DOCUMENT RESUME

ED 128 928	88	EA 008 701

AUTHOR	Linebarger, Dean; And Others
TITLE	ENSIM II: A Second Generation, Land Use-Based Enrollment Simulation Model. User's Manual. Research Report of Project Simu School: Santa Clara County Component.
INSTITUTION	Santa Clara County Office of Education, San José, Calif.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C.
BUREAU NO	S0010SW
PUB DATE	Aug 76
GRANT	OEG-9-72-0063(290)
NOTE	107p.; For a related document, see ED096?51; Some charts may not reproduce clearly
AVAILABLE FROM	Office of the Superintendent of Schools, Santa Clara County, 100 Skyport Drive, San Jose, California 95110 (free)
EDRS PRICE	MF-$0.83 HC-$6.01 Plus Postage.
DESCRIPTORS	*Computer Oriented Programs; Computer Programs; Educational Planning; Elementary Secondary Education; *Enrollment Projections; *Models; Prediction; *Program Descriptions; *Student Enrollment
IDENTIFIERS	Elementary Secondary Education Act Title III; *ENSIM II; ESEA Title III

ABSTRACT

This booklet presents the rationale and processes for ENSIM II, a computer-based long-range enrollment simulation model designed to consider variables not usually addressed in traditional projection techniques. The first ENSIM model was constructed for districts facing extensive residential growth and had as its primary focus the analysis of land-use data and the projection of new housing development. ENSIM II, which also generates residential development projections and which, like the original model, uses a basic cohort survival technique as its primary computational procedure, is appropriate for districts with either increasing or decreasing enrollments. Basic inputs into the model include historical enrollment information, census data, and land-use data. ENSIM II produces enrollment forecasts by age, race/ethnic composition, and geographical area. Separate sections of the booklet present a general description of ENSIM II, discuss data preparation procedures for the system, describe in detail specifications for input data sets, describe specifications for output data sets and a districtwide summary program, and list various subroutines and arrays and their uses. (Author/JG)

ENSIM II: A Second Generation, Land Use-Based Enrollment Simulation Model —User's Manual—

Dean Linebarger
Floyd Minana
Richard Cornish
Lester Hunt

Santa Clara County Component
100 Skyport Drive
San Jose, California 95110

PROJECT SIMU-SCHOOL

ENSIM II: A Second Generation, Land-Use Based Enrollment Simulation Model

User's Manual

RESEARCH REPORT

OF

PROJECT SIMU SCHOOL: SANTA CLARA COUNTY COMPONENT

Sponsored by a Grant From
Department of Health, Education and Welfare
U.S. Office of Education
Title III, Section 306
Elementary and Secondary Education Act

Grant OEG 9-72-0063 (290)
Project No. S0010SW

Office of the Superintendent of Schools
Santa Clara County
100 Skyport Drive
San Jose, California 95110

August, 1976

CONTENTS

Page

List of Figures

FOREWORD

Projecting future enrollments is one of the more arduous tasks of the educational planner. Methodologies and procedures for making forecasting less troublesome and more accurate have been a major interest of Project Simu School: Santa Clara County Component since its inception. Late in 1974, the Project produced a document which described ENSIM, an enrollment simulation model based on the analysis of land use data. The Foreword of that paper explained that the model had been developed for use in a semi-rural but rapidly growing school district and that its usefulness was thus limited to the prediction of increasing student populations. A further limitation was ENSIM's inability to forecast the race-ethnic composition of future enrollments. The Foreword concluded with a promise that future efforts of the Project would be aimed at adapting the simulation model to address a wider range of school planning settings.

This paper presents the rationale and procedures of the adapted enrollment simulation model. ENSIM II was developed during a master planning effort undertaken by Project Simu School in Alum Rock Union School District, a district in the older, east San Jose area and one whose enrollment has been on the decline in recent years. With the largest concentration of minority students in the county, Alum Rock was an especially appropriate site for adapting the original ENSIM model to produce projections of future race/ethnic composition.

Working with the Center for Urban Analysis, a project of the County Executive's Office with extensive and sophisticated computer capabilities, Project staff added to the land use analysis component of the earlier model a series of computer subroutines that would simulate migrational patterns, income changes and changes in residential vacancy rates. The final output of the Alum Rock projection effort was enrollments by age, race in small geographic areas one, five, seven and ten years into the future.

Although the ENSIM II Model has not been used since the Alum Rock effort, its designers feel the reasonableness of its forecasts in that field tests clearly warrant continued use and testing. This user's manual contains the procedures and documentation necessary for using the computer model in a variety of planning settings. The authors recognize that modifications in the model may be necessary, particularly insofar as input requirements are concerned, and an effort will be made to pr ide assistance to users who wish to install the system in their district. Tapes of the model's computer programs are available.

The Center for Educational Planning continues to function as the Santa Clara County Component of Project Simu School under local funding. Federal funding for the project expired June 30, 1975, but user assistance can be secured by calling (408) 299-4247.

Lester W. Hunt, Director
Center for Educational Planning

The project presented or reported herein was performed pursuant to a grant from the U.S. Office of Education, Department of Health, Education and Welfare. However, the opinions expressed herein do not necessarily reflect the position or policy of the U.S. Office of Education, and no official endorsement by the U.S. Office of Education should be inferred.

I. INTRODUCTION

A. Background

Projecting school enrollments over the long term has become an important, if somewhat hazardous, component of the educational planning process. Important because as resources grow more and more scarce, their prudent and equitable distribution becomes necessary to the survival of public instruction; and hazardous because the entire range of factors--social, economic and political--which tend to affect enrollment from year to year are in an accelerating state of change.

It is precisely because those factors are in continual flux and because their impact on student populations is far from direct that traditional methods of forecasting have, for the most part, ignored them. Most school enrollment projection techniques share at least four basic shortcomings: they rely extensively on past trends or lack clear guidelines for estimating future trends; they place too much emphasis on the determinants of enrollments that are easiest to deal with (i.e., retention rates, dropout rates) and lack detail on the major determinants (i.e., migration rates, new housing); they do not utilize or even consider socio-economic data on students, despite the fact that many school districts regularly collect such data; and they make no use of geographic data in distributing projections over space. When used to formulate plans and develop strategies, these traditional projection techniques can translate into empty or overcrowded classrooms.

This paper presents the rationale and processes for ENSIM II, a computer based long-range enrollment simulation model designed to consider variables not usually addressed in traditional projection techniques. The first ENSIM model was constructed for districts facing extensive residential growth and had as its primary focus the analysis of land use data and the projection of new housing development. ENSIM II, which also generates residential development and which, like the original model, uses a basic cohort survival technique as its primary computational procedure, is appropriate for districts with either increasing or decreasing enrollments. Basic inputs into the model include historical enrollment information, census data and land use data. ENSIM II produces enrollment forecasts by age, race/ethnic composition and geographical area.

The simulation model is based on two central hypotheses: first, that in a district nearing residential saturation, the major factor in enrollment change is household mobility; and second, that four household characteristics influence migration--income, size, race/ethnic makeup and tenure (own/rent). Further, a major assumption which guided development of the model was that the process of school planning is never completed. Conditions which affect enrollment patterns, facility needs and program requirements are constantly changing and must be monitored regularly if planning information is to be kept current. For this reason, ENSIM II was designed to facilitate easy update and annual "reforecasting" by user district staff.

The ENSIM II Model was developed for and field tested in Alum Rock School District, San Jose, California. Because its primary data source was a special school census, implementation of ENSIM II will be limited to those areas where similar kinds of data are available. Despite this limitation, however, it is believed that the model will significantly advance the cause of accurate long-term enrollment projections wherever it is used.

B. General Description of ENSIM II

1. Logic Flowchart

The purpose of this section is to provide the reader with a general overview of the various processes that comprise the ENSIM II Model. The model's Logic Flowchart, which is presented in Figure 1, will be described in terms that will give the non-programmer a sense of the model's parameters. What are the required inputs? When are they entered into the model? What are the ENSIM II outputs?

It should be noted that, for purposes of explanation, the "steps" or "sub-processes" required for completion of the major program processes (i.e., Out-Migration Process, etc.) are described as though a proper sequence existed. In the actual operation of the program, however, many of the "steps" are, for all practical purposes, done simultaneously. Moreover, many input data processes are done "outside" the model; these have been depicted as such in the flowchart.

a. Out-Migration Process

The Out-Migration Process is the first of the seven ENSIM II processes depicted in the Logic Flowchart. The purpose of this process is to project the numbers and characteristics of households and children who will leave the district under study during a given year of the projection period. Houses left vacant because of out-migrating households are also a product of this process.

Figure 2

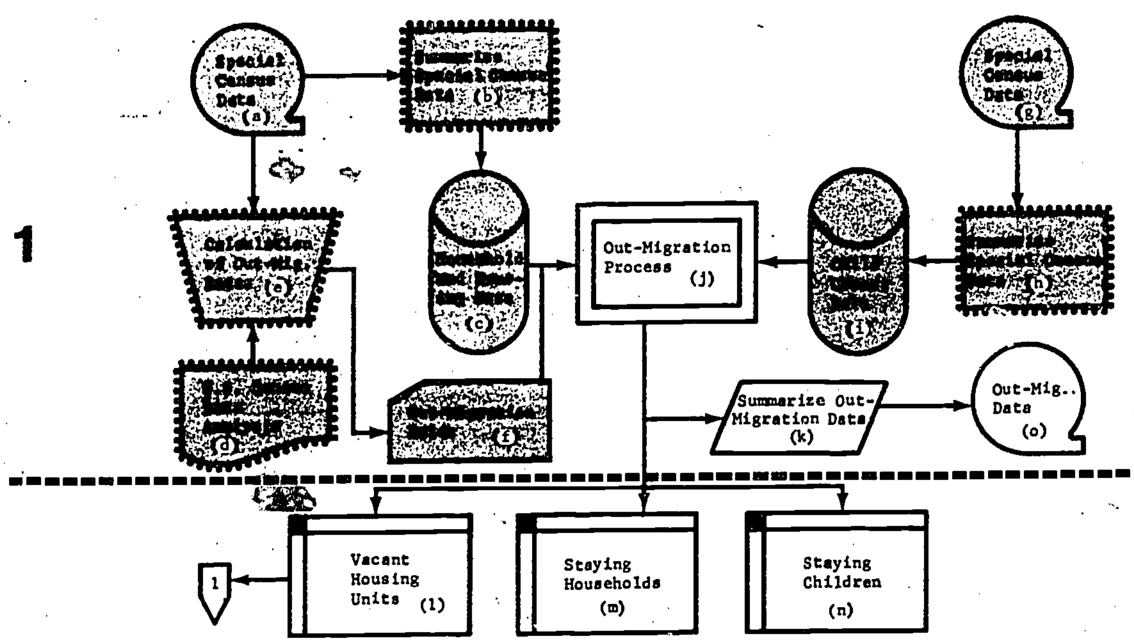
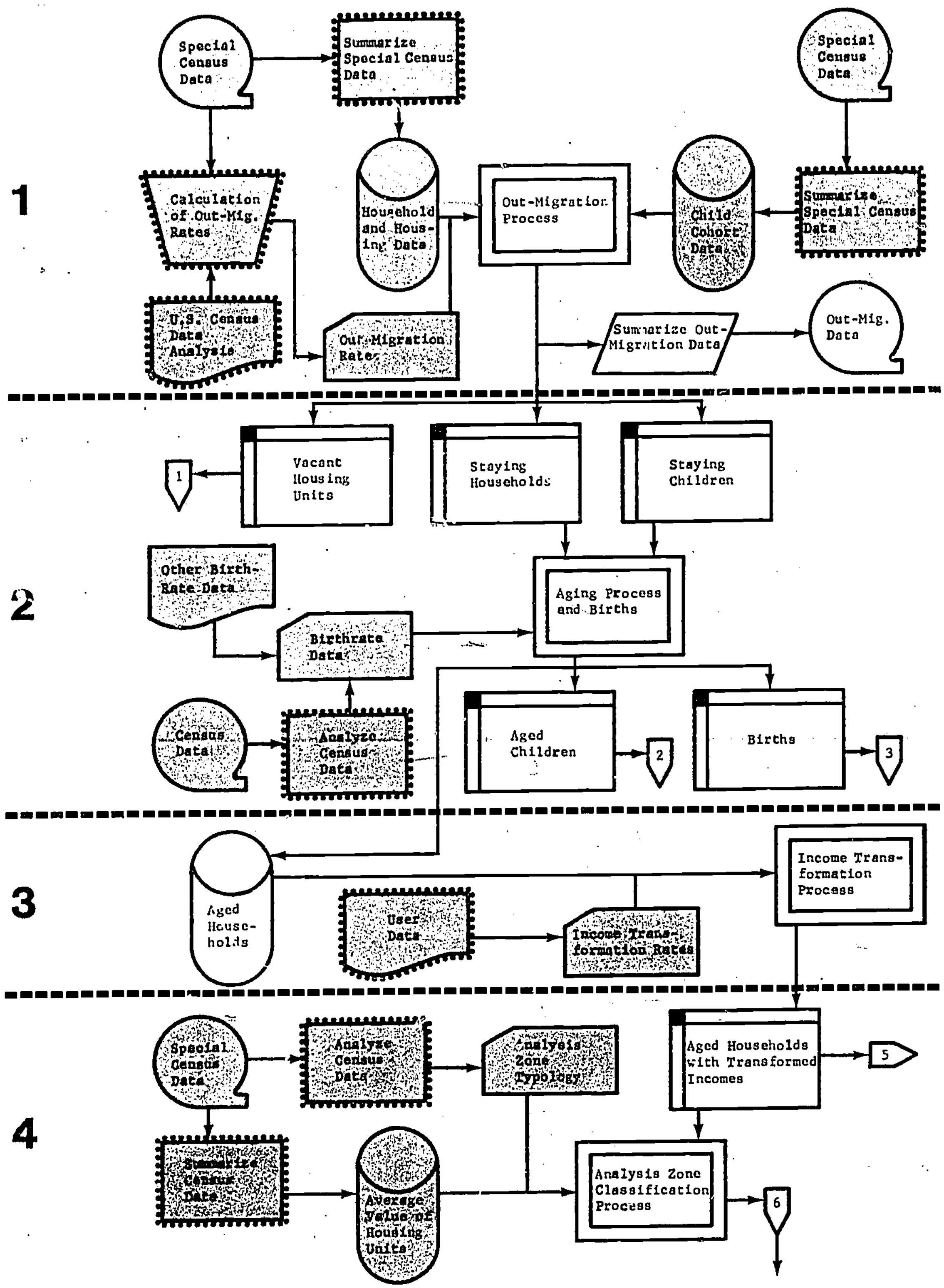


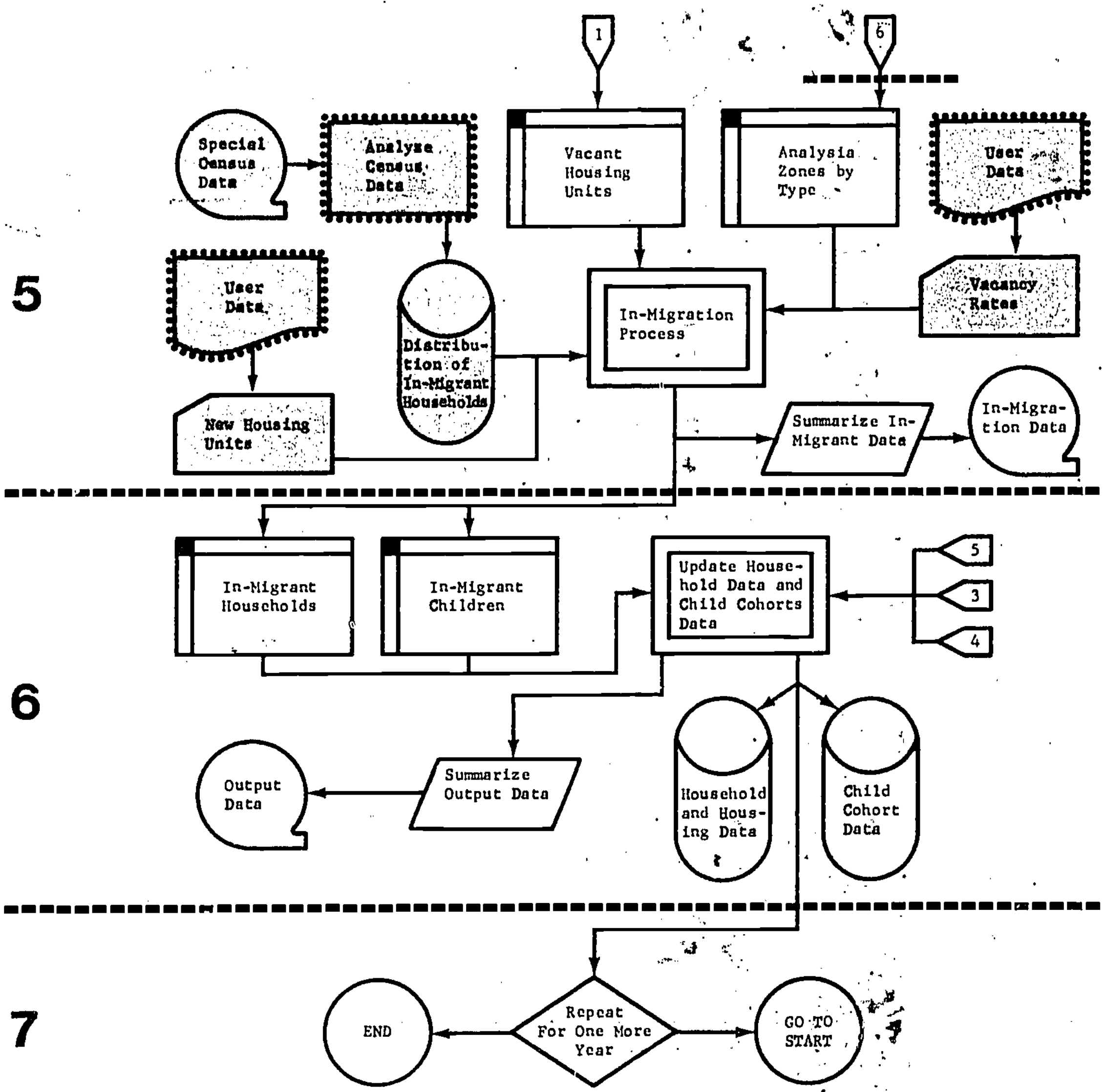
Figure 1

ENSIM II Logic Flow Chart



11

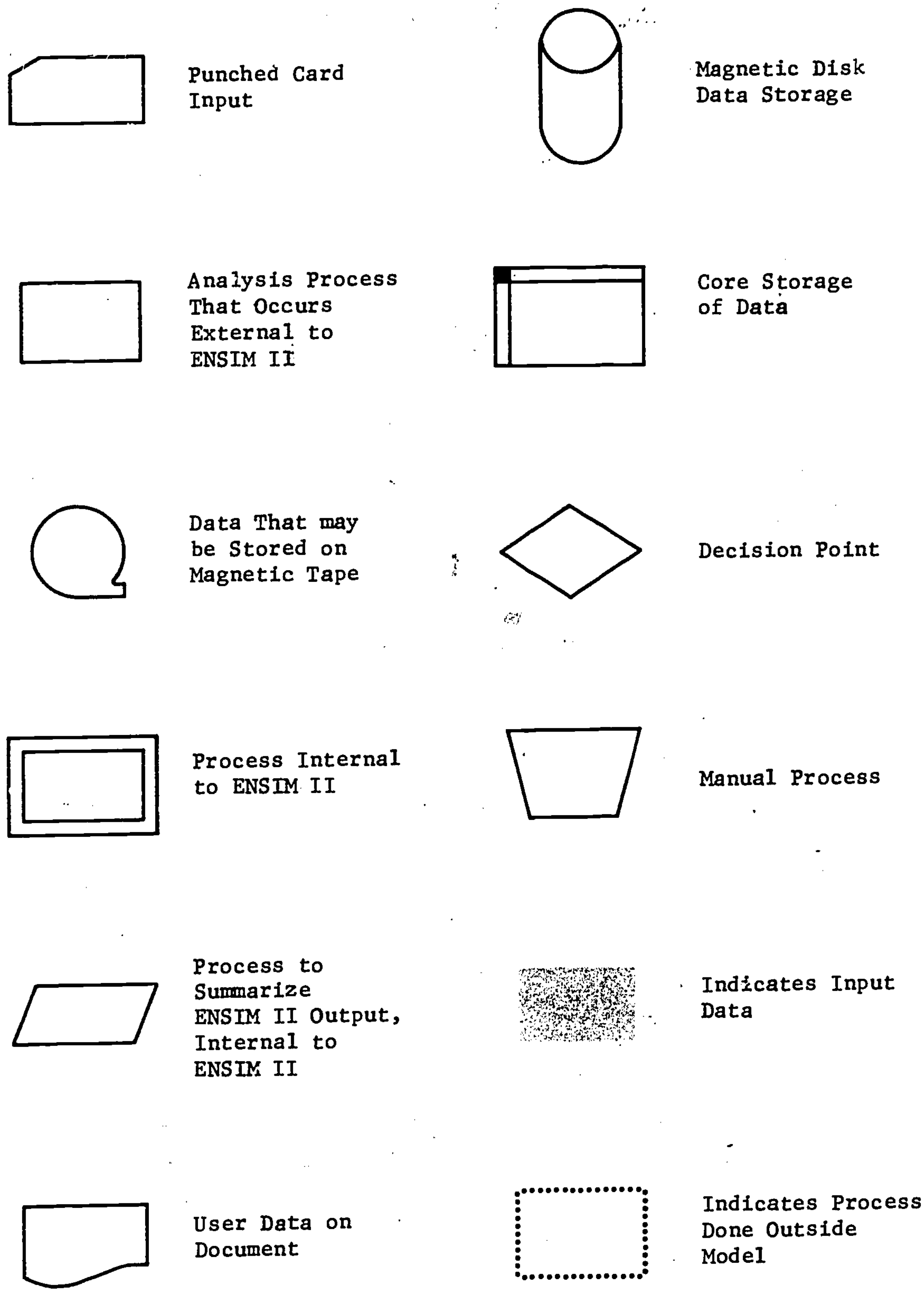
Figure 1 (cont.)



12

ENSIM II

Logic Flowchart Legend



As shown in Figure 2, Special Census Data (a) is summarized (b) into household and housing data (c). In this sub-process, households and housing units are geographically located by analysis zone and typed according to the characteristics below:

	Characteristic	# of Values	Suggested Values
HOUSING			
	kind of dwelling	2	single family multiple family

Thus, there are two housing unit types.

	Characteristic	# of Values	Suggested Values
HOUSEHOLDS			
	race/ethnic status	3	others Spanish-American Negro
	structure type of household residence	2	single family multiple family (owner/renter preferable if available)
	income	3	0-$10,000 $10,000-$15,000 $15,000 and up
	number of children members of household	4	0 1 2 3 or more
	age of female (spouse)	4	15-19 20-29 30-44 45 and up

Thus, there are 288 household types.

With the completion of this sub-process, each analysis zone in the district under study will have been identified as having had a given number of single and multiple family housing units as well as a given number of families whose characteristics placed them in one of the 288 household types at the time of the census. The product of this sub-process, Household and Housing Data (c), flows into the Out-Migration Process (j).

Next, length of residence data from the Special Census (a) and the U.S. Census (d) are used to calculate (e) out-migration rates (f) for each household type. Out-migration rates (f) flow into the Out-Migration Process (j).

In the next sub-process, Special Census Data (g) is summarized (h) into child cohort data (i). This child cohort data (i.e., the numbers of children by age, by race, who lived in each zone at the time of the special census, flows into the Out-Migration Process (j).

Outputs of the Out-Migration Process (j) are vacant housing units (l), staying households (m) and staying children (n). Also, out-migration data is summarized (k), and this out-migration data (o) becomes the first of the model's printed output. A sample of the out-migration data printout, produced at the end of a complete run, appears below:

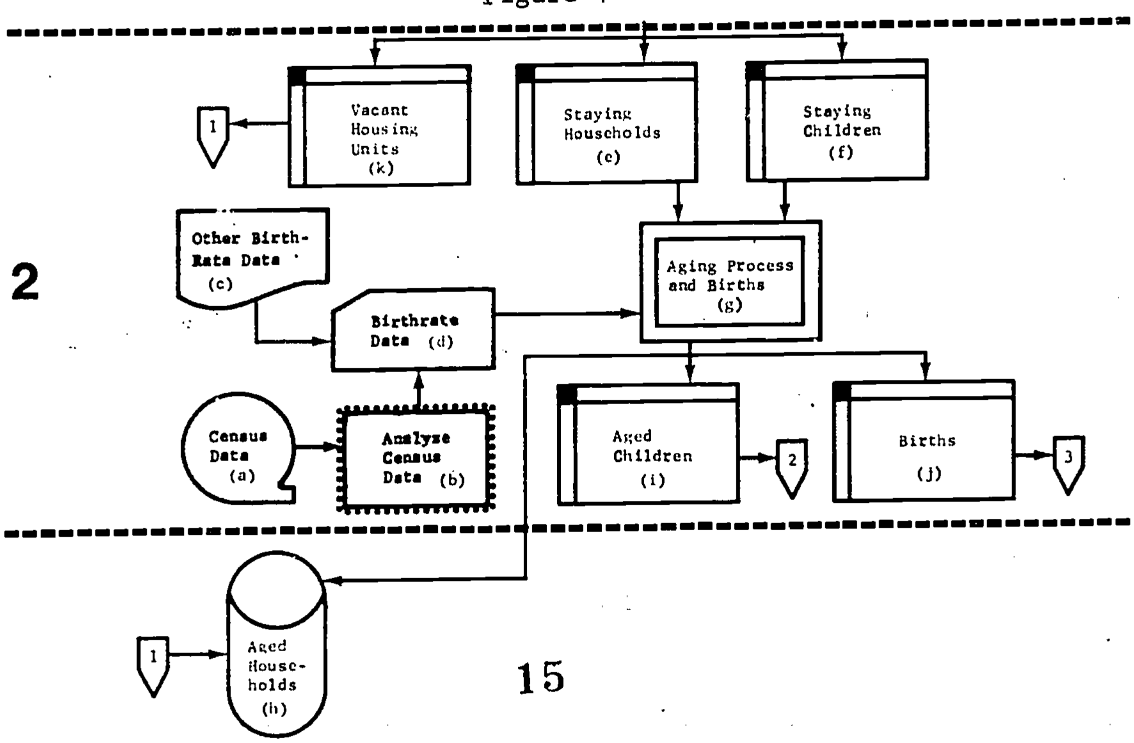
Figure 3

	Year 1972	1973	1974	1975
Out-Migration Others	0	1721	1756	1781
Out-Migration Spanish	0	2012	2106	2166
Out-Migration Negro	0	335	346	354
Out-Migration Total	0	4068	4208	4301
Out-Migration Ages 0-4	0	1194	1264	1314
Out-Migration Ages 5-8	0	1249	1301	1337
Out-Migration Ages 9-11	0	987	1002	1008
Out-Migration Ages 12-13	0	638	643	642

b. <u>Aging and Births Process</u>

The purpose of this process is to age certain members of the households (child-bearing aged females and children) one year for each year of the projection period and to forecast new births.

Figure 4



As shown in Figure 4, Special Census Data (a) is analyzed (b) and used, along with other birthrate data (c) such as that which can be obtained from local health departments, to produce final input birthrate data (d). Birthrates are developed for each of 48 female types. Characteristics and their suggested values used in female typology include:

Characteristics	# of Values	Suggested Values
race/ethnic	3	other Spanish-American Negro
parity (refers to number of previous children born into household	4	0 1 2 3 or more
age	4	15-19 20-29 30-44 45 and up

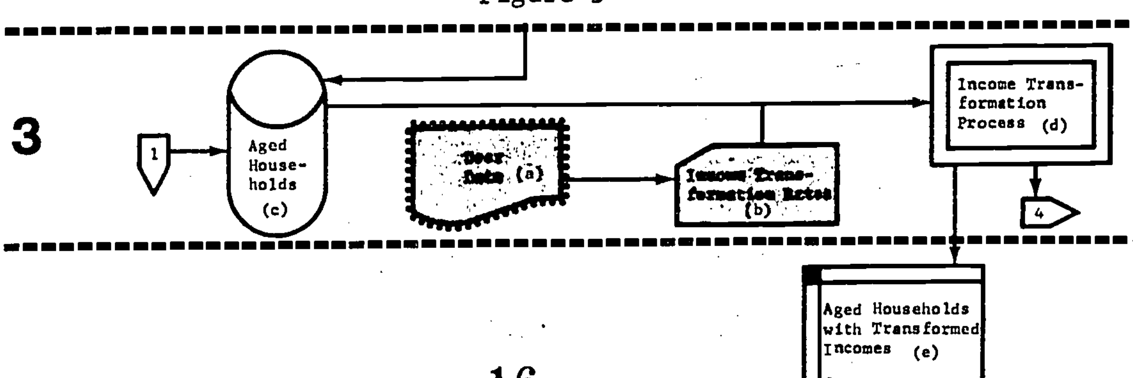
Thus, there are 48 classifications for females.

Next, birthrates (d) are applied to staying households (e) and births are projected. Staying children (f) also flow into the aging and birth process (g) and are aged with staying households one year. Thus, outputs of this process are aged households (h), aged children (i), and births (j). Note that aged children and births are set aside and not reintroduced until later in the model. Vacant housing (k), which was an output of the preceding process, is also stored for later use.

c. Income Transformation Process

The Income Transformation Process is used to simulate the movement of households from one income status to another over time. Households are moved on the basis of their income in the previous time period, their race/ethnic status, and the age of the household head's spouse (used as a proxy for the household head's age). The percentage of households moved depends upon factors input by the user of the model.

Figure 5



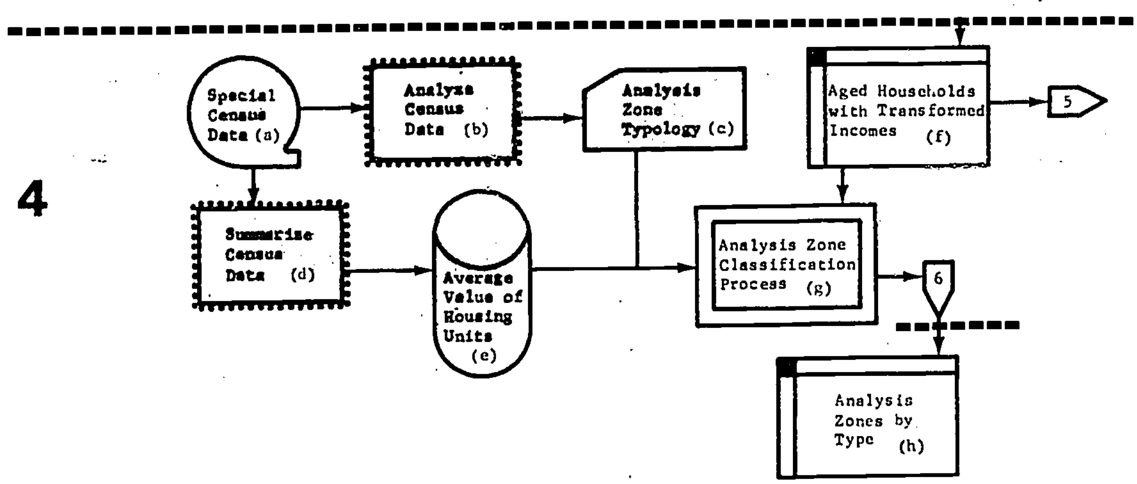
As shown in Figure 5, user data (a) is used to establish income transformation rates (b), and those rates, along with aged households (c), flow into the Income Transformation Process (d) to produce aged households with transformed incomes (e). The supply of vacant units is also passed on.

It should be noted that the Income Transformation Process is optional and is not required for completion of the simulation.

d. Analysis Zone Classification Process

In this process, all analysis zones are typed so that the characteristics of households moving into the zones can be predicted.

Figure 6



As shown in Figure 6, Special Census Data (a) is analyzed (b) to establish analysis zone typology (c). The characteristics and their values used to type analysis zones with existing residential development include:

Characteristics	# of Values	Suggested Values
average value of housing	3	low middle high
average number of children in households	4	1 or less 2 3 4 or more

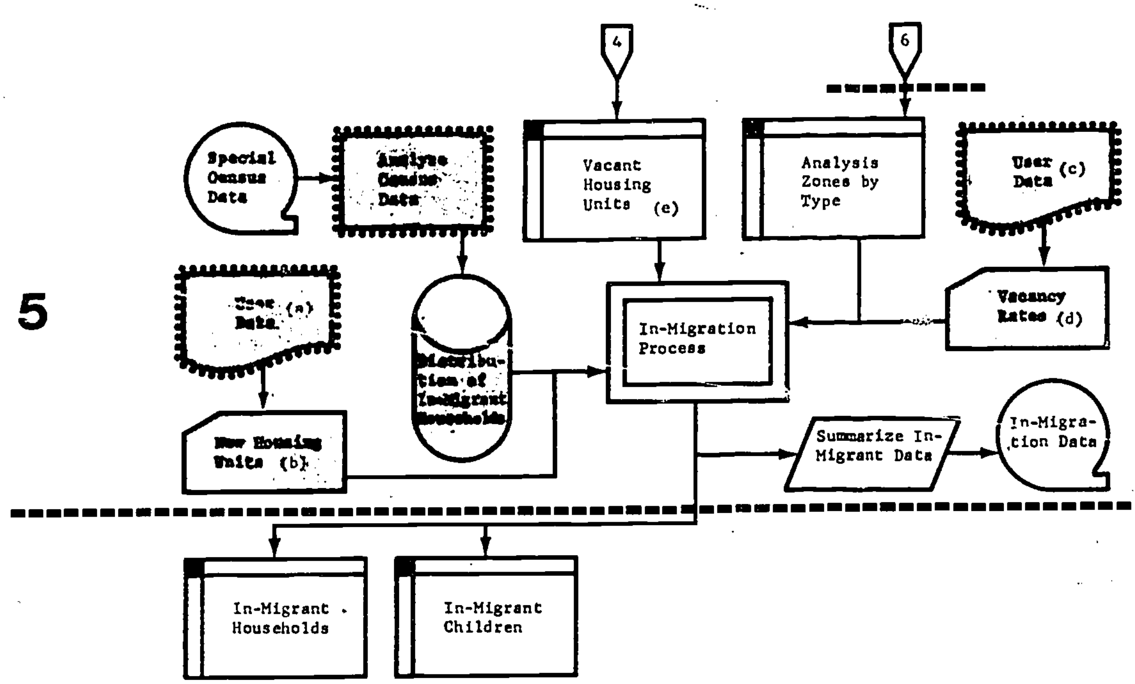
Three additional types, based on projected average housing values (high, middle and low) are used to classify analysis zones which have no existing residential development. Thus, there are 15 analysis zone types.

Once the analysis zone typology has been established, Special Census Data (a) is summarized (d) to determine the average value of housing (e) in each zone. Finally, aged households (f) are examined to determine the average number of children per household, by zone. This data flows with analysis zone typology (c) and average value of housing (e) into the Analysis Zone Classification Process (g) to determine analysis zone type (h). Thus, all analysis zones are classified into the 15 zone types according to average housing values and average numbers of children per household.

e. <u>In-Migration Process</u>

The purpose of the In-Migration Process is to predict the numbers and types of households, by analysis zone, that will be moving into the district under study. And, since a characteristic used in typing households is number of children, in-migrant children is also an output of this process.

Figure 7



18

The In-Migration Process can best be understood as two separate sub-processes: the first predicting the numbers of households to in-migrate into each analysis zone; and the second predicting the types of households to move into each analysis zone.

(1) Prediction of Numbers of In-Migrants

Numbers of in-migrating households are predicted on the basis of available housing. Thus, key inputs into this sub-process are:

--vacant housing units, or the number of housing units left vacant after the Out-Migration Process by analysis zone;

--new housing units, or the number of new dwellings predicted to become available in the projection year by analysis zone (new units are assumed filled); and

--vacancy rates, or the percentage of the total number of units that is predicted to remain vacant even after in-migration.

As shown in Figure 7, user data (a) is used to forecast new housing units (b), by analysis zone and by type, and the new housing unit data flows into the In-Migration Process. Next, user data (c) is used to establish vacancy rates (d), and the vacancy rates flow into the In-Migration Process. Finally, vacant housing units (e) flow directly into the In-Migration Process. Thus, the formula for predicting the numbers of in-migrating households is:

vacant housing units, by analysis zone	+	new housing units, by analysis zone	-	units vacant
	=	total numbers of available housing units, by analysis zone	=	total numbers of in-migrating households, by analysis zone

This calculation is done for single family units and again for multiple family units. Thus, the output of the sub-process is in-migrant households by single and multiple units.

(2) Prediction of In-Migrant Household Types

Special Census Data is analyzed and distributions of in-migrant households are established and flow into the In-Migration Process. Analysis zones by type also flow into the In-Migration Process.

19

The underlying assumption made in predicting the types of households that will in-migrate into a particular analysis zone is that:

> HOUSEHOLDS MOVING INTO A PARTICULAR AREA (ANALYSIS ZONE) WILL TEND TO SHARE THE CHARACTERISTICS OF HOUSEHOLDS ALREADY RESIDING IN THAT AREA.

Thus, key inputs into this sub-process are:

--analysis zone by type - all of the analysis zones classified into one of 15 types according to average house values and average numbers of children per household.

--distributions of in-migrant households - arrays of households that have recently in-migrated into the analysis zones are compiled. These arrays, one for each of the 15 analysis zone types, are then used as a probability vector for predicting the distribution of in-migrating households across household types.

Figure 8

Sample Probability Vector

HOUSEHOLD DISTRIBUTION VECTOR
FOR ANALYSIS ZONE TYPE "X"

Household type	1	2	3	4	5	6	...	140	141	142	143	144	
Single Family	6%	0	3%	.7%	4%	0	...	1%	2%	0	0	2%	= 100% of HH living in SF

Household type	1	2	3	4	5	6	...	140	141	142	143	144	
Multiple Family	0	5%	1%	.3%	.3%	0	...	3%	1%	6%	2%	0	= 100% of HH living in MF

As shown in Figure 8, the sample vectors array recently in-migrated households according to the characteristics used for household typology. Since the assumption is made that households moving into a particular area will tend to share the characteristics of households already residing there, each cell in the sample vector contains the percent of in-migrating households that can be expected to be a particular type. For example, in all analysis zones classified as "X", 6% of households moving into the zones' single family dwellings would be expected to have household type 1 characteristics. Similarly, it would be expected that 1% of the households moving into single family dwellings would have household type 140 characteristics.

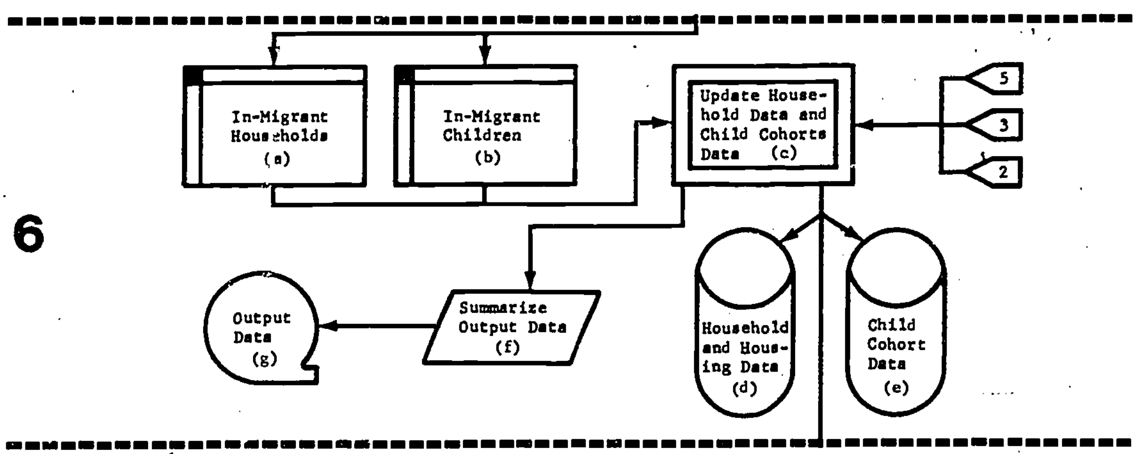
Once these percent distributions are established, they flow into the in-migration process and are applied to the projected numbers of in-migrating households to project in-migrant households and in-migrant children. Also, in-migration data is summarized and becomes the second of the model's printed output. A sample of the in-migration data printout appears below. As in the previous example, the printout would not be produced until all calculations were completed.

	1972	1973	1974	1975
In-Migration Others	0	1809	1865	1752
In-Migration Spanish	0	2292	2282	2311
In-Migration Negro	0	360	358	341
In-Migration Total	0	4461	4505	4404
In-Migration Ages 0-4	0	1633	1651	1615
In-Migration Ages 5-8	0	1390	1401	1375
In-Migration Ages 9-11	0	908	916	893
In-Migration Ages 12-13	0	530	535	520

f. Update Process

The purpose of the update process is to derive household and child cohort data.

Figure 9



6

As shown in Figure 9, in-migrant households (a) and in-migrant children (b) flow directly into the Update Process (c). Next, aged households with transformed incomes (5), and aged children (2) and births (3), both of which were output in the Aging and Births Process, flow into the Update Process (c). In this process, in-migrant households are added to existing households to produce households and housing data (d), and in-migrant children and births are added to aged children to produce child cohort data (e). Also in this process, output data is summarized (f) to produce output data (i.e., out-migration data) (g). A sample of the output data printout appears in Figure 10.

Figure 10

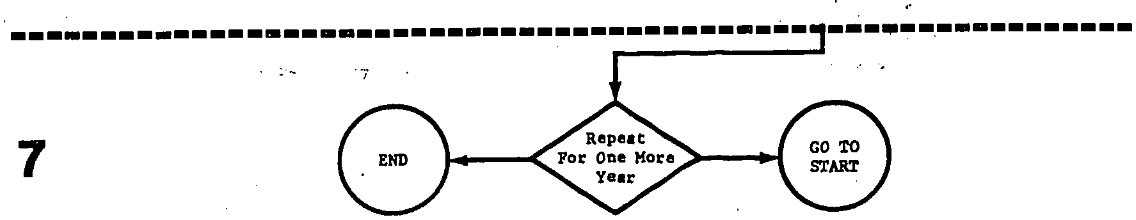
	1972	1973	1974	1975
Out-Migration Others	0	1720	1756	1781
Out-Migration Spanish	0	2013	2106	2166
Out-Migration Negro	0	335	346	354
Out-Migration Total	0	4068	4208	4301
Out-Migration Ages 0-4	0	1194	1264	1314
Out-Migration Ages 5-8	0	1249	3101	1337
Out-Migration Ages 9-11	0	987	1002	1008
Out-Migration Ages 12-13	0	638	643	642

g. Repeat/End Process

In this the final process, a determination is made to either repeat the computer run or to end it.

Figure 11

7



C. Choosing Basic Analysis Zones

The ENSIM II projection model has a two-level spatial dimension. Calculations are made for the entire school district and also for subareas of the district termed basic analysis zones. Analysis at this small geographical level allows enrollment forecasts to be made in aggregations useful for individual attendance area planning, and is thus a key feature of the ENSIM II model which sets it apart from more traditional forecasting techniques. The purpose of this section is to outline the criteria for selecting basic analysis zones.

1. Homogeniety

Analysis zones should partition the school district into zones with relatively homogeneous housing stock and population characteristics. The more homogeneous the zones, the more accurate the generalizations made about them in later steps of the analysis.

2. Size and Number of Zones

The ENSIM II program was written with an arbitrary limit of 287 zones. Thus, if a greater number of zones are desired, the program's parameters will need to be revised. It should be noted that the cost of the ENSIM II analysis is in proportion to the number of analysis zones used (i.e., the fewer the zones, the less the cost).

In general, zones should contain enough housing units to characterize the development trends in a given area. One house, for example, could not accurately suggest a zone's future residential character. However, ten houses, in an area with the potential for 20 housing units, could be seen as an indication of the zone's future character. Finally, zones must be drawn at a level that allows the aggregation of census data. In Figure 12, analysis zones were established at a level generally consistent with census blocks. Naturally, the availability and nature of census information will dictate the minimum analysis zone size in a given situation.

3. Zone Boundaries

School attendance area boundaries should be considered wherever possible in drawing basic analysis zone boundaries. If the district under study has firm plans for changing boundaries in the near term, these lines should also be considered. Attendance boundaries can be very important in establishing zones because they provide a point of reference for comparing current and projected data which is familiar and useful to district planners. In effect, attendance areas are split into smaller zones, providing a starting point for planning future staffing and altering school boundaries for the purpose of balancing enrollment levels and racial compositions.

Within the attendance area, zones should conform to other useful physical or political boundaries. Streets are the most obvious and desirable,

Figure 12


GADS ZONE BOUNDARY
GADS ZONE NUMBER
CAPITOL
RD
JACKSON
KING
McKEE
ROUTE
SAN ANTONIO
STORY
RD
WHITE
HOCK
ALUM ROCK AVE.
TULLY
Reid-Hillview
Airport
Penitencia
Creek
Rd

24

since they make mapping and data collection much simpler. If streets cannot be followed, property lines may be used since the analysis of individual housing units, and ultimately of households, is a major objective of the ENSIM II Model. Prominent topographic features such as transmission lines, may be used where settlement is very sparse and zones are large.

Political boundaries such as city limits may be useful in instances in which the presence of two separate political jurisdictions will have an effect on the rate of growth, social composition, tax base, or other issues of importance to the district. This will almost always be the case with districts covering more than one local governmental unit. In instances in which it is expected that a city's boundaries will change frequently, it may be useful to make some assumptions about where the city is expected to grow when drawing basic analysis zones. If official "spheres of influence" have been established among two or more cities which define the area within which annexation will be allowed for each jurisdiction, the creation of study boundaries is greatly simplified.

In summary, analysis zones should be homogeneous, they should be drawn to conform with all other boundaries which will aid in the analysis of future growth or which can be used to examine the impact of growth on the district under study, and they should be approximately the same size, except in areas with little or no development.

II. DATA PREPARATION

This chapter provides a general guide for the preparation of data necessary for the ENSIM II Model. In actual use of the model, the user will undoubtedly find it necessary to make some modifications based on data availability, computer hardware and unique district features.

A. Land Use Data

The purpose of this section is to determine the data elements, data sources and analysis techniques required in the residential development projection phase of the ENSIM II Model.

1. Inventory of Vacant Land

The first step in the land use analysis is the identification of all vacant land within the district under study. It is important that vacant land be located on a good quality, convenient scale map, and it is essential that the map include property lines. Generally speaking, a 1,000 scale map (1" = 1,000') is the smallest size on which small urban lots can be accurately measured for land area. The map should show all existing and proposed roads, railroads, water courses and, preferably, utility lines. Topographic features are important to the analysis, but may need to be obtained on another map. Suitable maps can be found at public works and planning departments, the assessor's office, or perhaps the parks and recreation department.

The first place to turn for information on vacant land is to the public works and planning departments of the local jurisdiction. These agencies often have land use inventories which identify vacant land and which, once updated with field checks or aerial photos, will greatly simplify this step in the analysis. Should such inventories not be available, determination of vacant parcels can be made using aerial photos alone. Interpretation of aerial photographs is an accomplished skill and should be done by competent personnel. Sources for the photographs include public works and planning departments, utilities companies, and in some instances, the U.S. Geological Survey. If aerials are not available, or are too out of date to provide accurate data on vacant land, a district may choose, to have its own aerial photography done. This would only be feasible, however, if the district under study was extremely large.

Barring the availability of vacant land inventories or aerial photographs, a field survey, done by driving or walking the entire district, must be conducted. Such an effort is made more manageable by enlisting the aid of each school in inventoring the vacant land in its own attendance area by using a clear and sufficiently detailed map and by briefing all participants to ensure accuracy and conformity to a common format.

Once vacant land has been located, the size of each parcel must be calculated using the dimensions of the land and the appropriate mathematical formula, provided the parcel is regularly shaped, or a planemeter if it is irregular. (A planemeter is an instrument which calculates area by registering the number of square units traveled by a small pointer over the boundary lines on a map.) Vacant land should be tabulated by analysis zone which in turn requires that the zone's boundaries be superimposed on the vacant land map when calculating area.

The sizes of large parcels of land which will not be developed for many years or which will not be developed at all can be approximated. This will ensure that all vacant land in the district under study will be accounted for.

2. Development Policy

In the next step of the analysis, the policies regulating the use of vacant land in the district are identified. Plans developed by appropriate jurisdictions must be ascertained and their applicability to vacant parcels determined.

The most common starting point for this effort is at the municipal general plan. This document defines the land uses which are desired for the community. The general plan may also include a timetable for development as well as a capital improvement plan for the service infrastructure required to support growth. With this information, a determination of how vacant land is planned to develop within the municipality can be made.

It is essential to check the general plan against the zoning ordinances of the municipality. Zoning ordinances define the allowable uses of land at present, whereas the general plan defines land uses which are desired for the future. The two land use policies may or may not be the same. (In California, zoning and general plans must agree.) If zoning is considerably different than the general plan, particularly in regard to residential land uses, it must be given equal weight in short-term projections.

In addition to the municipal policies for specific pieces of land, the more general policies, which may apply to large sectors of the city, must be considered. For example, a sector of the community may be placed in reserve status where development will be prohibited for a given number of years.

The development policy toward each vacant parcel of land designated for residential use must be identified. Policies will translate into allowable densities, as expressed in housing units per acre. A suitable percentage of residentially designated vacant land must be subtracted for streets and utilities. (This is normally 20-25%.) The remaining vacant land (in acres) is multiplied by the density (in dwelling units per acre) to arrive at the capacity of the land in dwelling units. In this way, the residential development potential of the district is calculated and made ready for tabulation by analysis zone.

3. Development Issues

Once all major policies have been identified, an analysis of key development issues must be accomplished. Development issues are generally political questions whose resolution will significantly influence the course of growth, or physical constraints or stimulants which will affect development. A few examples should serve to clarify the meaning of development issues here.

--Approval of major new commercial center

A major new shopping center, consisting of a large department store, specialty shops, restaurants, a theater and professional offices, is proposed for a site just outside of the school district. It is expected to generate other commercial ventures in the area and will attract considerable residential growth, especially apartments. Several environmentalist and homeowners groups have expressed opposition to the proposal, and early indications are that it may face stiff opposition in the Planning Commission.

--Lack of sewerage lines

The presence or lack of many municipal services can greatly influence the course of development. The lack of sewerage lines, and consequently, of access to a centralized sewage treatment facility, is a very common service deficiency and one that may prohibit the residential development of vacant land. If no funds are available, the jurisdiction having control could choose to disallow further development. More often, however, densities are altered and septic tanks are used for sewage treatment. This invariably necessitates larger lots to accommodate leach lines and perhaps a water well, since lack of municipal water often accompanies lack of sewerage lines. The larger the lot size, the lower the density, the fewer the new homes, and the more scattered the growth pattern. All are important considerations for the ENSIM II analysis.

--Transportation networks

Improved access through the completion of a new freeway or mass transit system will have an important effect on both the timing and extent of residential development. All information about such projects should be obtained, and contact should be maintained with the agencies which are responsible for its implementation.

Other factors which could have significance for the land use analysis include flood plain areas, topographic features, open space proposals, airport noise zones, redevelopment projects, agricultural preservation measures and water supply projects. The list of possibilities is endless, but the common theme is the same--impact which affects the future course of residential development and ultimately the future enrollments of the school district under study.

4. Growth Rate

At the heart of the ENSIM II projections is the forecast of annual housing development. The assemblage of development policies and issues described in the two previous sections should provide a good indication of the relative magnitude of future growth. Often, specific projections of future growth are made available by local or regional governmental agencies, and these should be utilized when possible. However, even when such projections are in existence, they often do not coincide with the geographic area of the district; too, such projections quite often prove to be inaccurate. In any event, growth rates must be selected for their reasonableness given what is known about the development potential of the district.

In order to make that reasonable selection, a growth history, going back at least five years, should be assembled. Ideally, the information gathered would include the actual number of houses and apartments constructed as well as the number of mobile homes installed. This information is normally available at the jurisdiction having control over building in the district. Should data on housing units constructed not be available, the number of building permits issued, or the number of applications for residential development submitted, may be useful.

Once the data is collected, it should be translated into an average annual growth rate, in housing units. The standard compound interest formula, shown below, can be used to calculate average annual growth.

$$i = \left(\frac{P2}{P1}\right)^{1/n} - 1$$

where i = average annual growth (expressed as fraction; i.e., 0.06=6%)
P2 = number of housing units at end of period
P1 = number of housing units at beginning of period
N = number of years between P2 and P1.

Another source of projection data might be the backlog of approved residential developments kept by many governmental agencies. This source frequently provides the best indication of growth expected in the short term. In fact, it is recommended that greatest emphasis be placed on developments already approved in establishing short-term growth rates (construction may lag behind approval by months or even years).

It is important that more than one growth rate be established so that a range of development possibilities is anticipated. We see at the national level a number of rapid economic and political changes that can drastically affect housing construction at the local level. So too, local issues can often result in high or low growth rates depending on their resolution. A high, medium and low growth projection provides the kind of flexibility and utility necessary for responding to the ever-changing economic climate.

The growth rates, expressed in annual percentage growth and in dwelling units, should be applied against the district's current number of housing units, and the increase, in units, added on each year. The table which follows represents a simplified example of this process.

CURRENT HOUSING: 5000 UNITS

Yearly Growth Rate	New Units 1st Yr.	Total Units After 1st Yr.	New Units 2nd Yr.	Total Units After 2nd Yr.	New Units 3rd Yr.	Total Units After 3rd Yr.
High Growth Rate - 8%	400	5400	430	5830	465	6295
Medium Growth Rate - 6%	300	5300	320	5620	340	5960
Low Growth Rate - 3%	150	5150	154	5304	160	5464

Two further refinements to the growth rate process are possible. First, different rates can be used for single and multiple family units; and second, the yearly growth rate may be varied over time. The latter is particularly useful since, as can be seen in the table above, the same growth rate will produce an increasing number of units each year. A more practical approach might be to reduce the rate of growth in later years so that the number of new units each year does not become unrealistically high.

5. Distribution of Growth

The previous step of the analysis established the magnitude of district-wide growth. Below, two techniques for distributing that growth, in housing units, through the district will be described. The first technique provides a basis for the second.

The object here is to develop a logical and rational basis for determining how much growth an area will receive. The "area" in this context is the individual school attendance area, of which there will be several throughout the district under study. (Later, a quick method for apportioning growth to zones within the attendance areas will be discussed.) The rationale for distributing growth is that an area will attract development in proportion to: (a) the amount of vacant land available for residential construction; (b) the total number of housing units that could be built on that land; and (c) the size of vacant parcels. The size of the parcels is important because one large parcel is more likely to attract development than several small ones.

In the second step of the land use analysis, the amount of vacant land available was tabulated by attendance area. In the third step, the number of dwelling units that could be built on that land, as defined by the general plan or zoning ordinances, was determined. In addition to these data collections, it is necessary to tabulate the total areas of vacant land by size of parcel. The sizes most appropriate for this tabulation depend on the range of parcel sizes that exist in the district. If, for example, there are many parcels of more than 20 acres, and a significant number at sizes ranging downward, the total acres of vacant land in parcels exceeding 20 acres would be most appropriate for the tabulation. If, however, no vacant parcel exceeds 15 acres, and many are under 10 acres, a different range of sizes would be tabulated. The idea here is to rank each attendance area according to the amount of land it contains in large parcels.

For each item of data--total vacant land, number of dwelling units and vacant parcel size--the attendance area's share of the district total is required. For example, if there are 10,000 acres of vacant land in the district and an attendance area has 280 acres of vacant land, the area's share of this data item would be 28%. So too with dwelling units and vacant land. Another example: a district has a large amount of vacant land, say 550 acres, in parcels of greater than 15 acres. An attendance area has 175 acres of land in parcels greater than 15 acres. Thus, the attendance area would have 32% under this data item. The example below provides an example of all the data items arrayed in tabular form.

Attendance Area	Vacant Land: Total Vacant	Vacant Land: % of Dist. Total	Dwelling Unit Capacity: Total Units	Dwelling Unit Capacity: % of Dist. Total	Acreage in Parcels of 15 Acres or More: Total Acres	Acreage in Parcels of 15 Acres or More: % of District Total
A	500	50%	2,000	40%	400	56%
B	200	20%	1,000	20%	100	14%
C	175	17.5%	500	10%	170	24%
D	125	12.5%	1,500	30%	50	6%
Total	1,000	100%	5,000	100%	720	100%

To distribute growth over attendance areas, a composite of the three data items is calculated for each area by adding the percentage value in each category and dividing by three. Thus, in the above example, attendance area A has 50% of all vacant land, 40% of the total possible dwelling units which could be constructed on this land, and 50% of all land greater than 15 acres. Adding the percentage values and dividing by three results in 47%. Therefore, attendance area A would be allocated 47% of the district's total growth in any one year.

Although this technique for apportioning growth offers a clearly defined and easily duplicated methodology, development in real life is not distributed by a mathematical formula. Rather, growth happens where it happens and when it happens for many and varied reasons. Therefore, the allocation of growth based on the method described above must be modified on the basis of actual residential growth history in the district under study. So then, for example, even though area D in the table above seems to have fewer factors to attract growth, it may be better located to attract development, and consequently, should be allocated a greater percentage of construction activity than would be indicated by the mathematical formula. Modifying growth in area D to better reflect real life would, of course, require lowering the growth in one or more of the other areas. This two-stepped process for apportioning growth districtwide is recommended because it offers an objective and easily defined point of departure for the more subjective approach which considers actual conditions.

The final task in this stage of the land use analysis is to distribute residential growth through zones within each attendance area. The technique described above should be followed, with one exception: total vacant land should be used as a criteria for allocation, with suitable modification based on any subjective data available (i.e., specific information about housing developments, etc.). This task should be addressed with a less formal, shorthand approach, since the accuracy of allocation at this level is not high. This is partly due to the fact that there are too many local factors which can intervene and upset the apportionments. As with all projection efforts, the larger the area projected for, the greater the accuracy of the results.

6. Synthesis of Projection Data

To clarify the steps of the land use analysis discussed in this section and to briefly address how data produced by the analysis may be applied to a continuous planning process, a short summary is presented below.

Inventory of vacant land - All vacant land is located and depicted on a map of the district. The map indicates where new development can occur. As development proceeds after the initial inventory has been completed, future projection efforts can update the inventory simply by eliminating already developed land from possible future development.

Development policy - The amount of vacant land that could be developed for residential uses is determined by defining the municipality's land use policies as they are reflected in the general plan and zoning ordinances. This determination allows the calculation of the number and type (single or multiple family) of housing units which are possible in the district at a given point in time by multiplying the acres of vacant land by the allowed housing units per acre. The notion of "at a given point in time" is important here because general plan and zoning policies can and do change. It is for this reason that all vacant land must be located in this step of the analysis, regardless of whether or not its current zoning allows residential development.

Development issues - Those important factors which will influence the rate and location of residential growth in the district are identified in this step. These can be political issues, such as a controversy over a growth-stimulating industrial development, or a parks and open space proposal which could constrain housing construction. Other factors might be physical constraints on development, such as steep topography, flood hazards, lack of sewers and water, etc. All these factors, even those of a political nature, may cause the modification of the development policies identified earlier.

The identification of development issues provides a means of understanding the dynamics of growth in the district through careful and ongoing analysis of the forces which stimulate and constrain development.

Growth rate - A range of high, medium and low growth rates is established for the district by examining the projection rates established by other agencies, if available, and by analyzing past and recent growth trends. These rates may relate directly or indirectly to the development issues identified earlier. The rates may thus vary over time, as issues are resolved and as new ones emerge. The application of yearly growth rates to the current number of housing units yields the yearly increase in new units. The rate is applied until all vacant land is exhausted.

Distribution of growth - Growth is distributed through attendance areas (or through any grouping of zones desired) on the basis of each area's share of vacant land, its potential for new housing, and the acres of vacant land in large parcels. Subjective data regarding actual growth in the district provides a means for altering distributions toward a more realistic outlook should the need arise. When an attendance area has exhausted all its vacant land, it is dropped from consideration in future growth distributions. Distributing new housing through analysis zones can be a less than formal process since too many variables can intervene at such a small level of detail. A distribution based on available vacant land and subjectively modified by knowledge of specific development activity is a workable approach.

B. Household Related Data

The purpose of this section is to describe the requirements, sources and preparation of household related data for the ENSIM II Model.

1. Special Census Data Requirements

ENSIM II utilizes special census data as its primary data input source. The special census data must contain records of individual households within the user district. The records must contain the following information for each household:

a. Household race/ethnic status;

b. Household population;

c. Type of dwelling unit;

d. Age, sex and school status of each member of the household;

e. Length of residence at the household's current address; and

f. A geographical code relating to the location (by analysis zone) of the household's residence.

Should this information be unavailable or incomplete, the user must estimate values for the missing data.

2. Household Data File

The ENSIM II Model requires that households be indexed by their characteristics. Each characteristic is assigned one of several alternative values. Figure 13 lists the househ ld characteristics used in the model and suggests possible values to be ed for each. It should be noted that the user is restricted to the alues listed in Figure 13 in assigning the number of children characte istic and the age of household head's spouse characteristic. For households with no female present (i.e., households with unmarried male heads), age of the household head is substituted for age of the female (spouse).

Since each household is assigned one value for each of the characteristics, the total number of household types is 288. These 288 household types are arrayed in Figure 14. The total number of households in each cell of the array is tabulated for each of the analysis zones within the user district.

For the purpose of computer processing, the tabulated households are stored in a vector which decomposes the array shown in Figure 14 into a linear record 288 units long. The values in each cell of the array indicate the order in which they appear in the household vector. The cells are first divided on the basis of race/ethnic status: the first 96 designate "others;" the second 96, Spanish Americans; and the last 96, Negroes. Each race/ethnic grouping of 96 cells is then divided into

Figure 14

288 HOUSEHOLD CLASSIFICATIONS*

Income	Age	Race/Ethnic	Others								Spanish American								Negro							
		Housing Type	Single Family				Multiple Family				Single Family				Multiple Family				Single Family				Multiple Family			
		Size	0	1	2	3+	0	1	2	3+	0	1	2	3+	0	1	2	3+	0	1	2	3+	0	1	2	3+
0-10K	15-19		1	5	9	13	49	53	57	61	97	101	105	109	145	149	153	157	193	197	201	205	241	245	249	253
	20-29		2	•	•	•	50	•	•	•	98	•	•	•	146	•	•	•	194	•	•	•	242	•	•	•
	30-44		3	•	•	•	51	•	•	•	99	•	•	•	147	•	•	•	195	•	•	•	243	•	•	•
	45-up		4	8	12	16	52	56	60	64	100	104	108	112	148	152	156	160	196	200	204	208	244	248	252	256
10-15K	15-19		17	21	25	29	65	69	73	77	113	117	121	125	161	165	159	173	209	213	217	221	257	261	265	269
	20-29		18	•	•	•	66	•	•	•	114	•	•	•	162	•	•	•	210	•	•	•	258	•	•	•
	30-44		19	•	•	•	67	•	•	•	115	•	•	•	163	•	•	•	211	•	•	•	259	•	•	•
	45-up		20	24	28	32	68	72	76	80	116	120	124	128	164	168	172	176	212	216	220	224	260	264	268	272
15K-up	15-19		33	37	41	45	81	85	89	93	129	133	137	141	177	181	185	189	225	229	233	237	273	277	281	285
	20-29		34	•	•	•	82	•	•	•	130	•	•	•	178	•	•	•	226	•	•	•	274	•	•	•
	30-44		35	•	•	•	83	•	•	•	131	•	•	•	179	•	•	•	227	•	•	•	275	•	•	•
	45-up		36	40	44	49	84	88	92	96	132	136	140	144	180	184	188	192	228	232	236	240	276	280	284	288

* Size relates to the number of children members of the household. The symbol K used in the income characteristic stands for $1,000. Income relates to total household income. Values in cells relate to order of appearance in household vector.

two groups of 48, this, on the basis of structure type of the housing unit occupied by households. The final divisions are by income, number of children and age of female, respectively. The program which compiles the vectors for each zone is essentially a cross-tabulation program classifying the households living in each analysis zone.

Figure 13

HOUSEHOLD CHARACTERISTICS AND THEIR VALUES

	Characteristic	# of Values	Suggested Values
HOUSING:			
	kind of dwelling	2	single family multiple family
HOUSEHOLDS:			
	race/ethnic status	3	Others Spanish-American Negro
	structure type of household residence	2	single family multiple family
	income	3	0-$10,000 $10,000-$15,000 $15,000 and up
	number of children, members of household	4	0 1 2 3 or more
	age of female (spouse)	4	15-19 20-29 30-44 45 and up

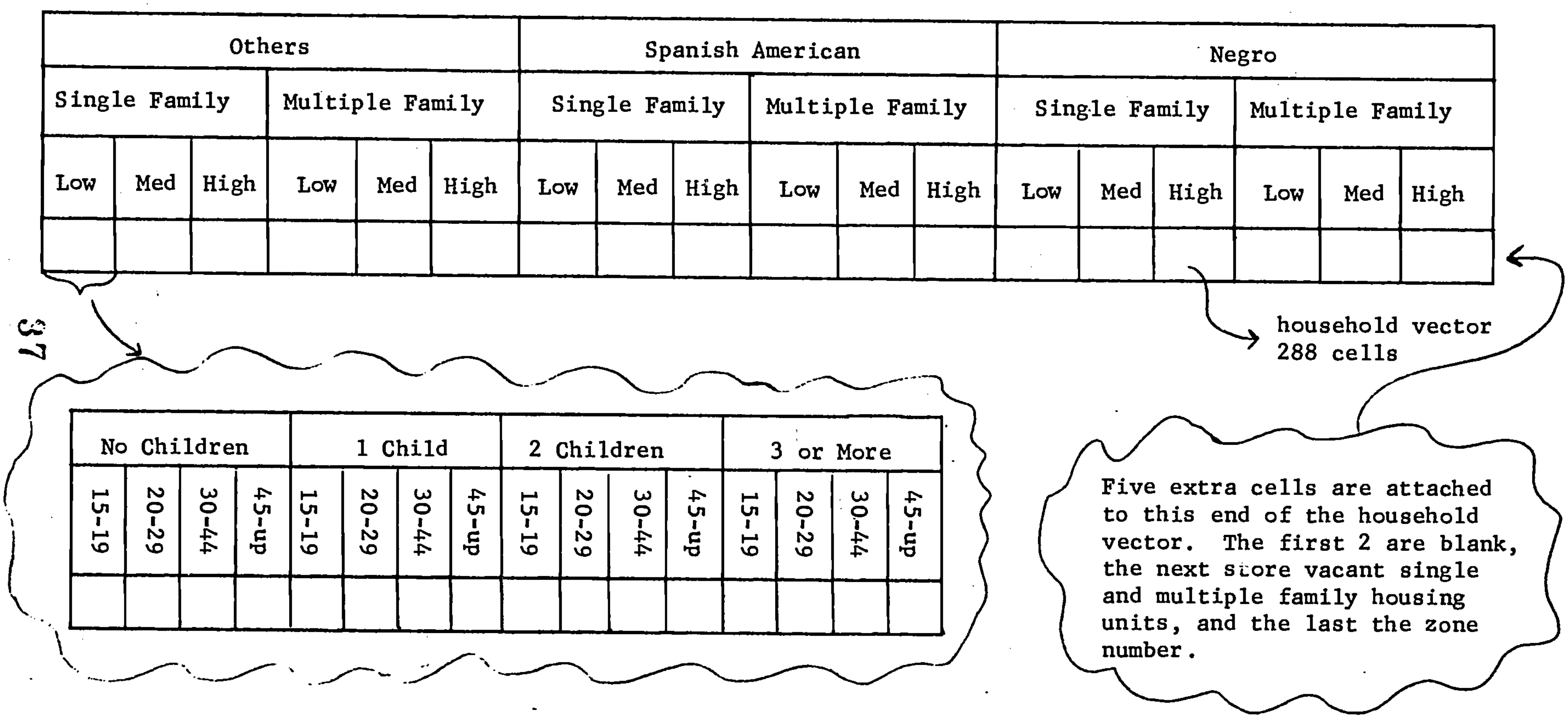
Figure 15 offers a sample layout of the household vector. The reader will note that five more cells are attached to the end of the vector. The first two are blank and are included for working space. The next two cells store the count of vacant single and multiple family housing units, respectively. The final cell stores the numerical designation of the geographical area (i.e., analysis zone) to which the data relate.

3. Average Number of Children per Household Vector

This vector is used to establish the average number of children in each type of household. The average number of children per household vector is compiled once for the entire school district. The reader will recall that one of the characteristics of households relates to the number of children members of the household. Thus, for most households, once the characteristics of the household are known, the number of children members

Figure 15

LAYOUT OF HOUSEHOLD INPUT VECTOR*



* Low, Med, and High refers to income classifications. Lower diagram illustrated in cloud is detailed layout of first 16 cells in vector. This layout repeats itself for each successive 16 cells.

in that household has also been established. This is not true for households with three or more children, however, and it is for the purpose of determining the average number of children in these households that this vector is employed.

The layout of the Average Number of Children per Household Vector is exactly the same as that of the Household Vector depicted in Figure 15. Each cell in the vector stores a number which relates the average number of children that are members of households stored in corresponding cells of the household vector. Figure 16 illustrates the layout of the Average Number of Children per Household Vector.

Data for this vector is compiled from the special census data. First, the total number of children members of each household type is found. These numbers are then divided by the total number of households in each respective type. As shown in Figure 16, the results for households with 0, 1 or 2 children are trivial.

4. Distributions of Child Age

As has been explained, the household vector supplies the ENSIM II Model with data on the number and types of households living in each analysis zone within the user district. The Average Number of Children per Household Vector supplies information about the number of children members of particular household types. The child age distributions described in this section, supplies data regarding the distribution of the ages of children in particular household types. With these three pieces of data, the number of children, by age, who are members of a diverse collection of household types can be estimated.

The ENSIM II Model utilizes 144 child age distributions, each of which is a probability vector that sums to a value of one. Since the model stores data on children age 0 (less than 1) through 19, each distribution is 20 units long. Each cell of a distribution stores a value relating to the probability that a child is of a particular age. Each distribution relates the probabilities that children from a particular type of household are of particular ages.

Figure 17 illustrates the layout of a single distribution. Various household characteristics influence the probabilities that children in a given household will be of a certain age; households with a spouse in the oldest age status present, for example, would be expected to have older children. A large number of children present in a household would tend to indicate a fairly even distribution of ages should be expected. ENSIM II allows all household characteristics, except structure type, to influence the distribution of children's ages. Thus, 144 separate distributions are compiled for the operation of the model. These distributions are stored as a vector 2880 cells long (20 x 144 = 2880). Figure 18 depicts the order of distributions in this vector.

The child age distribution vectors are compiled from the special census data. The number of children, by age, for each of the 144 combinations of household characteristics is first compiled. Next, the total number of children who are members of households in each of the 144 combinations

Figure 16

SAMPLE OF AVERAGE NUMBER OF CHILDREN PER HOUSEHOLD VECTOR*

Others						Spanish American						Negro					
Single Family			Multiple Family			Single Family			Multiple Family			Single Family			Multiple Family		
Low	Med	High	Low	Med	High	Low	Med	High	Low	Med	High	Low	Med	High	Low	Med	High

vector 288 cells

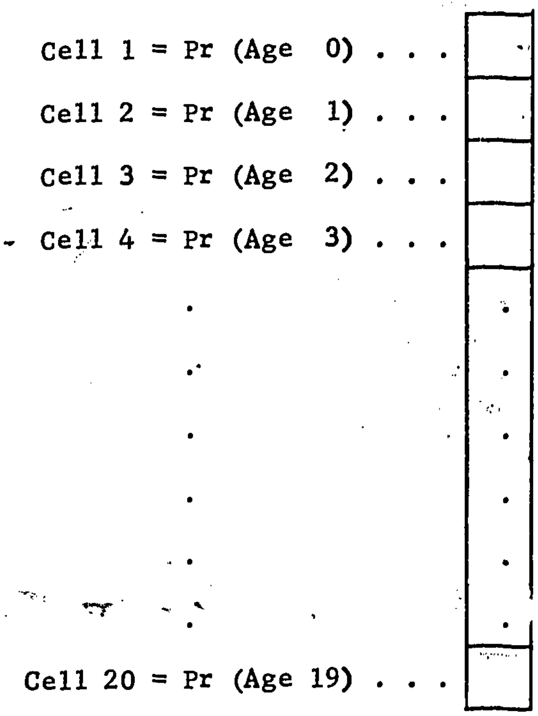
No Children				1 Child				2 Children				3 or More			
15-19	20-29	30-44	45-up	15-19	20-29	30-44	45-up	15-19	20-29	30-44	45-up	15-19	20-29	30-44	45-up
0	0	0	0	1	1	1	1	2	2	2	2	3+	3+	3+	3+

Numbers greater than 3 may have fractional parts (i.e., 3.25).

* Lower diagram illustrated in cloud displays values in cells in portion of vector. The exact number of children that are members of households with 0, 1, or 2 children is known. The 3+ numbers for households with 3 or more children indicate numbers greater than 3 (i.e., 3.25).

Figure 17

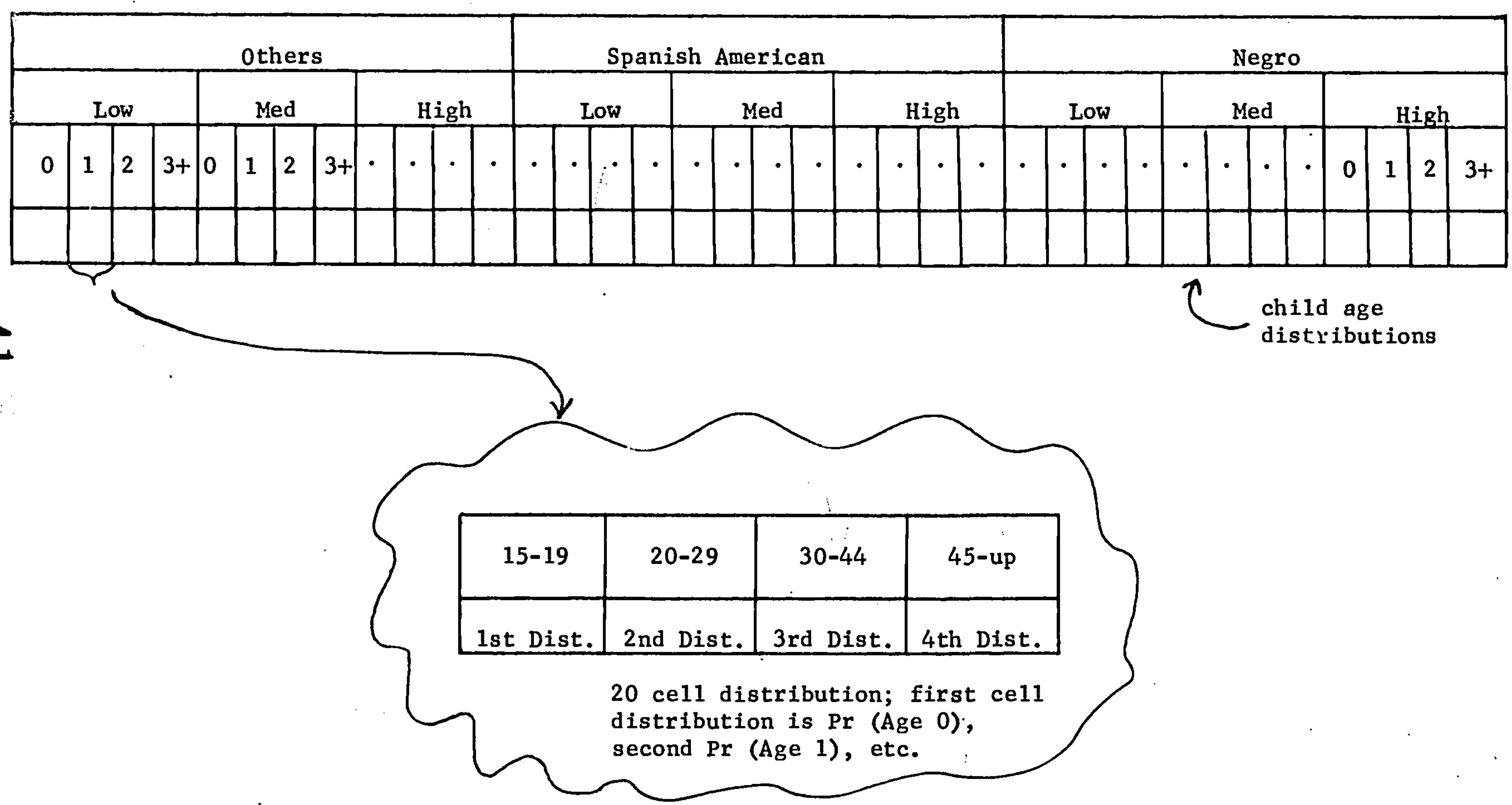
CHILD AGE DISTRIBUTION*



* Pr ( ) stands for the probability of the occurrence of the event in the brackets. In this case, a child from a particular household type being a particular age. The entries in the distribution sum to a value of one.

Figure 18

LAYOUT OF CHILD AGE DISTRIBUTIONS*



* Low, medium, high refers to income status of households that distributions relate to. 15-19, 20-29, 30-44, 45-up refers to the age of the household's spouse. Pr ( ) stands for the probability the event in the brackets will occur. Note that the distributions for households with no children will have zero entries.

is compiled. Then, the total number of children in each of the 144 combinations is divided into the number of children at each of the 20 age groups for each of the 144 combinations. The result is the desired distributions of child age. Child age distributions are compiled once for the entire school district under study.

5. Child Cohort Data

The child cohort data stores the counts of children living in each analysis zone. The numbers of children in each race/ethnic status and in each age group are compiled for each of the user district's analysis zones. Child cohort data for each zone is stored in a vector 60 cells long: the first 20 cells store children counts by age for the "others" race/ethnic status observed in the household vector; the second 20 stores the data for Spanish-American; and the third 20, for Negro. Figure 19 illustrates the child cohort vectors layout.

Like the other data sets, the child cohort data is derived from the special census data. This data set is essentially a cross-tabulation of the children living in the user school district. The child cohort data should be available for each basic analysis zone in the district.

6. Out-Migration Rates and Their Derivation

The ENSIM II Model uses migration rates to estimate which households will move out of the district during a given year. Migration rates are expressed by the percentage of households in a household type that will move in one year. The purpose of this section is to describe a methodology that can be used to derive these rates.

With the exception of number of children and structure type, all household characteristics used for the derivation of migration rates are defined here as they have been in the above sections. The number of children characteristic is expanded into a family size characteristic that assumes the following values:

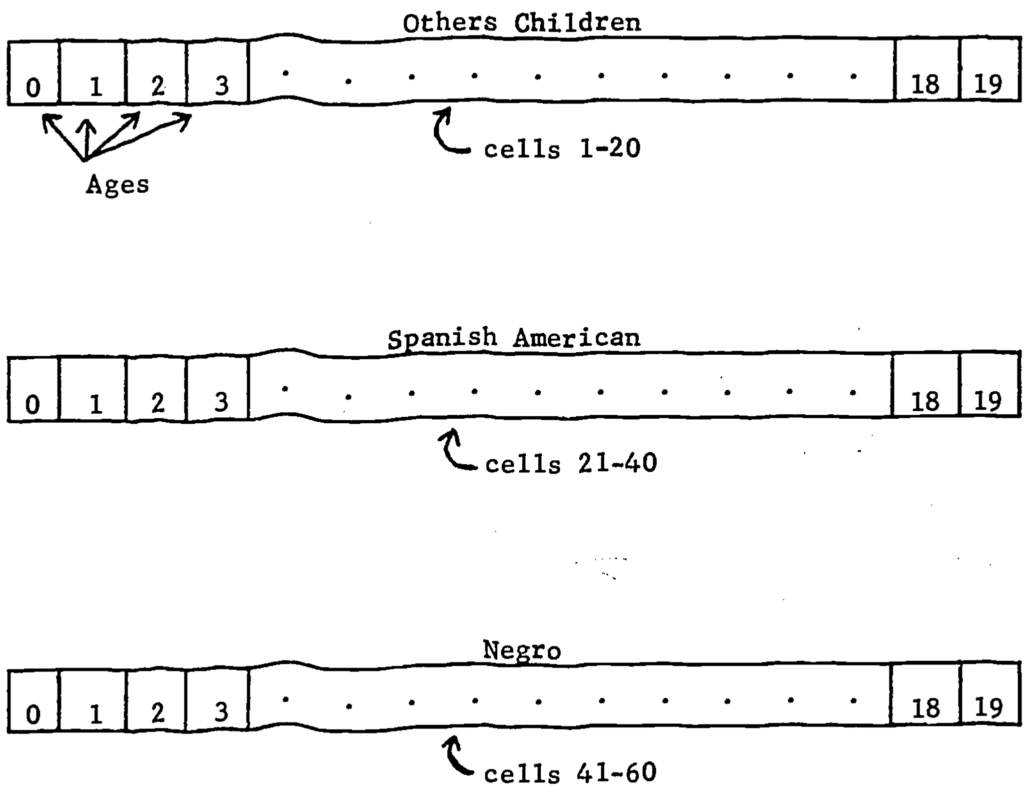
a. 1 person family;
b. 2 person family;
c. 3 person family;
d. 4 person family; and
e. 5 or more person family.

Type of structure is replaced by tenure (own/rent), probably a more reliable characteristic for projecting a household's movement.

Four household characteristics are hypothesized to influence out-migration. They are: income, size, race/ethnic status, and tenure. By selecting one value for each of these four characteristics, household chasses are generated. As a result, 90 household classes are created, and for them, 90 migration rates are estimated (3 income designations x 5 family size designations x 3 race/ethnic designations x 2 tenure designations = 90 household classes). Figure 20 arrays these 90 household classes.

Figure 19

LAYOUT OF CHILD COHORTS*



* Values in the cells relate to the age of the children whose total number are stored in the cells.

Figure 20

CLASSIFICATION SCHEME FOR 90 OUT-MIGRATION RATES

Size	Income	Race/Ethnic Housing Type	Others Own	Others Rent	Spanish Own	Spanish Rent	Negro Own	Negro Rent
1	Low							
	Medium							
	High							
2	Low							
	Medium							
	High							
3	Low							
	Medium							
	High							
4	Low							
	Medium							
	High							
5+	Low							
	Medium							
	High							

cells relate to the 90 out-migration *rates*

By dividing the count of households in each of the 90 classes in 1970 into those that remained until mid-year 1972, 2½ year retention rates are obtained. A transformation of these rates generates 1 year out-migration rates. A brief account of the methods used to prepare the 1970 and 1972 data for analysis, "backdating" the 1972 data for the calculation of the retention rates, follows.

In order to prepare the 1970 census data for analysis, it is necessary to obtain the following crosstabulations:

1) Households by size, income and tenure;
2) Households by race/ethnic class, size and income;
3) Households by race/ethnic class, size and tenure; and
4) Households by race/ethnic class, income and tenure.

A crosstabulation of households by race/ethnic class, size, income and tenure is not directly available and must, therefore, be estimated by utilizing the above data. To accomplish this, each cell in crosstab (2) is divided into owners and renters on the basis of the proportions observed in these categories at each income level for each race/ethnic category in crosstab (4). The result of this operation is used as a first approximation of the desired crosstabulation. The approximation is then adjusted proportionately by using the four tables mentioned above as expected marginal control totals. The final result is the crosstabulation of households by race/ethnic category, income, size and tenure.

The preparation of the 1972 data involves obtaining two crosstabulations of households by race/ethnic category, income, size, structure type (single family, multiple family) and length of residence. The first crosstabulation has all children less than two years of age subtracted from the household population variable; the second, all children less than three years of age. The count of households observed in each household category in each crosstabulation is then averaged. The overall effect of taking this average is to subtract all children less than two years of age from the household population as well as one half of those between two and three years of age. The final step in the preparation of 1972 data is backdating.

In this analysis, backdating 1972 data refers mainly to removing households that arrived between the two censuses from the household count in each category and removing recently born children from the household population variable. In the case of households with a two to three year length of residence, one-half were assumed to have been residents at the time of the 1970 census. Several attempts were made at backdating income for each household to its expected 1970 census value. The results of this process showed that stationary income may sometimes give the best results obtainable in light of the available information.

Because households in the census data available in 1972 were not indexed by tenure type, an indirect estimation procedure is used to estimate this dimension for persons who maintained their residence from the 1970 to the 1972 census. The first step of the procedure is to assume that tenure by race/ethnic category and structure type was the same as that of the area under study in 1970. This assumption provides data for a first approximation of retention rates. The second step involves assuming that tenure for each household cateogry in the backdated 1972 data (as indexed by race/ethnic category, size and income) would be the same as the proportions observed in 1970. This assumption implies equivalent migration rates for owners and renters. The number of owners and renters in each race/ethnic category resulting from this assumption is then compared to the number that would result from assuming the 1970 owner/renter proportions by race/ethnic category and structure type. The number of owners and renters in each race/ethnic category is then adjusted proportionally to agree with the number generated by assuming the 1970 owner and renter proportions. This final step completes the preparation of 1970 data--comparison of 1970 and 1972 data now allows the calculation of retention rates.

The use of these rates within the model requires a simplifying assumption. If it is assumed that on the average each household will have two adults present in addition to the number of children that are members of the household, these rates can be used directly. The rates for two person households would be used for households with no children; the rates for three person households, for households with one child; the rates for four person households, for households with two children; and the rates for five+ person households for households with three or more children. An alternative to this process would be the calculation of out-migration rates on the basis of the number of children present in households. The methodologies used should be determined by the data available for each particular study and by the judgment of the user.

7. Birthrates and Their Derivation

For the ENSIM II analysis, birthrates are calculated for females by race/ethnic status, age, and the number of previous children born into the household. The classification scheme for birthrates is illustrated in Figure 21. Birthrates are derived from special census data. To derive them, a crosstabulation of households, by age of female, race/ethnic status, the number of children in the household that are age 0, and the number of children older than 0 must be obtained. From this crosstabulation, birthrates can be calculated for each of the classifications in Figure 21. The number of children age 0 that are members of households in each classification should be compared to the number of households in each classification.

A certain number of households will have unmarried, male household heads. The direct calculation of births on the basis of the estimated birthrates will overestimate births. Two remedies to this situation can be employed. The first is to estimate births on the basis of the entire count of households in each cell of the birthrate classification scheme, with or without married male heads. The use of these rates would then assume that the levels of households with unmarried male heads, observed at the time the special census was conducted, will continue into the future. The second alternative is to simply not include these households and their housing units in the input data of

Figure 21

CLASSIFICATION SCHEME FOR BIRTHRATES*

Race/Ethnic	Parity	Age	15-19	20-29	30-44	45-up
Others	0					
	1					
	2					
	3+					
Spanish	0					
	1					
	2					
	3+					
Negro	0					
	1					
	2					
	3+					

matrix of birthrates

* Parity refers to the number of previous children born into a household.

the model. Where their numbers are small, this second alternative is simplest. A large number of households with unmarried male heads will have no children, and these are not of any particular interest to the user. The exclusion of these households simply implies that their present numbers will be maintained and will continue to occupy the present housing units. If this second alternative is used, any children in the excluded households should be entered into the child cohorts.

If birthrates are not calculated from special census data, those used should be adjusted to account for child mortality at birth. In addition, the rates should apply to both married and unmarried females.

A factor to raise or lower the overall rate of births can be entered into the ENSIM II Model. This factor will express the level of births in a given time period as a function of the birthrates in the previous time period. A factor of .95 will lower the birthrates 5% for each time period of the projection. A factor of 1.05 will raise births 5% for each time period of the projection.

8. Distribution of In-Migrant Households

The reader will recall that ENSIM II uses an analysis zone typology and a set of probability distributions of households that corresponds to this typology in deriving the characteristics of in-migrant households. Fifteen different types of analysis zones exist for the purposes of in-migration: the first 12 relating to the in-migration of households into housing units located in analysis zones with some existing development; and the remaining three to zones of new development. The analysis zone typology is based upon the average number of children present in households resident in zones before in-migration occurs, and the average value of housing units in zones. In working with the model, the following values have been used in the zone typology: analysis zone types 1 through 3 relate to zones where the average number of children per household is less than or equal to one; types 4 through 6, less than or equal to two children; types 7 through 9, less than or equal to 3 children; and 10 through 12, greater than three children. Within each of the divisions formed by the average number of children present in households, the lowest analysis zone type relates to zones with the lowest average value of housing, the next, to zones with the middle range valued housing; and the least to zones with the highest average value of housing (see Figure 22). The analysis zone types that relate to zones of new development are based upon the average housing unit value of the new development (see Figure 23).

For each of these analysis zone types, two distributions of in-migrant households must be derived. The first of these relates to households that will in-migrate to single family housing units, and the second, to multiple family units. Both of these distributions are probability vectors that sum to a value of one. The order that household types relate to successive cells of these vectors is similar to the order established for the household vector. Each of the distributions is 144 cells long, and both sum to a value of one. In each of these distributions, the first 48 cells relate to the first race/ethnic status in the

Figure 22

ANALYSIS ZONE TYPOLOGY AND ZONE CHARACTERISTICS
FOR ZONES WITH EXISTING DEVELOPMENT

Analysis Zone Type	Characteristics
	Zones where households average less than or equal to 1 child, and:
1	lowest valued housing
2	middle valued housing
3	highest valued housing
	Zones where households average less than or equal to 2 children, and:
4	lowest valued housing
5	middle valued housing
6	highest valued housing
	Zones where households average less than or equal to 3 children, and:
7	lowest valued housing
8	middle valued housing
9	highest valued housing
	Zones where households average more than 3 children, and:
10	lowest valued housing
11	middle valued housing
12	highest valued housing

Figure 23

ANALYSIS ZONE TYPOLOGY AND ZONE CHARACTERISTICS
FOR ZONES OF NEW DEVELOPMENT

Analysis Zone Type	Characteristics
13	lowest valued housing
14	middle valued housing
15	highest valued housing

50

household vector (i.e., Others), the second 48 to the second race/ethnic status (i.e., Spanish American), and the last 48 to the last race/ethnic status (i.e., Negro). Within each race/ethnic status, the first 16 cells relate to households in the lowest income status, the second 16 to households in the middle income status, and the last 16 to households in the highest income status. The next two divisions relate to number of children status and the age of female status respectively. Figure 24 displays the household classifications for the two basic types of in-migrant household distributions (i.e., single and multiple family housing unit vectors).

The processing of the in-migrant households involves the decomposition of the arrays shown in Figure 24 into two vectors 144 cells long. This decomposition is similar to the one performed for the household vectors. Figure 25 displays the layout of the in-migrant household vector for in-migrants into single family housing units. The values that appear in the cells of Figure 24 relate to the order of household characteristics displayed in Figure 25.

The user must input the values that define the analysis zone typology. Values are input for the three break-off points of the average number of children characteristic, and the two break-off points of the value of housing characteristic. The three average number of children break-off points will define four average number of children classes, while the two average housing unit value break-off points will define three average housing unit value classes. The combinations of these two sets of classes will define the typology. The estimation of in-migrant household distributions is based on the use of special census data. The break-off values that define the block typology are selected on the basis of those that help explain the average number of children that accompany in-migrating households.

In-migrant household distributions are calculated by compiling arrays of the households that have recently in-migrated into the analysis zones of the district under study. Once arrays have been compiled for recent in-migrants, into single and multiple family housing units for each type of analysis zone, the total number of households in each array is divided into the number of households in each cell of the array. The result of this operation is the desired distributions.

9. Average Value of Housing Units

For each analysis zone of the school district under study, data must be collected on the average value of housing units. This data, which can probably be obtained from the assessor's office, is used to classify analysis zones by comparing the housing unit values to the typology of analysis zones. It is important that housing unit values are those set at the time that the census being utilized was conducted. If average housing unit value data is not available, average household income can be used as a substitute, provided analysis zone typology is made to relate to household income levels. Also, value data collected on housing input data for new construction should then be divided by a fixed household income to housing unit value ratio. These ratios can normally be estimated from U.S. Census materials or other data sources.

Figure 24

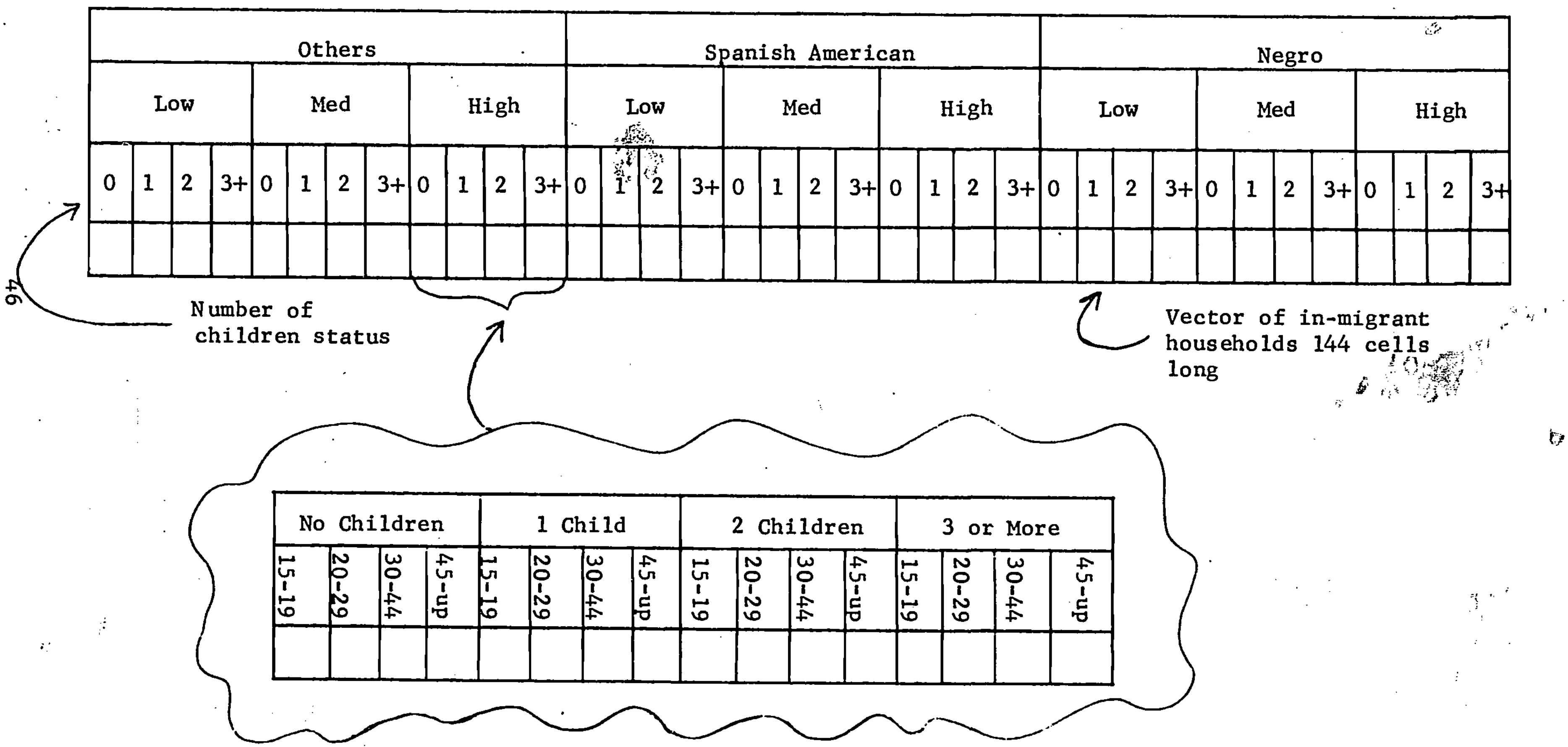
HOUSEHOLD CLASSIFICATIONS FOR THE SINGLE AND MULTIPLE FAMILY HOUSING UNIT DISTRIBUTIONS*

Income	Age	Distrib. Type	Single Family Housing Distribution												Multiple Family Housing Distribution											
		Race/Etnic	Others				Spanish				Negro				Others				Spanish				Negro			
		No. of Children	0	1	2	3+	0	1	2	3+	0	1	2	3+	0	1	2	3+	0	1	2	3+	0	1	2	3+
Low	15-19		1	5	9	13	49	53	57	61	97	101	105	109	1	5	9	13	49	53	57	61	97	101	105	109
	20-29		2	.	.	.	50	.	.	.	98	.	.	.	2	.	.	.	50	.	.	.	98	.	.	.
	30-44		3	.	.	.	51	.	.	.	99	.	.	.	3	.	.	.	51	.	.	.	99	.	.	.
	45-up		4	8	12	16	52	56	60	64	100	104	108	112	4	8	12	16	52	56	60	64	100	104	108	112
Med	15-19		17	21	25	29	65	69	73	77	113	117	121	125	17	21	25	29	65	69	73	77	113	117	121	125
	20-29		18	.	.	.	66	.	.	.	114	.	.	.	18	.	.	.	66	.	.	.	114	.	.	.
	30-44		19	.	.	.	67	.	.	.	115	.	.	.	19	.	.	.	67	.	.	.	115	.	.	.
	45-up		20	24	28	32	68	72	76	80	116	120	124	128	20	24	28	32	68	72	76	80	116	120	124	128
High	15-19		33	37	41	45	81	85	89	93	129	133	137	141	33	37	41	45	81	85	89	93	129	133	137	141
	20-29		34	.	.	.	82	.	.	.	130	.	.	.	34	.	.	.	82	.	.	.	130	.	.	142
	30-44		35	.	.	.	83	.	.	.	131	.	.	.	35	.	.	.	83	.	.	.	131	.	.	143
	45-up		36	40	44	48	84	88	92	96	132	136	140	144	36	40	44	48	84	88	92	96	132	136	140	144

* Values in cells relate to order of cells in in-migrant household vectors. Both of these distributions sum to a value of one.

Figure 25

DISTRIBUTION OF IN-MIGRANT HOUSEHOLDS*



* Display in cloud is a detailed illustration of the area of the in-migrant household vector indicated. Sum of all 144 cells should equal one. Low, Med, and High refers to income levels.

The user has the option of appreciating or depreciating average housing unit values. This is done through the use of a fixed factor input into the model: a factor of 1.05 would appreciate housing values 5% for each time period of the projection; .95 would lower values 5% for each period.

10. Income Transform Data

If desired, the ENSIM II user can change household income. Income changes are made by moving certain percentages of households from one income level to another. The percentage moved depends on factors which are entered once and operate for each time period of the projection. Households are moved from one income status to another on the basis of their income in the previous time period, their race/ethnic status and the age of the household head's spouse (used as a proxy for the household head's age).

Thus, income can be lowered or raised for households of various types. An array of the various factors used to transform income is illustrated in Figure 26.

11. Vacancy Rates

Vacancy rates for single and multiple family housing units are the last of the input data to be derived for use in ENSIM II. These rates are uniform for the entire school district and operate for all time periods of the projection. They cannot be changed during the projection period. The effect of these rates is to withhold specified percentages of the district's housing stock from in-migrant occupancy. The housing is withheld as it becomes available through the out-migration of households. A certain number of housing units will also be vacant at the beginning of the projection, these forming the initial stock of vacant units.

Figure 26

FACTORS TO TRANSFORM INCOME*

Present Income	Change to	Race/Ethnic	Others				Spanish				Negro			
		Age	15-19	20-29	30-44	45-up	15-19	20-29	30-44	45-up	15-19	20-29	30-44	45-up
Low	Medium													
	High													
Medium	Low													
	High													
High	Low													
	Medium													

72 cells

* 72 factors in all are used to transform household income.

II. SPECIFICATIONS FOR INPUT DATA SETS

A. Introduction

The following sub-sections describe the specifications for the input data sets whose preparation was described above. The description will be presented as follows: (a) data formats for the input data sets will be outlined first; (b) the data sets will be discussed in the order in which they were presented above; (c) a summary table of the physical characteristics of data sets residing on direct access data storage devices will follow; and finally (d) a description of the card deck set up required for operation of ENSIM II will be presented. Figure 27 presents a flowchart of the computer program structure of ENSIM II. The entire program is written in FORTRAN IV.

B. Specifications of Land Use Input Data

The land use data input into ENSIM II relates to the construction of housing units in each of the analysis zones within the district under study. Data summarizing all of the construction activity expected to occur within the yearly intervals of the projection is input for each zone. This data is stored on punched cards. The format of a single card is illustrated in Figure 28.

The first three columns of the housing input data should be coded with the zone number to which the data relates; columns 4 through 13 to the number of single family housing units to be constructed; columns 14-23 to multiple family housing units; and columns 24-33 to the average value of this housing. The remainder of the card may be coded to indicate in which time period the construction is to occur.

All of the housing units to be built within a given period of the projection are input at one time. The units to be constructed in the first year are input first; the second year, second; and so on for the remaining time periods. The cards relating to successive years of the projection are separated by a single card, with a value punched in the first three columns of the card. This value should equal the total number of analysis zones in the district plus one. Thus, if 120 analysis zones are used, a value of 121 is punched onto the separator cards. The FORTRAN format of the housing input cards is I3.3F10.4. Figure 29 illustrates this card deck set-up. This data deck must end with a separator card following the input for the final time period of the projection. The FORTRAN format of the separator cards is I3. If no housing is to be constructed in a particular year, only a separator card need be input for that year.

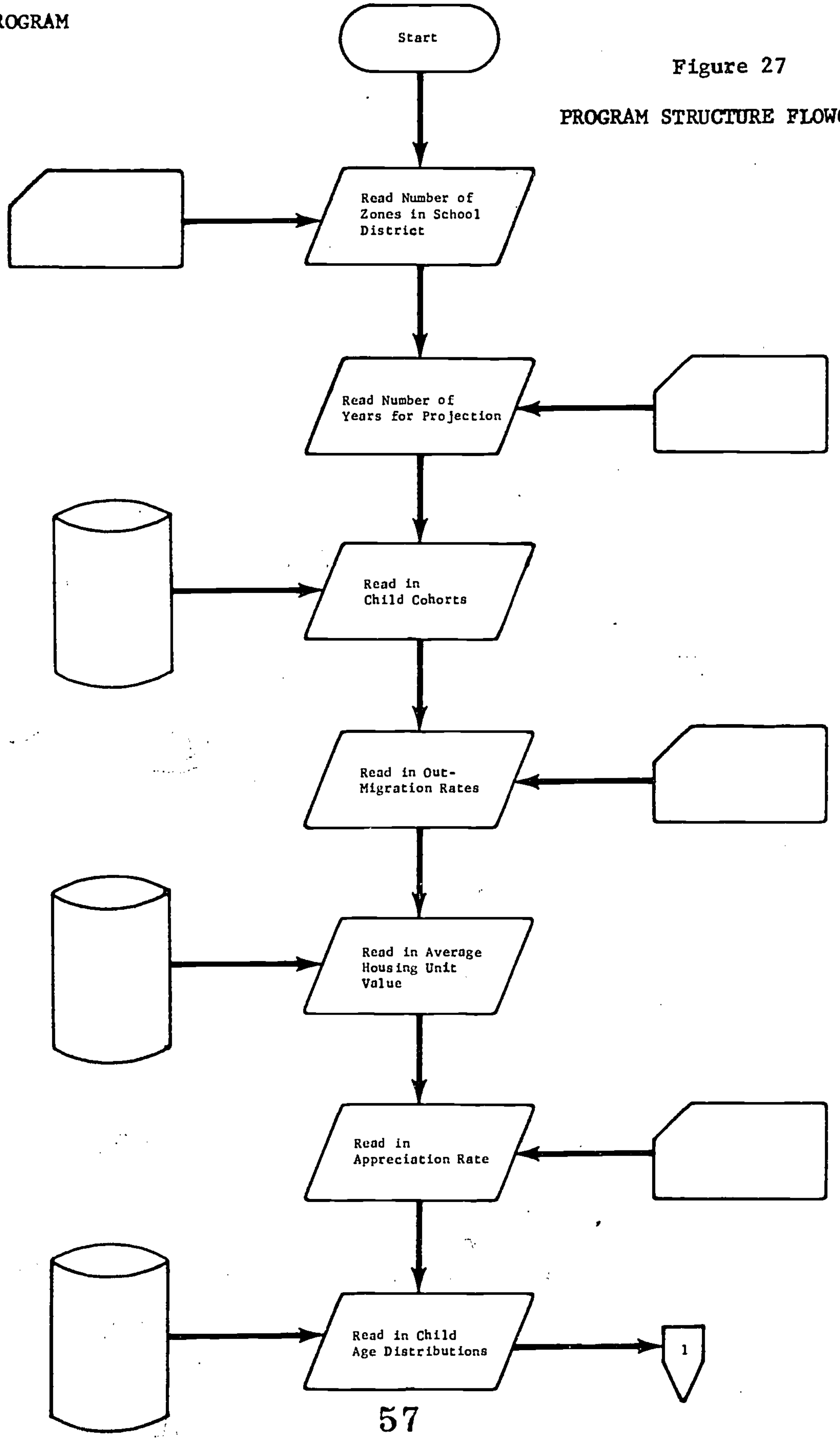
C. Household Data File

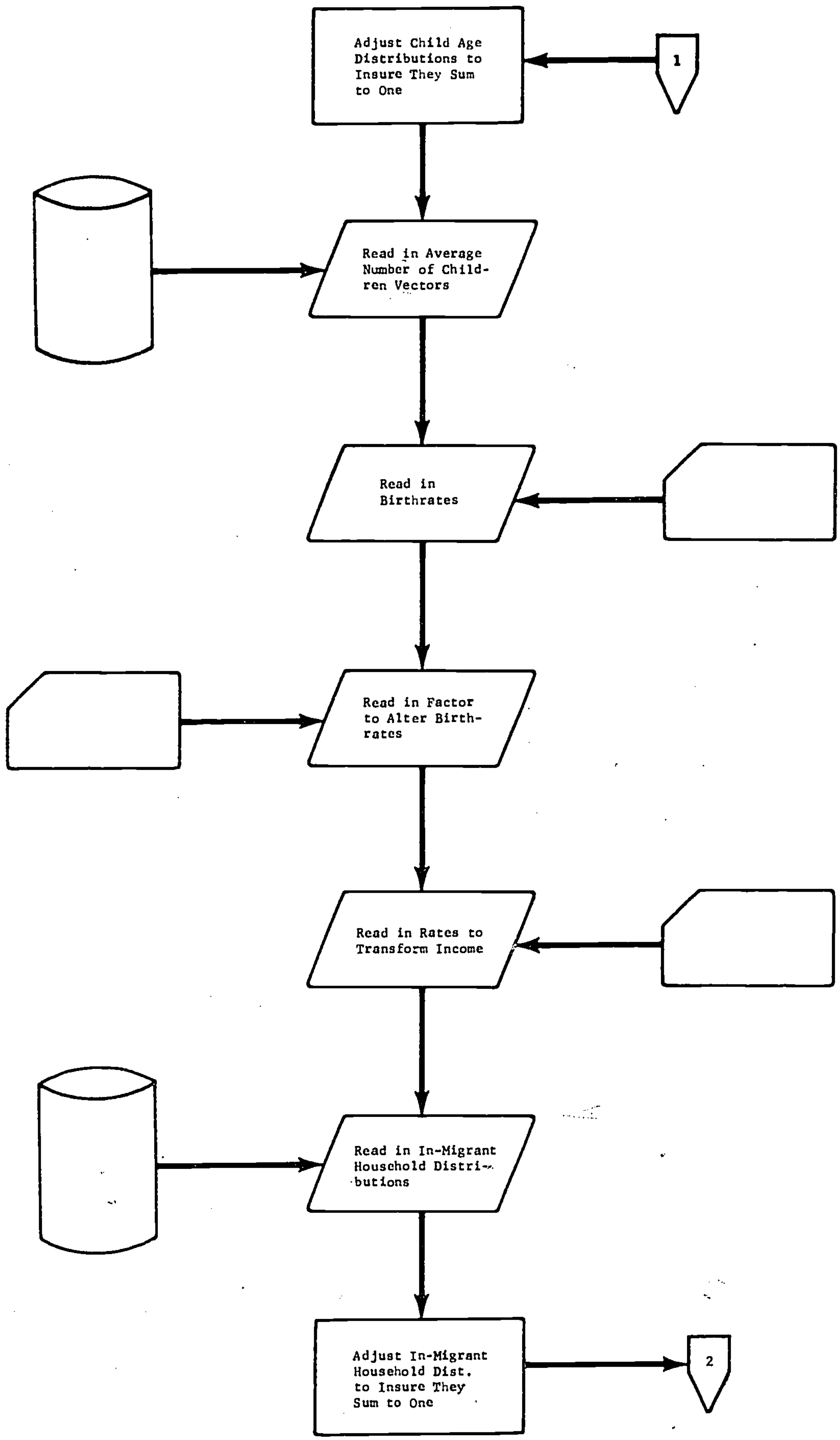
The household data file should be stored on magnetic disk or some other direct access data storage device. The reader will recall that this data set is composed of vectors 288 cells long, with five cells tacked onto the end. Before running the projection, this data set is copied for processing. A second data file, with the same physical characteristics as the household data set, must also be provided. This data file is used for working space.

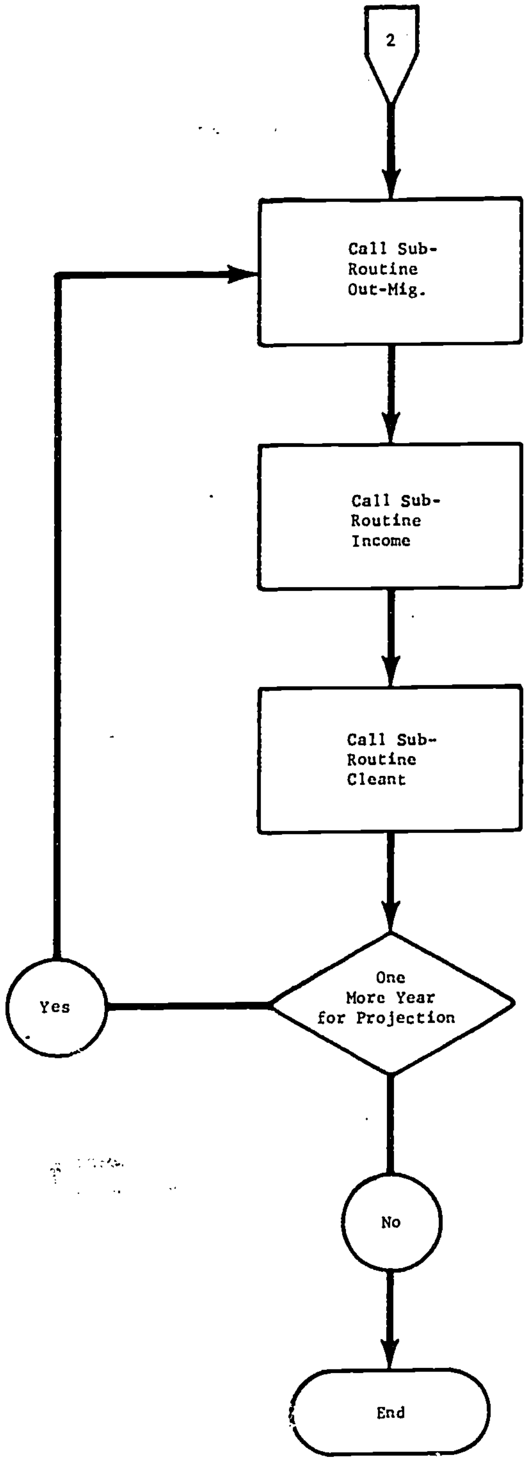
MAIN PROGRAM

Figure 27

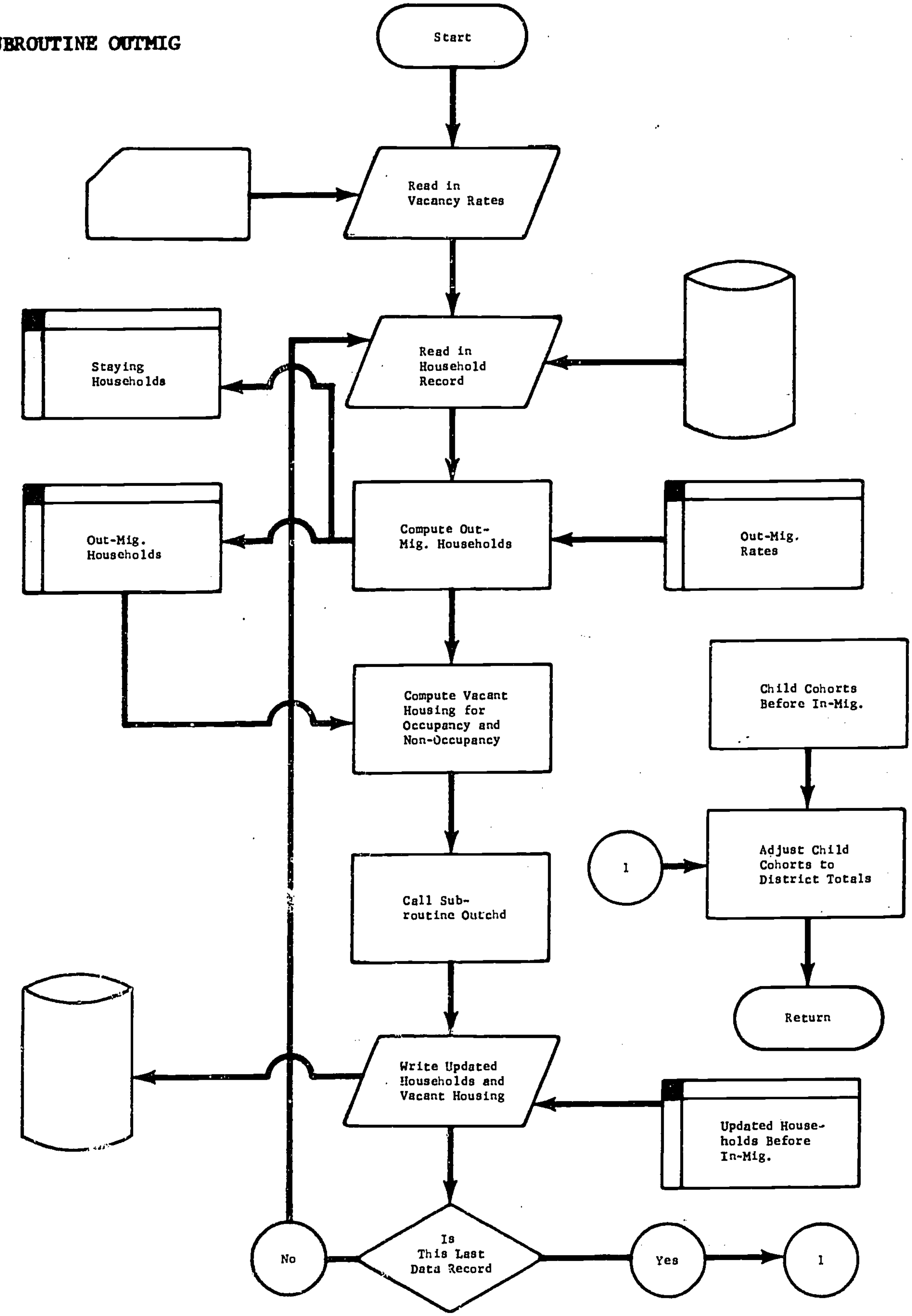
PROGRAM STRUCTURE FLOWCHART



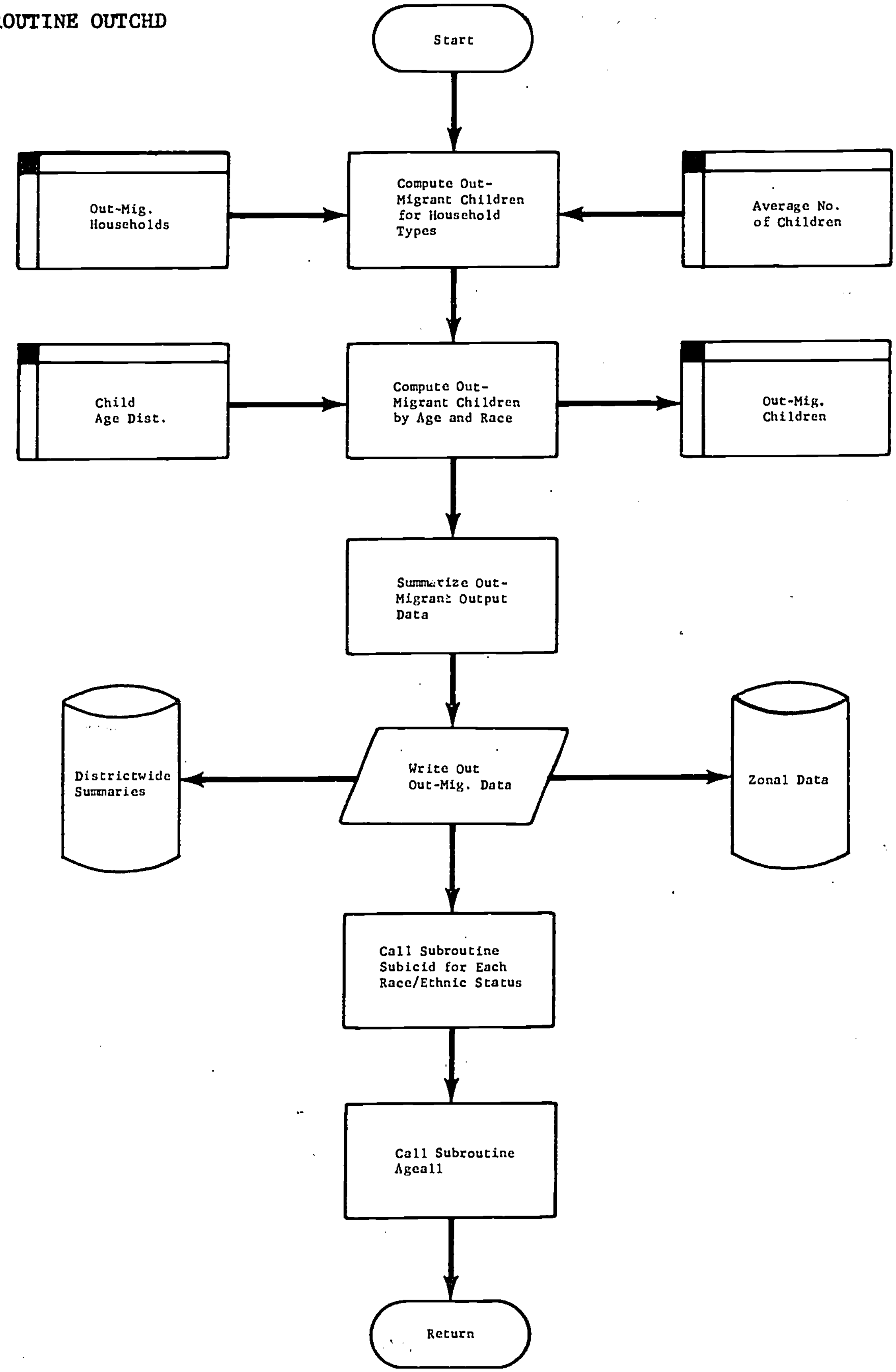

Adjust Child Age Distributions to Insure They Sum to One
1
Read in Average Number of Children Vectors
Read in Birthrates
Read in Factor to Alter Birthrates
Read in Rates to Transform Income
Read in In-Migrant Household Distributions
Adjust In-Migrant Household Dist. to Insure They Sum to One
2


2
Call Sub-
Routine
Out-Mig.
Call Sub-
Routine
Income
Call Sub-
Routine
Cleant
One
More Year
for Projection
Yes
No
End

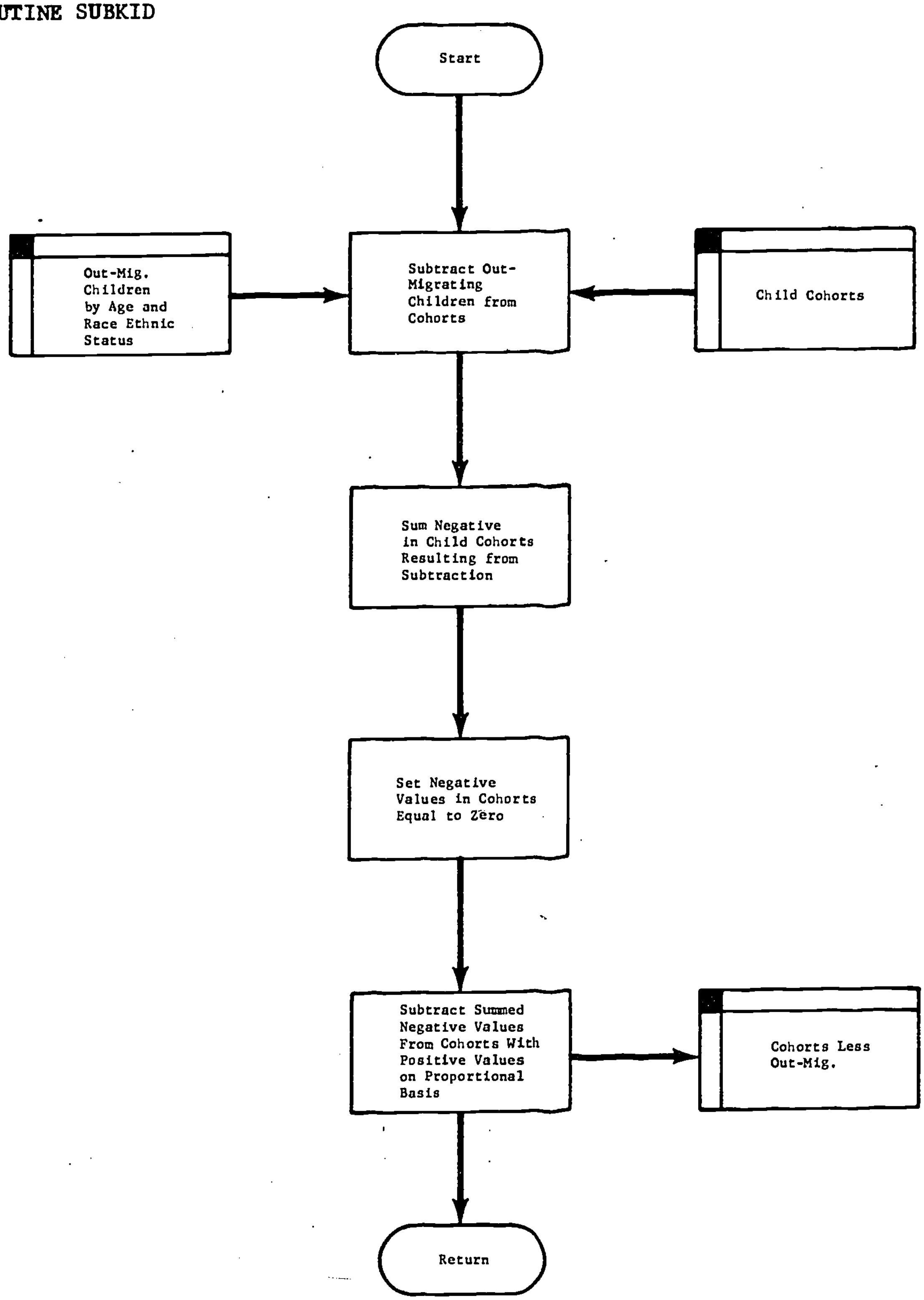
SUBROUTINE OUTMIG



SUBROUTINE OUTCHD

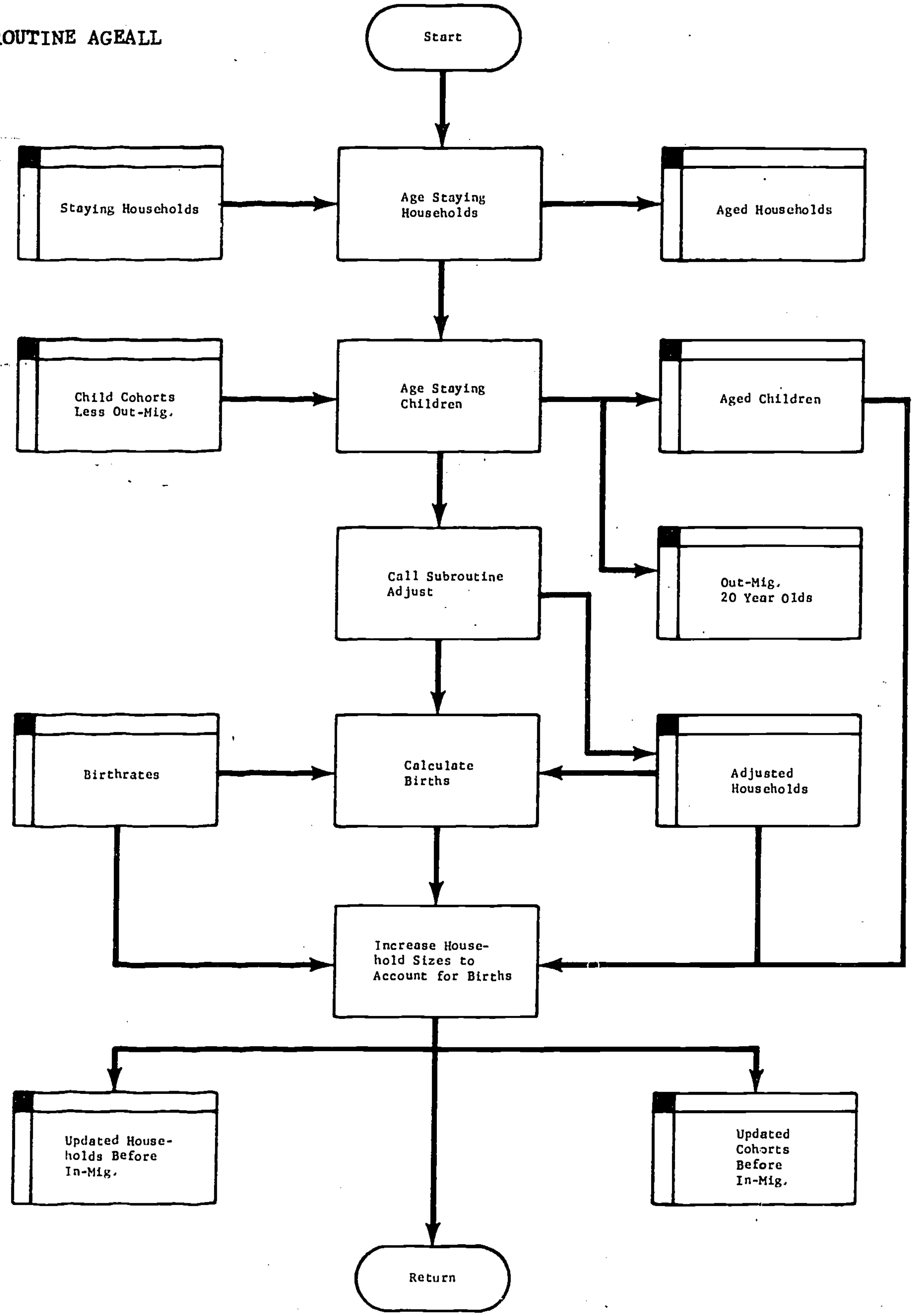


SUBROUTINE SUBKID



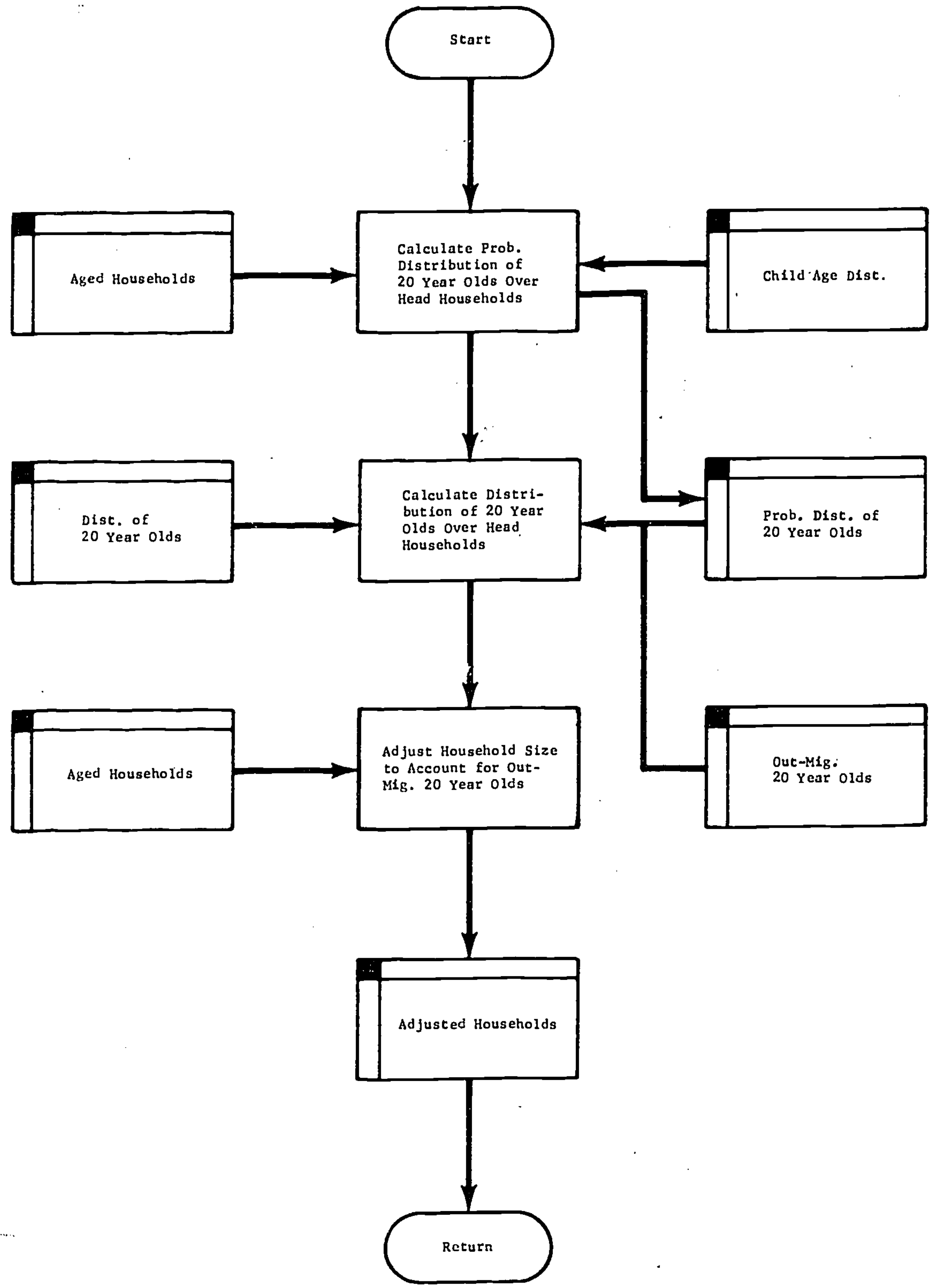
62

SUBROUTINE AGEALL

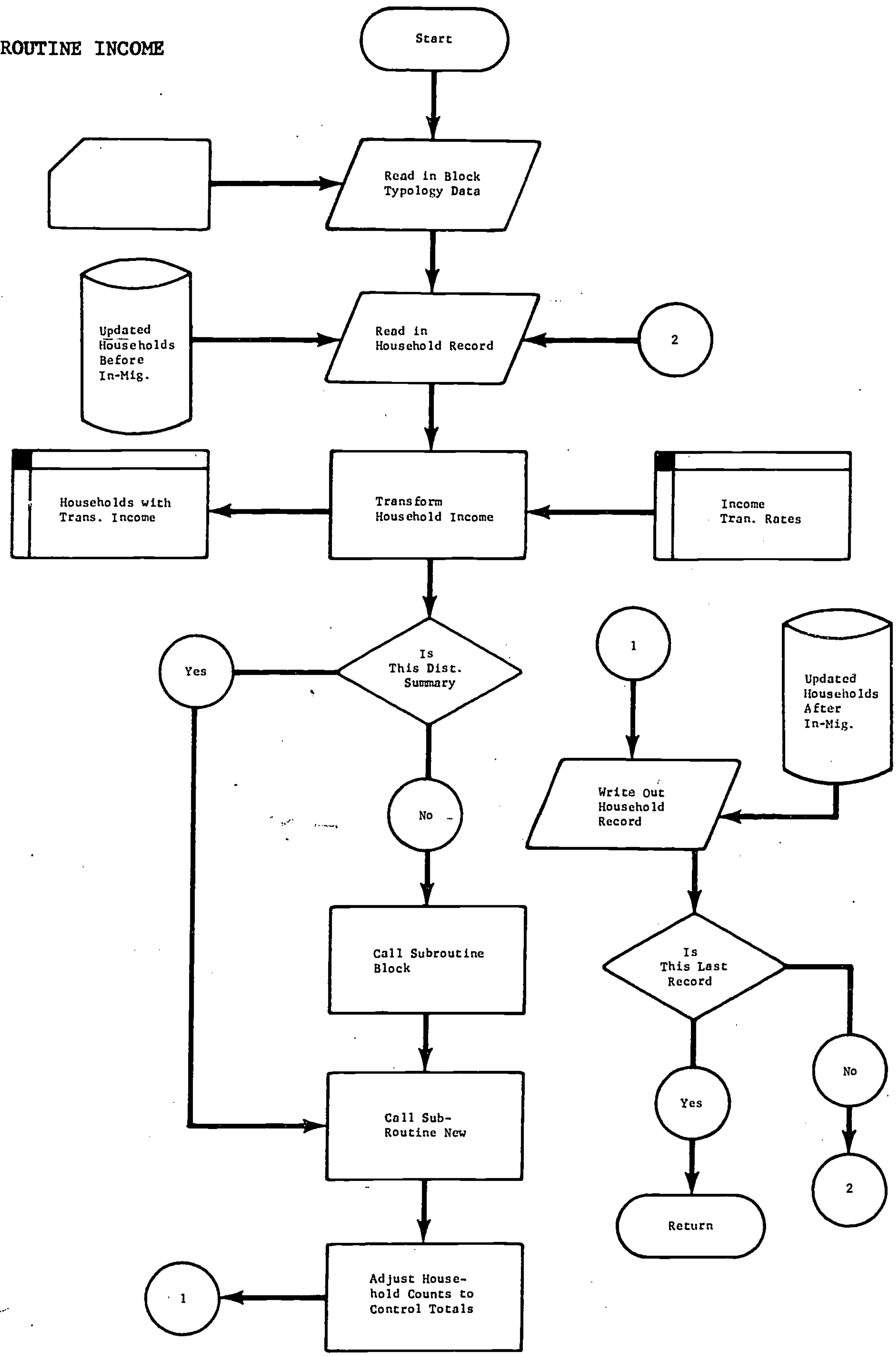


63

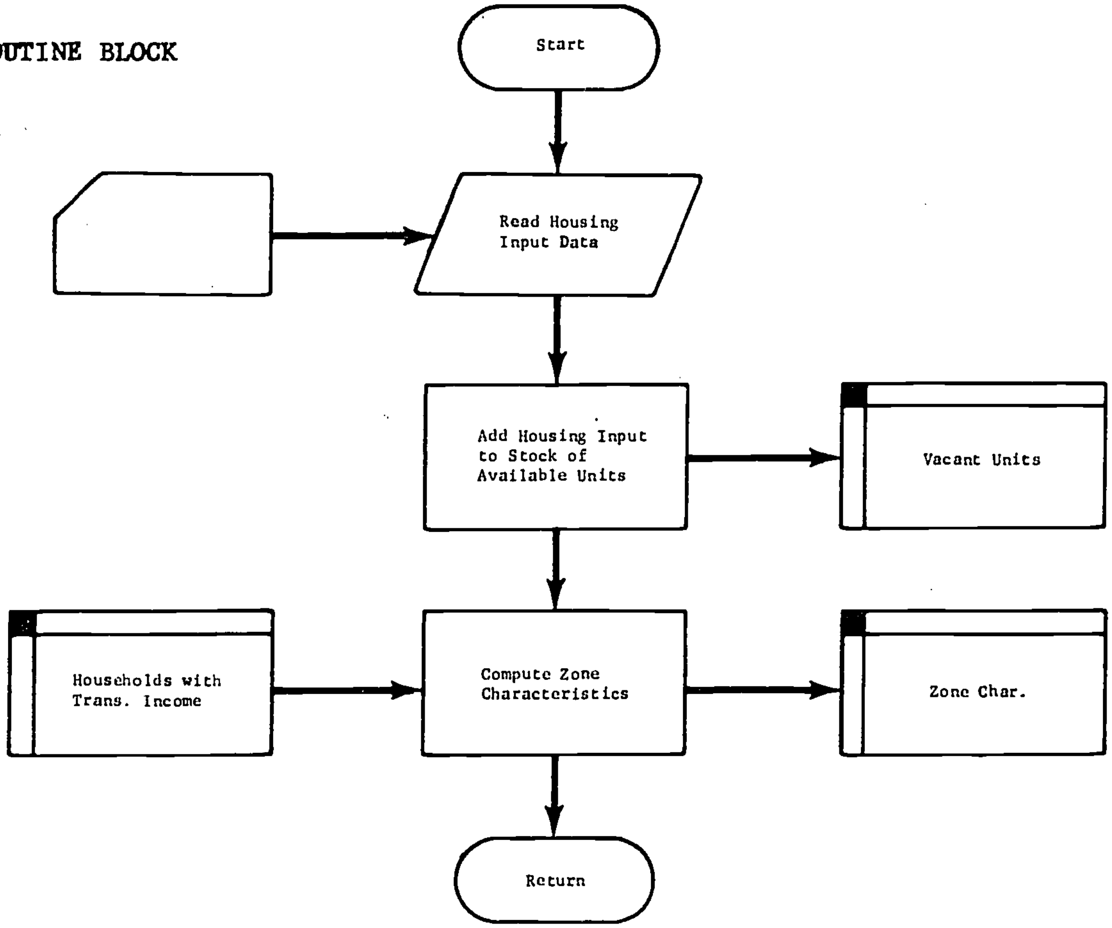
SUBROUTINE ADJUST



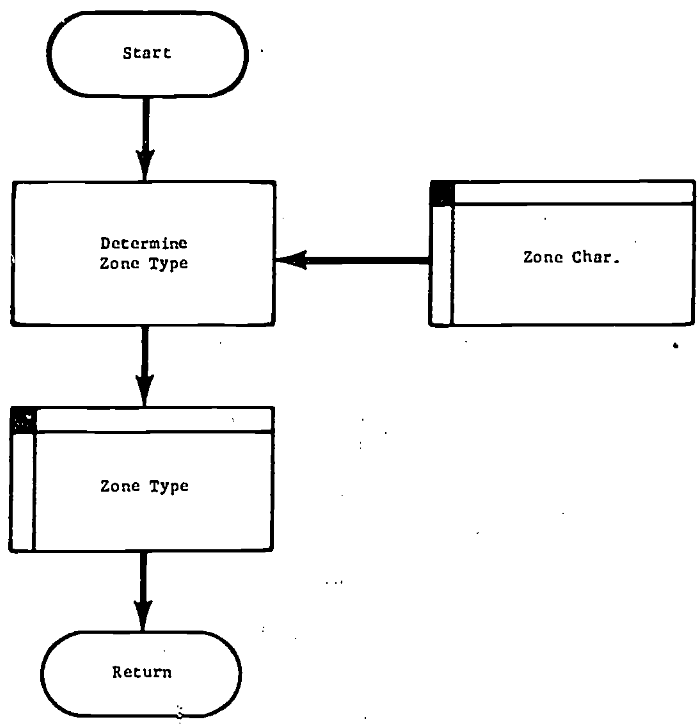
SUBROUTINE INCOME



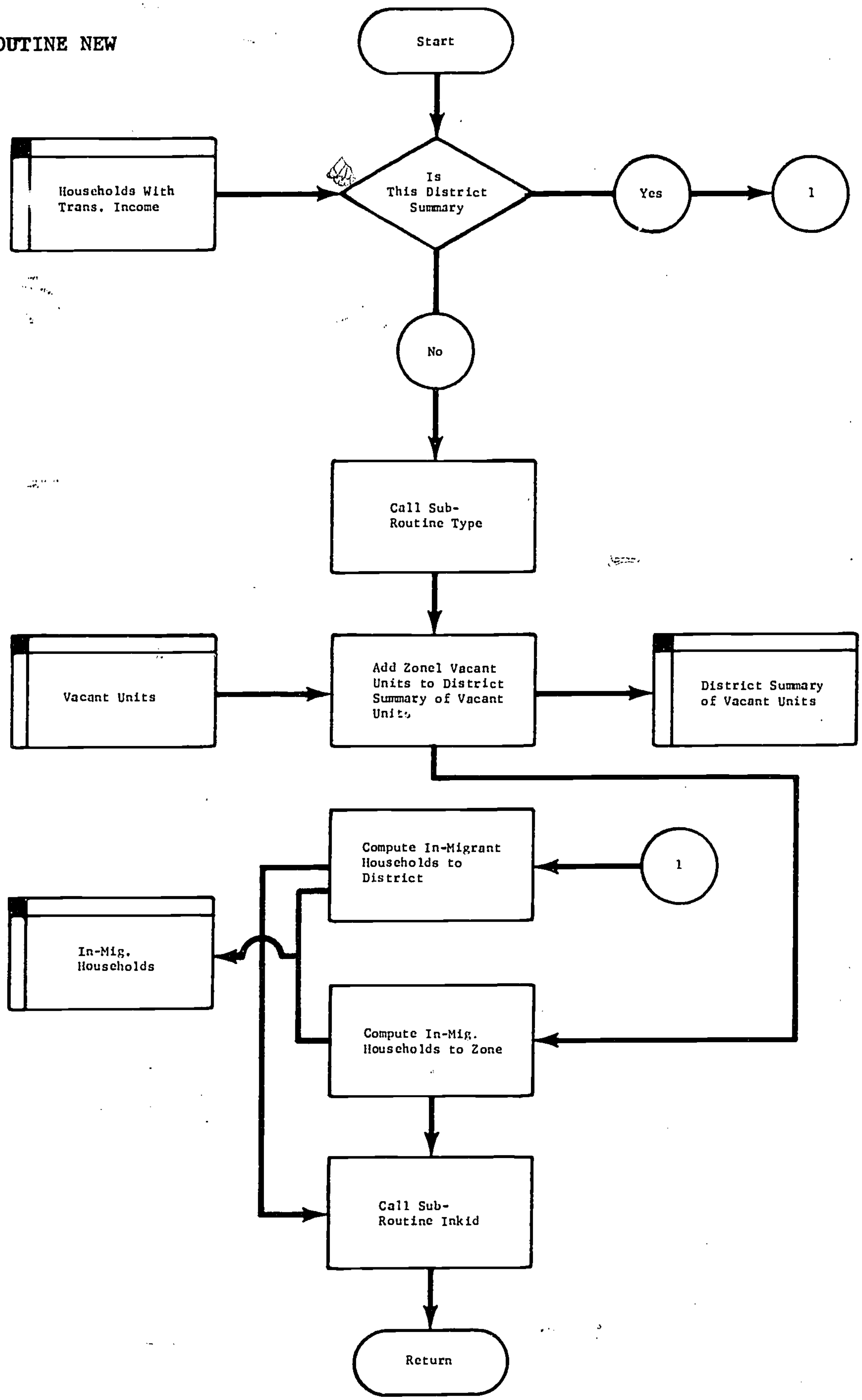
SUBROUTINE BLOCK



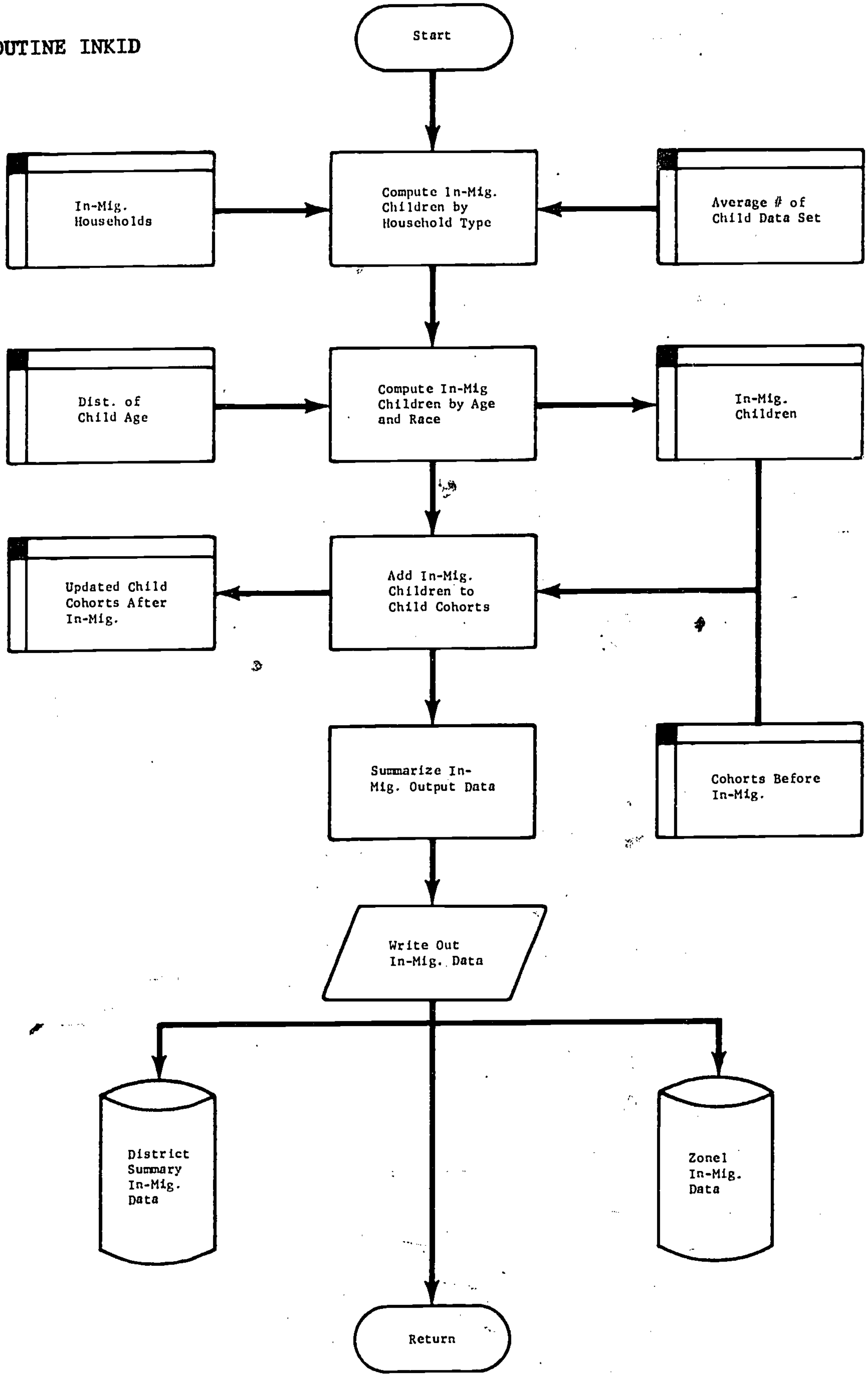
SUBROUTINE TYPE



SUBROUTINE NEW



SUBROUTINE INKID



SUBROUTINE CLEANI

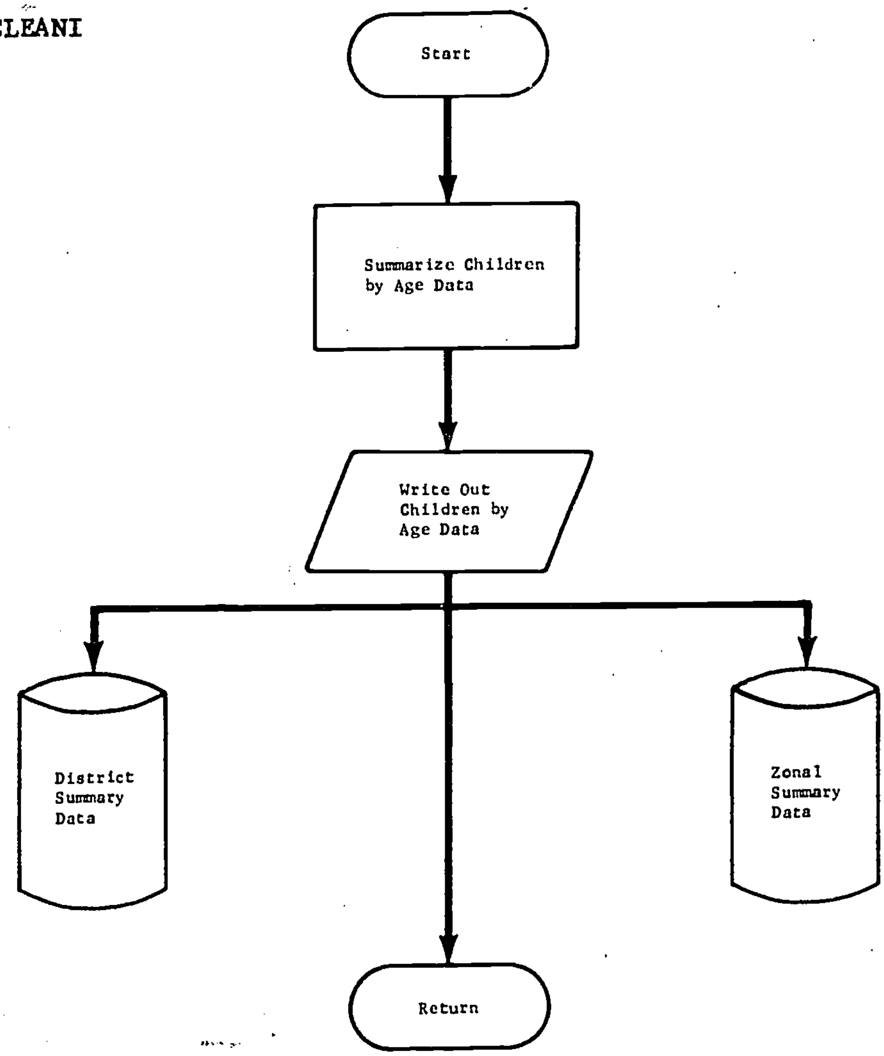
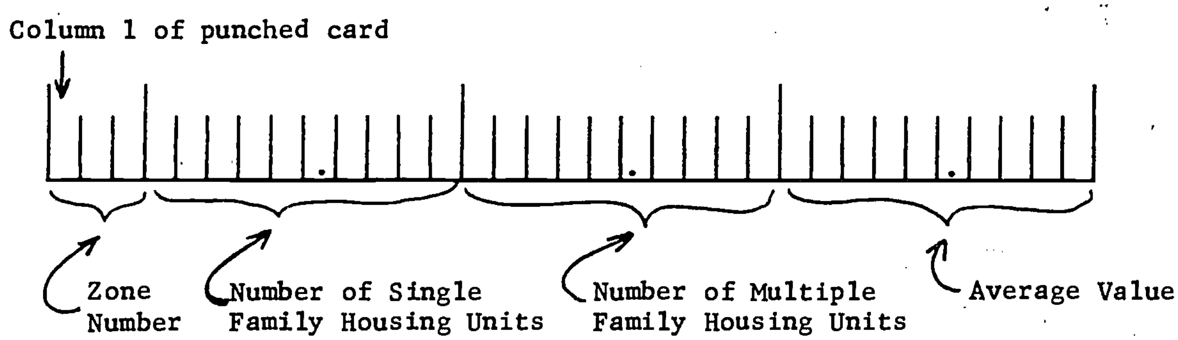


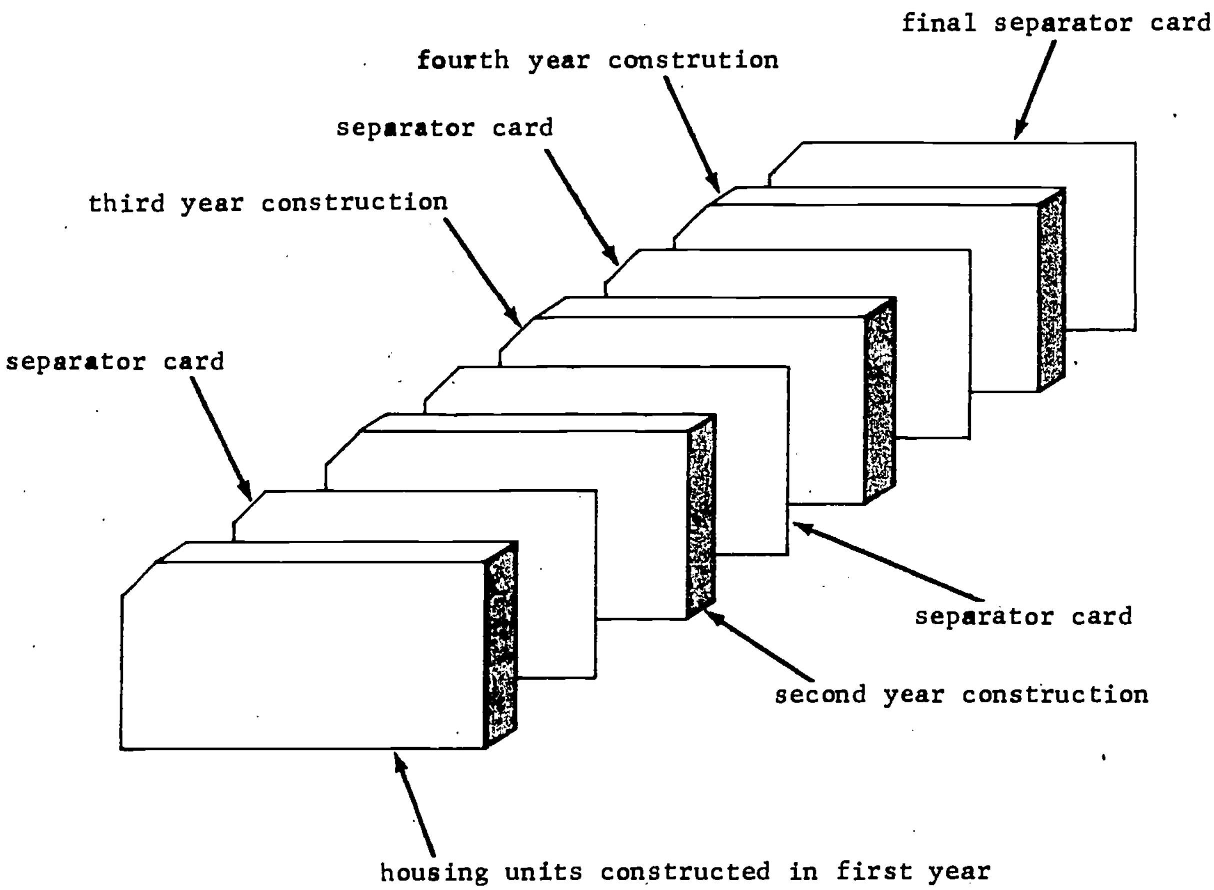
Figure 28

FORMAT OF HOUSING INPUT DATA CARD*



* Columns 1-3 store the zone number, 4-13 single family housing units, 14-23 multiple family housing units, and 24-33 average value of housing units. Zone numbers should be right justified. Format of each card is (I3,3F10.4). Dots in specific columns signify the placement of decimal points.

FIGURE 29
DECK SET UP FOR HOUSING INPUT DATA*



*Example of deck set up for four year projection.
Deck must end with a separator card.

The ENSIM II program reads and writes to both of these data sets throughout the entire projection. Thus, copying the household data set is necessary to preserve the original data--otherwise the household data would have to be recreated for each projection to be made. It is recommended that a system utility be used to copy the household data set and then pass the copied data set to the job step that contains the ENSIM II program.

The household data set is processed as a FORTRAN IV unformatted data set. The record format is variable, spanned and blocked (RECFM=VSB), with a logical record length and block size of 1172 (LRECL-1172, and BLKSIZE=1172). Thus, each household vector is stored as a logical record. FORTRAN device number 1 is reserved for this data set (GO.FT01F001). The physical characteristics of the working data set are the same as those of the household data file. FORTRAN device number 3 is reserved for this data set (GO.FT03F001).

The first vector in the household data set should relate data on housing units in the first analysis zone, the second vector, zone 2, the third vector, zone 3, and so on. The last vector in the household data file should be a districtwide summary of the household data of all the analysis zones. This summary is derived by taking the sum of the household vectors that store data on household residents in each analysis zone. This last vector stores data on the number of households by type of resident in the entire school district under study. The order of the zonal data and household summary is illustrated in Figure 30.

D. Average Number of Children Per Household Data Set

This data set relates information on the average number of children that are members of particular household types. The data should be stored as a single vector 288 cells long, with FORTRAN format of 288F8.5, on a direct access data storage device. The data set should be fixed and blocked (RECFM=FB), with a logical record length and block size of 2304 (LRECL=2304, and BLKSIZE=2304). The data is read once for each projection. Device number 8 (GO.FT08F001) is reserved for this data set.

E. Child Age Distributions

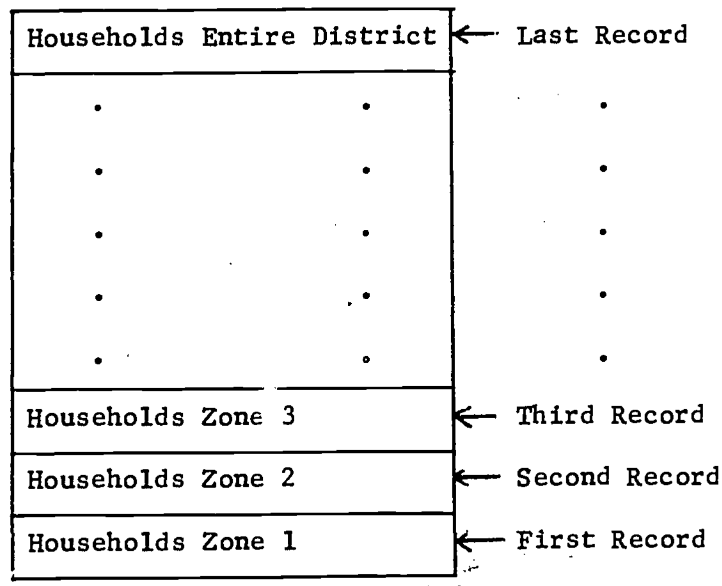
The 144 child age distributions should be stored on a direct access data storage device. Each distribution is stored as a single vector 20 cells long with FORTRAN format 20(F4.2,5x). This data should be fixed and blocked (RECFM=FB) with a logical record length of 180 (LRECL=180), and any convenient block size. The data set is read once for each projection and device number 9 (GO.FT09001) is reserved for it.

F. Child Cohort Data File

The child cohort data set is composed of vectors 60 cells long, each of which stores data on the child cohorts present in each analysis zone in the district. The first vector of this data set should store the data for analysis zone 1, the second vector for analysis zone 2, and so on. The last record in this

Figure 30

ORDER OF HOUSEHOLD VECTORS IN HOUSEHOLD DATA SET*



* Last Record is districtwide summary formed by taking sum of records for all analysis zones.

data set should store a districtwide summary of the data for analysis zones. This districtwide summary is formed by taking the sum of all the vectors for analysis zones--it stores data on children, by age and race/ ethnic status, resident in the entire school district.

The child cohort data set is processed as a FORTRAN IV formated data set. The format of the cohort vectors is 60F10.4. This data set is fixed and blocked (RECFM=FB) with a logical record length of 600 (LRECL=600) and any convenient block size. Thus, each child cohort vector is stored as a separate record. The data is read once for each projection and FORTRAN device number 10 is reserved for it (GO.FT10F001).

The first record in the child cohort data set should store the child cohort data for the first analysis zone, the second record for the second analysis zone, and so on. The last record is the districtwide summary. Figure 31 illustrates the set-up for the child cohort data set.

G. Format of Out-Migration Rates

Out-migration rates are input to ENSIM II through the use of punched cards. Four rates are punched on each card, and since there are 72 out-migration rates, 18 cards are input to the model. The FORTRAN format of each card is 4F5.4. Figure 32 illustrates the format of one card. Because out-migration rates relate to specific types of households, they must be placed on the punched card in a specific order as illustrated in Figure 33.

H. Format of Birthrates

Birthrates are also input on punched cards--four on each, in FORTRAN format 4F6.5. The individual rates are input in a form that relates to the percentage of women who are members of particular types of households and who will give birth to children within one time period of the projection. A rate of .15 would indicate that 15% of the women who are members of households of a particular type will give birth to children. Figure 34 illustrates the format of a single birthrate input card. Figure 35 illustrates the deck set-up and order of birthrates on punched cards. Sixty-two birthrates, and therefore 12 cards, are input to the model.

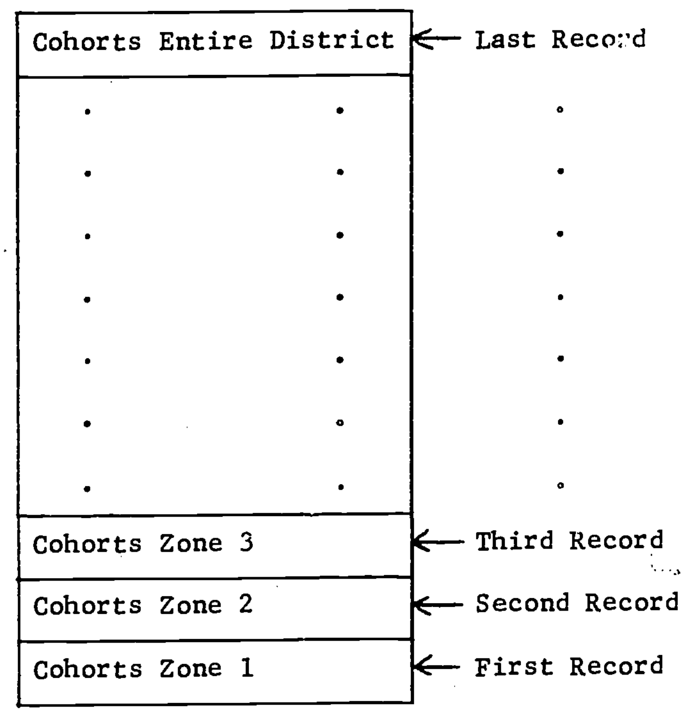
The factor to alter birthrates is input in a single card in FORTRAN format F5.3. Figure 36 illustrates the format of the card. It should be input after the birthrate input cards.

I. In-Migrant Household Data

The data utilized by ENSIM II to compute the characteristics of in-migrant households requires the use of a typology of analysis zones and distributions of in-migrant households that relate to this typology. The data used to define analysis zone typology is input on punched cards and defines the cut-off points for the two analysis zone characteristics used in the block typology. A single card is input for the cut-off points that relate to the

Figure 31

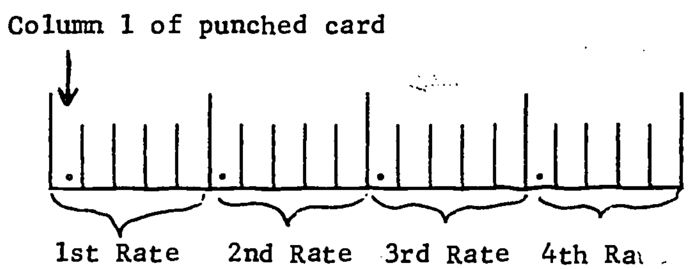
ORDER OF CHILD COHORT VECTORS IN CHILD COHORT DATA SET*



* Last Record is districtwide summary formed by taking sum of records for all analysis zones. Format of these records is 60F10.4.

Figure 32

FORMAT OF OUT-MIGRATION RATE CARDS*



* First rate is stored in columns 1 through 5, second 6-10, third 11-15, and fourth 16-20. Dots in specific columns signify the placement of decimal points. This format is used for all 18 out-migration rate cards. The format of these cards is 4F5.4.

76

Figure [illegible]

CARD ORDER OF OUT-MIGRATION RATES*

Card No.	Race/Ethnic	Income	Structure Type	No. of Children	0	1	2	3+	
18	Negro	High	Multiple Family						← Card 18
17	Negro	Med	Multiple Family						.
16	Negro	Low	Multiple Family						.
15	Negro	High	Single Family						.
14	Negro	Med	Single Family						.
13	Negro	Low	Single Family						.
12	Spanish	High	Multiple Family						.
11	Spanish	Med	Multiple Family						.
10	Spanish	Low	Multiple Family						.
9	Spanish	High	Single Family						.
8	Spanish	Med	Single Family						.
7	Spanish	Low	Single Family						.
6	Others	High	Multiple Family						.
5	Others	Med	Multiple Family						.
4	Others	Low	Multiple Family						.
3	Others	High	Single Family						.
2	Others	Med	Single Family						← Card 2
1	Others	Low	Single Family						← Card 1

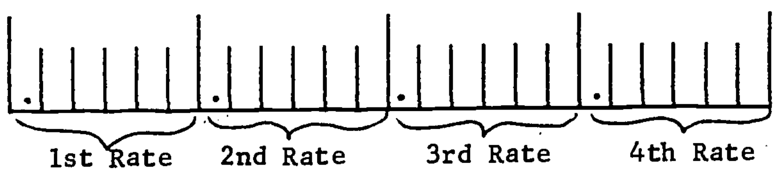
↖72 Out-Migration Rates

* The 72 out-migration rates indexed by their respective household characteristics.

Figure 34

FORMAT OF BIRTHRATE CARDS*

Column 1 of punched card



* Dots in specific columns relate to the placement of decimal points. Column 1 through 6 state first birthrate, 7-12 second, 13-18 third, and 19-24 fourth. This format is used for all 12 birthrate cards. The format of this card is 4F6.5.

Figure 35

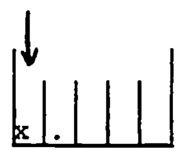
CARD ORDER FOR BIRTHRATE INPUT DATA*

Card No.	Race/Ethnic	Parity	Age 15-19	Age 20-29	Age 30-44	Age 45-up	
12	Negro	3+					← Card 12
11	Negro	2					•
10	Negro	1					•
9	Negro	0					•
8	Spanish	3+					•
7	Spanish	2					•
6	Spanish	1					•
5	Spanish	0					•
4	Others	3+					•
3	Others	2					•
2	Others	1					← Card 2
1	Others	0					← Card 1

* Birthrates are indexed by the characteristics of the households they relate to.

Figure 36

FORMAT OF FACTOR TO ALTER BIRTHRATES*



* This card should be input after birthrate cards. The format of this card is F5.3.

average value of housing characteristic and the average number of children characteristic. The FORTRAN format of this card (5F6.2) is illustrated in Figure 37. The average number of children characteristics are coded onto the card first, with the average value of housing characteristics following. In both cases, the lowest break-off points are coded first.

The distributions of in-migrant households are stored on magnetic disk. This data set is a FORTRAN formated data set in 144F10.4 format. Its record format is fixed and blocked (RECFM=FB) with a logical record length of 1440 (LRELL=1440) and any convenient block size. Thus, each distribution is stored as a separate record. The first record in this data set should be the distribution of households that will in-migrate into single family housing units located in analysis zones classified as type 1. The second record should be the distribution of households that will in-migrate to multiple family units in analysis zone type 1. The distributions for analysis zone type 2 follow these two distributions, then zone 3, and so on up to zone type 15. Figure 38 illustrates the order of these records. The distributions are read once for each projection and FORTRAN device number 14 (GO.FT14001) is reserved for them.

J. Average Housing Unit Value Data Set

The average housing unit value data set stores data on the value of housing units located in analysis zones. The data is stored in a single vector 287 cells long. The first cell should store data on the average value of units located in analysis zone 1, the second cell in analysis zone 2, and so on up to the last zone. Since the vector is restricted to a length of 287 cells, those cells not used should be coded to zero. If, for example, an application of ENSIM II involved the use of 120 zones, cells 121 through 287 would be coded to zero.

This data set is FORTRAN formated (287F10.4). FORTRAN device number 4 (GO.FT04F001) is reserved for it.

It should be noted that an appreciation factor for housing unit value can also be input by the user of the model. This appreciation factor is stored on a single card in FORTRAN format F5.3. Figure 39 illustrates the format of this card.

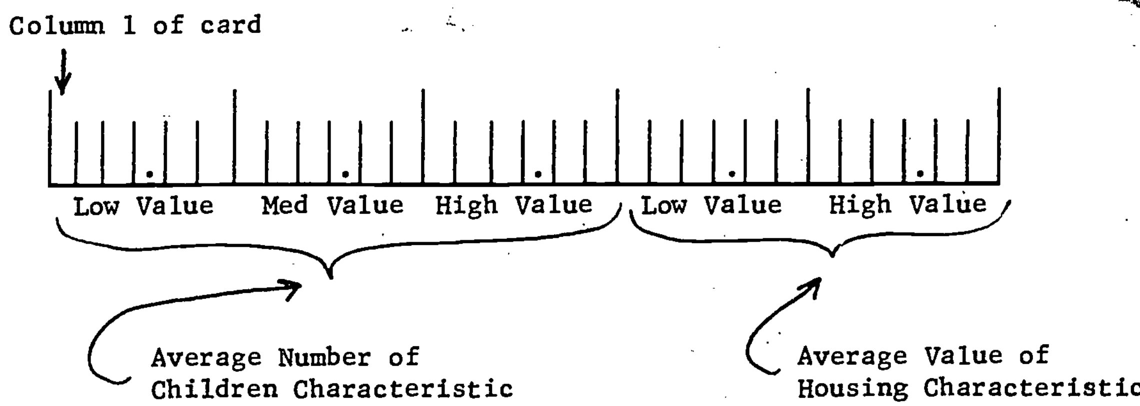
K. Income Transform Data

The ENSIM II user has the option of altering household income. The rates used to transform household income are input on punched cards--9 cards with 8 rates per card total the necessary 72 rates. The FORTRAN format of these cards is 8F5.3. Figure 40 illustrates the format of one of the income transform cards.

Since the income transform rates are used to move specific types of households from one income status to another, the placement of the rates on the input data cards must be in a particular order. This order is illustrated in Figure 41. The values in the cells shown in the diagram are the order in which the rates should be placed on the input cards. Rates 1 through 8

Figure 37

ANALYSIS ZONE TYPOLOGY INPUT CARD*



* Columns 1 through 6 store low value of average number of children characteristic, 7-12 med value, 13-18 high value. Columns 19-24 store low value for average losing value characteristic, 25-30 high value. Dots in specific columns relate to the placement of decimal points. Format of this card is 5F6.2.

Figure 38

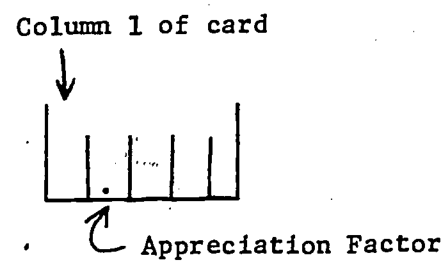
DATA FILE LAYOUT OF DISTRIBUTIONS OF IN-MIGRANT HOUSEHOLDS*

In-migrants to multiple family units, zone type 15	← 30th Record
In-migrants to single family units, zone type 15	← 29th Record
. . . . .	. . . . .
In-migrants to multiple family units, zone type 2	← 4th Record
In-migrants to single family units, zone type 2	← 3rd Record
In-migrants to multiple family units, zone type 1	← 2nd Record
In-migrants to single family units, zone type 1	← 1st Record

* Fortran Format is 144F10.4.

Figure 39

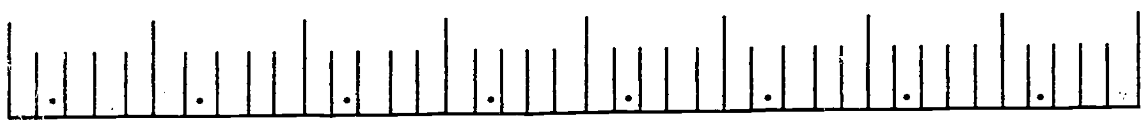
APPRECIATION FACTOR FOR HOUSING VALUE*



* Dot in specific column represents placement of decimal point. Format of card is F5.3.

Figure 40

FORMAT OF INCOME TRANSFORM DATA CARDS*



* Dots in specific columns represent the placement of decimal points. The FORTRAN format of this card is 8F5.3.

Figure 41

ORDER OF INCOME TRANSFORM RATES ON INPUT CARDS*

Present Income	Change to	Race/Ethnic	Others				Spanish				Negro			
		Age	15-19	20-29	30-44	45-up	15-19	20-29	30-44	45-up	15-19	20-29	30-44	45-up
Low	Med		1	2	3	4	25	26	27	28	49	50	51	52
	High		5	6	7	8	29	30	31	32	53	54	55	56
Med	Low		9	10	11	12	33	34	35	36	57	58	59	60
	High		13	14	15	16	37	38	39	40	61	62	63	64
High	Low		17	18	19	20	41	42	43	44	65	66	67	68
	Med		21	22	23	24	45	46	47	48	69	70	71	72

* Values in [illegible]ls relate to order of rates on cards. E[illegible]ght rates are placed on a card in FORTRAN format 8F5.3. The rates [illegible] this chart are indexed by the characteristics of the households they relate to and the changes they incur.

will go on the first card, 9 through 16 on the second card, and so on up to rates 64 through 72 on the 9th card. The rates shown in Figure 41 are indexed by the characteristics of the household they relate to and the changes they incur. Figure 42 illustrates upon which cards the rates should be placed.

L. Vacancy Rates

Vacancy rates for single and multiple family housing units are input on a punched card, FORTRAN format 2F5.4. The rate for single family units should appear first on the card, as shown in Figure 43.

M. Number of Years for Projection, Number of Zones in District

Two cards are input to the model to furnish data on the total number of analysis zones in the district under study and the total numbers of years for which the projection is being done. Total number of analysis zones is coded onto the first 3 columns of a punched card. If the total number of zones is less than 100, this number should be right justified. A second input card contains the number of years of the projection. This number is placed in the first 2 columns of its card. If the total number of years is less than 10, this number should be right justified. The FORTRAN formats of these cards are I3 and I2, respectively. Figure 44 illustrates these card formats.

N. Card Deck Set Up

Figure 45 illustrates the card deck set-up for the ENSIM II operation. Cards should be input in the following order:

1) Number of zones card
2) Number of years for projection
3) Out-migration rates
4) Appreciation rate for housing
5) Birthrates
6) Factor to alter birthrates
7) Rates to transform income
8) Vacancy rates
9) Block typology
10) Housing input data

Figure 46 summarizes some useful data for the input files stored on the direct access data storage device.

Figure 42

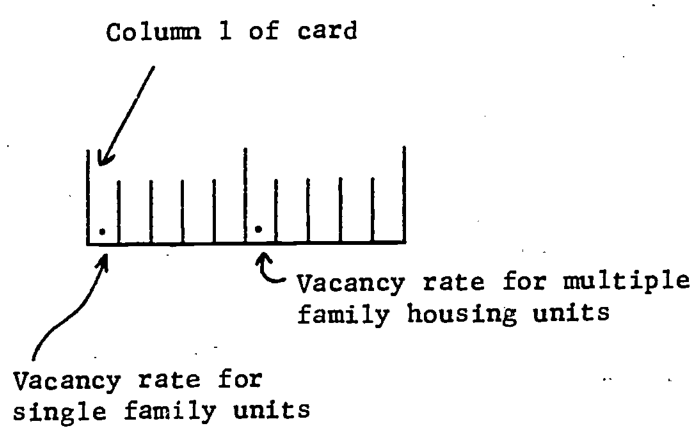
PLACEMENT OF RATES IN FIGURE 41 ON INPUT CARDS*

Present Income	Change to	Race/Ethnic	Others				Spanish				Negro			
		Age	15-19	20-29	30-44	45-up	15-19	20-29	30-44	45-up	15-19	20-29	30-44	45-up
Low	Med High		8 rates on first card				8 rates on fourth card				8 rates on seventh card			
Med	Low High		8 rates on second card				8 rates on fifth card				8 rates on eighth card			
High	Low Med		8 rates on third card				8 rates on sixth card				8 rates on ninth card			

* The values in the cells of Figure 41 indicate the placement of rates on cards. The rates with the lowest values in Figure 41 should be entered first on cards. Thus, the seventh card should start with rate 49 in Figure 41.

Figure 43

FORMAT OF VACANCY RATE CARD*

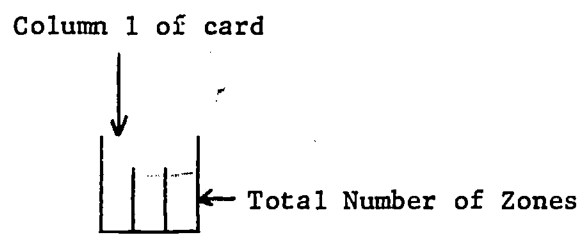


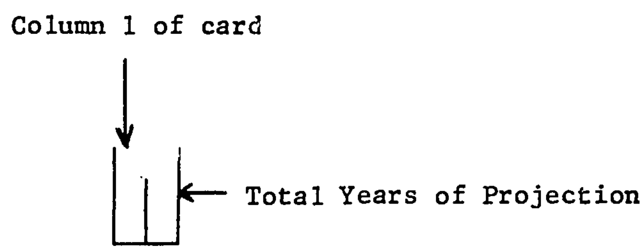
* Dots in specific columns relate to the placement of decimal points. The FORTRAN format of this card is 2F5.4.

89

Figure 44

FORMATS OF TOTAL NUMBER OF ZONES CARD
AND TOTAL NUMBER OF YEARS CARD*

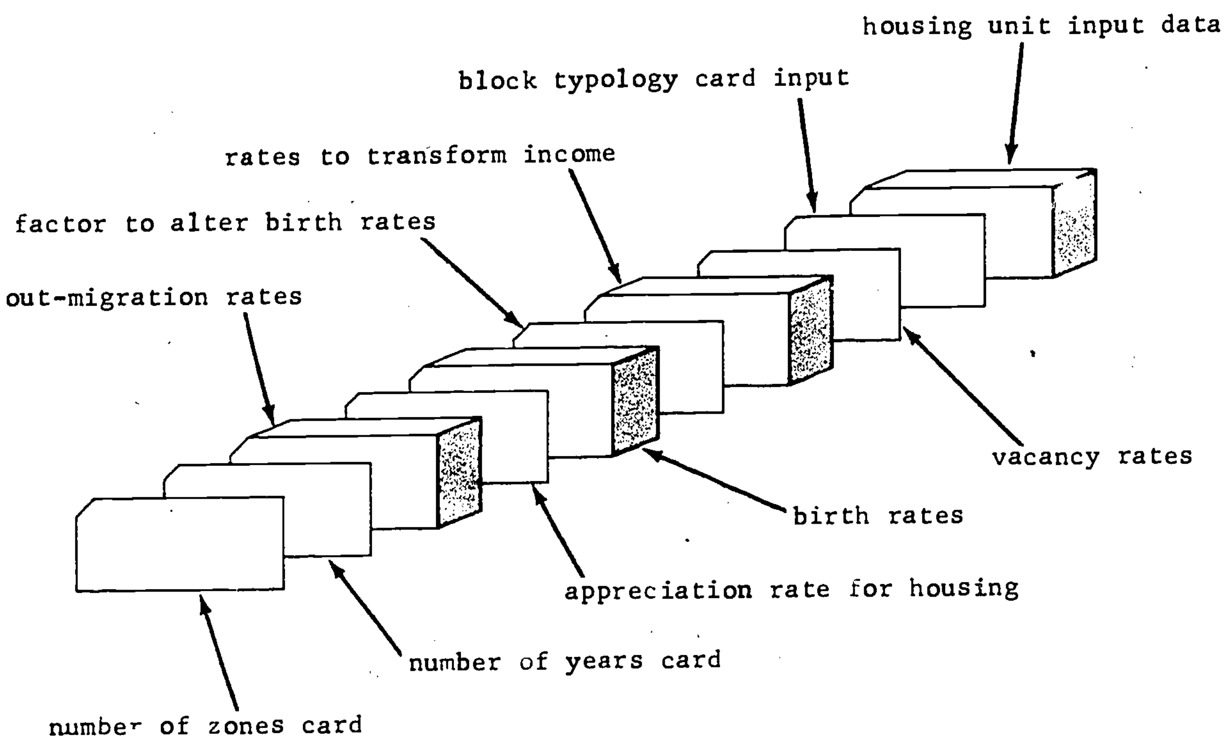




* Numbers should be right justified
and in I3 and I2 format.

90

FIGURE 45
CARD INPUT DECK SET UP FOR ENSIM II*



*Cards are placed in back of ENSIM II program deck in the order illustrated above.

Figure 46

SUMMARY DATA ON INPUT FILES STORED ON DIRECT ACCESS DEVICES*

Name of Data Set	Number of Records	RECFM	LRECL	BLKSIZE	Device Number	Format
Household Data File	N+1	VSB	1172	1172	GO.FT01F001	Unformatted
Average Number of Children Data File	1	FB	2304	2304	GO.FT08F001	288F8.5
Child Age Distributions	144	FB	180	Any Multiple of 180	GO.FT09F001	20F5.3, 4X
Average Housing Unit Value	1	FB	2870	2870	GO.FT04F001	287F10.4
Child Cohorts	N+1	FB	600	Any Multiple of 600	GO.FT10F001	60F10.4
Distributions of In-Migrant Households	30	FB	1440	Any Multiple of 1440	GO.FT14F001	144F10.4

* N = number of zones in district
VSB = variable spanned and blocked
FB = fixed and blocked
LRECL = logical record length
BLKSIZE = block size

SPECIFICATIONS OF OUTPUT DATA SETS AND SUMMARY PROGRAM

A. Introduction

Three data sets are output by the ENSIM II Model. The first of these stores data on the in and out-migration of children--this data relates to individual analysis zones. The second data set stores data on the number of children resident in each analysis zone and the final set stores districtwide summaries of the data stored in the first two sets. The following sub-sections will review the file layouts and physical characteristics of these data sets.

B. Analysis Zone Data on Migration

ENSIM II outputs data on children of junior high school age and younger (i.e., 0 through 13 years old). Data on migration into and out of analysis zones contains the following information:

1) Total out-migrating Spanish children
2) Total out-migrating black children
3) Out-migrating children by the following age groups:
    - a) ages 0-4
    - b) ages 5-8
    - c) ages 9-11
    - d) ages 12-13
4) Total in-migrating Spanish children
5) Total in-migrating black children
6) In-migrating children by the following age groups:
    - a) ages 0-4
    - b) ages 5-8
    - c) ages 9-11
    - d) ages 12-13

The migration data is stored in a sequential form. For the first time period of the projection, all of the data that relates to out-migration is stored first, beginning with zone 1, then zone 2, and so on up to the last analysis zone. The in-migration data follows and is stored in the same manner. Data for successive time periods of the projection follows the data for the first time period. Figure 47 illustrates this file layout, and Figure 48 illustrates the data layout of the out-migration and in-migration records. Data for each zone is stored in a separate record.

This data set is FORTRAN formated (6F10.4) and is fixed and blocked (RECFM=FB) with a logical record length of 60 (LRECL=60) and any convenient block size. FORTRAN device number 16 (GO.FT16F001) is reserved for this set.

C. Zonal Data on Children by Age

Zonal data on children by age includes:

1) The number of children by age for ages 1 through 13;

Figure 47

MIGRATION DATA FILE LAYOUT

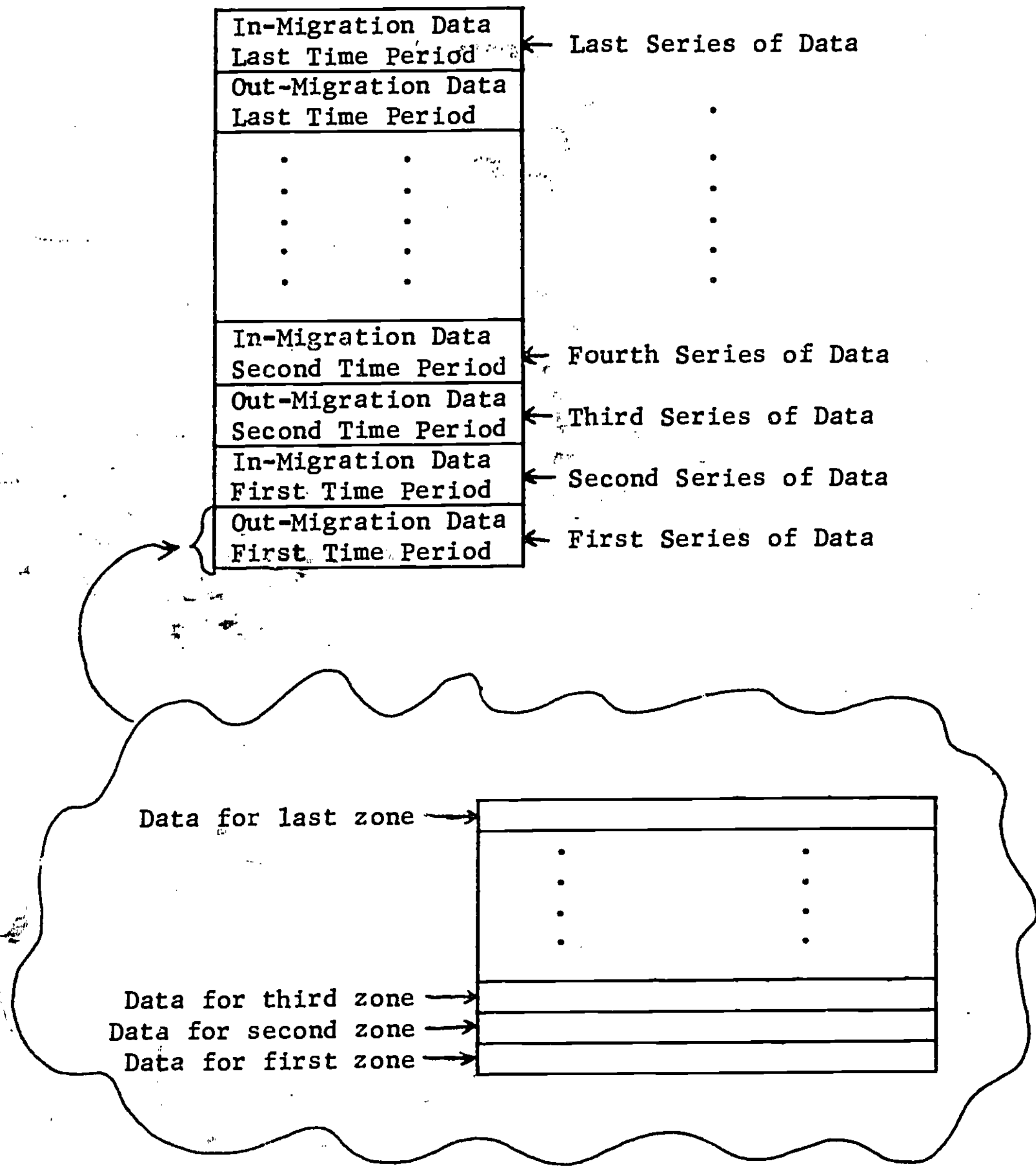


Figure 48

OUT-MIGRANT AND IN-MIGRANT CHILDREN RECORDS*

Out-Migrating Spanish Children	Out-Migrating Negro Children	Out-Migrating Children Ages 0-4	Out-Migrating Children Ages 5-8	Out-Migrating Children Ages 9-11	Out-Migrating Children Ages 12-13

Out-Migrant Record

In-Migrating Spanish Children	In-Migrating Negro Children	In-Migrating Children Ages 0-4	In-Migrating Children Ages 5-8	In-Migrating Children Ages 9-11	In-Migrating Children Ages 12-13

In-Migrant Record

* Record format of these records is 6F10.4.

2) The number of Spanish children by the following age groups:
   a) ages 0-4
   b) ages 5-8
   c) ages 9-11
   d) ages 12-13
3) The number of Negro children by the age group shown in 2 above.

Records for the first time period of the projection are stored first, the data for the second time period, second, and so on up to the last time period. Within each time period, the data for zone 1 is stored first, zone 2, second, and so on up to the last zone. This file layout is similar to the one illustrated in Figure 48. Figure 49 illustrates the location of the data elements in each record.

The format of this data set is 22F10.4. It is fixed and blocked (RECFM=FB) with a logical record length of 220 (LRECL=220) and any convenient block size. Device number 15 (GO.FT15F001) is reserved for this data set.

D. Output File for Summary Program

The output data file that provides districtwide summaries contains both migration data and summaries of children, by age, that are in the school district at the end of each time period of the projection. The records written to this data set are in exactly the same layout as the records written to the zonal data sets. Formats at 6F10.4,160X for the migration data and 22F10.4 for the children by age data. The reader should refer to Figures 48 and 49 for illustrations of these data records.

The records in this data set are fixed and blocked (RECFM=FB) with a logical record of 220 (LRECL=220) and any convenient block size. The first record in the set relates to the districtwide out-migration data for the first time period, the second contains districtwide in-migration data for the first time period, and the last record, the children by age summaries for the first time period. This pattern is repeated for each time period of the projection (see Figure 50). Device number 33 (GO.FT33F001) is reserved for this data set. The summary program uses device number 55 (GO.FT55F001) to read this data set.

The districtwide summary program also requires card input data containing data on children by age living within the school district at the beginning of the first time period of the projection. In addition, the base year and the number of years for which projections were made are entered by card. Figure 51 illustrates the data that the card input to the summary program should contain. Figure 52 illustrates the formats of these cards.

The first card should contain the total number of years the projection was run, plus a value of one to the initial input data on cards. The second card should contain the year the initial input data was collected, minus a value of one. Formats of these cards are I2 and I4 respectively. The following two cards, containing data on children by age, have formats of 11F6.0 and 8F6.0 respectively. The last card contains data on the race/ethnic composition of children in the district at the beginning of the first projection time period. The format for this card is 3F5.4.

Figure 49

LAYOUT OF CHILDREN BY AGE RECORDS*

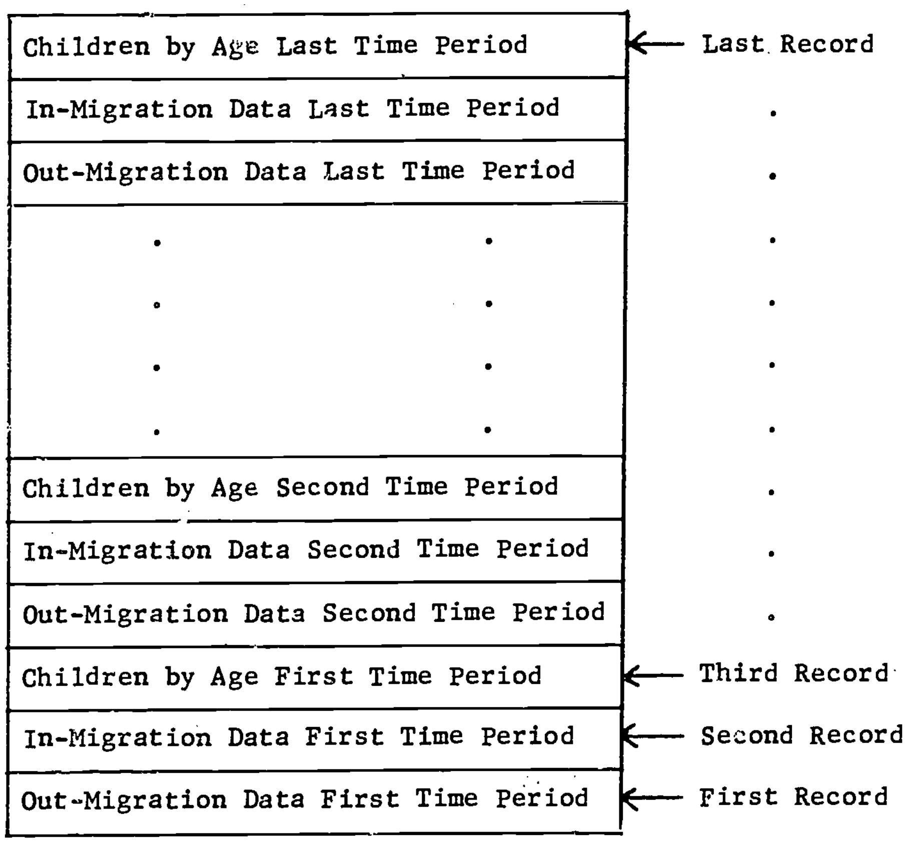
Total Children by Age														Spanish Children by Age Group				Negro Children by Age Group			
0	1	2	3	4	5	6	7	8	9	10	11	12	13	0-4	5-8	9-11	12-13	0-4	5-8	9-11	12-13

Record of 22 cells

* Format of this record is 22F10.4.

Figure 50

FILE LAYOUT OF DISTRICTWIDE SUMMARY DATA*



* Format of migration data is 6F10.4,160X. Format of children by age data is 22F10.4.

Figure 51

DATA ON SUMMARY PROGRAM INPUT CARDS

Card Number 1: [ ] ← Number of years projected for plus 1

Card Number 2: [ ] ← Initial year of data minus 1

Card Number 3

Children by Age										
0	1	2	3	4	5	6	7	8	9	10

Card Number 4

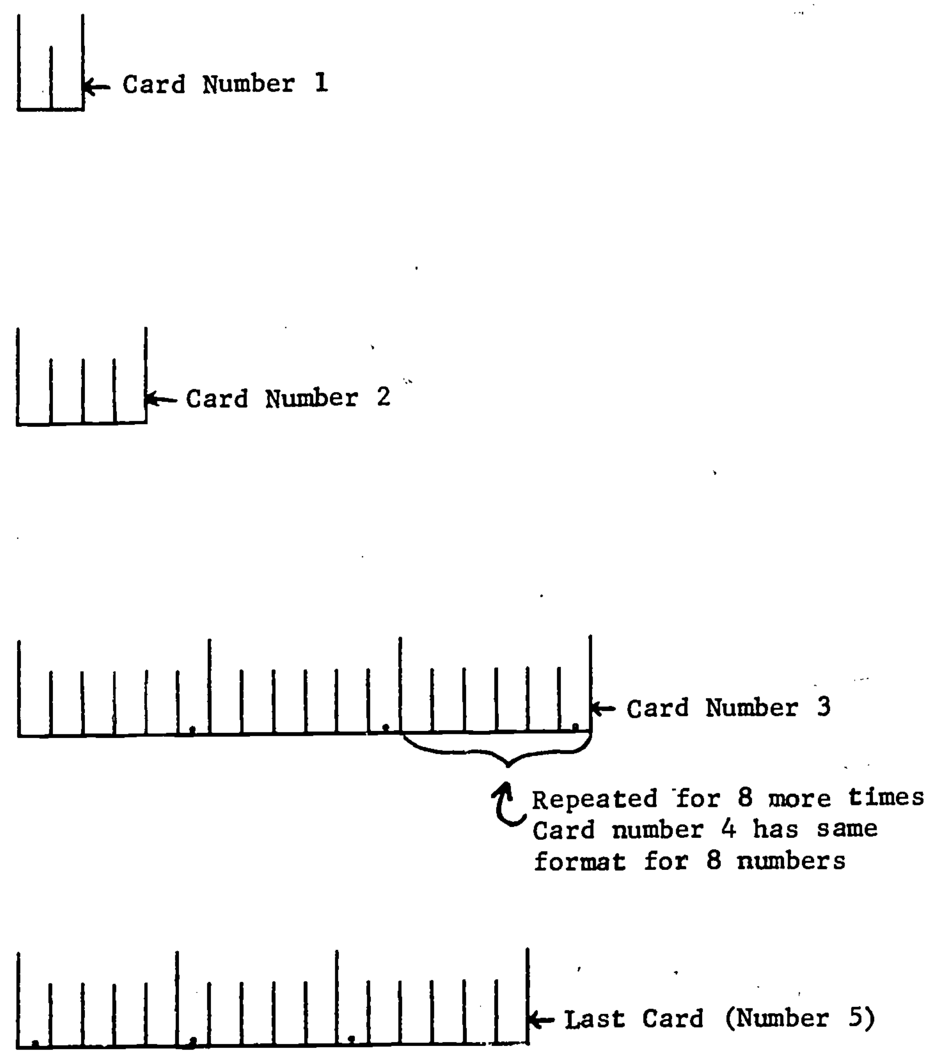
Children by Age				Children by Age Group			
11	12	13	Total	0-4	5-8	9-11	12-13

Card Number 5

% Children by Race		
Others	Spanish	Negro

Figure 52

FORMATS OF INPUT CARDS FOR OUTPUT SUMMARY PROGRAM*



* Card formats are: I2; I4; 11F6.0; 8F6.0; and 3F5.4 respectively. Dots in specific columns relate to the placement of decimal points.

V. PROGRAM OPERATIONS

A. Computer Environment

ENSIM II is designed to operate in an IBM 360/370 environment. Both the ENSIM II program and the summary program that accompanies it are written in FORTRAN IV. The physical characteristics of the data sets stored on magnetic disk, or some other direct access data storage device, are based on IBM standards. The use of direct access storage is suggested primarily for efficiency considerations. In particular, the storage of the household data file on such a device, in unformated form, incurs a considerable savings in processing times. This data set is repeatedly read and written by the ENSIM II program and must necessarily be processed in the most efficient manner. If necessary, all of the other data sets, including the output sets, can be conveniently stored on magnetic tape. Tapes are inexpensive and can later be processed for extraction of data relating to particular sub-areas of the district under study.

Conversion of the ENSIM II program to a non-IBM environment should present no difficulties. It is recommended that the user skim the listings of the program's subroutines to insure that the FORTRAN statements used in the program are compatible with available computer facilities. The current version of the ENSIM II program has been processed exclusively with a FORTRAN G compiler. In addition, the job control language needed to execute the program will have to be tailored to the standards of particular installations. The information reviewed above on the physical characteristics of the data sets used by ENSIM II should be sufficient to allow these modifications.

The names of the subroutines in the ENSIM II program are listed in Figure 53. Figure 54 illustrates the deck set-up used to process ENSIM II in a batch environment. Because the ENSIM II program is lengthy, it may be desirable to catalog it for batch processing. The main program must be the first program in the ENSIM II deck. The main program must preceed the ENSIM II subroutines. Figure 53 lists the subroutines in the order they are called for in processing.

B. Job Control Language

ENSIM II has been executed through the use of the FORTGCLD cataloged procedure. Both of these procedures are standard at most IBM installations. The physical characteristics of the data sets used by ENSIM II are reviewed in Figure 46. The necessary JCL used to process ENSIM II can be derived from the data in this summary. With the exception of the household data file, the initial disposition of all ENSIM II data sets is OLD. The final disposition of all ENSIM II data sets should be KEEP (DISP=OLD, KEEP).

The household data file is copied for processing by ENSIM II. The copied data set is then passed to the ENSIM II program. The final disposition of this data set is DELETE. The only other data set passed from one job step to another is the districtwide summary data set. This data set can also be deleted after its data is extracted by the summary program.

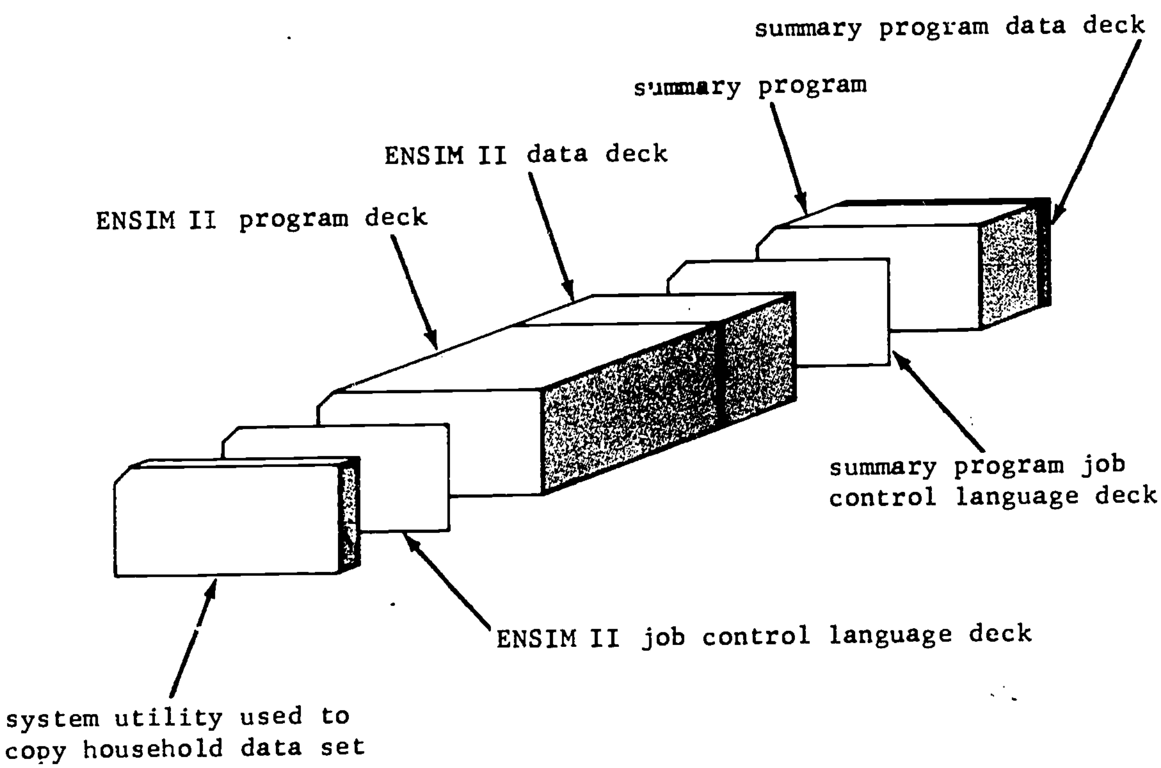
Figure 53

ENSIM II SUBROUTINES*

Order in Deck	Subroutine
(1)	MAIN (MAIN is main program)
(2)	OUTMIG
(3)	OUTCHD
(4)	SUBKID
(5)	AGEALL
(6)	ADJUST
(7)	INCOME
(8)	BLOCK
(9)	NEW
(10)	TYPE
(11)	INKID
(12)	CLEANI

* Order of subroutines in program deck relates to order in listing provided.

FIGURE 54
CARD DECK SET UP FOR ENSIM II BATCH PROCESSING



103

The districtwide summary program is also written in FORTRAN. This program is also executed through the use of the FORTGCLD or FORTGCLG procedures.

C. Storage and Processing Times

ENSIM II is designed to run at a region of 256K. This parameter is input to the cataloged procedure to insure that the proper amount of space is available for processing. At the present time, no overlay structure is used with the program.

In its test application, the ENSIM II model was run on an IBM 370/158 computer. Experience has shown that, with a system of 287 zones, the processing times run about one minute per each time period of the projection. Naturally, this figure will vary depending upon the computer used.

'I. DETAILED PROGRAM DESCRIPTION

A. List of Subroutines and Their Functions

Subroutine Name	Functions
(1) Main Program	Reads the following data items: number of years for projection, number of zones in school district, out-migration rates, child cohorts, average value of housing, factor to appreciate housing value, birthrates, factor to alter birthrates, distributions of in-migrant households, child's age distributions, average number of children per household data, and rates to transform household income. Calls subroutines OUTMIG, INCOME and CLEANI. Rewinds and closes all data sets.
(2) OUTMIG	Reads in vacancy rates. Reads household data and computes vacant housing units. Writes household data for further processing. Appreciates housing value, and adjusts the zonal child cohorts to districtwide marginal control totals. Computes the numbers and types of out-migrant households. Calls subroutine OUTCHD.
(3) OUTCHD	Computes the number and ages of out-migrant children. Summarizes out-migrant children data and writes this data to an output file. Calls subroutine SUBKID and AGEALL.
(4) SUBKID	Subtracts out-migrant children from child cohorts then returns to subroutine OUTCHD.
(5) AGEALL	Ages all households and children. Computes births and increases household sizes to account for these births. Passes the number of 20 years olds leaving households to subroutine ADJUST.
(6) ADJUST	Adjusts the household size characteristic to account for 20 year old children leaving home.
(7) INCOME	Reads household data processed by previous 6 subroutines. Reads in analysis zone typology data. Transforms household income. Calls subroutines BLOCK AND NEW. Adjusts household data to zonal control totals composed of counts of occupied housing units.

Subroutine	Functions
(8) BLOCK	Calculates block characteristics to be compared to analysis zone typology. Reads in new housing data.
(9) NEW	Computes numbers and types of new households. Calls subroutine TYPE and INKID.
(10) TYPE	Determines the analysis zone type of various analysis zones.
(11) INKID	Computes the numbers and ages of in-migrant children. Summarizes the in-migrant children data and writes this data to an output file.
(12) CLEANI	Summarizes children by age data and writes this data to an output file.

B. List of Arrays and Their Uses

Array	Dimension	Use
A	72	Stores out-migration rates.
H1, H2 and V3	288	Working space for household data.
VAL	287	Stores average housing value.
AVE	288	Stores average number of children data.
CHD	(20, 144)	Stores child's age distributions.
OUTKID & SKIDS	(20, 3)	Working space to compute out-migrant and in-migrant children.
RATES	48	Stores birthrates.
CHANGE	72	Stores income transformation rates.
BMAX	(144, 30)	Stores distributions of in-migrant households.
CHAR	(287, 6)	Stores analysis zone characteristics.
ITYPE	287	Stores analysis zone type code for various analysis zones.

Array	Dimension	Use
DBLC	(2, 15)	Stores summary data on the number of vacant housing units by analysis zone type.
AGE	22	Working space for summary data on children by age.
CHILD	(20, 864)	Stores child cohorts.

107