material for micro-level Monte Carlo studies involving list processing. The studies of Orcutt et al. (1961, 1976), Wilson and Pownall (1976), Kreibich (1979), Bonsall (1979) and others have employed such representations, adopting subsets of this information according to the study at hand.

The current activity pattern (which might include information pertaining to residence, workplace, shopping location, etc) and financial state of a household may be seen as the outcome of a particular set of decision processes, in which choices are made subject to a set of constraints. (The latter may of course render the choice set so small such that only one feasible alternative exists.) The work of Hagerstrand has served to emphasise the role of constraints on the actions of individuals in space and time, and this now forms one of the principal themes in research into travel behaviour.

We can conveniently summarize this behavioural perspective in the following terms (Williams and Ortuzar, 1979):

$$\{d_j^*, \underline{B}(\underline{s},\ \underline{Z})\ ;\ \underline{Y};\ Q\}_k$$

in which we denote

$\underline{d}^*$: descriptors (attributes) of individual decision maker k.

$\underline{B}$: the set of alternatives out of which a selection is considered by individual k. Each alternative might be characterized by state descriptors $\underline{s}$ and each state by the set of attributes $\underline{Z}$.

$\underline{\Psi}$: the set of constraints to which the actor k is subject.

$\underline{Q}$: the set of objectives motivating the choice process.

If the classical economic paradigm is invoked for the formation of models, the consumption bundles (expenditure patterns) are treated as continuous variables, and are readily amenable to the conventional econometric treatment. In this paper we shall examine the expenditure pattern through the use of the consumption function C_q, for each relevant good or service q

$$C_q = C_q(\underline{m}^{\underline{h}}, \underline{h}), \tag{2.1}$$

in which $\underline{m}^{\underline{h}}$ refers to the income of a household of type $\underline{h}$.

Activity patterns, on the other hand, usually command a discrete representation, being the outcome of a quantal choice process (McFadden, 1978). In this case the probability P_q^k that a given decision unit selects an activity (eg. residential location) choice A_q, may be expressed as

$$P_q^k = P(\underline{Z}^k, \phi \mid \underline{B}^k, \underline{\Psi}^k) \tag{2.2}$$

in which $\underline{Z}^k$ is the set of attribute vectors for all choices in the feasible set determined by $\underline{B}^k$ and $\underline{\Psi}^k$. P might be expressed as a parametric function, such as a member of the logit family, involving the set of coefficients $\underline{\phi}$, or an algorithm derived from the set of objectives $\underline{Q}$, and an explicit decision mechanism (Williams and Ortuzar, 1979).

Through the nature of available alternatives, and the economic and social organization of individuals, an event, corresponding to a transition between two states may well trigger off one or more others. The initial trigger inducing change may well take the form of demographic changes (eg. life cycle changes), economic impulses impinging on the household income, or other policy issues which modify the attributes $\underline{Z}$. The path through which these changes are propogated may, as Heggie (1978) has discussed, be through

(i) the available set of alternatives $\underline{B}$

(ii) the constraints $\underline{\Psi}$

(iii) indices through which preferences are recorded

Wilson and Pownall (1976) recognized the importance of the micro-representation implemented through the processing of samples of individuals as the key to understanding interdependence or correlation between transitions - in this case the influence of anobligatory relocation of residence on the total activity pattern of the household was sought.

A wide range of interdependencies may be specified in a model. The effect of marriage on housing market participation; that of birth (or pregnancy) on female labour force participation; and of migration on both labour and housing markets are evident examples of demographic and migratory triggers. Income changes which may stimulate car or household ownership changes with the subsequent modification of the household budget and activity pattern are clear examples of a straightforward economic impulse. The influence of death on the labour force participation of a spouse who strives to maintain income is a clear example of an interaction between demographic and economic processes. An extension to a consideration of the correlation of the complex $\{\underline{d}, \underline{m}, \underline{e}, \underline{A}\}$ in space and/or time is immediately apparent.

The dynamics of a household may be formally expressed in the following differential equations

Demographic Processes.

$$\frac{\Delta N_H(\underline{c},t)}{\Delta t} = \sum_{\underline{c}'}\{\rho(\underline{c}' \to \underline{c};\ t)\ N_H(\underline{c}',t) - \rho(\underline{c} \to \underline{c}';\ t)\ N_H(\underline{c},t)\} \qquad (2.3)$$

$\underline{c}, \underline{c}' \equiv \underline{c}_D, \underline{c}'_D$

Labour Supply

$$\frac{\Delta S_I(\underline{c},t)}{\Delta t} = \sum_{\underline{c}'} \{\eta(\underline{c}' \rightarrow \underline{c};t)\, N_I(\underline{c}',t) - \eta(\underline{c} \rightarrow \underline{c}';t)\, N_I(\underline{c},t) \qquad (2.4)$$

$\underline{c}, \underline{c}' \equiv \underline{c}_L, \underline{c}'_L$

Housing Demand

$$\frac{\Delta D_h(\underline{c},t)}{\Delta t} = \sum_{\underline{c}'} \{\xi(\underline{c}' \rightarrow \underline{c};t)\, N_H(\underline{c}',t) - \xi(\underline{c} \rightarrow \underline{c}';t)\, N_H(c,t)\} \qquad (2.5)$$

$\underline{c}, \underline{c}' \equiv \underline{c}_h, \underline{c}'_h$

$\Delta N_H(\underline{c},t)$, $\Delta S_I(\underline{c},t)$ and $\Delta D_h(\underline{c},t)$ are respectively the changes in the number of households in category $\underline{c}_D$, individuals in category $\underline{c}_L$ and households in $\underline{c}_h$, in the time interval $(t, t + \Delta t)$. The three sets of categories $\underline{c}_D$, $\underline{c}_L$ and $\underline{c}_h$ are each subsets of the total categories $\underline{k}$ identified with each household.

Monte Carlo simulation may be used to solve these differential equations. Because the demographic and economic characteristics associated with a household interact both through the existence of 'overlapping' category sets $\underline{c}_D$, $\underline{c}_L$ and $\underline{c}_h$, and in the presence of related decision makers H and I, this method is particularly appropriate. In order to achieve a solution by forward iteration, it is necessary to specify the transition rate matrices $\underline{\rho}$, $\underline{\eta}$ and $\underline{\xi}$ and the initial conditions. This specification process will be treated within the context of a regional application in Section Four.

In order to consider the interaction between demographic, social, and economic processes it is necessary to provide some interface between the household and an economic 'environment'. Orcutt et al. (1976) have considered such an interface at the national level, using a small econometric model. In a regional model currently under construction at Leeds, we are adopting Orcutt's general philosophy with appropriate modifications to our application. Most noteable among these is the provision of allocation mechanisms linking the supply and demand sides of segmented labour and housing systems. An outline of the regional model, which we describe further in Section Four, is shown in Figure One.

The individual transitions in the labour and housing markets are both responsible for and are subject to the aggregate properties of the allocation process. A matching of individuals (households) to jobs (houses) is central to the provision of a micro-level description of the dynamics and response of the systems. The strategy for interfacing macro- and micro- properties involves three simple stages

(i) the aggregation of behavioural units and stock items into pools of <u>potential</u> demand and supply. This process is denoted by E.

(ii) the solution of the allocation or matching process M at an aggregate level.

(iii) the matching of individual behavioural units to stock items through a process of <u>sampling</u> S from the aggregate allocation functions derived at stage (ii).

We turn now to consider various aspects of the allocation process M which is central to the transmission of economic impulses through the labour market into the housing system.

3. Allocation and Matching in the Labour and Housing Systems.

Elsewhere (Clarke et al. 1979) the authors have developed models of the allocation process, based on a stock-flow representation, in which the internal moves are distinguished from moves by new entrants to the systems. This representation is particularly useful for examining the relationship between mobility and external supply (see also Byler and Gale, 1978). Through the use of conservation or accounting relations developed at the macro-level, the notion of chaining, or more specifically, the dependence of internal moves on the availability of moves to vacant stock may readily be accomodated.

Because the logic of transitions in both systems is identical, a common representation of the accounting relations may be developed. These relations are now considered using the following notation:

$X(\underline{\alpha}, \underline{\beta} \rightarrow \underline{\beta}')$: the number of transitions, or moves, in a chosen time interval by persons (or households) of type $\underline{\alpha}$ between stock units (jobs or houses) $\underline{\beta}$ and $\underline{\beta}'$. $\underline{\alpha}$, $\underline{\beta}$, and $\underline{\beta}'$ are vectors of attributes characterizing the stock and flow elements.

$X(\underline{\alpha}, * \rightarrow \underline{\beta}')$: will be specifically used to distinguish new entrants into the system from internal moves.

$S^N(\underline{\beta})$, $S^V(\underline{\beta})$: new and vacant stock of type $\underline{\beta}$ available in the given time interval. $S^V(\underline{\beta})$ is derived from a previous time period.

$D^N(\underline{\alpha})$, $D^M(\underline{\alpha})$: the total potential demand for stock by units of type $\underline{\alpha}$. New entrants (N) and internal movements (M) are distinguished.

It should be noted that the allocation variable $\underline{X}$ corresponds to transitions which are actually achieved.

The demand and supply relations for both labour and housing systems may now be expressed. The moves to new stock and 'second hand' stock are distinguished. The accounting equation for new stock are as follows:

$$\sum_{\underline{\alpha}}\sum_{\underline{\beta}'} X(\underline{\alpha}, \underline{\beta}' \rightarrow \underline{\beta}) + \sum_{\underline{\alpha}} X(\underline{\alpha}, * \rightarrow \underline{\beta}) \leqslant S^N(\underline{\beta}) \quad (3.1)$$

The satisfied demand for new stock in class $\underline{\beta}$ is composed firstly, of all the internal moves by persons of type $\underline{\alpha}$ from stock units of type $\underline{\beta}'$. Summation over person types $\underline{\alpha}$ and stock types $\underline{\beta}'$ generates the total demand by internal movers. The second term in Equation (3.1) corresponds to all new entrants moving to stock in class $\underline{\beta}$.

The internal moves themselves generate the supply of stock which induces further transitions. The demand for and supply of 'second hand' stock may thus be written as follows:

$$\sum_{\underline{\alpha}}\sum_{\underline{\beta}'} X(\underline{\alpha}, \underline{\beta}' \rightarrow \underline{\beta}) + \sum_{\underline{\alpha}} X(\underline{\alpha}, * \rightarrow \underline{\beta}) \leqslant S^V(\underline{\beta}) + \sum_{\underline{\alpha}}\sum_{\underline{\beta}'} X(\underline{\alpha}, \underline{\beta} \rightarrow \underline{\beta}') \quad (3.2)$$

The right hand side of this equation is composed of the vacancies or the start of the period, and the moves <u>from</u> stock of type $\underline{\beta}$. The left hand side corresponds to that in Equation (3.1).

The inequalities in these equations (the accomodated demand for moves must be less than or equal to the available supply) may be converted into equalities simply by adding slack variables $\Delta S^N(\underline{\beta})$ and $\Delta S^V(\underline{\beta})$ to (3.1) and (3.2) respectively. As a result of the allocation process these may of course turn out to be zero.

The satisfied demand for moves must also be bounded from above by the total <u>potential</u> demand for moves. If this information, or an estimate of it, is available for both new entrants and internal movers the following relations must be satisfied:

$$\sum_{\underline{\beta}} X(\underline{\alpha}, * \rightarrow \underline{\beta}) \leq D^N(\underline{\alpha}) \tag{3.3}$$

$$\sum\sum_{\beta\beta'} X(\underline{\alpha}, \underline{\beta} \rightarrow \underline{\beta}') \leq D^M(\underline{\alpha}) \tag{3.4}$$

which may similarly be converted into equalities through the use of slack-variables. Mis-matches in the housing and labour systems and 'hidden demand' are partially manifested through these slack variables associated with Equations (3.1) - (3.4).

The relations (3.1) - (3.4) are simply the logical consequences of a specification of the allocation process within the chosen stock-flow representation. In order to generate a model of that process, some theoretical rationale for achieving the matching, reflected in the allocation matrix $\underline{\underline{X}}$, must be supplied. Clarke et al. (1979) have considered the generation of allocation models from extremal principles based on both economic and information theory concepts. We outline first the latter approach which employs the Efficient Information Adding (EIA) concepts discussed by Snickars and Weibull (1977), which, in the allocation model considered here, involves minimization of the quantity

$$Z = \sum_{\underline{\alpha}} \sum_{\underline{\beta}} \sum_{\underline{\beta}'} X(\underline{\alpha}, \underline{\beta} \rightarrow \underline{\beta}') \ln \frac{X(\underline{\alpha}, \underline{\beta} \rightarrow \underline{\beta}')}{\overline{X}(\underline{\alpha}, \underline{\beta} \rightarrow \underline{\beta}')}$$

$$+ \sum_{\underline{\alpha}} \sum_{\underline{\beta}} X(\underline{\alpha}, * \rightarrow \underline{\beta}) \ln \frac{X(\underline{\alpha}, * \rightarrow \underline{\beta})}{\overline{X}(\underline{\alpha}, * \rightarrow \underline{\beta})} \tag{3.5}$$

Subject to the relations (3.1) - (3.4). The quantity $\overline{X}$ in the objective function may be taken as the allocated flows in the previous time period. The specific model developed will be dependent on exactly what information is available. If, for example, the values of certain slack variables in the previous period are known, these may also be used to provide a basis for the discrimination of the prior information from the (projected) allocation.

An alternative strategy for the construction of a forecasting or response model is to make explicit the economic principles governing the allocation process and to express these in terms of an extremal principle characterizing the interaction between demand and supply. Associated with each combination of behavioural units characterized by $\underline{\alpha}$ and stock items characterized by $\underline{\beta}$, it is possible to define a matching function $B(\underline{\alpha}, \underline{\beta})$. The allocation may then be achieved by maximizing the overall 'benefit' of allocation. This must, of course, be interpreted within the context of the particular rules of allocation which the extremal principle (through the Kuhn-Tucker conditions characterizing the optimal solution of the mathematical programs) seeks to represent. Different sectors of the markets may operate under different mechanisms - in the private sector of the housing system $B(\underline{\alpha}, \underline{\beta})$ might have the conventional interpretation of a bid-rent, while in the public sector $B(\underline{\alpha}, \underline{\beta})$ might be based on a points system reflecting welfare considerations.

To allow for heterogeneity in the household or individual groups which compete in the markets, a dispersion term is entered into the objective function, as in the expression

$$Z = \sum_{\underline{\alpha}} \sum_{\underline{\beta}} \sum_{\underline{\beta}'} X(\underline{\alpha}, \beta \rightarrow \underline{\beta}') \cdot \{\bar{B}(\underline{\alpha}, \underline{\beta}') + G(\underline{X}, \underline{\theta})\} \tag{3.6}$$

Here $G(\underline{X}, \underline{\theta})$ may be a parametric function of the allocation variable. Under certain circumstances discussed by Williams and Senior (1978), Wilson et al. (1979), G may take the form

$$G = \frac{1}{\theta} \log X \tag{3.7}$$

and in the present context

$$G = \frac{1}{\theta} \log X(\underline{\alpha}, \underline{\beta} \rightarrow \underline{\beta}') \tag{3.8}$$

in which the parameter θ characterizes the dispersion of $B(\underline{\alpha}, \underline{\beta}')$ over the group $\underline{\alpha}$, for which the mean value is $\overline{B}(\underline{\alpha}, \underline{\beta})$.

Through this formulation an economic model of the matching process may be developed within a stock-flow representation, which is the counterpart of the extensions of the Herbert-Stevens-Harris model considered by Senior and Wilson (1974) and Williams and Senior (1978). In those extensions demand and supply were assumed to be given exogenously. Here, the definition of the allocation variables $X(\underline{\alpha}, \underline{\beta} \rightarrow \underline{\beta}')$ allows internal movement to be endogenously determined, and a range of interactions between different segments of the markets to be consistently documented.

We should point out that the stock-flow representation, which contains both information about the job or house which an individual or household moves <u>from</u> and that which it moves <u>to</u>, is potentially of high dimension. The process of forming the dual program may however be invoked to vastly reduce the diminsions of the calculation. For example, the dual of the EIA program (3.1) - (3.5) may be expressed as a concave maximization program subject only to non-negativity conditions on the dual variables.

4. Household Dynamics Within A Macro-Economic Environment - A Regional Application.

In this section we present an example of the interaction between demographic and economic processes, as outlined in Figure 1, involving the embedding of a micro-model of household dynamics within a regional macro-econometric model, which is itself 'plugged-in' to a national model along the lines suggested by Klein (1969). The general field of enquiry is the determination of the repercussions, for individuals and multiply classified households, of demographic trends in conjunction with alternative scenarios for labour demand in the Yorkshire and Humberside region. In the projections to 1982 we shall be primarily concerned with events associated with the labour market, that is, with operations portrayed on the left hand side of Figure 1.

The strategy for calculation is one of successive approximation. In the econometric models, which are used to generate labour demand, the household categorization is considerably less detailed than that employed in the micro-model, in which the variability which exists between individuals and households is maintained at a level appropriate to the relevant application, and is constantly subject to the external influences of the macro-model through the allocation models.

We shall discuss, in turn, four aspects of the application: the solution of the household dynamics model; the transmission of national forecasts to the required regional estimates of labour demand; the interaction between labour demand and supply through the allocation process; and, finally, the organisation of the solution process itself. The results of this application will be discussed in Section 5.

The strategy adopted for solving the coupled differential equations of household dynamics is, as we noted in Section 2, one of forward iteration from an initial state. In the Monte Carlo approach, the specification of this initial state involves the formation of a sample of households which is, in terms of the attributes considered relevant to the application, representative of the population of Yorkshire and Humberside. This sample, of size N^S, may be derived from a survey of individual households, or synthesised from either primary or secondary data sources. In the present application, the number of households $N(\underline{k}, t = o)$ with characteristics $\underline{k}$ was generated by sampling from a distribution $p(\underline{k})$, which represents the joint probability of a household with categories $\{k_1, k_2, \ldots k_N\}$.

For sufficiently high values of the sample size N^S, the expectation value of $N(\underline{k}^*, t=o)$ is given by

$$\langle N(\underline{k}^*, t=o)\rangle \simeq N^S . \, p(\underline{k}^*) \tag{4.1}$$

in which $\underline{k}^*$ represents <u>combinations</u> of the attributes in the complete set $\underline{k}$. What is considered to be a sufficiently high value for N^S is dependent on a number of factors including: the particular categories included in $\underline{k}^*$; the structure of the distribution $p(\underline{k})$; and the nature of the required output of the model. In the design and solution of the model, and particularly in the selection of a sample size, the statistical aspects must, as ever, be considered in conjunction with computational restrictions.

The probability distribution was itself synthesized from secondary data sources by decomposing the joint distribution into the product of conditionals

$$p(\underline{k}) = p(k_1)\, p(k_2|k_1) \ldots . \, p(k_n|\, k_1 \ldots . k_{n-1}) \tag{4.2}$$

and exploiting the independence between categories where reasonably justified. The method of sample synthesis is further discussed by Wilson and Pownall (1976), and Pownall (1977). In the present application the sample contained the following information associated with individuals and households:

<u>Individuals</u> : Label; age; age cohort; sex; race; education; marital status; job; number of weeks worked; wage; wage tragectory.

<u>Household</u> : Label; summary characteristics of the household head; number of individuals in the household; number of children; total income; tenure of house.

In the Monte Carlo approach, events or transitions in the model for household dynamics (eg. births, deaths, marriage, divorce, entry to the labour market, retirement, relocation, etc.) which are embodied in the matrices $\underline{\rho}', \underline{\xi}$ and $\underline{\eta}$, are induced in a straightforward manner. A comparison of a random number $0 \leqslant x \leqslant 1$ with a matrix element (for example, ρ_{ab} the probability of a transition from the state $a \rightarrow b$) determines whether the corresponding event is recorded. If $x \leqslant \rho_{ab}$ the transition is considered to occur, otherwise it is not.

The relationship between the matrix elements and the attributes associated with the various states, which constitute transition models may be underpinned by theoretical statements or simply empiricly derived expressions. These relationships are either in the form of analytic functions or 'look up tables' - the latter approach was adopted in the present model.

The occupational categories relevant to the Job label are as follows: (i) Managerial/professional; (ii) Intermediate/Non manual; (iii) Junior non-manual; (iv) Skilled/semi-skilled manual; (v) unskilled manual. Four further categories involved which may specify the individual are: unemployed; not in the labour force; retired; in educational establishment.

The labour market allocation model is <u>driven</u> from the supply side by the various demographic processes associated with ageing; and by the demand for labour which is itself derived from national economic growth predictions. We have adopted the regional forecasts of labour demand (up to 1982) derived by the University of Warwick Manpower Research Group (Lindley, 1976; Koegh and Elias, 1978) who have employed the large national Econometric model developed in the Department of Applied Economics at Cambridge (CPRG, 1978). The strategy for deriving regional forecasts is of the type discussed by Klein (1969) and Glickman (1977). The specific models adopted for the projection of employment $E_R^i(t)$ in industry i in region R at time t were of the form †

$$E_R^i(t) = f(E_*^i(t),t) \tag{4.2}$$

in which $E_*^i(t)$ is the number of employees in employment in the i^{th} industry at time t for the U.K. as a whole. The Warwick group have, in addition, derived regional totals of employment by occupation $E_R^{occ}(t)$. For the particular occupation groups considered in the present application (classes (i) - (v) above) we have developed independent estimates of $E_R^{occ}(t)$ using an "occupation by industry" matrix.

From the household dynamics model and the exogenous forecasts of labour demand the transitions generated by demographic and economic events are computed. The latter include any redundancies classified by occupation group. Although we feel that a rigorous treatment of matching and mobility requires the stock-flow representation outlined in Section 3 in which internal moves are functions of moves into and out of the labour (or housing) market, we have in this stage of the application adopted the traditional device of directly matching jobs and individuals from demand and supply pools for each category. That is, the state previously occupied by an individual making an internal transition is not considered.

† a series of eight models were actually tested including three which involved as the dependent variable, the employees in the i^{th} industry and r^{th} region as a proportion of the total in the i^{th} industry.

The Efficient Information Adding philosophy was invoked to bring about the matching of individuals, identified by sex and education to jobs classified by sex and occupation. Prior probabilities of matching were taken from the census. The resultant allocation was converted into integer form before the sampling process 3) could be employed to match a particular individual to a particular job.

In concluding this section it is appropriate to comment on various aspects of the computing process, the overall strategy of which is outlined in Figure 2. In addition to the handling of input and output of information from disc files the suite of programmes embody the following processes:

(i) ageing individuals

(ii) testing for and executing transitions (both primitive and induced)

(iii) matching items

(iv) general accounting and book-keeping.

We noted above that the selection of sample size N^s was determined partly by statistical and partly by computational considerations. In the later context the operations in (i) and (ii) give rise to CPU times proportional to N^s, while the matching algorithm, which may be used to account for interpersonal competition, gives rise to a contribution approximately proportional to $N_d \times N_s$, where N_d and N_s are respectively the number of demand and supply categories upon which items (eg. individuals and jobs) are matched. The latter multiplicative dependence can in fact be reduced by employing dual matching algorithms (see Section 2).

The efficient handling of data files is of paramount importance to the overall efficiency of the approach. Temporary files of individuals and households created within the program can be stored in binary form and this improves input/output manipulations. There is the usual trade off to be made between efficiency and system dependence. At present the suite of programs is run on a VAX-VMS interactive system.

5. The Results of a Regional Simulation Model. (First Draft).

In this section we present some preliminary results of the regional application discussed in Section 4. These results will be refined in the course of model and program development. At the current stage of specification the system entails a direct interaction between the labour and housing systems achieved through the household budget which is computed from wages and transfers, and state pensions (where applicable). The household tenure is allocated on the basis of household structure and income alone.

One reason perhaps why the micro-simulation approach has not been more generally applied is the rather specific and detailed surveys required for its implementation. In the present example, the demographic and economic characteristics of the household were generated from Secondary data sources (FES, GHS, Census, NES, etc.) which are themselves subject to sampling error. Although matrix "filling-in" techniques were used, the application must be considered limited by the information available in these sources.

Using the sample generation method outlined in the last section, a base year (1975) population for Yorkshire and Humberside was generated from a sample size N^S = 2676 (0.67%). Several of the relevant marginal probability distributions associated with this sample are summarized in Table 1. Discrepancies between survey and model results may be attributed to: direct sampling errors; the specification of the $p(\underline{k})$ matrix used to generate the initial population; and errors induced in intermediate stages in the generation of results. The discrepancies in Table 1 can in fact all be traced to the inadequacies of the particular secondary data sources adopted. This should be borne in mind when interpreting the results from the forecasts.

Two projections of labour demand were employed, based on the work of Keogh and Elias (1978); each of which involved assumptions about: world production and prices; the exchange rate; public expenditure; taxation and transfers; and labour market variables. While the Cambridge and Warwick groups examined the effect of six different policies on employment, our results are limited to a "standard view" (SV), and a policy (E) which employs assumptions involving higher productivity and investment (relative to SV) and a lowering of the exchange rate to reduce registered unemployment to a U.K. figure of 1.5 millions. These authors found the regional unemployment as follows: 11% (SV) and 7% (E) which correspond to a +4% and 0% shift from the 7% base value (1976).

It is important to note that, while the demographic events form an important "drive" to the labour market, in the present application the reciprocal causation is absent - the projections of households are independent of the labour demand forecasts.

The results from these initial model runs are summarized in Table 2. We have concentrated on two particular features of the end-period state - namely, the unemployment and associated tenure characteristics - although we would stress that any relevant cross-tabulated information relating to household and individual attributes, including their time dependence, could be easily accessed.

The unemployment trends associated with both forecasts are shown in Table 2a. These indicate very similar changes from the base year to those of Koegh and Elias (see above) and the discrepancies in their absolute values can once again be traced to misspecification of the individual attributes in the initial population. The incidence of unemployment among selected age groups is shown in Table 2b. The widely documented concentration of unemployment among young people and particularly school leavers is reflected in these figures. In terms of absolute numbers a large portion of the difference in unemployment accompanying Policy E and the standard view falls on the two age cohorts selected.

It is important to note that the measure of unemployment adopted here is that associated with a cross sectional sample. With the micro-simulation approach, it is a straightforward matter to record different features relating to the labour sector, for example: duration of unemployment, turnover within the labour market, etc.

The relative economic deprivation of households containing an unemployed worker is sensitive to the existence or otherwise of another participating worker in that household. The model results indicate that a large number (in the present version 67%) of the unemployed do in fact live in a household with another person in employment. This feature reflects itself in Table 2c in which the tenure split characteristics of households containing unemployed persons is shown. In the standard view propertionately more young people are unemployed. They are likely not to be heads of households, and hence take the tenure characteristics of their parents. In policy E on the other hand a greater proportion (relative to SV) of heads of households are unemployed and these tend to be concentrated in the rental sector of the housing market.

It should be appreciated that the tenure currently allocated to a household is made on the basis of income and household structure alone, and is unconstrained by supply conditions in the sector. It is our intention to develop the model in a symmetric fashion such that the new build housing stock, derived from a local econometric model, may be used to influence market conditions in a similar way to the labour demand projections. The addition of supply side restrictions is necessary for a realistic consideration of the mutual dependency of the labour and housing systems.

6. Directions of Further Research.

It has not been our intention in this paper to indulge in a comprehensive survey of the micro-simulation of socio-economic systems, nor to present a highly refined application, but to discuss some methodological issues in model development, and to convey a belief for the potential of the micro-approach in addressing certain types of questions. Our current research is broadly aimed at examining the dynamics of individuals and households within their economic environment and activity space.

It is in itself an important research topic to discern where the micro-simulation approach would provide the most satisfactory approach to problem analysis and solution. Indeed the whole issue of model design relates essentially to this question. Broadly speaking the type of approach adopted must be geared to the sort of question asked, the information which is available, and the type of answers sought. Where these answers demand rather detailed information about spatial actors the micro-approach, possibly in conjunction with macro-concepts may well prove to be a suitable candidate.

As Green (1964) has reminded us the aggregation process is intimately related to the processing and manipulation of information. In the context of model development it is perhaps useful to distinguish between issues relating to

(i) aggregation and information loss with respect to the attributes of individuals in their household context.

(ii) aggregation over the functional relations associated with the individuals in a group.

(iii) the treatment of complex interactive systems containing many spatial actors,

and to further differentiate between representational issues, theoretical statements, the structure of the system equations, and solution methods.

In the analysis of socio-economic systems and particularly when considering household dynamics within an economic environment it is often necessary to confront all three types of aggregation issue and to exploit the assumptions which accord with available information and requisite model outputs. In the present example, as in the work of Orcutt et al. (1976), the extended micro-representation has been used to retain important information pertaining to individuals within their household context, while the complexities of the economic system have been addressed at the macro-level.

In the example given the main interest was in the analysis of economic effects operating through the household budget and implicit assumptions about the role of space have been made. Space has been treated as a passive backcloth against which the economic and demographic processes have been mapped at the regional level.. One of the current interests is to try to understand how space should be effectively considered and represented in the micro-approach when addressing different problems at different spatial scales.

It will not have escaped the reader's notice that many of the equations with which we have been concerned are conservation equations (accounting relations) specified in a particular state-space representation. In the representation, theoretical statements may be contained within transition matrices. In the example considered in Section 5 no explicit behavioural assumptions have been made. This reflects an interest in natural dynamics based on existing trends. In future work we intend to build explicit transition rate models to more closely link demographic, social and economic processes.

In spite of the detailed data requirements we believe that an extended micro-approach has much to commend itself as a tool for systems analysis.

In the context of a number of social policy issues relating to current urban problems the variability between individuals <u>is</u> very considerable and extended over many attributes: constraints <u>are</u> highly selective and choice sets <u>seldom</u> identical. Where distributional aspects are of paramount interest and where the planning approach is orientated towards the individual/household rather than spatial groups or similar aggregates, the micro-simulation approach should, at the very least, be considered as a method of analysis. Although existing data sources might often exclude its use this situation should not be allowed to be perpetuated. The design of problem-oriented models must itself go hand in hand with the design and collection of appropriate data sets.

Acknowledgements.

We wish to thank G.T. Keogh and D.P.B. Elias for providing us with regional employment data. We alone are responsible for the preliminary results presented in this paper. In addition we acknowledge the invaluable help provided by Mr Jim Swift, manager of the VAX-VMS system, Dept. of Mechanical Engineering, University of Leeds.

References.

Bonsall, P.W. (1979) "Micro-simulation of mode choice: a model of organized car sharing". PTRC Proceedings, Summer Annual Meeting, University of Warwick.

Byler, J.W. and S. Gale. (1978) "Social accounts and planning for change in urban housing markets". Environment and Planning A 10, 247-266.

Cambridge Policy Review Group (1978) Economic Policy Review, Number 4, University of Cambridge.

Clarke, M., P. Keys, and H.C.W.L. Williams (1979) "Models of matching and mobility in the Labour and Housing Systems". Working Paper, School of Geography, University of Leeds (in preparation).

Cyert, R.M. and J.G. March (1963) "A Behavioural Theory of the Firm". Prentice-Hall, Inc.: Englewood Cliffs, N.J.

Department of Employment (1976), "Family Expenditure Survey" 1975. HMSO, London.

Glickman, N.J. (1977) "Econometric Analysis of Regional Systems", Academic Press, New York.

Green, H.A.J. (1964) "Aggregation in Economic Analysis". Princeton University Press, Princeton, New Jersey.

Hagerstrand, T. (1970) "What about people in regional science?" Papers and Proceedings of the Regional Science Association 24, 7-21.

Kain, J.F. and W.C. Apgar (1977) "Simulation of the Market Effects of Housing Allowances. Vol. II: Baseline and Policy Simulations for Pittsburgh and Chicago". Department of City and Regional Planning, Harvard University.

Keogh, G.T. and D.P.B. Elias (1978) "A Model for projecting regional employment in the U.K.". Discussion Paper, Manpower Research Group, University of Warwick, Coventry.

Klein, L.R. (1969) "The specification of Regional Econometric Models", Papers and Proceedings of the Regional Science Association 23, 105-115.

Kreibich, V. (1979) "Modelling car availability, modal split and trip distribution by Monte Carlo simulation : a short way to integrated models". Transportation, 8, 153-166.

Leigh, R. and D.J. North (1978) "The potential of the Microbehavioural approach to Regional Analysis". London Papers in Regional Science 8 (Ed. P.W.J. Batey) Pion Press, London.

Lindley, R.M. et al. (1978) "Britain's Medium-term Employment Prospects", Manpower Research Group, University of Warwick.

McFadden, D. (1978) "On the use of Probabilistic Choice models in Economics". Presented at the meeting of the American Economic Association, Chicago, Illinois, August 1978.

O.P.C.S. (1978) General Household Survey 1976, H.M.S.O. London.

Orcutt, G. M. Greenberger, A. Rivlin, and J. Korbel (1961) "Microanalysis of Socioeconomic Systems: a Simulation Study". Harper and Row, New York.

Orcutt, G., S. Caldwell, and R. Wertheimer (1976) "Policy Explorations through Microanalytic Simulation", The Urban Institute, Washington.

Pownall, C.E. (1978) "A new representation of the urban system and for the study of micro-level interdependence". M.Phil. Thesis, School of Geography, University of Leeds.

Senior, M.L. and A.G. Wilson (1974) "Explorations and Syntheses of linear programming and spatial interaction models of residential location". Geographical Analysis 6, 209-237.

Shubik, M. (1960) "Simulation of the Industry and the Firm", American Economic Review 50, 908-919.

Snickars, F. and J. Weibull (1977) "A minimum information principle - theory and practice ", Regional Science and Urban Economics, 7, 137-168.

Thrift, N. and H.C.W.L. Williams (1979) "On the development of behavioural location models." Economic Geography (in press).

Tversky, A. (1972) "Elimination by aspects: a theory of choice", Psycological Review 79(4), 281-299.

Williams, H.C.W.L. and J. Ortuzar (1979) "Behavioural Travel Theories, model specification and the Response Error problem". P.T.R.C. proceedings, Summer Meeting, University of Warwick.

Williams, H.C.W.L. and M.L. Senior (1978) "Accessibility, Spatial interaction and the spatial benefit analysis of land use - Transportation Plans". Chapter II in 'Spatial Interaction Theory and Planning Models'. (Ed. A. Karlquist, L. Lundquist, F. Snickars and J.W. Weibull) North Holland, Amsterdam.

Wilson, A.G. (1974) "Urban and Regional Models in Geography and Planning". John Wiley, London.

Wilson, A.G., J. Coelho, S.M. Macgill and H.C.W.L. Williams (1979) "Optimization in Locational and Transport Analysis, mimeo School of Geography, University of Leeds.

Wilson, A.G. and C.E. Pownall (1976) "A new representation of the urban system for modelling and for the study of micro-level interdependence". Area 8, 246-254.

		0-15	16-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	6[illegible]
1975 Males	Model	26	7	10	9	8	7	6	7	6	6	3	[illegible]
	Observed	25	8	7	8	6	6	6	6	7	5	6	1[illegible]
1975 Females	Model	24	7	10	9	7	6	7	6	6	6	4	8
	Observed	22	7	7	7	6	6	5	6	6	6	6	1[illegible]

Age - Sex Distributions (% of total population).

		Public Rented	Private Rented	Owner Occupied (mortgage)	Owned Outright
1975	Model	29	19	35	16
	Observed	27	13	34	22

Tenure Split (% of total households)

		0	1	2	3	4 +
1975	Model	18	35	25	14	8
	Observed	23	33	33	8	3

No. of workers per household (including unemployed) (% of total households)

		Professional Managerial	Intermediate Non-manual	Junior Non-manual	Skilled and semi-skilled manual	Unskilled
1975	Model	16	27	8	42	7
	Observed	17	22	9	45	7

Distribution of Employment by Occupation (% of those in employment)

Table 1. A comparison of selected base year Demographic and Employment Characteristics.

	1975	1978	1980	1982
Standard view	5%	6.9%	7.5%	8.8%
Policy E	5%	4.3%	4.3%	5.5%

Unemployment Rate

	1975	1982(standard view)	1982(policy E)
16-19 male	12%	26%	19%
20-24 male	9%	15%	5%
16-19 female	1%	18%	10%
20-24 female	1.5%	10%	3%

Percentage of all young people unemployed (1975 and 1982)

	1975		1982 Standard View		1982 Policy E.	
council rented	21	29.5	29	26	43	25.6
private rented	9	19	13	14	15	15.8
owner occupier (mortgage)	57	35.5	45	46	30	44.8
owned outright	9	16	10	14	12	14
	Unemp.	Total	Unemp.	Total	Unemp.	Total.

Tenure split characteristics of households containing unemployed persons compared with Total Tenure split.

Table 2. Unemployment rates and some characteristics of the unemployed for two policy alternatives 1975-1982.

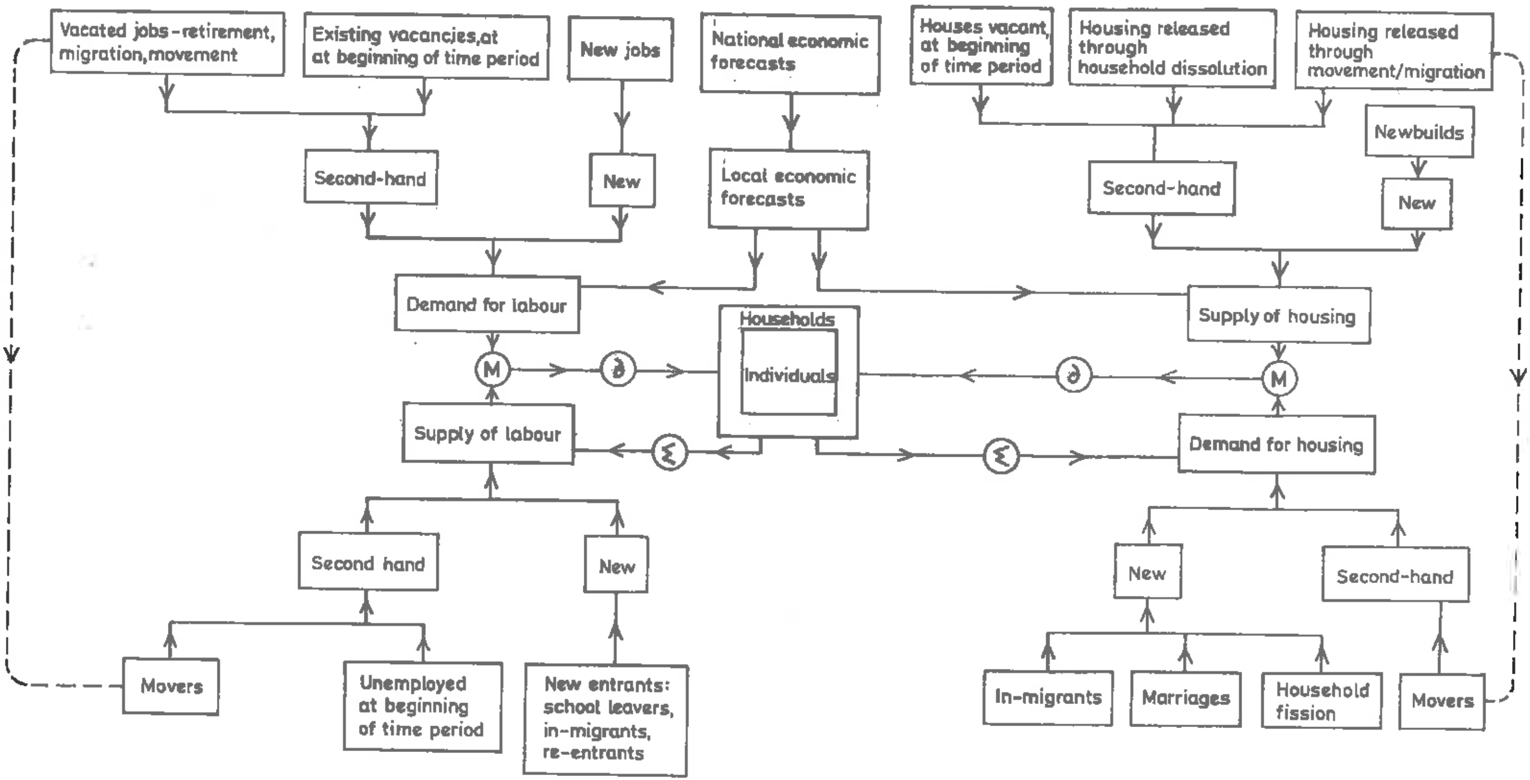

Fig.1 : Structure of the Labour and Housing Market Models

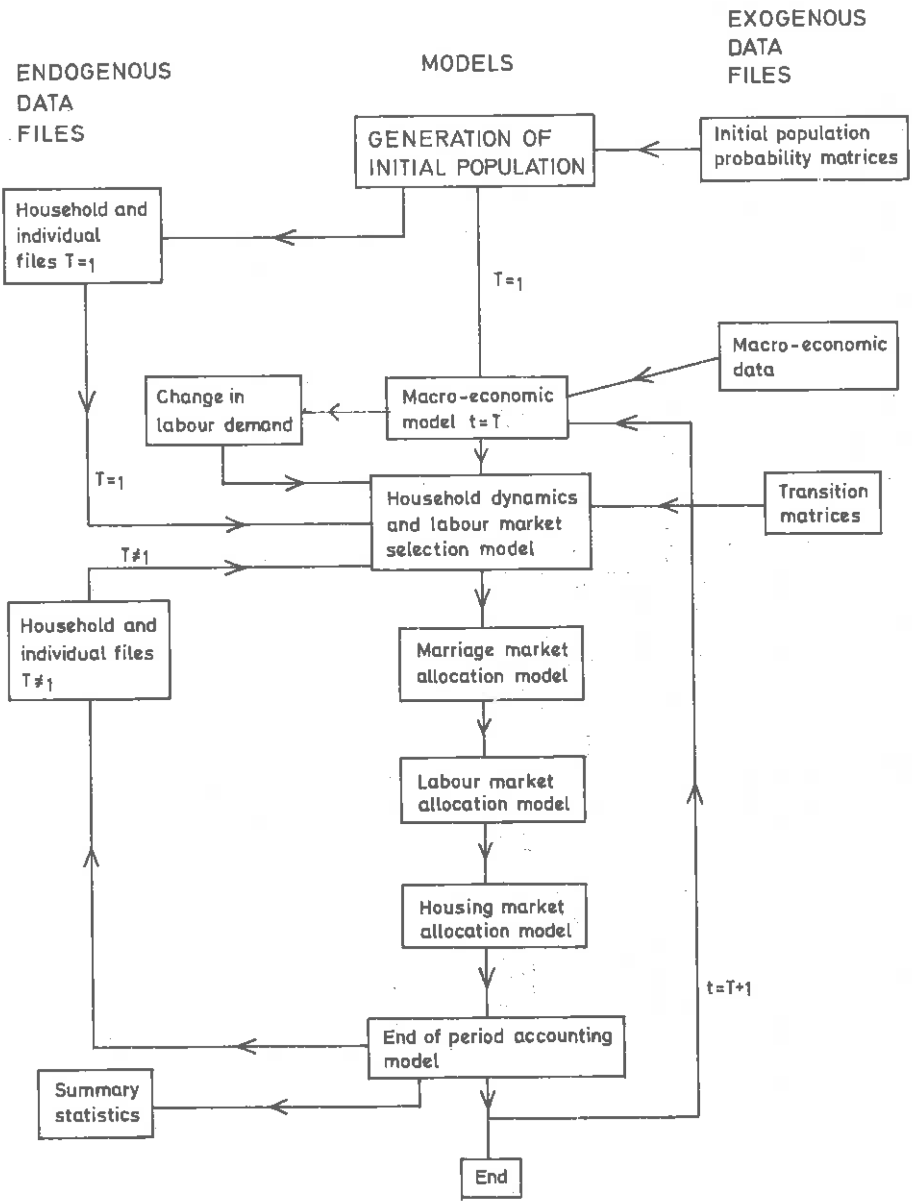

Fig. 2 : Overview of the Computational Procedure