

Chapter 3

POPULATION SYNTHESIZER

CONTENTS

Chapter Outline	1
3.1 Overview	3
3.2 Terminology	5
3.3 Key Concepts.....	8
3.4 Major Data Inputs	10
3.4.1 Obtaining STF-3A Data.....	10
3.4.2 PUMS data.....	13
3.4.3 Obtaining MABLE/Geocorr Data:.....	17
3.4.4 Forecast Marginal File.....	21
3.4.5 Land Use Data File.....	22
3.4.6 Network data	22
3.4.7 Vehicle population.....	22
3.5 Major Data Outputs	25
3.6 Module Interfaces	30
3.7 Configuration Files	31
3.8 Algorithms	34
3.8.1 Introduction	34
3.8.2 IPF Algorithms.....	36
3.8.2.1 Traditional IPF Procedure for One Block Group.....	36
3.8.2.1.1 General Approach	36
3.8.2.1.2 Mathematical Statement for Traditional IPF Procedure	41
3.8.2.1.3 Practical Example in 3-Dimension	44
3.8.2.2 Two-Step IPF Procedure.....	50
3.8.3 Update Households for Forecast.....	67
3.8.4 Match a Household in Cross-Classification Table to a sample Household in PUMS	71
3.8.5 Locate the Households on the Network	74
3.8.6 Assign Vehicle Ownership for the Households.....	75
3.8.7 Assign a Unique Number to Each Household and Person in the Population.....	76
Appendix A	77
Appendix B	86
References.....	117

POPULATION SYNTHESIZER

Chapter Outline

The Population Synthesizer module in TRANSIMS discusses the creation of a synthetic regional population whose demographics closely match that of the real population in a study area. Figure-3.1 shows the main functions of the Population Synthesizer within the TRANSIMS framework and its relation to the other modules. The Population Synthesizer module creates synthetic households that represent every real individual household in the metropolitan region under study. These synthetic households have synthetic persons that match the Census Block Group population demographics obtained from the census bureau. The synthetic households are located on the network using land use data, so as to approximate regional population distribution. The household locations consequently determine some of the travel origins and destinations. In addition, the synthetic population's demographics form the basis for individual and household activities requiring travel.

The chapter discusses the following main functions of the Population Synthesizer:

1. The generation of synthetic households from census data at the block group level.
2. The development of associated demographic characteristics (e.g. age, sex, income, etc.) for each synthetic household.
3. The placement of each synthetic household on a link in the transportation network.
4. The assignment of vehicles to each household including vehicle emission type and their initial vehicle location on the network.

This chapter begins with an overview of the Population Synthesizer, definition of the terminology, key concepts, major data inputs and outputs, module interfaces, Iterative Proportional Fitting (IPF) algorithm, and values for configuration files.

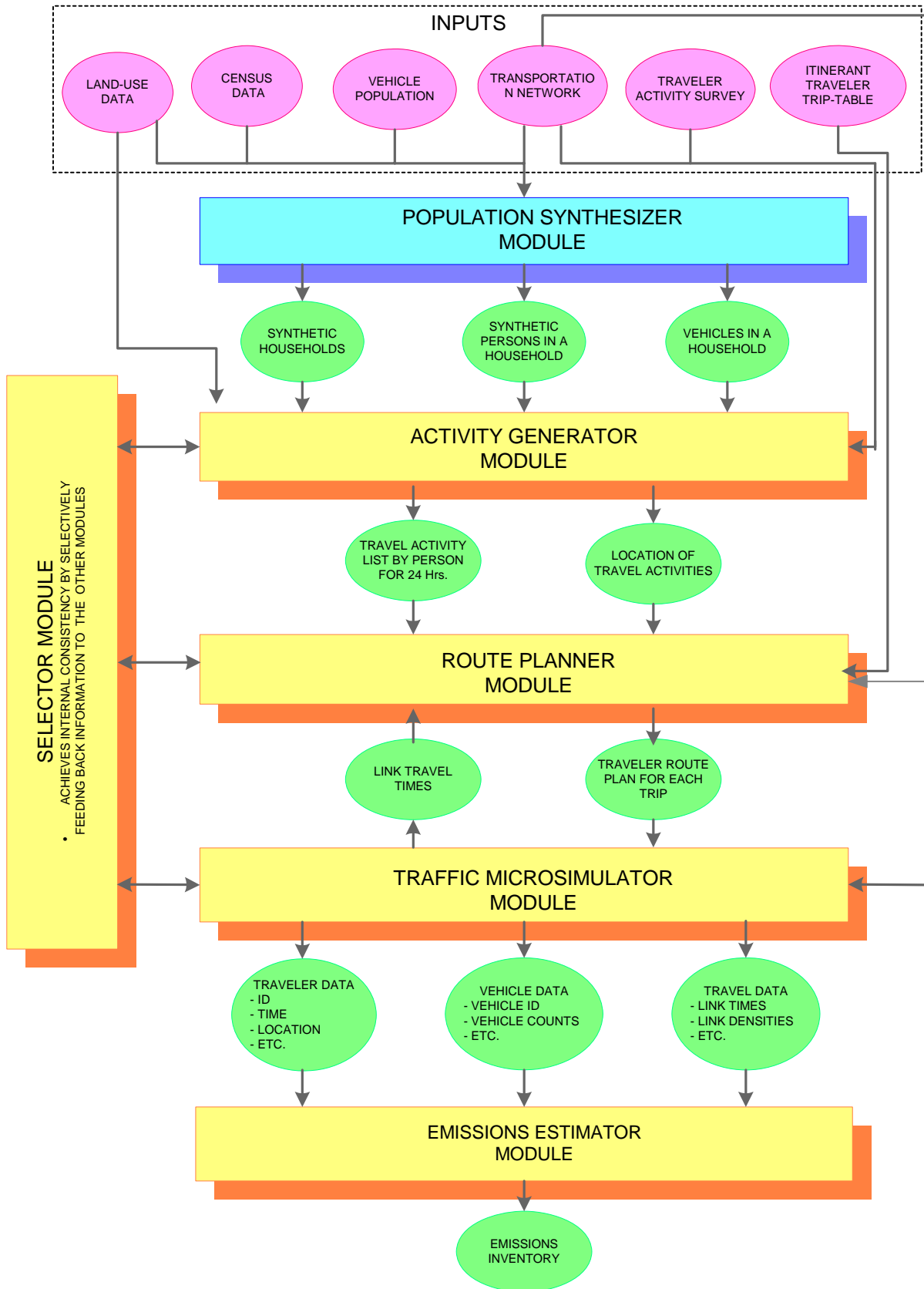


Figure-3.1: The Population Synthesizer module within the TRANSIMS framework

3.1 Overview

TRANSIMS is a microscopic model that simulates every single person and vehicle in a region. It then generates activities for them based on the demographic data available, and also simulates their movements on the transportation network. But when dealing with large areas, acquiring and inputting such large amounts of data proves quite infeasible. The population data available from the census bureau is not sufficient to duplicate the demographics of every single person living in the households present in the region. To address this problem TRANSIMS uses a synthetic population system that is designed to produce populations (family households, non-family households, and group quarters) that are statistically equivalent to actual populations when compared at the level of census block groups or higher. The Population Synthesizer uses the census data available and extrapolates it to create detailed synthetic households.

Figure 3.2 shows how the Population Synthesizer uses census data, land use data, and network data to generate a synthetic population of households that contain persons with demographic characteristics that resemble that of the Census data. It also shows how it locates the households within the TRANSIMS network, and how it assigns vehicles to these households using the vehicle population data.

Using the three census data sets (STF-3A, PUMS and MABLE/Geocorr data) a synthetic population is generated according to the **iterative proportional fitting (IPF)** algorithm developed by Beckman, Baggerly, and McKay [1996]. Individual households are placed at activity locations on the transportation network using land use data.

TRANSIMS uses a two-step **IPF** procedure to determine the proportion of households for each block group's demographic classification. Also there is a Household Locator (**BlockGroupLoc**) program in TRANSIMS that generates home locations for synthetic households on a transportation network and a Vehicle Assignment (**Vehgen**) program that creates a TRANSIMS vehicle file that contains information about the household's vehicles and their starting locations.

In TRANSIMS, travel results when a person engages in an activity at one location and time and must partake in another activity at a different location later in the day. The activities one engages in and their locations are often driven by household and personal demographics. For example, demographics such as the age of the individual, his/her income, gender, and whether he/she is employed greatly influence the person's daily itinerary. Therefore, in the methodology outlined here, not only are the individual households generated, but the household and individual demographics are also estimated.

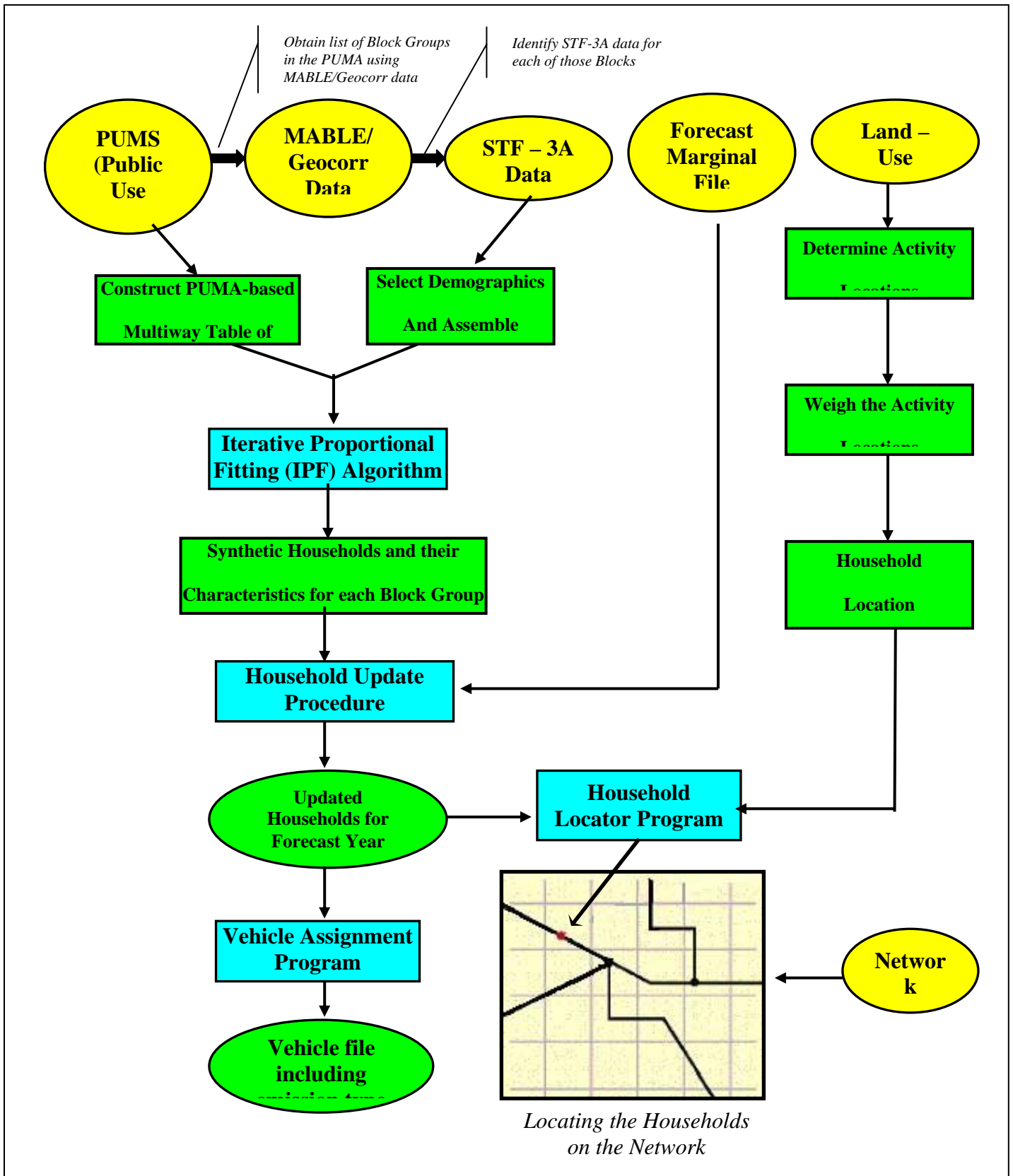


Figure-3.2: The Population Synthesizer takes in various types of census data to generate synthetic households, individuals, and vehicles

3.2 Terminology

- **Synthetic Population:** The creation of a model population that is statistically identical to the real population in the metropolitan area under study is termed as a Synthetic Population.

- **Census Tract:** A census tract is a small statistical subdivision of a county defined by the Census Bureau. The approximate size of a Census Tract is about 1500 households or 4000 persons. Census tract data allows a user to find population and housing statistics about a specific part of an urban area. A single community may be composed of several census tracts.

- **Census Blocks:** Census blocks are the smallest geographic units for which basic demographics are available from the Census Bureau. They are small areas bounded on all sides by visible features such as streets, roads, streams, and railroad tracks, and by invisible boundaries such as property lines, legal limits, and short imaginary extensions of streets and roads.

- **Census Block Groups:** Census block groups are areas, which include a variable number of census blocks and are used as the summary level for much of the Census Bureau's demographic data. Block groups typically have a population of about 1,000 people.

- **Zones:** Zones can be defined as geographic aggregations of individual households and business establishments in the study area.

- **Land Use Data:** Land use data are stored in TRANSIMS in the network activity location file. This file contains the locations of those places on the network in which activities may take place. Associated with these locations is a set of land use characteristics that indicate the type of activities that may take place at the location. These characteristics are also used to place each household within a block group at an activity location.

- **Household:** Each household in a synthetic population is classified as family, non-family, or individuals living in group quarters such as dorms. At a minimum, each household must contain at least one person. Family households contain one or more adults and possibly children. Household demographics vary in accordance with source data and study needs.

- **STF-3A data:** These files contain demographic summary tables from the Census data for small geographic areas such as, census tracts, or census block groups. These one-dimensional summary tables contain information on the distribution of the age of the householder, the number of workers in the family, and other demographic variables. However, these variables are not cross-classified against one another. Tables 3.1 and 3.2 show typical STF-3A data.

Table-3.1: Number of Family Households by Number of Workers for Census Tract-1036, Block Group-1, of Arlington County, VA






Number of Family households, <i>N</i>, with number of workers in households				
(1989)				
Workers	0	1	2	3 or Greater
<i>N</i>	30	57	224	14

Table-3.2: Number of family households by age of householder for Census Tract 1036, Block Group 1, of Arlington County, VA

Number of Family Households, <i>N</i>, by age of Householder in the Given Range							
Age	15 to 24	25 to 34	35 to 44	45 to 54	55 to 64	65 to 74	75 and over
<i>N</i>	32	82	72	54	33	24	28

- **PUMA and PUMS:** The Public Use Microdata Sample (PUMS) file consists of a 5% representative sample of almost complete census records (Long Form) from those households contained in a collection of census tracts or other small geographic census areas, which collectively is called a Public Use Micro Area (PUMA). A PUMA is constructed so that it contains approximately 100,000 individuals. These files are edited to protect the confidentiality of all individuals, but they have the information necessary to conduct effective research and analysis. An example household from PUMS is shown in Table-3.3.

Table-3.3: Example Household from PUMS

			
<u>Age</u>	26	26	7
<u>Income</u>	\$25k	\$18k	\$0
<u>Status</u>	Worker	Worker	Student
<u>Automobile</u>			

- **MABLE/Geocorr:** MABLE is an acronym for Master Area Block Level Equivalency file. MABLE refers to the name of the database that is used by the Geocorr engine to obtain the block groups that compose a PUMA.

Once we have identified an appropriate PUMA, we use the MABLE/Geocorr data to obtain a list of block groups within that PUMA. Then, for these block groups, we obtain summary statistics from STF-3A to construct a multidimensional table from the PUMS data. “Block” here refers to 1990 census blocks, the smallest geographic units used in tabulating the 1990 census. Block was chosen as the base unit for the application because the Census Bureau used these blocks as their “atomic unit” for all other census-based geographies in 1990. The MABLE/Geocorr search engine supports data maintained by the Center for International Earth Science Information Network of Columbia University (CIESIN), a non-profit, non-government organization.
- **Geocorr Engine:** The essential function of Geocorr is to produce geographic correspondence files, i.e. files that define the relationship between two sets of geographic code-groups. This geographic correspondence file generator is an application that will run on a cluster of powerful UNIX computers. And it can be accessed on the World Wide Web using a graphical user interface.

The data and search engine are available at the following Internet site:
<http://plue.sedac.ciesin.org/plue/geocorr>

The combined data set and the search engine yield a correspondence between PUMAs and census block groups.

Ref: MABLE/GEOCORR Tackling The Geographic Babel Problem

- **Forecast Marginal File:** When creating a forecast population, the Population Synthesizer requires a “Forecast Marginal File” as input. This file contains the forecast marginal distributions, like those given above in Table-3.1, for household size, householder age, and household income for each census tract and block group. This file must be created outside the TRANSIMS framework. Future forecasts can typically be obtained from transportation planning agencies.
- **Itinerant Travelers:** Each vehicle driver must have an entry in the synthetic population. Therefore, fictitious individuals are added to the population to represent those who travel on the network but do not live in the area undergoing study. They are known as “itinerant travelers” and are added to the synthetic population as single-person households, each with one vehicle.

3.3 Key Concepts

- **Iterative Proportional Fitting (IPF):** In the case of Census data, information of only certain characteristics is collected for the entire population (STF-3A). But, detailed characteristics are collected for a sample of the population (PUMS). Hence, in order to have detailed characteristics for the entire population the tabulations of this sample data need to be **adjusted** for the complete count. The adjustment procedure is carried out by the IPF algorithm, which uses the sample population data as an **estimate** of what would have been obtained by tabulating the characteristics for the entire population. The IPF routine estimates and refines the entries in the detailed sample PUMS data to make them exactly match the known STF-3A summary data while maintaining the correlation structure of the sample table.
- **Determining Household/Activity Locations:** TRANSIMS uses a modified IPF algorithm to generate synthetic households from the four data sets (STF-3A, PUMS, MABLE/Geocorr, Forecast Marginal File). Then, the land use data is used to place each synthetic household at an activity location along the transportation network. Each activity location is assigned to a link on the network and is associated with the land use characteristics that surround it. Multiple activities can be assigned to one activity location.
Example: An activity location on a link can represent the land use characteristics on one side of a street, whereas a second activity location could represent activities taking place on the other side of the street. Moreover, activity locations could represent individual buildings on a street, or one activity location could represent all activities that take place on the street. Determining a household location is depicted graphically in Figure-3.3.

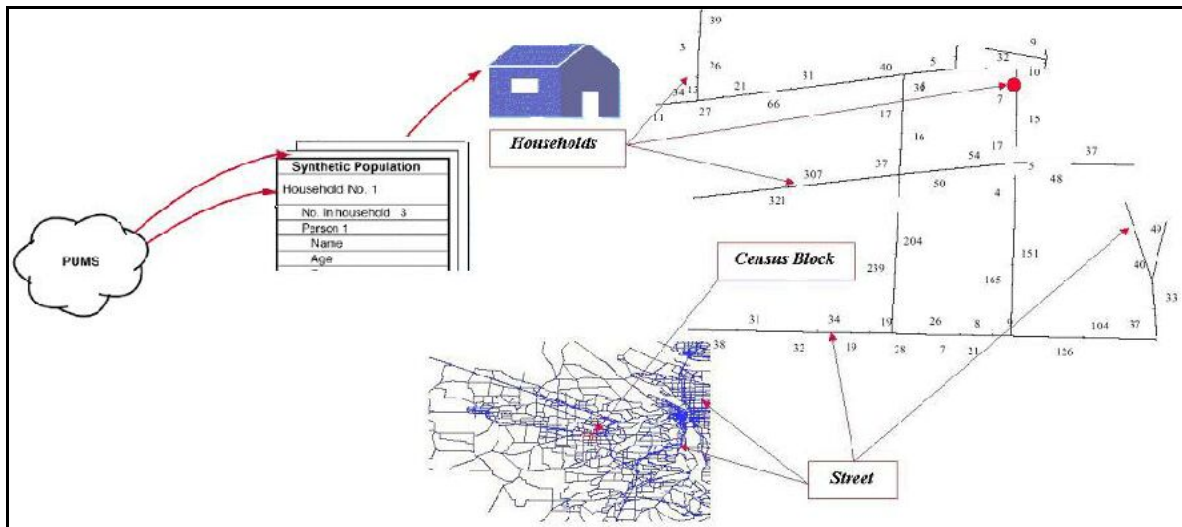


Figure-3.3: Creating households and placing them on the network involves randomly selecting actual PUMS households in accordance with the proportions derived from the Iterative Proportional Fitting method

Synthetic populations in essence “drive” the next module, the “Activity Generator,” which uses the data to create individual travel activities, such as work, school and shopping mall trips. Population data also can be used to categorize and filter population subsets used for various types of equity analyses.

- Vehicle Population:** Each synthetic household is created with a number of vehicles assigned to it from the PUMS data. These vehicles are given a unique number and identified as belonging to the household. One of the vehicle emission types, as shown in the module output section, is assigned to each vehicle in the population. This assignment is done at random according to a national distribution of vehicle emission types. Additionally, a starting location, one of the parking locations on the driving network, is assigned to each vehicle. Traditionally, this location has been the parking location that is closest to the household location. This information is written to the vehicle file.

Each vehicle represents an entry in a vehicle file. This file contains the TRANSIMS vehicle identification number, the household to which it is assigned, and the vehicle emissions type. The emissions type is used in the Emissions Estimator module to determine emissions. Emissions type reflects the operating condition of the vehicle, its type, and age.

To assign vehicle ownership, TRANSIMS currently uses the number generated from the synthetic population procedure using the PUMS. A more refined vehicle ownership model based on population demographics and network characteristics could be implemented and called in the framework after the synthetic population has been generated.

- Forecasting:** The population generated using the census data is a baseline synthetic population. Hence, TRANSIMS uses forecast marginal files to update this population to a future year for which the user is conducting the analysis. At this time, for

TRANSIMS Version 1.1, the update methodology requires forecast demographics to be in the exact format as used by the Portland Metropolitan Planning Organization. This will change in future releases of the program.

Iterative proportional fitting uses the correlation structure of the generated block group demographic tables and the STF-3A type forecast demographics for the update.

3.4 Major Data Inputs

To create and locate a virtual population on a network, the Population Synthesizer requires the following types of source data as shown in Figure-3.4:

- Census Data such as:
 - U.S. Census Bureau Summary Tape File 3A (STF-3A) data
 - U.S. Census Bureau Public Use Microdata Samples (PUMS)
 - Master Area Block Level Equivalency/Geographic Correspondence Engine (MABLE/GEOCORR), and
 - Forecast marginal demographic data.
- Land use data such as:
 - Activity Location and Characteristics (Discussed in Chapter 2)
- Network data such as:
 - Node and Link Files (Discussed in Chapter 2)
- Vehicle population data

3.4.1 Obtaining STF-3A Data

The PUMS and STF-3A data are available on CD-ROM from the Census Bureau. The STF-3A data for 1990 is also available to download from the Internet. The following steps are performed to generate an STF-3A data file:

1. Go to the US Census Bureau's 1990 Census Lookup web site: <http://venus.census.gov/cdrom/lookup>
2. From the *Choose a database to browse* click on the link for *STF-3A*
3. Select state(s) from the menu and click on *submit*.
4. In the *Choose a data retrieval option* select *Choose TABLES to retrieve* and click on *submit*.
5. Select the tables (summary tables for each category are mentioned below) that you wish to retrieve and then click on *submit*.
6. Choose HTML format in the *Data retrieval option* and click on *submit*.
7. This will open an output of the selected tables in the browser. After the data is downloaded, use your browser's *Save File* feature to save the data to disk in .txt

format. It is advisable to save the file with a meaningful name, such as *ArlingtonCounty_stf3a.txt*

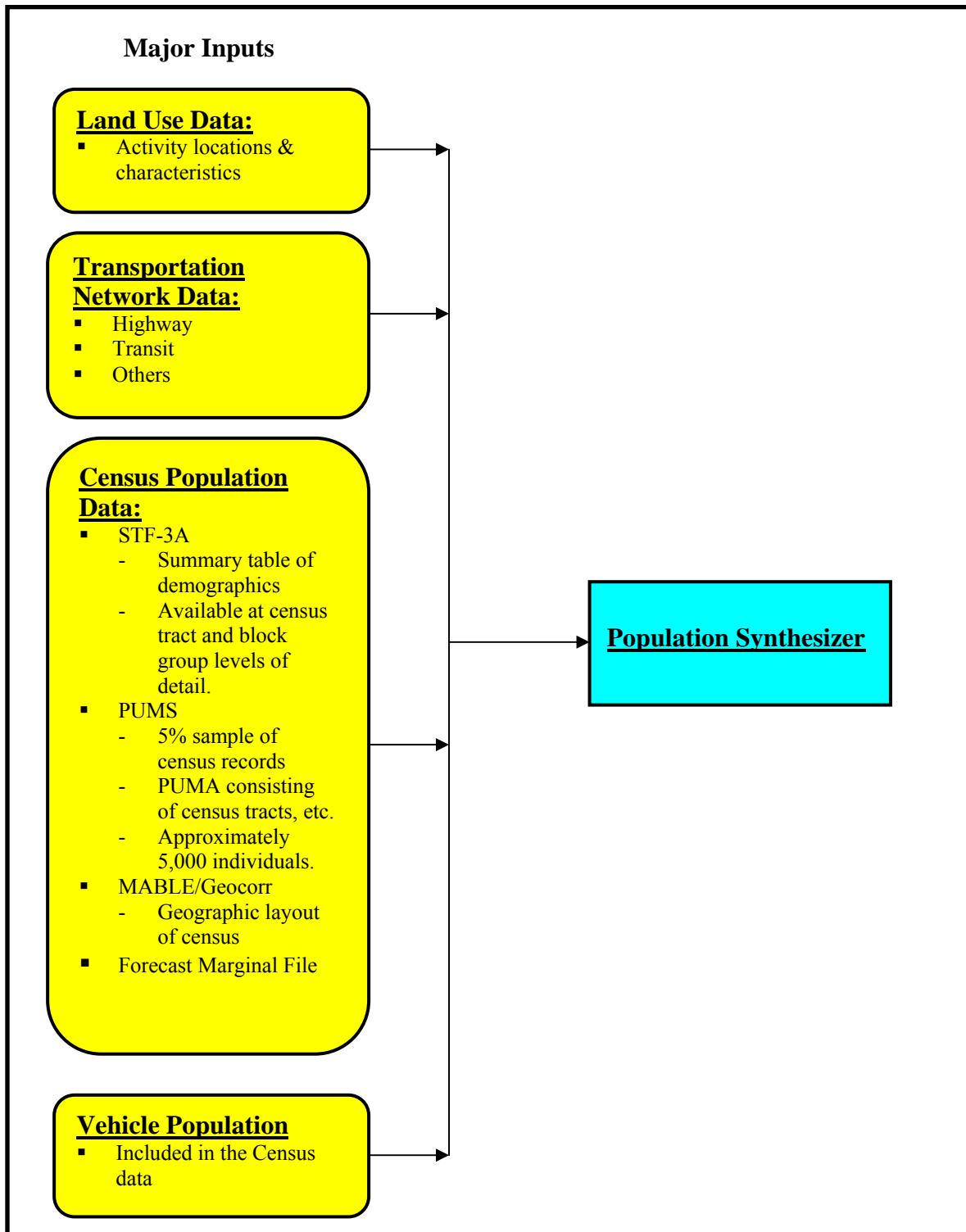


Figure-3.4: The major data inputs for the Population Synthesizer

Multiple demographics from STF-3A are used to create the synthetic population. Synthetic households in TRANSIMS are divided into three categories:

- 1 Family households—those households with two or more related persons
- 2 Non-family households—persons living alone or unrelated persons living together
and
- 3 Group quarters—dwellings such as prisons or college dormitories

Since travel activities depend on the type of household and its demographics, the population synthesizer generates households and group quarters for each of the three types. Family households are considered first.

The summary tables in STF-3A that concern family households are:

- 1) P24: Age of the Householder
- 2) P107: Family Income
- 3) P112: Number of Workers in the Family
- 4) P124 A&B: Poverty Status in 1989 (which is not used here), race of the householder, Family Type, and Presence and Age of Children.

Note: Not all categories that are given for the above STF-3A tables are used in the procedure. For example, there are 25 categories of income in table P107. These are collapsed to only seven categories in TRANSIMS. The Census STF-3A summary table P124 (A & B) is distributed by race of the householder. This table is used to create a single race by family class summary table. Data in the categories of below the poverty level and above the poverty level are added to give the resulting 12 family types. These are:

1. Married Couple: Children under age 5 only
2. Married Couple: All children between 5 and 17
3. Married Couple: Children under 5 and 5 to 17
4. Married Couple : No children under 18
5. Male Householder-No Wife Present: Children under age 5 only
6. Male Householder-No Wife Present: All children between 5 and 17
7. Male Householder-No Wife Present: Children under 5 and 5 to 17
8. Male Householder-No Wife Present: No children under 18
9. Female Householder-No Husband Present: Children under age 5 only
10. Female Householder-No Husband Present: All children between 5 and 17
11. Female Householder-No Husband Present: Children under 5 and 5 to 17
12. Female Householder-No Husband Present: No children under 18

Summary tables in STF-3A for non-family households are:

- 1) P17: Household Type and Relationship
- 2) P20: Race, Household Type, Presence and Age of Children. (The race of non-family householders can be derived from this table.)
- 3) P24: Age of Non-family Householder
- 4) P110: Non-family Household Income
- 5) P127: Poverty Status (not used here), Age of Householder, Household Type

There are only two summary tables for group quarters in STF-3A. These are:

- 1) P40: Group Quarters,
- 2) P41: Group Quarters, Age

*An Example output file obtained by following the above steps and selecting the data to collect for Block group #1, Census Tract #1036 of Arlington County in Virginia State is included in **Appendix A**.*

(Note: The summary tables for Age of Non-family and Family Householder are listed under the same field)

3.4.2 PUMS data

The PUMS data is available on the CD-ROM that can be obtained from the Census Bureau. The PUMS 'A' Sample contains records for 5 percent of the population, and the PUMS 'B' Sample contains records for 1 percent of the population. The PUMS 'A' sample is used for analysis in TRANSIMS.

The PUMS* contains housing unit and person records drawn from the full census sample. Each record is 231 characters in length and contains information coded from the 'long form' questionnaire collected by the census bureau. A housing unit record has an 'H' in the first character position, and a person record has a 'P' in the first position.

The file is hierarchical in structure. That is, a housing unit record is followed in sequence by records for the persons in that housing unit. In the example given below we can see that the first line starts with the character 'H' which represents a housing unit record, followed by two 'P' records which indicates that there are two persons living in that housing unit. In case we have a vacant unit, there will be no associated person records with it. Each housing unit and person record contains a weight field (HOUSWGT or PWGT1), which has a value that represents the record's relationship to the entire population. For person records in-group quarters, each have a housing unit record with zero weight.

Given in Table-3.4a and Table-3.4b is a list of the demographic options that are available for the households and for the persons living in each households with the allowed values for each demographic item that is used by the Population Synthesizer module.

(*: 1990 Census of Population and Housing; Public Use Microdata Samples (PUMS))

Data	Record Size	Beginning At Field	Description of Data	Allowed Values	Description of Values
RECTYPE	1	1	Record Type	H	Housing Record
PUMA	5	13	Public use microdata area (state dependent)	00100... 99999	PUMA code.
RHHINC	7	141	Household income	0000000	N/A* (GQ**/vacant/no income)
				-999999... 9999999	Total household income in dollars
RWRKR89	1	148	Workers in family in 1989	0	N/A*
				1	No Workers
				2	1 Workers
				3	2 Workers
				4	3 Workers
R18UNDR	1	162	Presence of person under 18 years in household	0	N/A* (No person under 18 in household/GQ**/vacant)
				1	1 or more person under 18 in household

* *Not Applicable*

** *Group Quarters*

Table-3.4b: Demographic options and allowed values available in the Household record

Data	Record Size	Beginning At Field	Description of Data	Allowed Values	Description of Values
RECTYPE	1	1	Record Type	P	Person Record
RELAT1	2	9	Relationship	00	Householder
				01	Husband/Wife
				02	Son/Daughter
				03	Stepson/Stepdaughter
				04	Brother/Sister
				05	Father/Mother
				06	Grandchild
				07	Other relative
				Non Related	
				08	Roomer/boarder/foster child
				09	Housemate/roommate
				10	Unmarried partner
				11	Other non-relative
				Group Quarters	
12	Institutionalized person				
13	Other person in group quarters.				
SEX	1	11	Sex	0	Male
				1	Female
AGE	2	15	Age	00	Less than 1 year
				01... 89	Age in years
				90	90 or more years old
WORK89	1	122	Worked last year (1989)	0	N/A (less than 16 years old)
				1	Worked last year
				2	Did not work last year

3.4.3 Obtaining MABLE/Geocorr Data:

Once the PUMA is identified we use the MABLE/Geocorr data to obtain a list of block groups within the PUMA. The MABLE/Geocorr data can be obtained from the Internet. Perform the following steps to generate a MABLE data file:

1. Go to the MABLE/Geocorr web site <http://plue.sedac.ciesin.org/plue/geocorr/>
2. In the Input Options select a state.

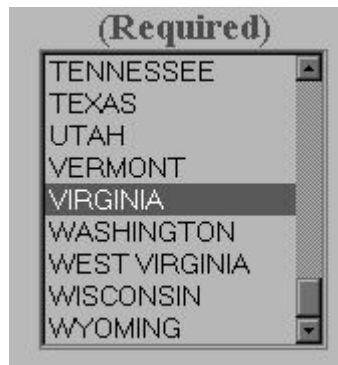


Figure 3.5: The Select State window

3. Select the source and destination Geocode data.

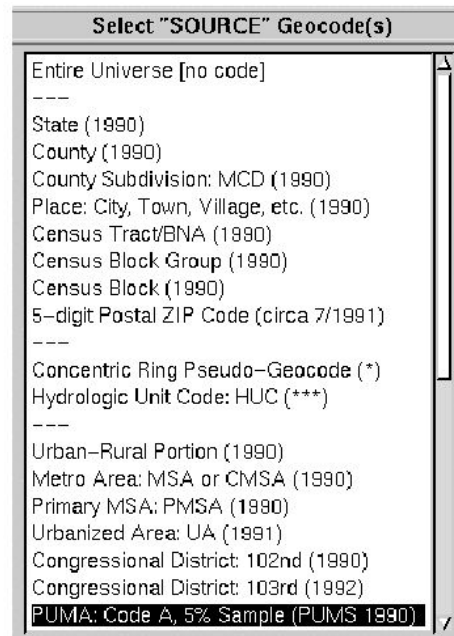


Figure-3.6: The Select "Source" Geocode(s) window

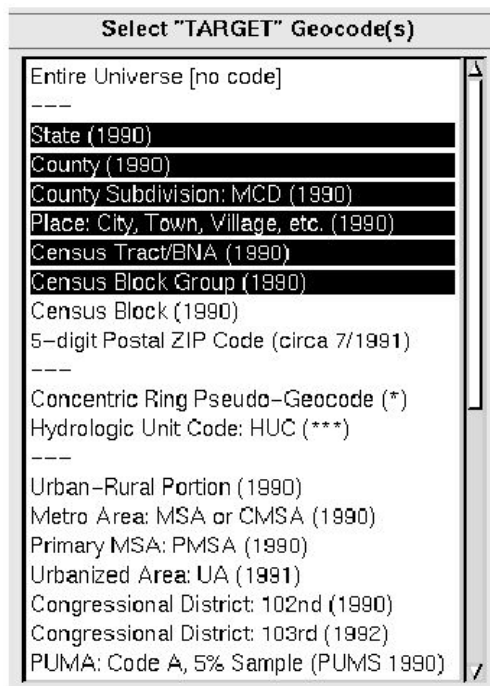


Figure-3.7: The Select “TARGET” Geocode(s) window

4. For weighting variable, select Population (1990 Census). Do **not** check the box *Ignore Census Block*.

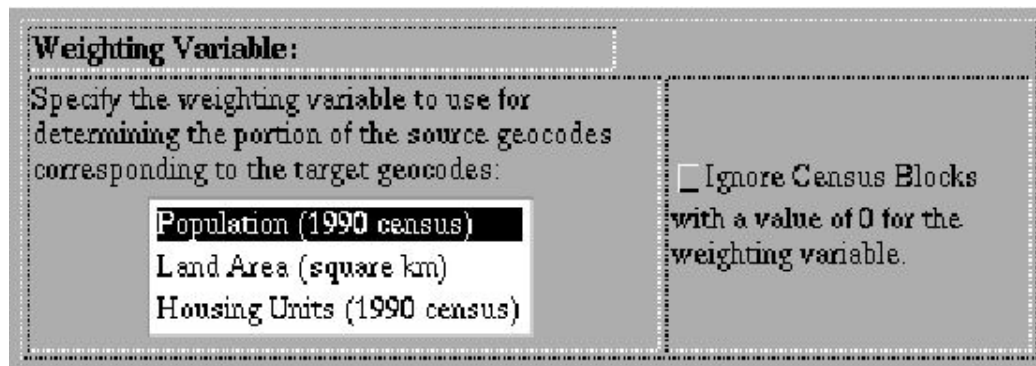


Figure-3.8: The Weighting Variable window

5. Select *Comma Separated Value File = Codes and Names*. No additional output options are required.

Comma Separated Value File

Generate a CSV file

Just Codes (No Names)
Codes and Names
 Just Names (No Codes)

Use tabs (not commas) as delimiter

• Process time for large areas may be several minutes.

Reset Defaults Run Request

Figure-3.9: The Comma-Separated Value File window

6. Click the [*Run Request*] tab. The web server will require several minutes to generate the file. When it is finished, a page with a *geocorr.csv* link will be displayed.
7. Click on the link to view the MABLE data.
8. After the data is downloaded, use your browser's *Save File* feature to save the data to disk. It is advisable to save the file with a meaningful name, such as *Virginia_MABLE.csv*. The file name extension must be *.csv*

An example of a Mable/Geocorr file obtained following the above steps is shown below and Table-3.5 gives the variables in the output file:

Note: *Variables in the output file are displayed from the left to the right order. ('C' indicates a character variable and 'N' a numeric)*

Table-3.5: Variables in the MABLE/Geocorr file

#		Variables	Description
1	C	STATE	State Code (FIPS)
2	C	APUMA	PUMA code for 1990 5% (A) Sample
3	C	COUNTY	ID for County
4	C	COUSUBFP	County Subdivision (FIPS)
5	C	PLACEFP	Place Code (FIPS)
6	C	TRACT	Census Tract / BNA
7	C	BG	Block Group
8	C	STAB	State Postal Code
9	C	CNTYNAME	County Name
10	C	COUSUBNM	County Subdivision Name
11	C	PLACENM	Place Name
12	N	POP	1990 Populations
13	N	AFACT	APUMA to block group allocation factor

Example file:

STATE	APUMA	COUNTY	COUSUBFP	PLACEFP	TRACT	BG	STAB	CNTYNAME	COUSUBNM	POP	AFACT
51	800	51013	99013	3000	1036	1	VA	Arlington VA	Arlington	1960	0.011
51	800	51013	99013	3000	1036	2	VA	Arlington VA	Arlington	690	0.004
51	800	51013	99013	3000	1036	3	VA	Arlington VA	Arlington	814	0.005

Diagram illustrating the mapping of variables in the example file to their descriptions:

- State Code (51)
- PUMA Code (800)
- County ID (51013)
- County Subdivision (99013)
- Place Code (3000)
- Census Tract (1036)
- Block Group (1, 2, 3)
- State Postal Code (VA)
- County Name (Arlington VA)
- County Subdivision Name (Arlington)
- '90 Population (1960, 690, 814)
- APUMA to block group allocation factor (0.011, 0.004, 0.005)

3.4.4 Forecast Marginal File

A file containing the marginal distributions of household size, householder age, and household income as a function of block group must be created or obtained if a population is to be forecast. This file has no header and each line represents a single census block group. Table-3.6 gives the fields that must be present in the given order in the Forecast Marginal file.

Table-3.6: Fields in the Forecast Marginal File

Data	Description
TRACT	The census tract number in a decimal format – the tract must be separated from the block group number by a decimal point; leading or trailing zeros are allowed.
BG	The census block group number.
H1	The number of households containing one person.
H2	The number of households containing two people.
H3	The number of households containing three people.
H4	The number of households containing four or more people.
A1	The number of households whose head is less than 24 years old or younger.
A2	The number of households whose head is between 25 and 54 years old.
A3	The number of households whose head is between 55 and 64 years old.
A4	The number of households whose head is 65 years old or older.
I1	The number of households with annual income at or below \$17,999
I2	The number of households with annual income between \$18,000 and \$28,999.
I3	The number of households with annual income between \$29,000 and \$40,499.
I4	The number of households with annual income at or above \$40,500.

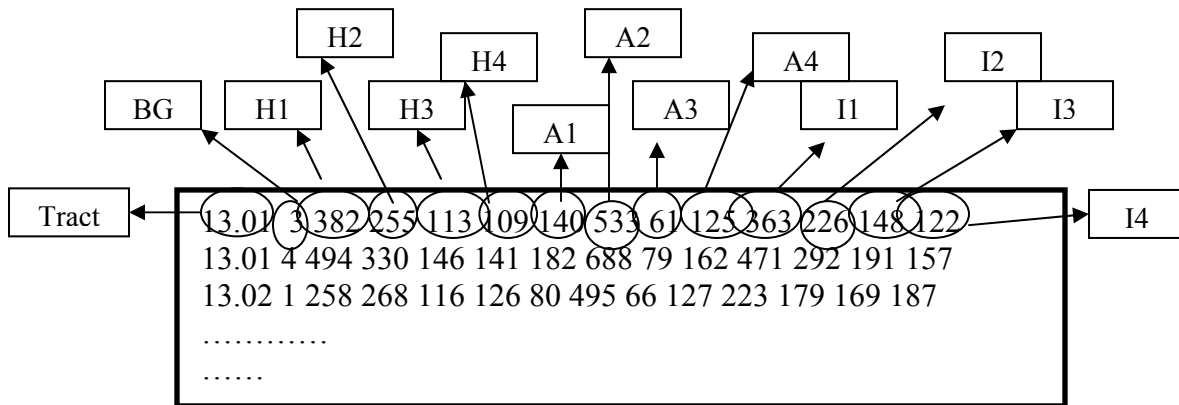


Figure-3.10: Example of a Forecast Marginal File

3.4.5 Land Use Data File

As discussed in Chapter 2, land use information is used to place each household within a block group at an activity location on a walk link. Land use data is stored in the network activity location file.

3.4.6 Network data

Also discussed in Chapter 2, network data, mainly the link and node files, are used by the Population Synthesizer to place each household on the transportation network.

3.4.7 Vehicle population

The number of vehicles owned by each household is given in the PUMS and is thus included in the synthetic population generated by the Population Synthesizer. Hence, each synthesized household will be assigned a certain number of vehicles according to the demographic information.

The Emissions Estimator module requires that the vehicles be identified by vehicle subtypes and by emissions type. This information is not present in the demographic data. Hence, the Population Synthesizer uses the data obtained from the national distribution of vehicle subtypes and emissions types in a model, which assigns vehicle types to each vehicle in the population. Table-3.8a contains the distribution of the 23 light-duty vehicle subtypes, and Table-3.8b gives the different categories of distribution of vehicle emission types for light-duty vehicles. These are discussed in detail in the Emissions Estimator module, which follows later.

Table-3.8a: Fields in vehicle type distribution file

Field	Description
FRACTION1	Fraction of LDVs pf subtype 1.
FRACTION2	Fraction of LDVs pf subtype 2.
FRACTION3	Fraction of LDVs pf subtype 3.
FRACTION4	Fraction of LDVs pf subtype 4.
FRACTION5	Fraction of LDVs pf subtype 5.
FRACTION6	Fraction of LDVs pf subtype 6.
FRACTION7	Fraction of LDVs pf subtype 7.
FRACTION8	Fraction of LDVs pf subtype 8.
FRACTION9	Fraction of LDVs pf subtype 9.
FRACTION10	Fraction of LDVs pf subtype 10.
FRACTION11	Fraction of LDVs pf subtype 11.
FRACTION12	Fraction of LDVs pf subtype 12.
FRACTION13	Fraction of LDVs pf subtype 13.
FRACTION14	Fraction of LDVs pf subtype 14.
FRACTION15	Fraction of LDVs pf subtype 15.
FRACTION16	Fraction of LDVs pf subtype 16.
FRACTION17	Fraction of LDVs pf subtype 17.
FRACTION18	Fraction of LDVs pf subtype 18.
FRACTION19	Fraction of LDVs pf subtype 19.
FRACTION20	Fraction of LDVs pf subtype 20.
FRACTION21	Fraction of LDVs pf subtype 21.
FRACTION22	Fraction of LDVs pf subtype 22.
FRACTION23	Fraction of LDVs pf subtype 23.

Table-3.8b: Distribution of vehicle emission types* for Autos as used in the Emissions Estimator

Vehicle Technology Category	Number of Vehicle Subtypes	
	Normal Emitting	High Emitting
Cars		
No Catalyst	1	1
2-way Catalyst	1	1
3-way Catalyst, Carbureted	1	1
3-way Catalyst, FI, > 50K miles, low power/weight	1	1
3-way Catalyst, FI, > 50K miles, high power/weight	1	
3-way Catalyst, FI, < 50K miles, low power/weight	1	
3-way Catalyst, FI, < 50K miles, high power/weight	1	
Tier 1, > 50K miles, low power/weight	1	1
Tier 1, > 50K miles, high power/weight	1	
Tier 1, < 50K miles, low power/weight	1	
Tier 1, < 50K miles, high power/weight	1	
Total Cars	11	5
Light Duty Trucks (LDT)		
Pre – 1979 (<=8500 GVW)	1	1
1979 to 1983 (<=8500 GVW)	1	1
1984 to 1987 (<=8500 GVW)	1	1
1988 to 1993 <=3750 LVW	1	1
1988 to 1993 >3750 LVW	1	
Tier 1 LDT2/3 (3751-5750 LVW or Alt. LVW)	1	1
Tier 1 LDT4 (6001-8500 GVW, >5750 Alt. LVW)	1	
Total LDT	7	5

* Barth, M., T. Younglove, T. Wenzel, G. Scora, F. An, M. Ross, and J. Norbeck (1997), “Analysis of Modal Emissions for a Diverse in-use Vehicle Fleet.” Transportation Research Record, No. 1587, Transportation Research Board, National Academy of Science, pp 73-84.

3.5 Major Data Outputs

The major output of the Population Synthesizer module is a synthetic population of households containing a set of information associated with each household and household locations within the TRANSIMS network as shown in Figure-3.11.

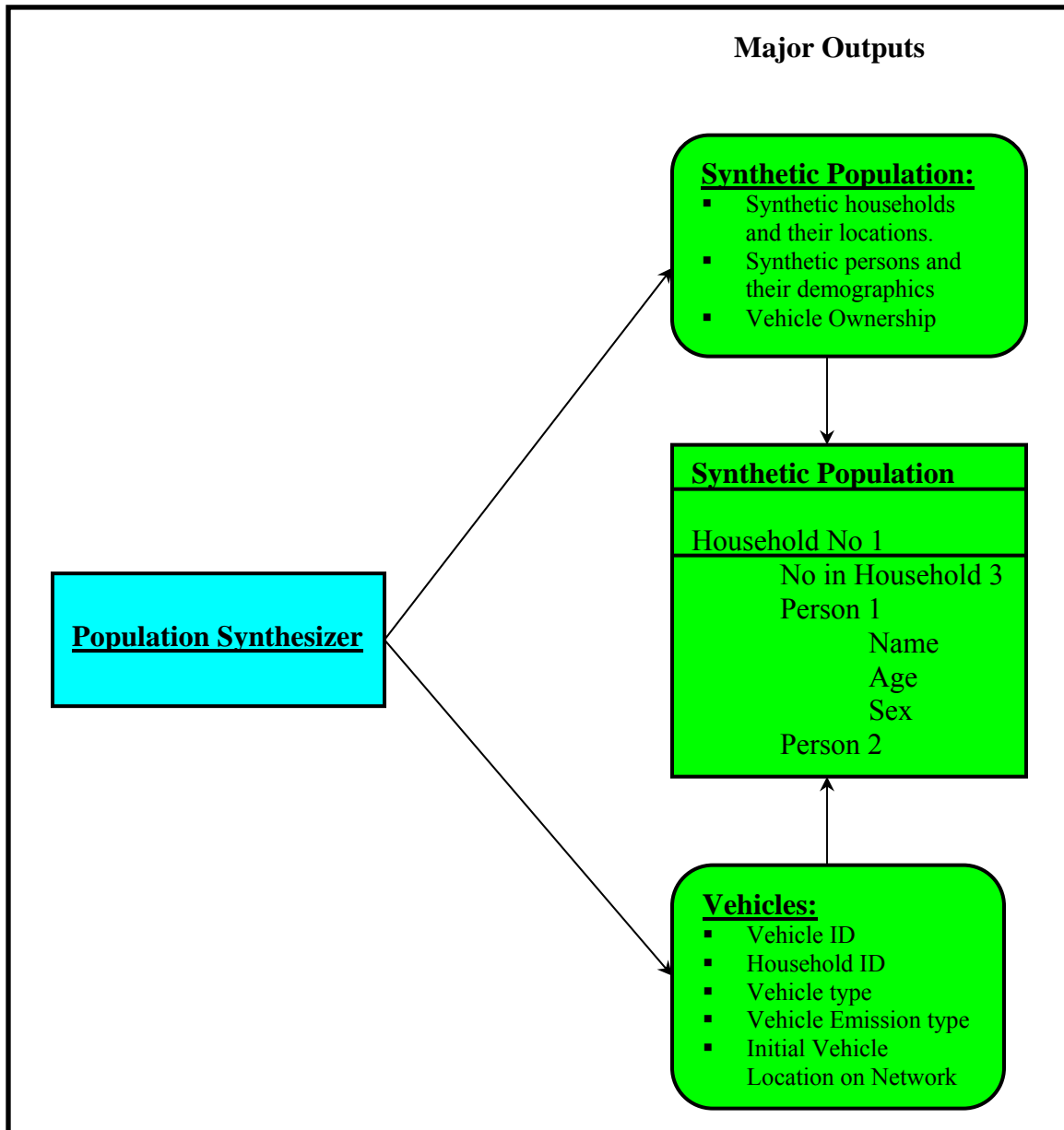


Figure-3.11: Major Data Outputs of the Population Synthesizer module in TRANSIMS

The format of these files is mainly composed of a household demographic line followed by a person demographic line for each person living in that household, as shown in Table-3.8, and in the example that follows. The first line of the file contains household demographic data such as tract ID, block group ID, household ID, number of persons living in the household, number of autos in the household, the home location ID (a value of -1 may be used if the home location is not yet known), the PUMS household ID, the presence of persons under the age of 18 in the household (0 = none; 1 = 1 or more present), number of workers in the family (0 = N/A; 1 = no workers; 2 = 1 worker; 3 = 2 workers; 4 = 3 or more workers) and the household income. The second line of the file contains person demographic information such as household ID, person ID, age of the person, relationship of that person to the householder, sex of the person (0 = male; 1 = female) and whether that person worked in the year 1989 (0 = N/A less than 16; 1 = worked; 2 = did not work). Table-3.8 below gives a description of the demographic characteristics of the population output file.

Table-3.8: Population File Specifications

Data	Description
TRACT	ID of the census tract
BG	Block Group ID
HHRD	Household record
HHID	Household ID
Persons Field	Number of persons in the household.
Autos Field	Number of autos in the household.
Home Location	Location of the household on the network.
PUMSHH	PUMS Household ID
R18UNDR	Presence of person under 18 years in household
RWRKR89	Workers in family (Household) in '89
RHHINC	Household Income
HHID	Household ID
Person ID	ID of the person in that household
AGE	Age of the Person
RELAT1	Relationship with other members in the family (Household)
SEX	Sex of the Person
WORK89	The person worked or did not work last year (1989)

Example of a household Output File

Household 1000 of tract 00001, block group ID 00002 and PUMS ID 17643, at location 1253 has 4 persons (one or more of them are under the age of 18), 2 autos, and two people who worked in 1989 and made an income of \$38800. The head of the household is a male of age 38 and has a wife of age 36. They both worked in 1989. They have two daughters of ages 7 and 4 respectively.

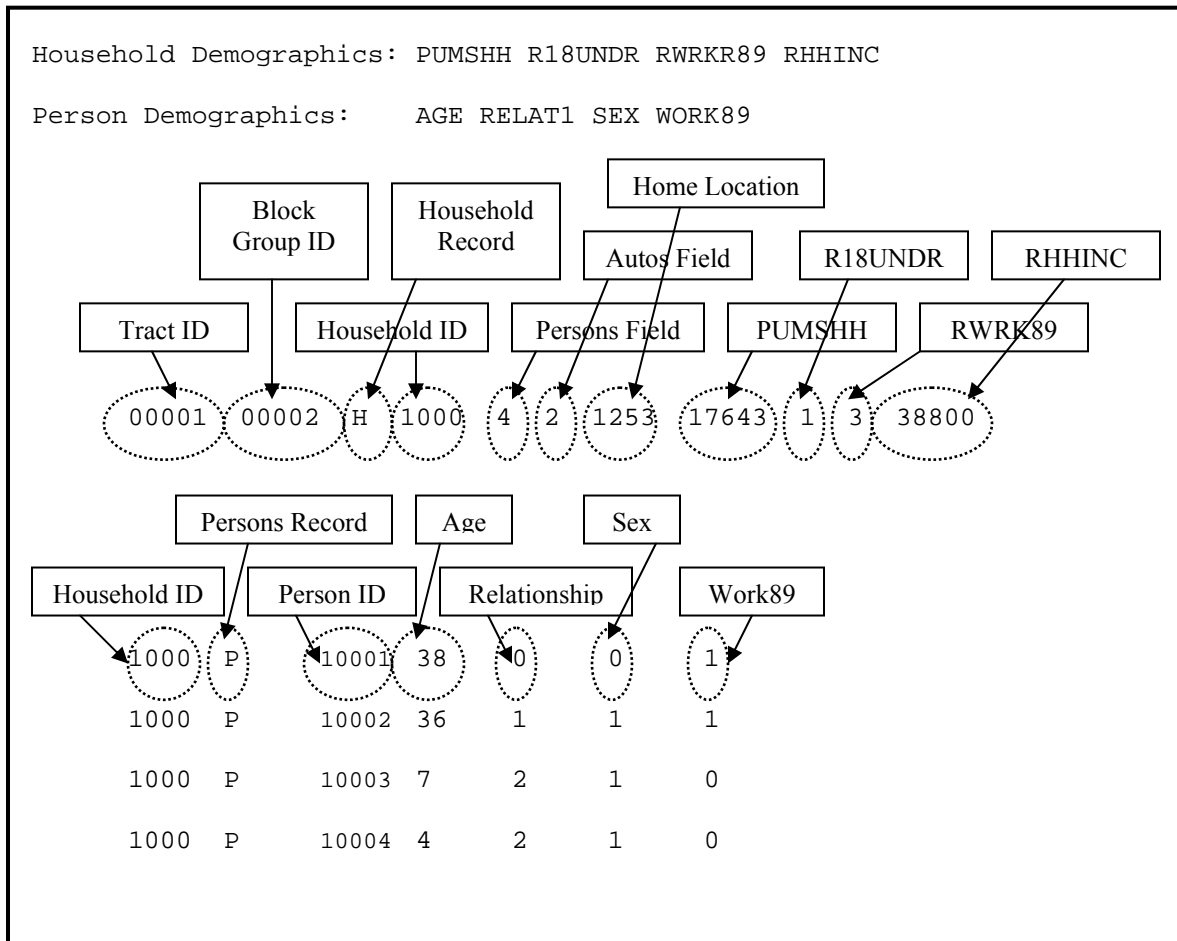


Figure-3.12: Example of a household Output File

Vehicle File

The TRANSIMS Population Synthesizer generates and assigns private vehicles to households while the Activity Generator assigns a set of possible vehicles to each member of a household. Freight and transit vehicles (and the plans for their drivers) are generated by separate utilities, but these must be included in the vehicle database. The vehicle IDs assigned by these utilities must be unique.

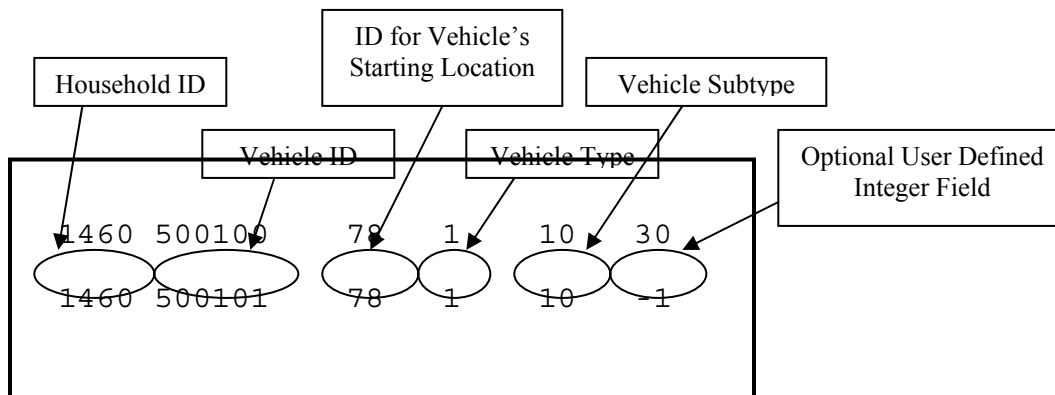
The vehicle file specifications are shown in Table-3.9. It gives the household ID, vehicle ID, the location of the vehicle on the network, the vehicle type and the vehicle subtype if desired.

Table-3.9: Vehicle File Specification

Data	Description
HHID	Household ID
VEHICLE	Vehicle ID
LOCATION	ID of vehicle's starting location
VEHTYPE	Vehicle type, which is one of the following: 1 = auto, 2 = truck, 4 = taxi, 5 = bus, 6 = trolley, 7 = streetcar, 8 = light rail, 9 = rapid rail, 10 = regional rail.
VSUBTYPE	Vehicle subtype, used for emissions. The vehicle subtype must correspond to a vehicle subtype specified in the vehicle prototype file.

Example of Vehicle File

In the example shown below, household 1460 has two vehicles (500100 and 500101); both start at the home location (78) and are of vehicle type auto (1). The user-defined emissions vehicle subtype (10) is the same for both vehicles. The optional user-defined integer field is present in this file. One integer field is an indicator of the maintenance level of the vehicle. Note that the second vehicle (500101) has unknown/unused value (-1) for the optional integer field.



3.6 Module Interfaces

An outline of the module interfaces for the Population Synthesizer is shown in Figure-3.13.

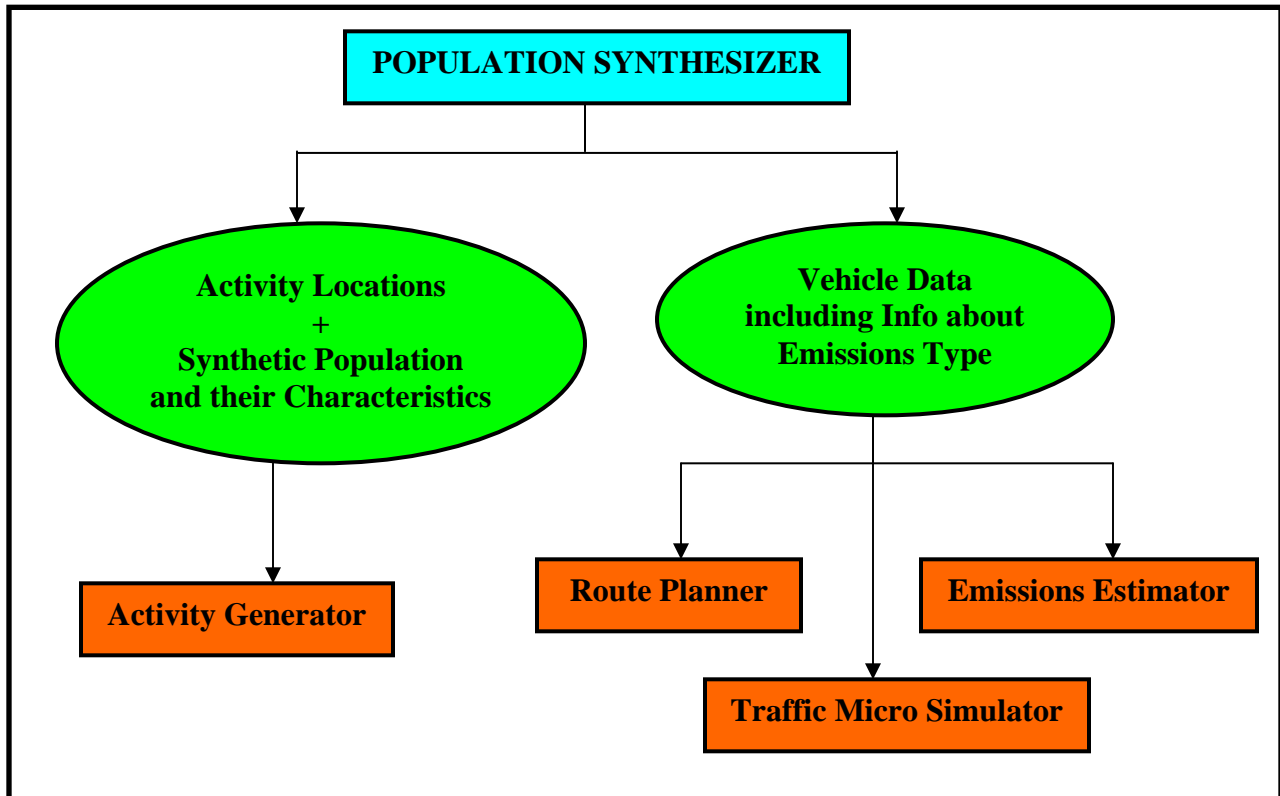


Figure-3.13: Flowchart of the module interfaces for the Population Synthesizer

Once households, work places, schools, stores and shops, etc., are placed at “activity locations” along the transportation network, the synthetic population is ready to “interact” with other TRANSIMS modules.

The output from the Synthetic Population module in essence “drives” the Activity Generator, which uses the synthetic households and their demographic characteristics as shown in Table-3.8, and the activity location data, to create individual travel activities, such as work, school, or shopping mall trips. Matching each household’s demographics against those from household travel and activity surveys, the Activity Generator builds an activity list for each individual in the household.

The Route Planner makes use of the Vehicle information from the Population Synthesizer, as shown in Table-3.9, and the activity information from the Activity Generator to find each traveler’s quickest (or minimal cost) route to each activity during the day.

Also the details of the vehicles generated and assigned to households by the Population Synthesizer are used by the Traffic Microsimulator module to execute each individual's travel plans throughout the transportation network, including their use of vehicles such as cars or buses.

The Emissions Estimator is designed to take data acquired during a traffic simulation, conducted by the Microsimulator, to calculate vehicle emissions. It uses the output Vehicle File (Table-3.9), generated by the Population Synthesizer in which each vehicle is represented as an entry in that file. This file contains the TRANSIMS vehicle identification number, the household to which it is assigned, and the vehicle emissions type. The emissions type is used in the Emissions Estimator module to determine emissions. It reflects the operating condition of the vehicle, its type and age.

3.7 Configuration Files

The following tables list all files used to run the Population Synthesizer and their descriptions, including the default values.

Synthetic Population File Configuration Keys

Configuration Key	Description
POP_NUMBER_HH	The number of households to be generated.
POP_BASELINE_FILE	Baseline synthetic population files—not located on a transportation network.
POP_LOCATED_FILE	Synthetic population file containing population located on a specific transportation network.
POP_STARTING_VEHICLE_ID	The first vehicle ID to be assigned. Other vehicles will be numbered sequentially from this starting number.
POP_STARTING_HH_ID	The first household ID to be assigned. Other synthetic households will be numbered sequentially from this number.
POP_STARTING_PERSON_ID	The first person ID to be assigned. Other synthetic persons will be numbered sequentially from this number.
POP_TRIPTABLE_FILE	The name of the file that contains the synthetic population generated by the Trip Table Activity Generator.
POP_NEAREST_BG_FILE	The name of the Tract/Block Group Substitution file that contains information about the nearest tract/block group for block groups that have no activity locations on the transportation network. Used by the Block Group Locator.

Population Synthesizer Configuration Keys

Configuration Key	Description
SYNPOP_BASE_DIRECTORY	<i>\$TRANSIMS_HOME</i>
SYNPOP_STF_INFO_DIRECTORY	<i>\$TRANSIMS_HOME/source/SYNPOP/Par ep2/stf</i>
SYNPOP_STF_DATA_DIRECTORY	The directory in which the STF3A dBase files is located.
SYNPOP_MABLE_FILE	The directory in which the MABLE output file is stored.
SYNPOP_PUMS_DIRECTORY	The directory in which the PUMS data are stored.
SYNPOP_MARGINALS_FILE	The location of the Forecast Marginals file.
SYNPOP_STATE	The two-letter abbreviation (lower case) of the state of interest.
SYNPOP_PUMAS	The list of five-digit PUMA numbers to be processed (separated by semicolons).
SYNPOP_HOUSEHOLD_DEMOGRAPHICS	The list of household PUMS fields to be placed in the population output files (separated by semicolons). See the file <i>\$TRANSIMS_HOME/data/synpop/docs/pumsusdd.txt</i> for a complete list of the possible fields.
SYNPOP_PERSON_DEMOGRAPHICS	The list of person PUMS fields to be placed in the population output files (separated by semicolons). See the file <i>\$TRANSIMS_HOME/data/synpop/docs/pumsusdd.txt</i> for a complete list of the possible fields.
SYNPOP_BASE_PREFIX	The file name prefix for the base-synthesized population output files. No base-year output will be generated if this key is blank.
SYNPOP_FORECAST_PREFIX	The file name prefix for the forecast-synthesized population output files. No forecast output will be generated if this key is blank.
SYNPOP_TEMP_DIRECTORY	The directory in which temporary working files will be placed.
SYNPOP_KEEP_TEMP_FILES	Whether to retain the working files after the population synthesis is complete (1 = yes, 0 = no).
SYNPOP_RANDOM_SEED	The random number seed (integer).

BlockGroupLoc Configuration Keys

Configuration Key	Description
POP_BASELINE_FILE*	The name of the file containing the baseline population.
POP_LOCATED_FILE*	The name of the file where the located population will be written.
ACT_HOME_HEADER	The user data column header in the network activity location file used to specify single family home locations. Default = <i>HOME</i> .
ACT_MULTI_FAMILY_HEADER	The user data column header in the network activity location file used to specify multifamily home locations. If not specified, multifamily user data from the activity location file is ignored.
ACT_TRACT_HEADER	The user data column header in the network activity location file used to specify the census tract. Default = <i>TRACT</i> .
ACT_BLOCKGROUP_HEADER	The user data column header in the network activity location file used to specify the block group. Default = <i>BG</i> .
POP_STARTING_HH_ID	The number from which the generated households will be sequentially numbered. Default = 1.
POP_STARTING_PERSON_ID	The number from which the generated persons will be sequentially numbered. Default = 101.
NET_DIRECTORY*	The directory where the network files reside.
NET_NODE_TABLE*	The network node table name.
NET_LINK_TABLE*	The network link table name.
NET_ACTIVITY_LOCATION_TABLE*	The network activity location table name.
POP_NEAREST_BG_FILE	The name of the Tract/Block Group Substitution file that contains information about the nearest tract/block group for block groups that have no activity locations on the transportation network.

* Configuration Keys required for *BlockGroupLoc*. All others are optional and will use default values.

3.8 Algorithms

3.8.1 Introduction

The Population Synthesizer module utilizes four main mathematical procedures, which are:

- a) The Iterative Proportional Fitting (IPF) procedure
- b) The Household Update procedure,
- c) The Household Locator program, and
- d) The Vehicle Assignment program.

The IPF algorithm utilizes the census data available in PUMS and STF-3A to create a synthetic population of households that imitate the demographics of the real households in a study area as shown in Figure-3.14. Once the synthetic households are updated for future forecasts using the Household Update procedure, they are individually placed on the network using a Household Locator program that makes use of the available land use and network data. Vehicles are assigned to each synthetic household through the IPF algorithm using the PUMS data. However, the Vehicle Assignment program randomly assigns to each vehicle a vehicle type and an emission type characteristic based on a national distribution of vehicle characteristics.

Before explaining the algorithms, it is important to revisit and clarify the relationships among PUMA, PUMS, Block Group, and STF-3A. A metropolitan area is comprised of many PUMA. Each PUMA is comprised of many block groups. The complete census records (long forms) are sampled for the households within the PUMA, and are called PUMS. The PUMS usually consists of a 5% representative sample of all the households in a PUMA as shown in Figure-3.15. Each block group has its own summary of marginal demographic data for all the households referred to as STF-3A obtained from census short form. The STF-3A does not provide the cross tabulation of the demographic data. The objective of the IPF algorithm is to use the PUMS data to create a detailed cross-tabulated demographic data, which is not provided in the STF-3A for each block group.

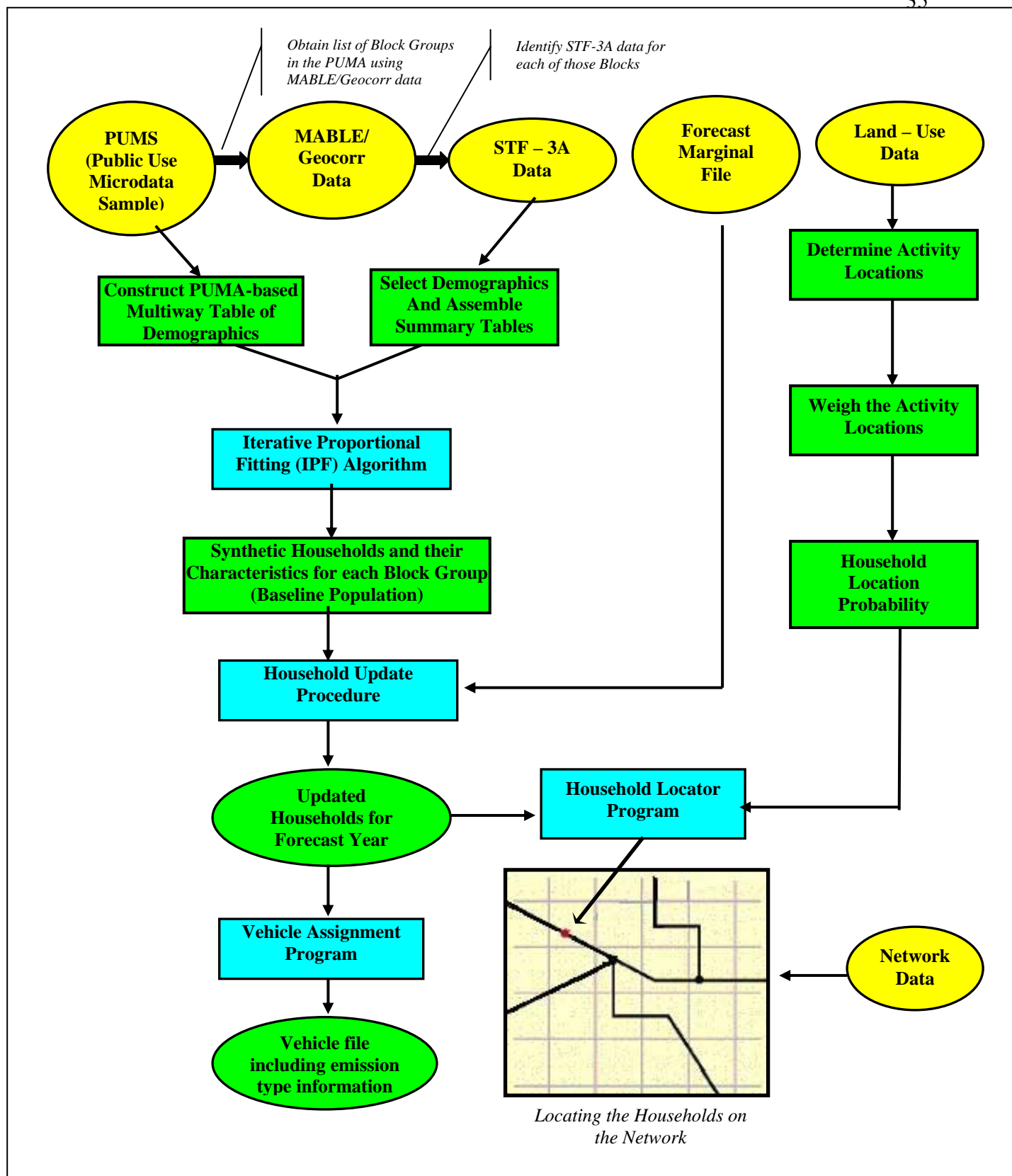


Figure-3.14: Location of the Algorithms (blue color) in the Population Synthesizer Framework

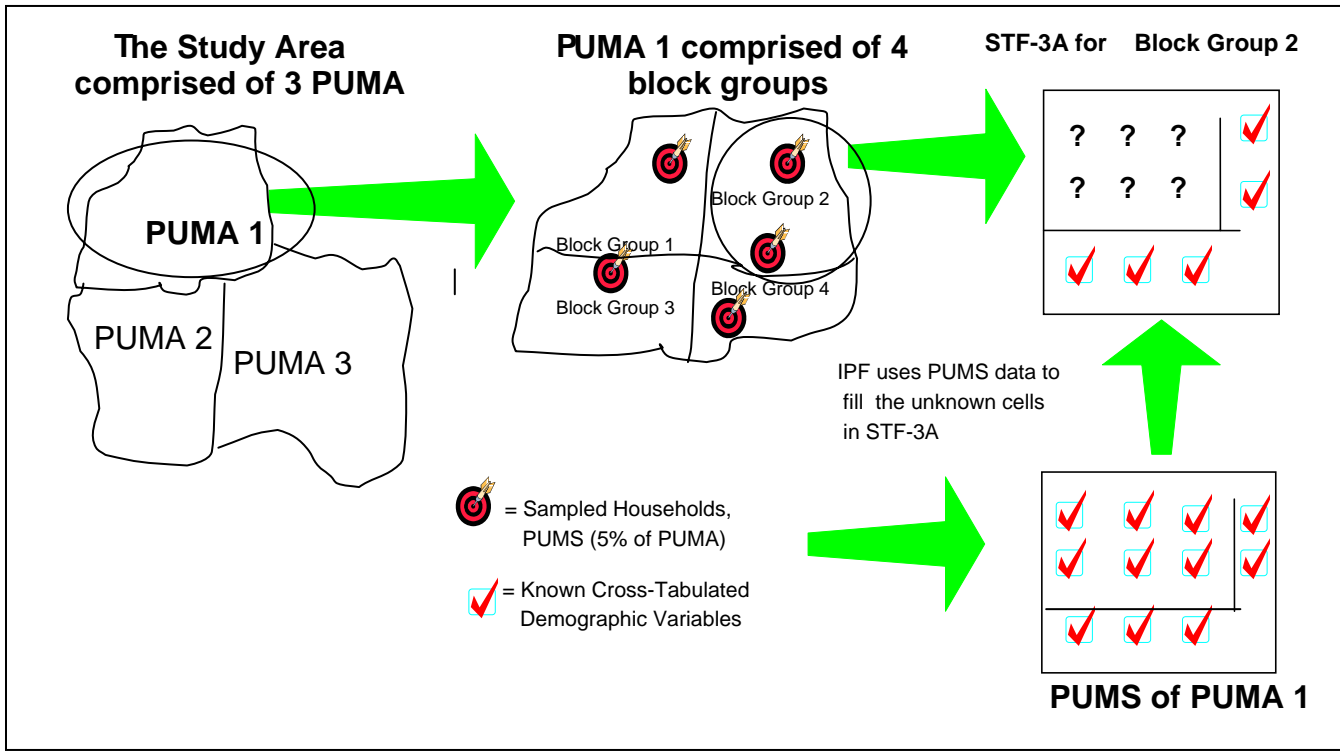


Figure-3.15: Illustrations of the concepts and relationships of PUMA, PUMS, Block Group, and STF-3A

3.8.2 IPF Algorithms

The TRANSIMS Population Synthesizer uses two types of IPF algorithms; the two-step IPF procedure developed by Beckman *et al.* [1996], and the traditional IPF procedure proposed by Deming and Stephan [1940]. The traditional IPF procedure fits **only one block group at a time** in a PUMA. Each block group in a PUMA is considered independently, which Beckman *et al.* [1996] finally showed that fitting only one block group at a time using the traditional IPF procedure is not entirely correct. Hence, the two-step IPF procedure can **simultaneously consider all block groups** that make up the PUMA. However, the **two-step** IPF procedure makes use of the traditional IPF procedure in its analysis. Therefore, we present the traditional IPF procedure first, and then we follow it by the two-step IPF procedure.

3.8.2.1 Traditional IPF Procedure for One Block Group

3.8.2.1.1 General Approach

The general approach, described in this section, provides the procedure of iterative proportional fitting (IPF) of one block group to the PUMS. The procedure begins by

selecting a PUMA in the study area. MABLE/Geocorr search engine is used to obtain a *list of block groups* within the PUMA. After that, the summary statistic STF-3A for each block group in the selected PUMA are obtained. Each of the block groups identified in the PUMA should have an appropriate STF-3A file that summarizes demographic data for the block groups. Table-3.10 gives an example of STF-3A data for block group 1 in Arlington County for two demographic variables (2 dimensions), which are the number of family households by *the householder's age*, and by the *number of workers in the household*. The table does not give the entries for any cross-classified demographics; that is, the proportion of households for each block group's classification is unknown. Moreover, there are many block groups in this PUMA in Arlington County. Table-3.10 shows only block group 1 information.

Table-3.10: An example of the summary data for census tract 1036, block group 1 of Arlington, VA, obtained from STF-3A file

Householder's Age								
Workers	15-24	25-34	35-44	45-54	55-64	65-74	>74	Total
0	?	?	?	?	?	?	?	30
1	?	?	?	?	?	?	?	57
2	?	?	?	?	?	?	?	224
>2	?	?	?	?	?	?	?	14
Total	32	82	72	54	33	24	28	325

Marginal values

Because PUMS contains complete household records for the identified PUMA, it could be used to generate the proportion of households in the cross-classified table for the block group under investigation. This matching between the block group STF-3A, and the PUMS is carried out by the traditional IPF algorithm, which is presented next.

The PUMS data can be used to construct a corresponding complete multi-dimensional demographic table whose marginal values correspond with STF-3A summary statistics. For example, Table-3.11 gives the PUMS complete multi-dimensional demographic that correspond to the STF-3A demographics, which are the number of family households classified by two demographic variables: *the householder's age*, and the *number of workers in the household*.

Table-3.11: A PUMS data for census tract 1036 containing block group 1, 2, and 3 of Arlington, VA

Householder's Age								
Workers	15-24	25-34	35-44	45-54	55-64	65-74	>74	Total
0	3	12	14	11	21	51	54	166
1	9	47	63	44	59	45	14	281
2	18	56	113	83	45	17	6	338
>2	2	17	70	88	55	17	3	252
Total	32	132	260	226	180	130	77	1037

Given Tables 3.10, and 3.11, we need to find the proportion of households in the marginal Table-8 that exactly match the known STF-3A summary data for this block group while maintaining the correlation structure of Table-3.11 constructed from PUMS. To determine this, Deming and Stephan [1940] developed a traditional iterative proportional fitting (IPF) procedure, which estimates the proportions of households for each block group's classification.

The following example demonstrates how the traditional IPF procedure works for the Arlington County **block group 1**, where the STF-3A table (Table-3.10) for the block group 1, and the PUMS data (Table-3.11) are given.

Start by computing the **adjustment factor** for each marginal value of a demographic variable in the STF-3A table, which in this case is the number of workers. The adjustment factor equals the sum of each marginal value in STF-3A divided by the corresponding marginal value in PUMS (or the updated table for subsequent iterations). In this example, for zero number of workers, the adjustment factor is $30/166 = 0.1807$ as shown in Table-3.12. This factor (0.1807) is then multiplied by all the values for the zero workers in the first row as shown in Table-3.13. The first entry in the zero workers row is 3, which is multiplied by 0.1807 to yield 0.5422. This procedure is replicated for the four rows of the demographic variable, *workers*, which has four values, 0, 1, 2, and >2. The results are shown in Table-3.13.

Table-3.12: Results of the adjustment factor for workers variable

Householder's Age									Recall sum of STF-3A	Adjustment factor
Workers	15-24	25-34	35-44	45-54	55-64	65-74	>74	Total		
0	3	12	14	11	21	51	54	166	30	$30/166 = \mathbf{0.1807}$
1	9	47	63	44	59	45	14	281	57	$57/281 = \mathbf{0.2028}$
2	18	56	113	83	45	17	6	338	224	$224/338 = \mathbf{0.6627}$
>2	2	17	70	88	55	17	3	252	14	$14/252 = \mathbf{0.0556}$
Total	32	132	260	226	180	130	77			

PUMS data

Table-3.13: Updated data from PUMS table

Householder's Age									Adjustment factor
Workers	15-24	25-34	35-44	45-54	55-64	65-74	>74	Total	
0	$3 \times 0.1807 =$ 0.5422	2.1687	2.5301	1.9880	3.7952	9.2169	9.7590	30	0.1807
1	$9 \times 0.2028 =$ 1.8256	9.5338	12.7794	8.9253	11.9680	9.1281	2.8399	57	0.2028
2	$18 \times 0.6627 =$ 11.9290	37.1124	74.8876	55.0059	29.8225	11.2663	3.9763	224	0.6627
>2	$2 \times 0.0556 =$ 0.1111	0.9444	3.8889	4.8889	3.0556	0.9444	0.1667	14	0.0556
Total	14.4079	49.7594	94.0859	70.8080	48.6412	30.5557	16.7419	325	

The second step is to use the marginal totals in the updated table to calculate the adjustment factor for the second demographic variable, *Householder's Age*, which has seven values, 15-24, 25-34, 35-44, 45-54, 55-64, 65-74, and >74 years. The results are shown in Table-3.14. For example, the adjustment factor for the number of households in the age group of 15 to 24 is $32/14.4079 = 2.2210$.

Table-3.14: Results of the adjustment factor for Householder's Age variable

Householder's Age								
Workers	15-24	25-34	35-44	45-54	55-64	65-74	>74	Total
0	0.5422	2.1687	2.5301	1.9880	3.7952	9.2169	9.7590	30
1	1.8256	9.5338	12.7794	8.9253	11.9680	9.1281	2.8399	57
2	11.9290	37.1124	74.8876	55.0059	29.8225	11.2663	3.9763	224
>2	0.1111	0.9444	3.8889	4.8889	3.0556	0.9444	0.1667	14
Total	14.4079	49.7594	94.0859	70.8080	48.6412	30.5557	16.7419	325
Recall sum of SFT-3A	32	82	72	54	33	24	28	
Adjustment factor	$32/14.4079 =$ 2.2210	$82/49.7594 =$ 1.6479	$72/94.0859 =$ 0.7653	$54/70.8080 =$ 0.7626	$33/48.6412 =$ 0.6784	$24/30.5557 =$ 0.7855	$28/16.7419 =$ 1.6725	

Update Table-14 by multiplying each cell in each column by the corresponding adjustment factor as shown in Table-3.15. At the end of this column factoring step, the sum of the row values will not exactly match the STF-3A marginal values. Therefore, **new adjustment factors** for the first variable (*Workers*) need to be calculated. Alternating between column and row factoring, this updating is repeated until all the adjustment factors are close to 1.0. Then, the IPF routine is finalized.

Table-3.15: Updated table

Householder's Age								
Workers	15-24	25-34	35-44	45-54	55-64	65-74	>74	Total
0	1.2042	3.5738	1.9362	1.5161	2.5748	7.2394	16.3215	34.3659
1	4.0547	15.7111	9.7795	6.8066	8.1195	7.1697	4.7495	56.3906
2	26.4943	61.1587	57.3083	41.9489	20.2327	8.8491	6.6502	222.6423
>2	0.2468	1.5564	2.9760	3.7284	2.0730	0.7418	0.2787	11.6011
Total	32.0000	82.0000	72.0000	54.0000	33.0000	24.0000	28.0000	325.0000

In the Arlington example of block group 1, the above two steps (row and column factoring) are conducted three more iterations to reach the final solution, which is shown in Table-3.16. The total marginal numbers of each column and row in Table-3.16 approximately equal the total values in the original SFT-3A table (Table-3.10). The final results are shown in Table-3.16. The fractional values of households are also adjusted to produce integer values of households, as shown in Table-3.17.

Table-3.16: Final results of the traditional IPF routine for Arlington block group 1

Householder's Age									Recall sum of STF-3A
Workers	15-24	25-34	35-44	45-54	55-64	65-74	>74	Total	
0	0.942	2.624	1.707	1.501	2.500	6.189	14.286	29.739	30
1	3.705	14.272	10.342	7.463	9.277	7.247	4.915	57.221	57
2	29.004	63.210	56.292	40.636	18.287	9.590	8.498	225.517	224
>2	0.308	1.894	3.660	4.400	2.945	0.974	0.301	14.481	14
Total	33.958	82.000	72.000	54.000	33.000	24.000	28.000	326.958	
Recall sum of STF-3A	32	82	72	54	33	24	28		325

Table-3.17: Final Table for Block Group 1, Arlington County after adjusting into integer numbers

Householder's Age (HHAGE)								Total
Workers	15-24	25-34	35-44	45-54	55-64	65-74	>74	
0	1	3	2	2	2	6	14	30
1	4	14	10	7	9	7	5	56
2	29	63	56	40	18	10	9	225
>2	0	2	4	4	3	1	0	14
Total	34	82	72	53	32	24	28	325

3.8.2.1.2 Mathematical Statement for Traditional IPF Procedure

Consider a PUMA comprised of some n census tracts or *blocks*, indexed by $b = 1, \dots, n$. Suppose that we are interested in estimating a multiway classification table for **each** of these n blocks independently based on cross-classifying some m -demographic measures. Then each classification table (for each block) is an m -dimensional matrix. For example where $m = 3$, there are three demographic measures considered, which could represent:

- 1) age of the head of a household (HHAGE),
- 2) total income (INC), and
- 3) number of employed household members (WORKERS).

If the i^{th} demographic variable has q_i categories, for $i = 1, \dots, m$, then each such classification table is an m -dimensional matrix of size $q_1 \times q_2 \times \dots \times q_m$. According to the previous example where $m = 3$, suppose that the **first** demographic variable, HHAGE (in years), has **seven** categories ($q_1 = 7$), which are,

- 1) $15 \leq \text{HHAGE} \leq 24$,
- 2) $25 \leq \text{HHAGE} \leq 34$,
- 3) $35 \leq \text{HHAGE} \leq 44$,
- 4) $45 \leq \text{HHAGE} \leq 54$,
- 5) $55 \leq \text{HHAGE} \leq 64$,
- 6) $65 \leq \text{HHAGE} \leq 74$, and
- 7) $\text{HHAGE} > 74$.

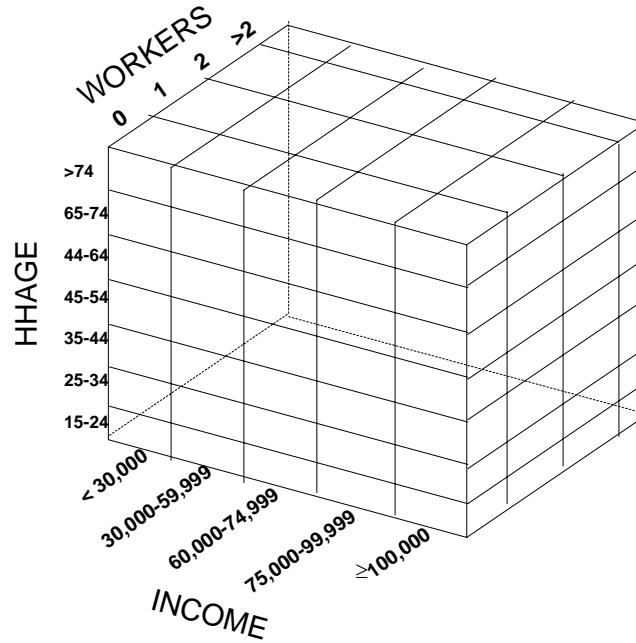
For the **second** demographic variable, INC (in \$), it has **five** categories ($q_2 = 5$), which are,

- 1) $\text{INC} < 30,000$,
- 2) $30,000 \leq \text{INC} < 60,000$,
- 3) $60,000 \leq \text{INC} < 75,000$,
- 4) $75,000 \leq \text{INC} < 100,000$, and
- 5) $\text{INC} \geq 100,000$.

And for the **third** demographic variable, WORKERS, it has **four** categories ($q_3 = 4$), which are,

- 1) $\text{WORKERS} = 0$,
- 2) $\text{WORKERS} = 1$,
- 3) $\text{WORKERS} = 2$, and
- 4) $\text{WORKERS} > 2$.

Then the classification table is a 3-dimensional matrix of size $7 \times 5 \times 4$, which is shown below.



The generalized mathematical formulation developed by Deming *et al.* [1940] for the above 3-dimensional matrix of $7 \times 5 \times 4$ can be stated as follows:

$$m_{ijk}' = n_{ijk} \frac{m_{i..}}{n_{i..}} \quad (\text{I})$$

$$m_{ijk}'' = m_{ijk}' \frac{m_{.j.}}{m_{.j.}'} \quad (\text{II})$$

$$m_{ijk}''' = m_{ijk}'' \frac{m_{..k}}{m_{..k}''} \quad (\text{III})$$

where

- n = In general, it refers to a cell or a marginal value in the PUMS data,
- m = In general, it refers to the marginal value of the STF-3A matrix,
- n_{ijk} = Sample frequency from PUMS data falling in the cell n_{ijk} (please see the figure below for illustration),
- $m_{i..}$ = Marginal data obtained from STF-3A file for the first (i) dimension, for example here the first (i) dimension is the HHAGE variable,
- $n_{i..}$ = Marginal data of the PUMS for the first (i) dimension,
- $m_{.j.}$ = Marginal data of the STF-3A file for the second (j) dimension, for example here the second (j) dimension is the INCOME variable,

- $m_{.j}'$ = Marginal data of the updated matrix (from PUMS) for the second (j) dimension,
- $m_{..k}$ = Marginal data of the STF-3A file for the third (k) dimension, for example here the third (k) dimension is the WORKERS variable, and
- $m_{..k}'$ = Marginal data of the updated matrix for the third (k) dimension.

Figure-3.16 illustrates the above notations graphically.

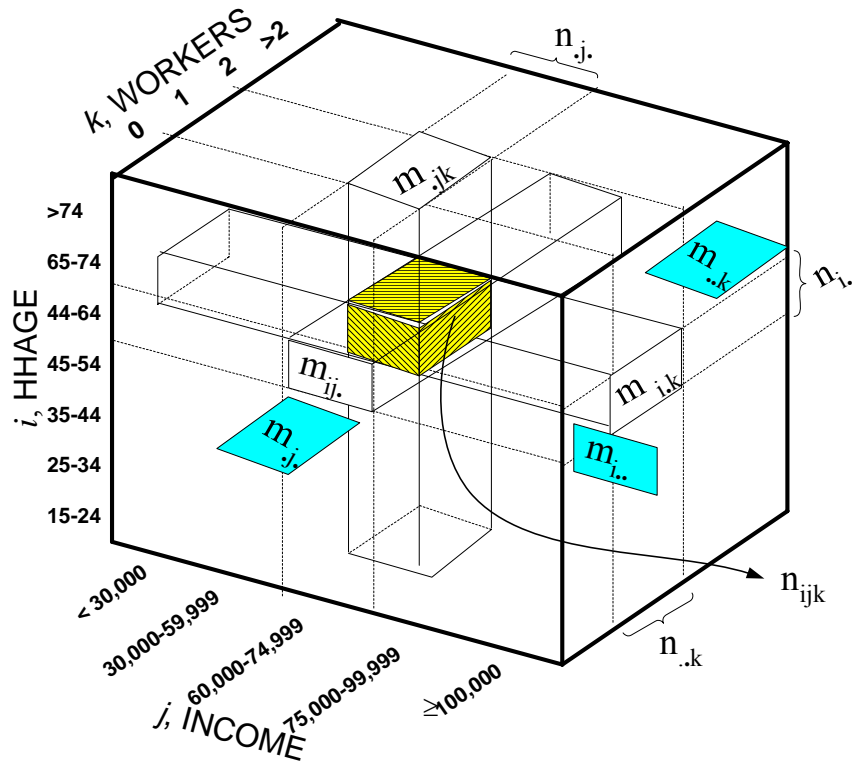


Figure-3.16: Representation of a 3-Dimension Matrix for IPF Calculation

Equation (I) is used to update each cell in the PUMS table for the first dimension (i , HHAGE), which is currently denoted as n_{ijk} . The cell is updated in the matrix by dividing the current value by an adjustment factor, $\frac{m_{i..}}{n_{i..}}$. The factor is computed by multiplying a

current STF-3A marginal value for the HHAGE demographic variable, $m_{i..}$ by a marginal data of the PUMS, $n_{i..}$. Equations (II), and (III) follow the same idea of updating cells for the second (j , INCOME), and third (k , WORKERS) dimensions respectively. Note that these equations (I) to (III) are repeated for each iteration. One needs to proceed to the next iteration by using the same equations to update the matrix (for each block group) until all adjustment factors for each dimension **are close to one**. The above equations could be

expanded to deal with 4-dimension measure (n_{ijkl}) or more ($n_{ijkl\dots}$) without any change to the procedure.

3.8.2.1.3 Practical Example in 3-Dimension

The same example from Arlington County, VA given in Tables 3.10 and 3.11 is used here for block group 1, but with three demographic variables instead of two. The new variable is INCOME, which is used in addition to HHAGE, and WORKERS variables.

Table-3.18 illustrates **both the 3-dimension STF-3A** marginal data for the block group 1 of Arlington County, **and the PUMS** data that includes this block group.

Table-3.18: 3-dimension STF-3A data for block group 1 of Arlington County, and the corresponding PUMS data

		Householder's Age (HHAGE)							Total PUMS	STF-3A
Workers	Income Group*	15-24	25-34	35-44	45-54	55-64	65-74	>74		
0	1	1	3	2	3	8	25	37	166	30
	2	2	5	8	7	9	16	15		
	3	0	2	0	0	2	5	2		
	4	0	1	2	1	1	2	0		
	5	0	1	2	0	1	3	0		
1	1	3	6	8	13	18	9	3	281	57
	2	4	24	27	6	12	18	8		
	3	1	12	18	2	9	7	1		
	4	0	2	5	12	11	7	2		
	5	1	3	5	11	9	4	0		
2	1	5	1	2	0	0	6	4	338	224
	2	12	33	21	17	12	5	2		
	3	1	9	56	8	7	1	0		
	4	0	10	19	25	6	1	0		
	5	0	3	15	33	20	4	0		
>2	1	0	3	6	4	3	0	0	252	14
	2	2	8	18	22	25	12	2		
	3	0	1	14	18	8	1	0		
	4	0	2	14	19	12	2	0		
	5	0	3	18	25	7	2	1		
Total PUMS		32	132	260	226	180	130	77	1037	
STF-3A		32	82	72	54	33	24	28		325

Income Group*	Group 1	Group 2	Group 3	Group 4	Group 5	Total
PUMS	173.000	352.000	185.000	156.000	171.000	1037.00
STF-3A	58	138	46	61	22	325

*Note that Income Group 1 is $INC < 30,000$, Income Group 2 is $30,000 \leq INC < 60,000$, Income Group 3 is $60,000 \leq INC < 75,000$, Income Group 4 is $75,000 \leq INC < 100,000$, and Income Group 5 is $INC \geq 100,000$.

The following procedure illustrates the 3-dimension application of the traditional IPF procedure to block group 1 of Arlington County.

STEP 1

Start with the **first** dimension, let us say *Workers*. Find the adjustment factors for all marginal totals in this dimension. The adjustment factor is obtained by dividing STF-3A value by the corresponding marginal value of PUMS as conducted in the previous example. Hence, there are four adjustment factors that need to be calculated here, which are adjustment factors for *Workers*=0, *Workers*=1, *Workers*=2, and *Workers*>2. Table-3.18.1 shows the calculated adjustment factors for the *Workers* variable. They are the same as in Table-3.13 presented earlier in the 2-dimension case.

Table-3.18.1: The adjustment factors for the WORKERS variable in the traditional IPF procedure

Workers	Householder's Age (HHAGE)				Total PUMS	STF-3A	Adjustment factor
	Income Group	15-24	...	>74			
0	1	1	...	37	166	30	30/166 = 0.1807
	2	2	...	15			
	3	0	...	2			
	4	0	...	0			
	5	0	...	0			
1	1	3	...	3	281	57	57/281 = 0.2028
	2	4	...	8			
	3	1	...	1			
	4	0	...	2			
	5	1	...	0			
2	1	5	...	4	338	224	224/338 = 0.6627
	2	12	...	2			
	3	1	...	0			
	4	0	...	0			
	5	0	...	0			
>2	1	0	...	0	252	14	14/252 = 0.0556
	2	2	...	2			
	3	0	...	0			
	4	0	...	0			
	5	0	...	1			
Total PUMS		32	...	77	1037		
STF-3A		32	...	28		325	

Income Group	Group 1	Group 2	Group 3	Group 4	Group 5	Total
PUMS	173.000	352.000	185.000	156.000	171.000	1037.00
STF-3A	58	138	46	61	22	325

Next, we update the table by multiplying each row of *Workers* variable by its corresponding adjustment factor. Note that this iteration is using **equation (I)**, which is

$$m_{ijk}' = n_{ijk} \frac{m_{i..}}{n_{i..}}. \text{ The updated table is show in Table-3.18.2.}$$

Table-3.18.2: Step 1 of the first iteration of the traditional IPF procedure

Workers	Householder's Age (HHAGE)								Total	STF-3A
	Income Group	15-24	25-34	35-44	45-54	55-64	65-74	>74		
0	1	0.181	0.542	0.361	0.542	1.446	4.518	6.687	30	30
	2	0.361	0.904	1.446	1.265	1.627	2.892	2.711		
	3	0.000	0.361	0.000	0.000	0.361	0.904	0.361		
	4	0.000	0.181	0.361	0.181	0.181	0.361	0.000		
	5	0.000	0.181	0.361	0.000	0.181	0.542	0.000		
1	1	0.609	1.217	1.623	2.637	3.651	1.826	0.609	57	57
	2	0.811	4.868	5.477	1.217	2.434	3.651	1.623		
	3	0.203	2.434	3.651	0.406	1.826	1.420	0.203		
	4	0.000	0.406	1.014	2.434	2.231	1.420	0.406		
	5	0.203	0.609	1.014	2.231	1.826	0.811	0.000		
2	1	3.314	0.663	1.325	0.000	0.000	3.976	2.651	224	224
	2	7.953	21.870	13.917	11.266	7.953	3.314	1.325		
	3	0.663	5.964	37.112	5.302	4.639	0.663	0.000		
	4	0.000	6.627	12.592	16.568	3.976	0.663	0.000		
	5	0.000	1.988	9.941	21.870	13.254	2.651	0.000		
>2	1	0.000	0.167	0.333	0.222	0.167	0.000	0.000	14	14
	2	0.111	0.444	1.000	1.222	1.389	0.667	0.111		
	3	0.000	0.056	0.778	1.000	0.444	0.056	0.000		
	4	0.000	0.111	0.778	1.056	0.667	0.111	0.000		
	5	0.000	0.167	1.000	1.389	0.389	0.111	0.056		
Total		14.408	49.759	94.086	70.808	48.641	30.556	16.742	325.000	
STF-3A		32	82	72	54	33	24	28		325

Income Group	Group 1	Group 2	Group 3	Group 4	Group 5	Total
PUMS	173.000	352.000	185.000	156.000	171.000	1037.00
STF-3A	58	138	46	61	22	325

STEP 2

For the second dimension (HHAGE), calculate the adjustment factors for all marginal totals in this dimension by using the updated matrix and the STF-3A data. There are seven adjustment factors that need to be calculated here, which are adjustment factor for $15 \leq \text{HHAGE} \leq 24$, $25 \leq \text{HHAGE} \leq 34$, $35 \leq \text{HHAGE} \leq 44$, $45 \leq \text{HHAGE} \leq 54$, $55 \leq \text{HHAGE} \leq 64$, $65 \leq \text{HHAGE} \leq 74$, and $\text{HHAGE} > 74$. Table-3.18.3 shows the calculations for obtaining the adjustment factors for HHAGE variable.

Table-3.18.3: The adjustment factors for the HHAGE variable in the traditional IPF procedure

		Householder's Age (HHAGE)								
Workers	Income Group	15-24	25-34	35-44	45-54	55-64	65-74	>74	Total	STF-3A
0	1	0.181	0.542	0.361	0.542	1.446	4.518	6.687	30	30
	2	0.361	0.904	1.446	1.265	1.627	2.892	2.711		
	3	0.000	0.361	0.000	0.000	0.361	0.904	0.361		
	4	0.000	0.181	0.361	0.181	0.181	0.361	0.000		
	5	0.000	0.181	0.361	0.000	0.181	0.542	0.000		
...	
>2	1	0.000	0.167	0.333	0.222	0.167	0.000	0.000	14	14
	2	0.111	0.444	1.000	1.222	1.389	0.667	0.111		
	3	0.000	0.056	0.778	1.000	0.444	0.056	0.000		
	4	0.000	0.111	0.778	1.056	0.667	0.111	0.000		
	5	0.000	0.167	1.000	1.389	0.389	0.111	0.056		
Total		14.408	49.759	94.086	70.808	48.641	30.556	16.742	325.000	
STF-3A		32	82	72	54	33	24	28		325
Adjustment factor		$32/14.408 = 2.221$	$82/49.759 = 1.648$	0.765	0.763	0.678	0.785	1.672		

Income Group	Group 1	Group 2	Group 3	Group 4	Group 5	Total
PUMS	173.000	352.000	185.000	156.000	171.000	1037.00
STF-3A	58	138	46	61	22	325

Again, we update the table by multiplying each column of *HHAGE* variable by its corresponding adjustment factor. Note that this iteration is using **equation (II)**, which is

$$m_{ijk}'' = m_{ijk}' \frac{m_{.j}}{m_{.j}'}$$

The updated results are shown in Table-3.18.4.

Table-3.18.4: Step 2 of the first iteration of the traditional IPF procedure

		Householder's Age (HHAGE)								
Workers	Income Group	15-24	25-34	35-44	45-54	55-64	65-74	>74	Total	STF-3A
0	1	0.401	0.893	0.277	0.413	0.981	3.549	11.183	34.366	30
	2	0.803	1.489	1.106	0.965	1.103	2.271	4.534		
	3	0.000	0.596	0.000	0.000	0.245	0.710	0.605		
	4	0.000	0.298	0.277	0.138	0.123	0.284	0.000		
	5	0.000	0.298	0.277	0.000	0.123	0.426	0.000		
1	1	1.352	2.006	1.242	2.011	2.477	1.434	1.018	56.391	57
	2	1.802	8.023	4.191	0.928	1.651	2.868	2.714		
	3	0.451	4.011	2.794	0.309	1.239	1.115	0.339		
	4	0.000	0.669	0.776	1.856	1.514	1.115	0.679		
	5	0.451	1.003	0.776	1.702	1.239	0.637	0.000		
2	1	7.360	1.092	1.014	0.000	0.000	3.123	4.433	222.642	224
	2	17.663	36.040	10.650	8.592	5.395	2.603	2.217		
	3	1.472	9.829	28.401	4.043	3.147	0.521	0.000		
	4	0.000	10.921	9.636	12.635	2.698	0.521	0.000		
	5	0.000	3.276	7.607	16.678	8.992	2.082	0.000		
>2	1	0.000	0.275	0.255	0.169	0.113	0.000	0.000	11.601	14
	2	0.247	0.732	0.765	0.932	0.942	0.524	0.186		
	3	0.000	0.092	0.595	0.763	0.302	0.044	0.000		
	4	0.000	0.183	0.595	0.805	0.452	0.087	0.000		
	5	0.000	0.275	0.765	1.059	0.264	0.087	0.093		
Total		32.000	82.000	72.000	54.000	33.000	24.000	28.000	325.000	
STF-3A		32	82	72	54	33	24	28		325

Income Group	Group 1	Group 2	Group 3	Group 4	Group 5	Total
PUMS	47.072	121.937	61.621	46.261	48.110	325.000
STF-3A	58	138	46	61	22	325

STEP 3

Next, the adjustment factors for the third dimension variable (Income) are calculated. Here the "Income" variable has five categories, which are $INC < 30,000$, $30,000 \leq INC < 60,000$, $60,000 \leq INC < 75,000$, $75,000 \leq INC < 100,000$, and $INC \geq 100,000$. Hence, five adjustment factors are calculated in this step. They are shown in Table-3.18.5.

Table-3.18.5: The adjustment factors for the INCOME variable in the traditional IPF procedure

		Householder's Age (HHAGE)								
Workers	Income Group	15-24	25-34	35-44	45-54	55-64	65-74	>74	Total	STF-3A
0	1	0.401	0.893	0.277	0.413	0.981	3.549	11.183	34.366	30
	2	0.803	1.489	1.106	0.965	1.103	2.271	4.534		
	3	0.000	0.596	0.000	0.000	0.245	0.710	0.605		
	4	0.000	0.298	0.277	0.138	0.123	0.284	0.000		
	5	0.000	0.298	0.277	0.000	0.123	0.426	0.000		
...				...						

>2	1	0.000	0.275	0.255	0.169	0.113	0.000	0.000	11.601	14
	2	0.247	0.732	0.765	0.932	0.942	0.524	0.186		
	3	0.000	0.092	0.595	0.763	0.302	0.044	0.000		
	4	0.000	0.183	0.595	0.805	0.452	0.087	0.000		
	5	0.000	0.275	0.765	1.059	0.264	0.087	0.093		
Total		32.000	82.000	72.000	54.000	33.000	24.000	28.000	325.000	
STF-3A		32	82	72	54	33	24	28		325

Income Group	Group 1	Group 2	Group 3	Group 4	Group 5	Total
PUMS	47.072	121.937	61.621	46.261	48.110	325.000
STF-3A	58	138	46	61	22	325
Adjustment factor	58/47.072 = 1.232	138/121.937 = 1.132	0.747	1.319	0.457	

Again, we update the table by multiplying each row of *Income Group* variable by its corresponding adjustment factor. Note that this iteration is using **equation (III)**, which is

$$m_{ijk}''' = m_{ijk}'' \frac{m_{..k}}{m_{..k}''}$$

Proceed to the next iteration by repeating **STEPS 1 to 3** and using the current updated table until all the adjustment factors reach close to **one** at the end of **STEP 3** (in the last iteration).

After we have conducted three more iterations, the final results for the Arlington County block group 1 are shown in Table-3.18.6.

Table-3. 18.6: Final results for block group 1 of Arlington County using the traditional IPF procedure

Workers	Income Group	Householder's Age (HHAGE)							Total	STF-3A
		15-24	25-34	35-44	45-54	55-64	65-74	>74		
0	1	0.359	0.857	0.316	0.472	1.143	3.601	10.560	29.724	30
	2	0.581	1.157	1.024	0.891	1.042	1.867	3.468		
	3	0.000	0.277	0.000	0.000	0.139	0.349	0.277		
	4	0.000	0.242	0.268	0.133	0.121	0.244	0.000		
	5	0.000	0.081	0.090	0.000	0.041	0.123	0.000		
1	1	1.582	2.519	1.859	3.005	3.782	1.906	1.259	56.963	57
	2	1.709	8.163	5.081	1.123	2.043	3.088	2.720		
	3	0.256	2.444	2.029	0.224	0.917	0.719	0.204		
	4	0.000	0.712	0.985	2.353	1.961	1.258	0.712		
	5	0.150	0.359	0.331	0.725	0.540	0.242	0.000		
2	1	9.512	1.515	1.676	0.000	0.000	4.584	6.055	223.935	224
	2	18.493	40.490	14.255	11.482	7.368	3.094	2.453		
	3	0.923	6.613	22.766	3.236	2.574	0.371	0.000		
	4	0.000	12.849	13.507	17.683	3.858	0.648	0.000		
	5	0.000	1.296	3.586	7.850	4.325	0.872	0.000		
>2	1	0.000	0.439	0.473	0.311	0.215	0.000	0.000	14.379	14
	2	0.306	0.979	1.187	1.433	1.499	0.750	0.254		
	3	0.000	0.075	0.562	0.714	0.292	0.038	0.000		
	4	0.000	0.261	0.984	1.320	0.767	0.133	0.000		
	5	0.000	0.132	0.427	0.586	0.151	0.045	0.046		
Total		33.871	81.461	71.406	53.542	32.779	23.934	28.006		

STF-3A	32	82	72	54	33	24	28		
Income Group	Group 1	Group 2	Group 3	Group 4	Group 5	Total			
PUMS	57.775	138.771	46.478	61.678	22.257	326.958			
STF-3A	58	138	46	61	22	325			

3.8.2.2 Two-Step IPF Procedure

As stated earlier, TRANSIMS uses a two-step IPF procedure, which provides a modification to the just described IPF procedure proposed by Deming and Stephan [1940]. The traditional procedure fits **only one block group at a time**. Beckman *et al.* [1996] showed that fitting only one block group at a time may not be entirely correct, because the sum of the blocks STF-3A should also have the same correlation structure as the PUMS data, which equally represents all the blocks in a PUMA. Hence, they proposed the two-step IPF procedure that can **simultaneously consider all block groups** that make up the PUMA.

Since the two-step IPF procedure makes use of the traditional IPF routine, the mathematical description of the traditional IPF is presented first.

The Notations used in the procedures are provided first:

M_b is the final cross-classification matrix for block group b ,

n is the total number of block groups in the PUMA under study,

$M_b^{\text{marginals}}$ is the marginal matrix of STF-3A for block group b , for which information in each cell is unknown; only its marginal values are known,

M_{PUMS} is the complete matrix of PUMS data, for which all cells are known,

m is the number of demographic variables, which denotes dimension of matrix, for example, HHAGE, INCOME, WORKERS, etc,

q_m is the categories of the m demographic variable, for example, in WORKERS variable ($m = \text{WORKERS}$), the categories of the variable are 0 WORKERS ,1 WORKERS, 2 WORKERS, and >2 WORKERS. Hence $q_{\text{WORKERS}} = 4$.

Procedure for Estimating a Cross-Classified Table for a Block Group from PUMS

Data

Consider a PUMA comprised of some n census tracts or *blocks*, indexed by $b = 1, \dots, n$. Suppose that we are interested in estimating a multiway classification table for **each** of these n blocks independently based on cross-classifying some m -demographic measures. If the i^{th} demographic variable has q_i categories, for $i = 1, \dots, m$, then each such classification table is an m -dimension matrix of size $q_1 \times q_2 \times \dots \times q_m$. The estimation problem addressed in this section can accordingly be stated as follows.

Given: Based on, say, a 5% representative sample of the PUMA, suppose that we have a corresponding complete census record file for this sample called PUMS. This PUMS file can be used to construct a corresponding complete m -dimensional cross-classification matrix M_{PUMS} of size $q_1 \times q_2 \times \dots \times q_m$. In addition, for each block $b = 1, \dots, n$, suppose that we are given the total count of entities belonging to each of the q_i categories, for each

demographic variable $i = 1, \dots, m$, where the **sum** of these q_i values for each $i = 1, \dots, m$ equals the total population count N_b for the given block (obtained from **STF-3A** files). Hence, for each block b having a total population count of N_b , we are given a partition of this value N_b into the q_1 categories corresponding to demographic variable 1, and into the q_2 categories corresponding to demographic variable 2, and so on. Let us denote by M_b the cross-classification m -dimensional matrix of size $q_1 \times q_2 \times \dots \times q_m$ for block b , for $b = 1, \dots, n$, where each such matrix M_b is similar in structure to M_{PUMS} . The foregoing available data for each block b then corresponds to the **marginals** of M_b pertaining to the totals

$$\{M_{\dots j \dots} \text{ for each } j = 1, \dots, q_i\} \text{ for each } i = 1, \dots, m. \quad (1)$$

↑
position i

We need to find the individual entries of the matrices M_b , for each $b = 1, \dots, n$, that satisfy these marginal totals and that reflect the relative proportions in M_{PUMS} as closely as possible in a least-square sense. Additionally, we would like the correlation structure of each of these n matrices to be the same, and that of their sums to match the correlation structure of M_{PUMS} , as indicated by the odds-ratios (correlation structure) referred by Beckman et al. [1996].

In the scheme described below, we employ the IPF tool originally proposed by Deming and Stephan [1940]. Given two conformable multidimensional matrices M_1 and M_2 , where M_1 is completely specified (for example, M_{PUMS}), and only certain specified marginals are available for M_2 (for example, M_b , which is constructed by using STF-3A file), IPF determines the entries of M_2 such that they satisfy the given marginal totals (STF-3A files), match the relative proportions exhibited by the entries of M_1 as closely as possible in a least-squares sense, and additionally, produce the same odd-ratios (correlation structure) for M_2 as for M_1 . Notationally, we will denote this operation as

$$M_2 = \text{IPF}(M_1, M_2^{\text{marginals}}), \quad (2)$$

where $M_2^{\text{marginals}}$ denotes that only the aforementioned marginals are available for M_2 as input into IPF. Note that these marginals might be just of type (1), or in general, could be with respect to any dimensional slices, having certain elements fixed in specified dimensions (such as element j in dimension i in (1)), and with the sum being taken over the remaining dimensions. The routine IPF would then sequentially perform updates of the elements of M_2 , based on each such specified marginal value. We assume that the operation implied by (2) conforms with such a general specification of $M_2^{\text{marginals}}$.

Procedure for Determining Tables for Multi-Blocks Simultaneously

TRANSIMS adopts the following two-step procedure for determining the required cross-classification matrices M_b , for $b = 1, \dots, n$.

Step 1

Let M_{TOTAL} denote the sum $\sum_{b=1}^n M_b$ of the n individual block matrices. Note that based on the given data, we know $M_{TOTAL}^{\text{marginals}}$, where

$$M_{TOTAL}^{\text{marginals}} = \sum_{b=1}^n M_b^{\text{marginals}} \quad (3)$$

and where $M_b^{\text{marginals}}$ denotes the unknown matrix M_b but with its individual marginals specified in each dimension. Using IPF according to the traditional IPF procedure as stated in equation (2), we determine

$$M_{TOTAL} = \text{IPF}(M_{PUMS}, M_{TOTAL}^{\text{marginals}}). \quad (4)$$

The resulting complete matrix M_{TOTAL} has the same odd-ratios (correlation structure) as M_{PUMS} .

Graphical Illustration of a Generic Example

Consider a generic situation in which we have PUMA with $n = 4$ block groups, and we have $m = 2$ -dimensional classification matrices, each of size 2×3 ($q_1 = 2, q_2 = 3$). Hence, we have the following data, where asterisks (*) denote known data, and ? denote unknown data.

$$M_{PUMS} = \begin{array}{c} \text{marginals:} \\ \begin{array}{|ccc|c} * & * & * & * \\ * & * & * & * \end{array} \\ \text{marginals: } * \quad * \quad * \end{array}$$

$$M_b = \begin{array}{c} \begin{array}{|ccc|} ? & ? & ? \\ ? & ? & ? \end{array} \quad \text{for } b = 1, \dots, 4 \text{ (four blocks)} \end{array}$$

$$M_b^{\text{marginals}} = \begin{array}{c} \text{marginals:} \\ \begin{array}{|c|} M_b \\ \hline \end{array} \quad \begin{array}{c} * \\ * \end{array} \quad \text{for } b = 1, \dots, 4 \\ \text{marginals: } * \quad * \quad * \end{array}$$

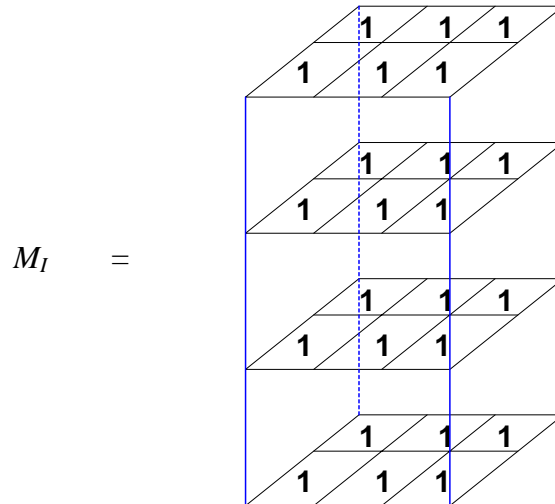
Step 1:

(a) Let $M_{TOTAL}^{\text{marginals}} = \sum_{b=1}^4 M_b^{\text{marginals}}$.

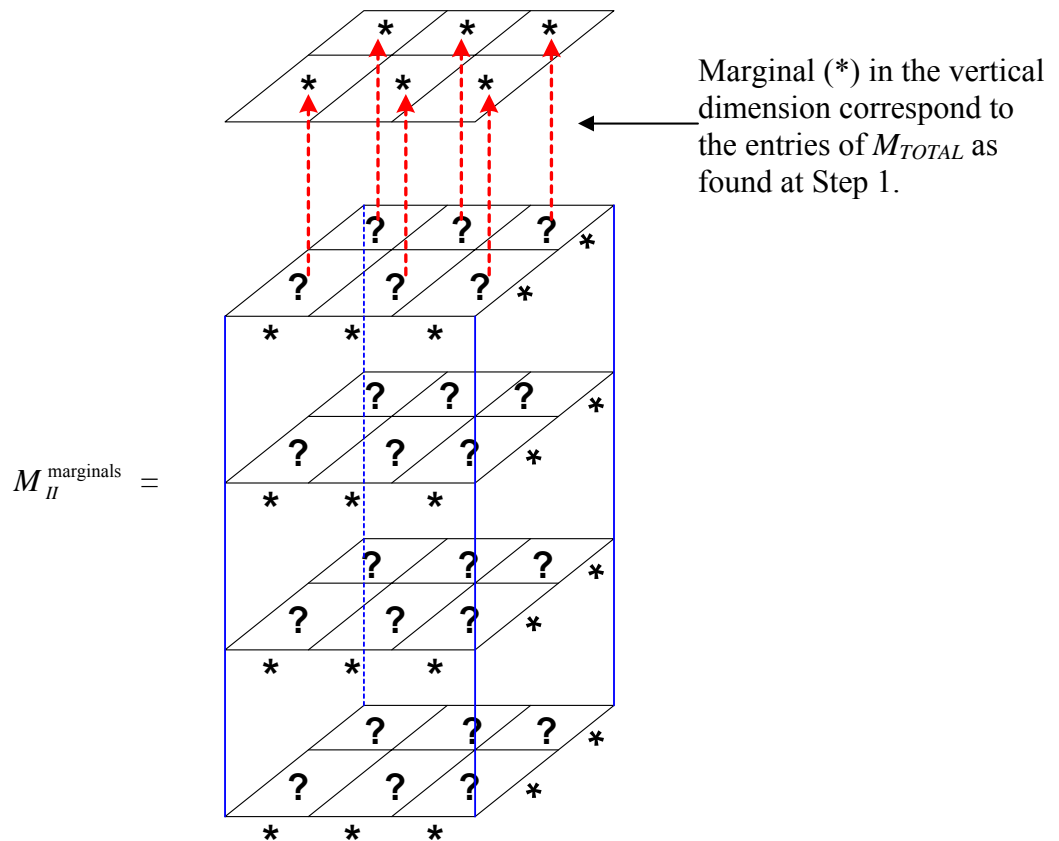
(b) Compute $M_{TOTAL} = \text{IPF}(M_{PUMS}, M_{TOTAL}^{\text{marginals}})$.

Step 2:

- (a) Compute M_I as the following three-dimensional $2 \times 3 \times 4$ matrix, having all elements equal to one.



- (b) Let $M_{II}^{\text{marginals}}$ be given by the following specified marginal (*) with respect to the corresponding matrix M_{II} .



(c) Compute $M_{II} = \text{IPF}(M_I, M_{II}^{\text{marginals}})$.

The required matrices M_b , $b = 1, \dots, 4$, would then reside at each of the (vertical) levels, $b = 1, \dots, 4$ of the resultant matrix M_{II} .

Example of the Two-Step IPF Procedure in 2-Dimensional Matrices

Again, the procedure begins after selecting a PUMA in the study area. Once a PUMA is selected, the MABLE/Geocorr search engine is used to obtain a *list of block groups* within the PUMA. Then the summary statistics are obtained for each block group in the selected PUMA. Each of the block groups identified in the PUMA should have an appropriate STF-3A file that summarizes demographic data for the block group.

IPF STEP 1

1. After obtaining a list of n -block groups in the PUMA and STF-3A file for each block group, sum the marginal values for each demographic variable for **all** the n -block groups.

To illustrate this step, suppose we are given a PUMA, which contains only **two block groups** (Beckman *et al.* [1996]), and STF-3A data for each block group as shown below in Tables 3.19 and 3.20. Each block group has **two** demographic variables, $v1$ and $v2$ and each variable has **two** values, 1 and 2. The sum of these two demographic marginals for the two block groups is shown in Table-3.21.

Table-3.19: STF-3A for Block Group 1

	$v2 = 1$	$v2 = 2$	Total
$v1 = 1$?	?	1700
$v1 = 2$?	?	1050
Total	1505	1245	2750

+

Table-3.20: STF-3A for Block Group 2

	$v2 = 1$	$v2 = 2$	Total
$v1 = 1$?	?	1405
$v1 = 2$?	?	905
Total	700	1610	2310

Table-3.21: Marginal sum of the two block groups (M_{TOTAL})

	$v2 = 1$	$v2 = 2$	Total
$v1 = 1$?	?	3105
$v1 = 2$?	?	1955
Total	2205	2855	5060

2. Given the PUMS table for this PUMA, and the marginal sum table of all the block groups, we conduct a traditional IPF fitting of the two tables as discussed in the previous section. This involves finding the adjustment factors for all values of the

demographic variables, and repeating the process till the adjustment factors become close to 1.

The given PUMS data for the previous example shown in Tables 3.19 to 3.21 is provided in Table-3.22. Note that the total number of households in the PUMS, which is 253, is 5% of the total number of households, 5060, in the Table-3.21, which also represents the total blocks in the PUMA.

Table-3.22: Corresponding PUMS data

	v2 = 1	v2 = 2	Total
v1 = 1	45	108	153
v1 = 2	63	37	100
Total	108	145	253

We conduct the traditional IPF procedure for the two Tables 3.21 and 3.22. There is no need to discuss again the steps of the procedure here, since it has been described in detail earlier. We only show the results of the first iteration including the adjustment factors, and the final results as shown in Tables 3.23 to 3.27.

Table-3.23: Values of the adjustment factors for v1

	v2 = 1	v2 = 2	Total	Recall sum of STF-3A	Adjustment factor
v1 = 1	45	108	153	3105	$3105 / 153 = 20.2941$
v1 = 2	63	37	100	1955	$1955 / 100 = 19.5500$
Total	108	145	253		

Table-3.24: Updated data after row factoring

	v2 = 1	v2 = 2	Total
v1 = 1	45×20.2941 = 913	2192	$913 + 2192 =$ 3105
v1 = 2	63×19.5500 = 1232	723	1955
Total	$913 + 1232 =$ 2145	2915	5060

Table-3.25: Values of the adjustment factors for v2

	$v_2 = 1$	$v_2 = 2$	Total
$v_1 = 1$	913	2192	3105
$v_1 = 2$	1232	723	1955
Total	2145	2915	5060
Recall sum of STF-3A	2205	2855	5060
Adjustment factor	$\frac{2205}{2145}$ = 1.028	$\frac{2855}{2915}$ = 0.979	

Table-3.26: Updated data after column factoring

	$v_2 = 1$	$v_2 = 2$	Total
$v_1 = 1$	$913 \times 1.028 =$ 939	$2192 \times 0.979 =$ 2146	$939 + 2146 =$ 3085
$v_1 = 2$	1266	708	1975
Total	$939 + 1266 =$ 2205	2855	5060

After conducting five more iterations, the results are shown in Table-3.27. Note that the results are verified to be correct because the values of each column and row are equal to the totals in Table-3.21, which are the *marginal sum of the two block groups*.

Table-3.27: Final results of the IPF STEP 1

	$v_2 = 1$	$v_2 = 2$	Total	Recall sum of STF-3A
$v_1 = 1$	949	2156	3105	3105
$v_1 = 2$	1256	699	1955	1955
Total	2205	2855	5060	
Recall sum of STF-3A	2205	2855	5060	

IPF STEP 2

1. From the given data in STF-3A table for each block group, construct new tables classified by demographic variable (instead of being classified by each block group like the original STF-3A table). Here, each demographic variable has its own table classified by all the block groups.

As an example, Tables 3.19 and 3.20 show the STF-3A data for demographic variables, v_1 and v_2 , and for two block groups. The conversion of these tables to satisfy the above rule that is classification by the demographic variables, v_1 and v_2 independently, are shown in Tables 22 and 23 respectively.

Table-3.28:New table classified by v_1 at IPF STEP2 **Table-3.29:New table classified by v_2 at IPF STEP2**

	Block group 1	Block group 2	Total
$v_1 = 1$	1700	1405	3105
$v_1 = 2$	1050	905	1955
Total	2750	2310	5060

+

	Block group 1	Block group 2	Total
$v_2 = 1$	1505	700	2205
$v_2 = 2$	1245	1610	2855
Total	2750	2310	5060

2. Similar to IPF STEP1, we proceed to find a marginal sum of all block groups in the PUMA. But at this step, we add one more dimension to the marginal sum table, which is the dimension of the **block group**.

Recall Table-3.21, the marginal sum of the two block groups. Add block groups 1 and 2 to each column in Table-3.21. From the re-classified Tables 3.28 and 3.29, use the values in the appropriate cells to obtain the one-more-dimensional table as shown in Table-3.30. The marginal values obtained for Table-3.30 are treated as the traditional STF-3A data in an IPF procedure.

(Recall) Table-3.21: Marginal sum of the two block groups

	$v_2 = 1$	$v_2 = 2$	Total
$v_1 = 1$?	?	3105
$v_1 = 2$?	?	1955
Total	2205	2855	5060

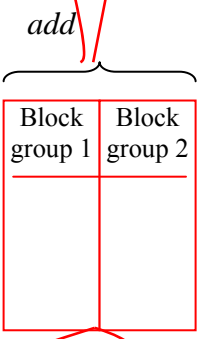


Table-3.30: One-more-dimensional table

	$v_2 = 1$		$v_2 = 2$		Total
	Block group 1	Block group 2	Block group 1	Block group 2	
$v_1 = 1$					3105
$v_1 = 2$					1955
Total					5060
Sum of STF-3A	1505	700	1245	1610	5060

3. Let all the cells in the one-more-dimensional table have a value of 1. Then, sum the total values of each row and column as shown in Table-3.31. The table of ones is treated in this case as if it is the PUMS data.

Table-3.31: One-more-dimensional table with all elements equals to 1

	$v_2 = 1$		$v_2 = 2$		Total
	Block group 1	Block group 2	Block group 1	Block group 2	
$v_1 = 1$	1	1	1	1	4
$v_1 = 2$	1	1	1	1	4
Total	2	2	2	2	8
Sum of STF-3A	1505	700	1245	1610	5060

4. Here, we apply the traditional IPF procedure to compute **adjustment factor** for each marginal value. The adjustment factor equals to the sum of each marginal value in STF-3A divided by the corresponding total value in the previous table.

For example from Table-3.31, we calculate first the adjustment factor for the first demographic variable, $v_1=1$, and repeat it for $v_1=2$. The results are shown in Table-3.32.

Table-3.32: Results of the adjustment factor for v_1

	$v_2 = 1$		$v_2 = 2$		Total	Sum of STF-3A	Adjustment factor
	Block group 1	Block group 2	Block group 1	Block group 2			
$v_1 = 1$	1	1	1	1	4	3105	$3105/4 =$ 776.250
$v_1 = 2$	1	1	1	1	4	1955	$1955/4 =$ 488.750
Total	2	2	2	2	8	5060	
Sum of STF-3A	1505	700	1245	1610	5060		

Before we proceed to calculate the adjustment factor for the next demographic variable, v_2 , we **update** the data in the table by multiplying each cell by the corresponding adjustment factor. For the example in Table-3.32, we multiply the adjustment factor **776.250** by every cell (except the Total column) in the **first row (when $v_1=1$)**. The updated data is shown in Table-3.33.

Table-3.33: Updated data

	$v_2 = 1$		$v_2 = 2$		Total	Sum of STF-3A
	Block group 1	Block group 2	Block group 1	Block group 2		
$v_1 = 1$	776.250	776.250	776.250	776.250	3105	3105
$v_1 = 2$	488.750	488.750	488.750	488.750	1955	1955
Total	1265	1265	1265	1265	5060	5060
Sum of STF-3A	1505	700	1245	1610	5060	

Next, we calculate the adjustment factor for the second demographic variable, v_2 , which also has two values, $v_2=1$ and $v_2=2$. The result is shown in Table-3.34.

Table-3.34: Results of the adjustment factor for v2

	v2 = 1		v2 = 2		Total	Sum of STF-3A
	Block group 1	Block group 2	Block group 1	Block group 2		
v1 = 1	776.250	776.250	776.250	776.250	3105	3105
v1 = 2	488.750	488.750	488.750	488.750	1955	1955
Total	1265	1265	1265	1265	5060	5060
Sum of STF-3A	1505	700	1245	1610	5060	
Adjustment factor	$1505/1265 =$ 1.190	$700/1265 =$ 0.553	$1245/1265 =$ 0.984	$1610/1265 =$ 1.273		

Again, we update the table by multiplying each cell by the corresponding adjustment factor. The results are shown in Table-3.35.

Table-3.35: Updated data

	v2 = 1		v2 = 2		Total	Sum of STF-3A
	Block group 1	Block group 2	Block group 1	Block group 2		
v1 = 1	776.25×1.19 = 923.5	429.5	764.0	988.0	3105	3105
v1 = 2	581.5	270.5	481.0	622.0	1955	1955
Total	1505	700	1245	1610	5060	5060
Sum of STF-3A	1505	700	1245	1610	5060	

5. Next, we **collapse** Table-33 by classifying data according to variable **v1** and **Block group** only. Then, we find all the adjustment factors corresponding to the collapsed table. The results are shown in Table-33.1.

Table-35.1: Collapsed Table-3. 35

	v1=1		v1=2	
	Block Group 1	Block Group 2	Block Group 1	Block Group 2
	$923.5+764$ = 1687.5	1417.5	1062.5	892.5
STF3A	1700	1405	1050	905
Adjustment factor	1.007	0.991	0.988	1.014

Again, we update Table-3.35 by multiplying each cell by the corresponding adjustment factor. The results are shown in Table-3.36.

Table-3.36: Updated data

	$v2 = 1$		$v2 = 2$		Total	Sum of STF-3A
	Block group 1	Block group 2	Block group 1	Block group 2		
$v1 = 1$	923.5×1.007 = 930.4	425.8	769.6	979.2	3105	3105
$v1 = 2$	574.6	274.2	475.4	630.8	1955	1955
Total	1505	700	1245	1610	5060	5060
Sum of STF-3A	1505	700	1245	1610	5060	

6. Next, we collapse the current table by classifying data according to variable $v2$ and $v1 = 1$, and $v1 = 2$ and Block group. Then, we find all the adjustment factors corresponding to the collapsed table. The results are shown in Table-3.36.1.

Table-3.36.1: Collapsed Table-3.36

	$v2=1$		$v2=2$	
	$v1 = 1$	$v1 = 2$	$v1 = 1$	$v1 = 2$
	BG1+BG2	BG1+BG2	BG1+BG2	BG1+BG2
	930.4+425.8= 1356.1	848.9	174 8.9	1106.1
STF3A	949	1256	2156	699
Adjustment factor	0.700	1.480	1.233	0.632

Again, we update Table-3.36 by multiplying each cell by the corresponding adjustment factor. Then, find all the adjustment factors. Results are shown in Table-3.37.

Table-3.37: Updated data

	$v_2 = 1$		$v_2 = 2$		Total	Sum of STF-3A	Adjustment factor
	Block group 1	Block group 2	Block group 1	Block group 2			
$v_1 = 1$	$930.4 \times 0.700 = 651.1$	297.9	948.8	1207.2	3105	3105	1
$v_1 = 2$	850.2	405.8	300.4	398.6	1955	1955	1
Total	1501.29	703.709624	1249.20022	1605.8	5060	5060	
Sum of STF-3A	1505	700	1245	1610	5060		

7. Once we find the adjustment factors for all demographic variables, we continue updating the table by recalculating all the adjustment factors for each demographic variable. **The process is repeated for steps 4 to 6 until all the adjustment factors for each variable reach close to one.**

After repeating the above steps for four more iterations, the final results are shown in Table-3.38.

Table-3.38: The final table of the 2-step IPF procedure

	$v_2 = 1$		$v_2 = 2$		Total	Sum of STF-3A
	Block group 1	Block group 2	Block group 1	Block group 2		
$v_1 = 1$	701	428	999	1157	3105	3105
$v_1 = 2$	804	452	246	453	1955	1955
Total	1505	700	1245	1610	5060	5060
Sum of STF-3A	1505	700	1245	1610	5060	

To break up Table-3.38 by **block group**, Tables 3.38.1, and 3.38.2 are provided. They produce the same marginals as Tables 3.19, and 3.20.

Table-3.38.1: Cross-Classification Table for BG1 **Table-3.38.2: Cross-Classification Table for BG2**

	$v_2 = 1$	$v_2 = 2$	Total
$v_1 = 1$	701	999	1700
$v_1 = 2$	804	246	1050
Total	1505	1245	2750

+

	$v_2 = 1$	$v_2 = 2$	Total
$v_1 = 1$	248	1157	1405
$v_1 = 2$	452	453	905
Total	700	1610	2310

(Recall) Table-3.19: STF-3A for Block Group 1 **(Recall) Table-3.20: STF-3A for Block Group 2**

	$v_2 = 1$	$v_2 = 2$	Total
$v_1 = 1$?	?	1700
$v_1 = 2$?	?	1050
Total	1505	1245	2750

+

	$v_2 = 1$	$v_2 = 2$	Total
$v_1 = 1$?	?	1405
$v_1 = 2$?	?	905
Total	700	1610	2310

To graphically illustrate, and summarize the difference between the traditional IPF, and the 2-step IPF, we use an example of a PUMA comprising of **two block groups**, with each block group having **two demographic variables** (v_1 and v_2), where v_1 is denoted as “HOUSEHOLD SIZE (HHSIZE)” and v_2 is denoted as “NUMBER OF WORKERS (#WORK)”. Each variable has only two possible values. For HHSIZE, they are $HHSIZE < 3$, and $HHSIZE \geq 3$. For #WORK, there are $\#WORK < 2$, and $\#WORK \geq 2$. Note that this is a hypothetical example. In real applications, we use more than two demographic variables and we assign more than two values for each variable. We have the following data, where correction marks (✓) denote known data, and question marks (?) denote unknown data that we wish to find out.

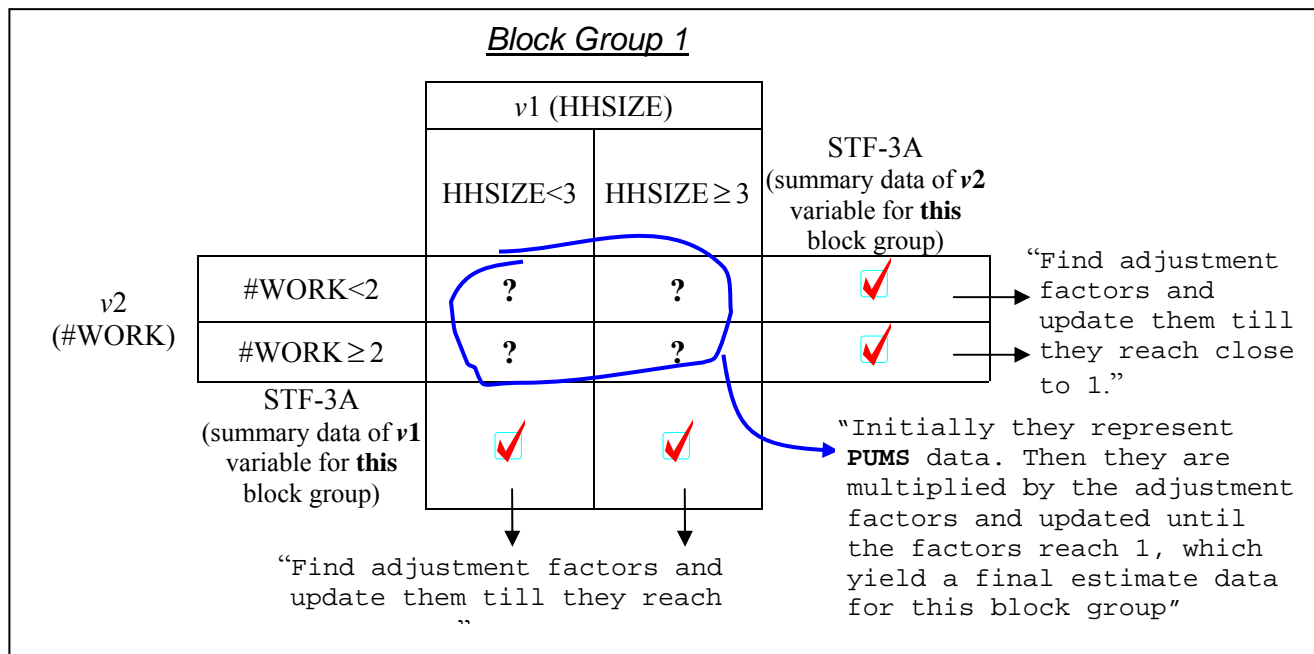


Figure-3.17: Concept of the traditional IPF procedure applied for block group 1

Figure-3.17 summarizes the traditional IPF procedure for block group 1. The same procedure needs to be repeated for **block group 2**. Finally when all the adjustment factors reach close to 1, we obtain the final estimate data for each block group, which have the marginal values equal to STF-3A. Figure-3.18 graphically summarizes the two-step IPF procedure, and displays the functions of each step. It also provides the contrast between it the traditional IPF.

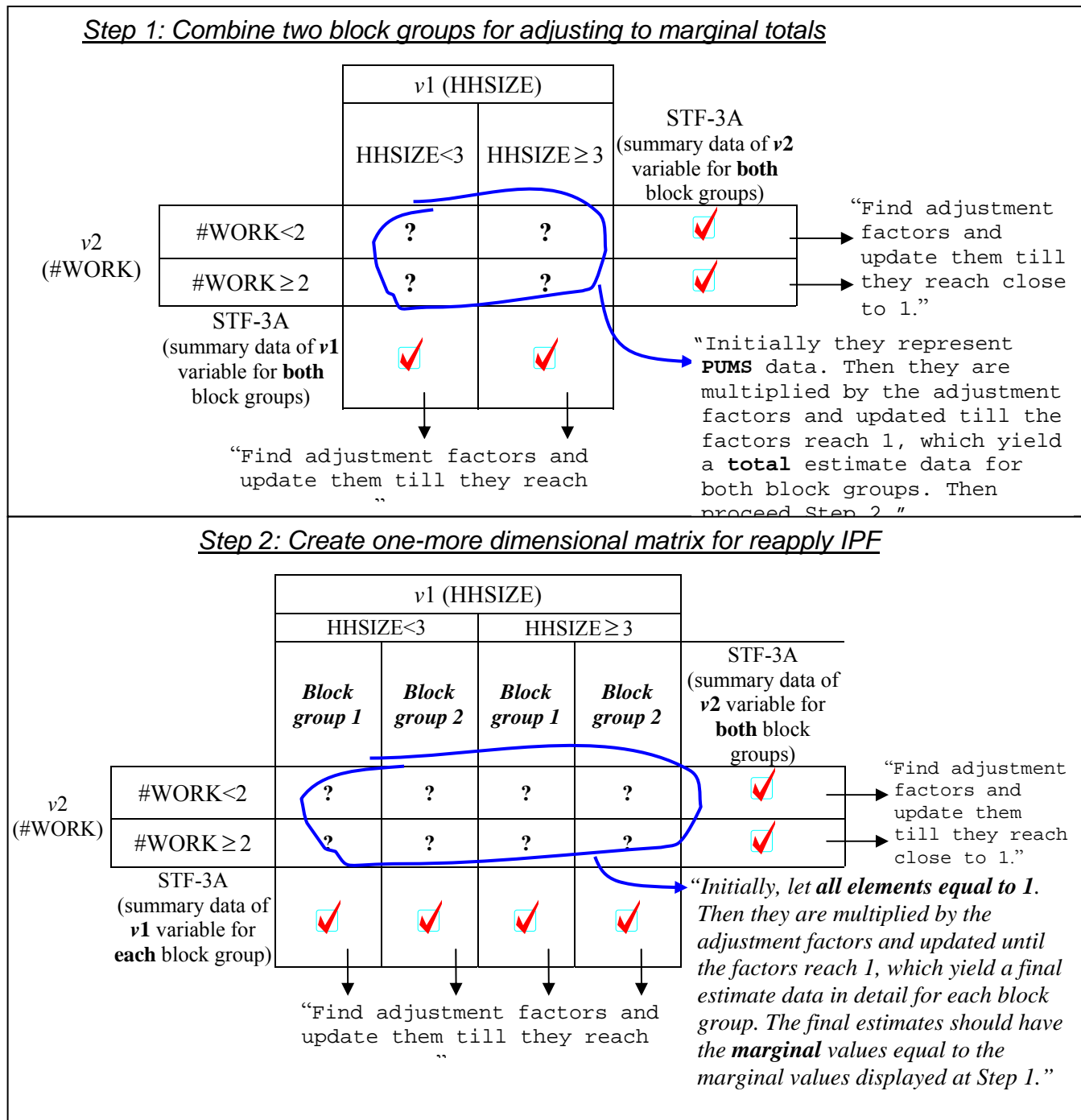


Figure-3.18: Concept of the two-step IPF procedure

3.8.3 Update Households for Forecast

After obtaining the baseline synthetic population using the 2-step IPF procedure in Section 8.2.2, the user still needs to forecast the baseline population to a future planning year. Forecasted populations for TRANSIMS are generated using a modification of the procedures given in the paper by Beckman *et al.* [1996].

Before starting the forecast procedure, the Population Synthesizer requires a **Forecast Marginal File** as input. This file contains the forecast marginals, like those given by the STF-3A files, for each census tract and block group. This input file is created outside TRANSIMS, and is typically obtained from transportation planning agencies.

To forecast a population, one starts again with an individual PUMA. Then the PUMS for that PUMA are projected to the forecast year. After that, the IPF procedure is used again to fit the STF-3A forecast marginal to the projected PUMS to obtain a cross-classification table to the future year.

At the end of the baseline IPF algorithm, three multidimensional arrays classified by household type have been estimated for each block group in a PUMA. The dimensions of the three tables or arrays (one for family households, one for non-family households, and one for those living in a group quarter) represent particular demographic types. The entries in the arrays are the proportions of households of the various demographic types generated in the complete baseline procedure.

For **family** (f) and **non-family** (n) households, the array for each type has five demographic dimensions:

- householder's age (i -dimension),
- family income (j -dimension),
- number of workers in the family (k -dimension),
- type of family (l -dimension), and
- householder's race (m -dimension).

For group quarters (g), the array has two dimensions:

- type of group quarters (i -dimension), and
- age of the individuals (j -dimension).

In general, let the proportions in the arrays be denoted by the following:

- $^f p_{i,j,k,l,m}$ for family array (f),
- $^n p_{i,j,k,l,m}$ for non-family array (n), and
- $^g p_{i,j}$ for group quarters (g).

The forecast procedure proceeds as follows. At each final cross-classification table for each block group obtained from the IPF procedure, calculate the proportionate number of each cell in the table by dividing the number of households in each cell by the total number of households in the block group. The block groups in the PUMA that have the maximum

number (non-zeros) in the estimated proportion arrays are identified in order to **forecast the future PUMS**. Of these, the block group with the largest number of households is selected. For this block group, the proportions of family households, non-family households, and group quarter residents are denoted by w^f , w^n , and w^s , respectively.

Next, an array H with dimensions equal to the number of forecast demographics is created. For example, these demographics may be the size of the household (P), the age of the householder (A), and the household income (I). Without loss of generality, we consider only these three demographics here. $H_{P,A,I}$ is computed in the following steps. For each sample in the PUMS:

Step 1:

Determine if the household is a family, non-family, or a group quarter resident. Here, we assume the household is a family household.

Step 2:

For this family household, ascertain the demographic indexes according to the notations of demographic dimension for a family household type. Since it is a family household, it has five demographic indexes, which are householder's age, family income, number of workers in the family, type of family, and householder's race. Assume that these are 1, 3, 2, 4, 5.

Step 3:

Determine the indexes for the forecast demographics P , A , I . For example here, we assume that these indexes are p , a , i .

Step 4:

Let WT be the household weight given in the PUMS, and.

Step 5:

Update the array value $H_{p,a,i}$ as follows:

$$H_{p,a,i} = H_{p,a,i} + w^f \times p_{1,3,2,4,5}^f \times WT.$$

To compute a set of forecast cross-classified proportions for each block group in the PUMA, the IPF procedure is used to fit the updated array $H_{P,A,I}$, as a future PUMS, to the input forecast marginal values of STF-3A.

Example of Households Forecast

We continue using the example from Arlington County block group 1 by assuming that this block group has the maximum number of non-zeros in the estimated proportion, and also has the largest number of households, 325 households, compared with other block groups within the PUMA. After obtaining the final cross-classification table for the family households in this block group as shown in Table-3.39 (here we used the results from the traditional IPF procedure), we proceed to calculate the proportions for each cell in the table. For the sake of illustration, we collapsed the three dimensional table into two-dimensional table, which has only two demographic variables that are *Householder's Age*, and *Workers*.

Table-3.39: Final Table for Block Group 1, Arlington County

Householder's Age (HHAGE)								Total
Workers	15-24	25-34	35-44	45-54	55-64	65-74	>74	
0	1	3	2	2	2	6	14	30
1	4	14	10	7	9	7	5	56
2	29	63	56	40	18	10	9	225
>2	0	2	4	4	3	1	0	14
Total	34	82	72	53	32	24	28	325

As an example of calculating a proportion, the cell that has Householder's Age between 15 and 24 years, and zero worker in Table-3.39, the proportion is $1/325 = 0.003$. The final estimated proportion table is shown in Table-3.40.

Table-3.40: Estimated proportions of family households for Block Group 1, Arlington County

Householder's Age (HHAGE)								Total
Workers	15-24	25-34	35-44	45-54	55-64	65-74	>74	
0	0.003	0.008	0.005	0.005	0.008	0.019	0.044	0.091
1	0.011	0.044	0.032	0.023	0.028	0.022	0.015	0.175
2	0.089	0.193	0.172	0.124	0.056	0.029	0.026	0.689
>2	0.001	0.006	0.011	0.013	0.009	0.003	0.001	0.044
Total	0.104	0.251	0.220	0.165	0.101	0.074	0.086	1.000

The PUMS table for this PUMA containing the block group 1 is show in Table-3.41.




Table-3.41: A PUMS data for Block Group 1, Arlington County

Householder's Age								
Workers	15-24	25-34	35-44	45-54	55-64	65-74	>74	Total
0	3	12	14	11	21	51	54	166
1	9	47	63	44	59	45	14	281
2	18	56	113	83	45	17	6	338
>2	2	17	70	88	55	17	3	252
Total	32	132	260	226	180	130	77	1037

Because there are only two demographic variables (*HHAGE*, and *Workers*), and all households in the block group are *family*-type households, the proportions in Table-3.40 represent the proportion variable $f_{HHAGE,Workers}$, which will be used in the equation of Step 5 discussed earlier.

Next, we select a sample household in the PUMS. Suppose that a household that we select is a *family*-type household, and the household's demographic information is shown in Table-3.42.

Table-3.42: Example Household from PUMS

			
<u>Age</u>	26	26	7
<u>Income</u>	\$25k	\$18k	\$0
<u>Status</u>	Head of Household	Spouse	Child
<u>Work Status</u>	Worker	Worker	Student
<u>Household Weight, WT</u>	201		

We are also given that within this block group, the proportion of family households (w^f) in the next ten years is 0.4 (meaning that 40% of families in the block group are *family*-type household, 60% left are non-family or group households).

Then, we can update each cell in the PUMS (Table-3.37) by using the following equation:

$$H_{HHAGE,Worker}^{new} = H_{HHAGE,Worker}^{current} + w^f \times f_{HHAGE,Workers} \times WT.$$

The example household in *Table 44* falls in the cell that has *HHAGE* between 25 and 34, and number of workers is two. The cell has 56 households in the PUMS. So the updated value is

$$56 + (0.4)(0.193)(201) = 71.5 \text{ or approximately equals to } 72 \text{ households.}$$

↑
↑
↙

(given)
(From Table3.40)
(From Table 3.42)

The procedure is repeated until all cells in the PUMS are updated/forecasted. Then using the predicted PUMS, and the given marginal values (STF-3A) for the projected year, we again apply the IPF procedure to obtain a forecast of a synthetic population in the format of a cross-classification table for each block group.

3.8.4 Match a Household in Cross-Classification Table to a sample Household in PUMS

At the end of the IPF procedure, a cross-classification table for each block group in a PUMA is determined. The table provides only the **number of households** in each cell classified by demographic variables. However, we need to match each household in the STF-3A cell to an actual-sampled household in the corresponding PUMS cell with similar demographics. Once the match is finalized, the full properties, and characteristics of the sampled household in PUMS are assigned to the synthetic household in STF-3A.

The matching could be conducted by randomly selecting a household from the PUMS that match the demographics of each of the synthetic households in the cells of the estimated multiway tables for each block group. However, Beckman *et al.* [1996] devised a matching procedure based on a weight for each household in the PUMS. The weight is computed based on the differences of demographic values between a PUMS household, and a household in the block group. A PUMS household with closer demographic values to a household in the block group is assigned higher weight/probability. The weight/probability is computed by considering a “distance” between the value of a demographic variable of a PUMS household, indicated by p , and the value of the corresponding demographic variable of a household characterized by a cell, c , in the cross-classification table for the block group. Beckman *et al.* [1996] use the following function to calculate the weight, $D(p, c)$.

$$D(p, c) = w_p \prod_{i \in J} \left(1 - \left| \frac{(d_i^p - d_i^c)}{r_i} \right|^k \right) \cdot \prod_{i \in J} (1 - \Delta(d_i^p, d_i^c)) \quad (9)$$

where

w_p is a weight associated with household p obtained from the PUMS,

J is a set of demographic variables, such as, income, age, number of workers for a household,

d_i^p is a value of the i^{th} demographic variable for household p from the PUMS,

d_i^c is a value of the i^{th} demographic variable for household from a cell c in the cross-classification table in the block group,

r_i is a range of the i^{th} demographic values in the PUMS,

k is an arbitrary positive (>0) constant, note that the less value of k , the more weight yielded for households that are closer in demographic values,

$\Delta(d_i^p, d_i^c) = \begin{cases} \alpha & , d_i^p = d_i^c \\ 1 - \alpha & , d_i^p \neq d_i^c \end{cases}$, where α is an arbitrary positive constant between $[0, 1)$.

The probability of selecting a household p in the PUMS to match with a household in the cell c in the block group is given by

$$Pr\{\text{Selecting Household } p\} = D(p, c) / \sum_{i=1}^n D(i, c). \quad (10)$$

Note that the PUMS is comprised of n households.

Example of Matching Households

Again let us use the example from block group 1 in Arlington County with 2-dimension variables, *householder's age*, and *number of workers*. The PUMS data is presented here in Table-3.43, which is a repeat to Table-3.11. Also sample household out of the 56 households in the cell (25-34, 2) is shown in Table-3.44. Also the IPF final table for this block group 1 is shown in Table-3.45, which is a repeat to Table-3.17.

Table-3.43: The PUMS data for Block Group 1, Arlington County

Householder's Age (HHAGE)								
Workers	15-24	25-34	35-44	45-54	55-64	65-74	>74	Total
0	3	12	14	11	21	51	54	166
1	9	47	63	44	59	45	14	281
2	18	56	113	83	45	17	6	338
>2	2	17	70	88	55	17	3	252
Total	32	132	260	226	180	130	77	1037



Table-3.44-Example Household from the PUMS




			
<i>Age</i>	26 (d_1^p)	26	7
<i>Income</i>	\$25k	\$18k	\$0
<i>Status</i>	Head of Household	Spouse	Child
<i>Work Status</i>	Worker	Worker (Hence $d_2^p = 2$)	Student
<i>Household Weight, w_p</i>	201		

Table-3.45: Final Table for Block Group 1, Arlington County

Householder's Age (HHAGE)								
Workers	15-24	25-34	35-44	45-54	55-64	65-74	>74	Total
0	1	3	2	2	2	6	14	30
1	4	14	10	7	9	7	5	56
2	29	63	56	40	18	10	9	225
>2	0	2	4	4	3	1	0	14
Total	34	82	72	53	32	24	28	325

From Table-3.45, there are 63 households in this block group with the householder's age between 25-34 years, and two workers in the family. The PUMS data has 56 corresponding households. We can randomly pick a household from the 56 households in the PUMS to match one of the 63 households in the block group. This could be repeated 63 times to satisfy all the households. This approach gives equal probability to each of the 56 households in the PUMS data. Beckman *et al.* [1996] want to give higher probabilities to those households with desired demographic characteristics. An example on how this is conducted is presented below.

In this example, there are only two demographic variables, which are HHAGE, and Workers. Hence, J in equation (9) is equal to 2, and HHAGE is assigned $i = 1$, and Worker is assigned $i=2$. Note that for the HHAGE variable ($i = 1$), the values in the cross-classification table for the block group 1 are provided in a **range** format. Hence, the d_1^c is obtained from a mid value in the range. For example in a HHAGE range of [25, 34], the d_1^c equals 29.5 years.

For the second demographic variable, Worker ($i=2$), the value of the corresponding household in the cross-classification table, d_2^c , equals simply to 2 (there are two workers in the household). It is **exactly the same number of** d_2^p in the PUMS. Then, $\Delta(d_2^p, d_2^c) = \alpha$ is activated. Suppose that we are given $k = 1$, and $\alpha = 0.7$, the weight for selecting the household is calculated by using the equation (9) as follows:

$$\begin{aligned}
 D(p, c) &= (201) \times \left(1 - \left| \frac{(d_1^p - d_1^c)}{r_1} \right|^1 \right) \left(1 - \left| \frac{(d_2^p - d_2^c)}{r_2} \right|^1 \right) \cdot (1 - \Delta(d_1^p, d_1^c)) (1 - \Delta(d_2^p, d_2^c)) \\
 &= (201) \times \left(1 - \left| \frac{26 - 29.5}{74 - 15} \right|^1 \right) \left(1 - \left| \frac{2 - 2}{2 - 0} \right|^1 \right) \cdot (1 - \Delta(26, 29.5)) (1 - \Delta(2, 2))
 \end{aligned}$$

$$\begin{aligned}
&= (201)(1 - 0.0593)(1)(1 - (1 - \alpha))(1 - \alpha) \\
&= (201)(0.9407)(0.7)(1 - 0.7) \\
&= \underline{\underline{39.7069}}.
\end{aligned}$$

Once all weights for other households in the PUMS are computed, we can calculate the probability for selecting the household p by dividing the weight for the household p , 39.7069, by a summation of all n households' weights in the PUMS.

After the probabilities for all the households in the PUMS are computed, we select a PUMS household to match with a household in a cross-classification table in a block group one at a time until all households in the block group are matched.

3.8.5 Locate the Households on the Network

The Population Synthesizer module places each synthetic household at **an activity location** on a **walk** link of the TRANSIMS network by using the land use data in the *Activity Location File*. The land use data contain the identity of the activity locations, their locations, the corresponding block group and census tract, and some indication of the activities that may be performed at these locations, for example, residential, commercial, and industrial property surrounding the link, number of employers and the total number of employees around the link, etc. A program called *BlockGroupLoc* is used to generate home locations for synthetic households on a transportation network by correlating census tract and block group user data values specified in the network *Activity Location File* with tract and block group data in the synthetic households. The user data in the activity location table in the TRANSIMS transportation network must contain tract, block group, and residential and commercial land use values. Candidate home locations must have the same census tract and block group as the synthetic household. They also must have residential land use values greater than zero.

When locating a population on a sparse transportation network, some census tract/block groups may not have any associated activity locations. The Tract/Block Group Substitution File enables users to specify a list of alternate census tract/block group pairs that can be used to identify candidate activity locations where the households in the missing tract/block group can be located.

The Block Group Locator uses the Tract/Block Group Substitution File to identify candidate home locations for households whose tract/block group is not found in any activity location on the transportation network. The home location is selected from the activity locations that match the alternative tract/block group pairs that are specified in the file.

Generally, households are placed on these activity locations by taking the steps outlined below. Note that we place households on the network for one **block group at a time**.

Step 1:

Identify all activity locations within a block group. Assume that there are n of them.

Step 2:

Determine a weight obtained from **Activity Location Table** in the **TRANSIMS Network Data**, w_i ; $i = 1, \dots, n$, for each activity location i . These weights are proportional to the probability that a household will be placed at the activity location. Note that the weights determined to the activity locations in a block group will depend on the quality and availability of land use data. For example, the activity locations that have greater residential area should have greater possibilities to be chosen to locate household locations.

Step 3:

Compute the probability, p_i , by dividing each individual weight for an activity by the total weight of all the activity locations in the block groups. The equation is shown below:

$$p_i = \frac{w_i}{\sum_{i=1}^n w_i}.$$

The resultant ratio is used as the probability of a household being located on a network.

Step 4:

Randomly assign each household in the block group to one of the n activity locations according to the probabilities, p_i ; $i = 1, \dots, n$. The household then is placed at that activity location. Notice that households need not to be placed at unique activity locations. Many households can share the same activity location.

3.8.6 Assign Vehicle Ownership for the Households

Each synthetic household is created with a number of vehicles assigned to it. The number of vehicles owned by each household is given in the PUMS and is therefore transferred to the synthetic households through the matching process. The *Vehgen* program creates a TRANSIMS vehicle file that contains information about the household's vehicles and their starting locations. These vehicles are given a unique number and identified as belonging to the household. The Emissions Estimator module (described later) requires that the vehicles be identified by emissions type. However, the emissions type of the vehicles in the household is unknown. The Population Synthesizer contains a model to assign vehicle types to each vehicle in the population. This assignment is done at random according to a national distribution of vehicle emission types. Additionally, a starting location, one of the parking locations on the driving network, is assigned to each vehicle. Traditionally, this location has been the parking location that is closest to the household location. This information is written to the vehicle file.

3.8.7 Assign a Unique Number to Each Household and Person in the Population

The last step to generate a synthetic population is to assign a unique number to each household and person in the population. **BlockGroupLoc** generates the households, and persons IDs. The person number is the unique identifier that is carried through the Route Planner to the Traffic Microsimulator. All of the outputs that are person oriented reference these numbers.

Appendix A

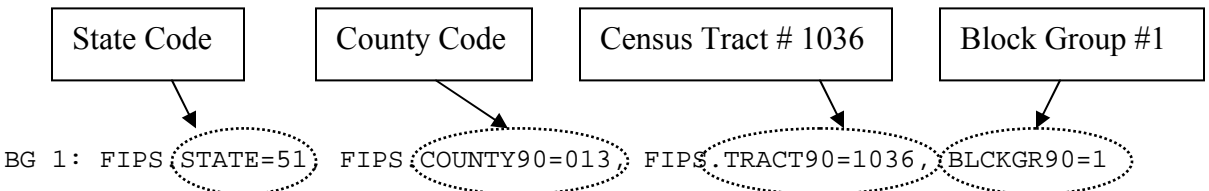
An Example output file obtained by following the steps given in the module and selecting the data to collect for Block group #1, Census Tract #1036 of Arlington County in Virginia State:

1990 Census Lookup

1990 US Census Data

Database: C90STF3A

Summary Level: State--County--County Subdivision--Place/Remainder--Census Tract--Block Group



The Summary Tables are as follows:

For Family Households:

HOUSEHOLD TYPE BY AGE OF HOUSEHOLDER (P24)

Universe: Households

Family households:

15 to 24 years.....	32
25 to 34 years.....	82
35 to 44 years.....	2
45 to 54 years.....	54
55 to 64 years.....	33
65 to 74 years.....	24
75 years and over.....	28

Nonfamily households:

15 to 24 years.....	97
25 to 34 years.....	263
35 to 44 years.....	116
45 to 54 years.....	79
55 to 64 years.....	96
65 to 74 years.....	88
75 years and over.....	52

FAMILY INCOME IN 1989 (P107)

Universe: Families

Less than \$5,000.....	0
\$5,000 to \$9,999.....	12
\$10,000 to \$12,499.....	0
\$12,500 to \$14,999.....	0
\$15,000 to \$17,499.....	0
\$17,500 to \$19,999.....	0
\$20,000 to \$22,499.....	8
\$22,500 to \$24,999.....	9
\$25,000 to \$27,499.....	11
\$27,500 to \$29,999.....	18
\$30,000 to \$32,499.....	14
\$32,500 to \$34,999.....	12
\$35,000 to \$37,499.....	11
\$37,500 to \$39,999.....	14
\$40,000 to \$42,499.....	3
\$42,500 to \$44,999.....	10
\$45,000 to \$47,499.....	23
\$47,500 to \$49,999.....	15
\$50,000 to \$54,999.....	22
\$55,000 to \$59,999.....	14
\$60,000 to \$74,999.....	46
\$75,000 to \$99,999.....	61
\$100,000 to \$124,999.....	22
\$125,000 to \$149,999.....	0
\$150,000 or more.....	0

WORKERS IN FAMILY IN 1989 (P112)

Universe: Families

No workers.....	30
1 worker.....	57
2 workers.....	224
3 or more workers.....	14

POVERTY STATUS IN 1989 BY RACE OF HOUSEHOLDER BY FAMILY TYPE AND PRESENCE AND AGE OF CHILDREN (P124-A)

Universe: Families above the poverty level in 1989

White:

Married-couple family:

With related children under 18 years:

Under 5 years only.....32

5 to 17 years only.....19

Under 5 years and 5 to 17 years.....23

No related children under 18 years.....150

Other family:

Male householder, no wife present:

With related children under 18 years:

Under 5 years only.....0

5 to 17 years only.....0

Under 5 years and 5 to 17 years.....0

No related children under 18 years.....11

Female householder, no husband present:

With related children under 18 years:

Under 5 years only.....5

5 to 17 years only.....6

Under 5 years and 5 to 17 years.....0

No related children under 18 years.....	31
Black:	
Married-couple family:	
With related children under 18 years:	
Under 5 years only.....	0
5 to 17 years only.....	0
Under 5 years and 5 to 17 years.....	0
No related children under 18 years.....	0
Other family:	
Male householder, no wife present:	
With related children under 18 years:	
Under 5 years only.....	0
5 to 17 years only.....	0
Under 5 years and 5 to 17 years.....	0
No related children under 18 years.....	0
Female householder, no husband present:	
With related children under 18 years:	
Under 5 years only.....	0
5 to 17 years only.....	0
Under 5 years and 5 to 17 years.....	0
No related children under 18 years.....	0
American Indian, Eskimo, or Aleut:	
Married-couple family:	
With related children under 18 years:	
Under 5 years only.....	0
5 to 17 years only.....	0
Under 5 years and 5 to 17 years.....	0
No related children under 18 years.....	0
Other family:	
Male householder, no wife present:	
With related children under 18 years:	
Under 5 years only.....	0
5 to 17 years only.....	0
Under 5 years and 5 to 17 years.....	0
No related children under 18 years.....	0
Female householder, no husband present:	
With related children under 18 years:	
Under 5 years only.....	0
5 to 17 years only.....	0
Under 5 years and 5 to 17 years.....	0
No related children under 18 years.....	0
Asian or Pacific Islander:	
Married-couple family:	
With related children under 18 years:	
Under 5 years only.....	6
5 to 17 years only.....	9
Under 5 years and 5 to 17 years.....	0
No related children under 18 years.....	15
Other family:	
Male householder, no wife present:	
With related children under 18 years:	
Under 5 years only.....	0
5 to 17 years only.....	0
Under 5 years and 5 to 17 years.....	0
No related children under 18 years.....	0
Female householder, no husband present:	
With related children under 18 years:	

Under 5 years only.....	0
5 to 17 years only.....	0
Under 5 years and 5 to 17 years.....	0
No related children under 18 years.....	0
Other race:	
Married-couple family:	
With related children under 18 years:	
Under 5 years only.....	0
5 to 17 years only.....	4
Under 5 years and 5 to 17 years.....	1
No related children under 18 years.....	0
Other family:	
Male householder, no wife present:	
With related children under 18 years:	
Under 5 years only.....	0
5 to 17 years only.....	0
Under 5 years and 5 to 17 years.....	0
No related children under 18 years.....	0
Female householder, no husband present:	
With related children under 18 years:	
Under 5 years only.....	0
5 to 17 years only.....	3
Under 5 years and 5 to 17 years.....	0
No related children under 18 years.....	0

POVERTY STATUS IN 1989 BY RACE OF HOUSEHOLDER BY FAMILY TYPE AND PRESENCE
AND AGE OF CHILDREN (P124-B)

Universe: Families below the poverty level in 1989

White:

Married-couple family:	
With related children under 18 years:	
Under 5 years only.....	0
5 to 17 years only.....	0
Under 5 years and 5 to 17 years.....	2
No related children under 18 years.....	0
Other family:	
Male householder, no wife present:	
With related children under 18 years:	
Under 5 years only.....	0
5 to 17 years only.....	0
Under 5 years and 5 to 17 years.....	0
No related children under 18 years.....	0
Female householder, no husband present:	
With related children under 18 years:	
Under 5 years only.....	0
5 to 17 years only.....	2
Under 5 years and 5 to 17 years.....	0
No related children under 18 years.....	0

Black:

Married-couple family:	
With related children under 18 years:	
Under 5 years only.....	0
5 to 17 years only.....	0
Under 5 years and 5 to 17 years.....	0
No related children under 18 years.....	0
Other family:	

Male householder, no wife present:	
With related children under 18 years:	
Under 5 years only.....	0
5 to 17 years only.....	0
Under 5 years and 5 to 17 years.....	0
No related children under 18 years.....	0
Female householder, no husband present:	
With related children under 18 years:	
Under 5 years only.....	0
5 to 17 years only.....	0
Under 5 years and 5 to 17 years.....	0
No related children under 18 years.....	0
American Indian, Eskimo, or Aleut:	
Married-couple family:	
With related children under 18 years:	
Under 5 years only.....	0
5 to 17 years only.....	0
Under 5 years and 5 to 17 years.....	0
No related children under 18 years.....	0
Other family:	
Male householder, no wife present:	
With related children under 18 years:	
Under 5 years only.....	0
5 to 17 years only.....	0
Under 5 years and 5 to 17 years.....	0
No related children under 18 years.....	0
Female householder, no husband present:	
With related children under 18 years:	
Under 5 years only.....	0
5 to 17 years only.....	0
Under 5 years and 5 to 17 years.....	0
No related children under 18 years.....	0
Asian or Pacific Islander:	
Married-couple family:	
With related children under 18 years:	
Under 5 years only.....	0
5 to 17 years only.....	0
Under 5 years and 5 to 17 years.....	0
No related children under 18 years.....	6
Other family:	
Male householder, no wife present:	
With related children under 18 years:	
Under 5 years only.....	0
5 to 17 years only.....	0
Under 5 years and 5 to 17 years.....	0
No related children under 18 years.....	0
Female householder, no husband present:	
With related children under 18 years:	
Under 5 years only.....	0
5 to 17 years only.....	0
Under 5 years and 5 to 17 years.....	0
No related children under 18 years.....	0
Other race:	
Married-couple family:	
With related children under 18 years:	
Under 5 years only.....	0
5 to 17 years only.....	0

Under 5 years and 5 to 17 years.....	0
No related children under 18 years.....	0
Other family:	
Male householder, no wife present:	
With related children under 18 years:	
Under 5 years only.....	0
5 to 17 years only.....	0
Under 5 years and 5 to 17 years.....	0
No related children under 18 years.....	0
Female householder, no husband present:	
With related children under 18 years:	
Under 5 years only.....	0
5 to 17 years only.....	0
Under 5 years and 5 to 17 years.....	0
No related children under 18 years.....	0

For Non-Family Households:

HOUSEHOLD TYPE AND RELATIONSHIP (P17)

Universe: Persons

In family households:

Householder.....	325
Spouse.....	290
Child:	
Natural-born or adopted.....	219
Step.....	6
Grandchild.....	0
Other relatives.....	29
Nonrelatives.....	38

In nonfamily households:

Male householder:

Living alone.....	292
Not living alone.....	89

Female householder:

Living alone.....	304
Not living alone.....	106

Nonrelatives.....252

In group quarters:

Institutionalized persons.....	0
Other persons in group quarters.....	11

Filler.....

RACE OF HOUSEHOLDER BY HOUSEHOLD TYPE AND PRESENCE AND AGE OF CHILDREN (P20)

Universe: Households

White:

Family households:

Married-couple family:

With own children under 18 years.....	76
No own children under 18 years.....	150

Other family:

Male householder, no wife present:

With own children under 18 years.....	0
---------------------------------------	---

No own children under 18 years.....	11
Female householder, no husband present:	
With own children under 18 years.....	13
No own children under 18 years.....	31
Nonfamily households.....	664
Black:	
Family households:	
Married-couple family:	
With own children under 18 years.....	0
No own children under 18 years.....	0
Other family:	
Male householder, no wife present:	
With own children under 18 years.....	0
No own children under 18 years.....	0
Female householder, no husband present:	
With own children under 18 years.....	0
No own children under 18 years.....	0
Nonfamily households.....	67
American Indian, Eskimo, or Aleut:	
Family households:	
Married-couple family:	
With own children under 18 years.....	0
No own children under 18 years.....	0
Other family:	
Male householder, no wife present:	
With own children under 18 years.....	0
No own children under 18 years.....	0
Female householder, no husband present:	
With own children under 18 years.....	0
No own children under 18 years.....	0
Nonfamily households.....	0
Asian or Pacific Islander:	
Family households:	
Married-couple family:	
With own children under 18 years.....	15
No own children under 18 years.....	21
Other family:	
Male householder, no wife present:	
With own children under 18 years.....	0
No own children under 18 years.....	0
Female householder, no husband present:	
With own children under 18 years.....	0
No own children under 18 years.....	0
Nonfamily households.....	48
Other race:	
Family households:	
Married-couple family:	
With own children under 18 years.....	5
No own children under 18 years.....	0
Other family:	
Male householder, no wife present:	
With own children under 18 years.....	0
No own children under 18 years.....	0
Female householder, no husband present:	
With own children under 18 years.....	3
No own children under 18 years.....	0
Nonfamily households.....	12

NONFAMILY HOUSEHOLD INCOME IN 1989 (P110)

Universe: Nonfamily households

Less than \$5,000.....	15
\$5,000 to \$9,999.....	24
\$10,000 to \$12,499.....	12
\$12,500 to \$14,999.....	28
\$15,000 to \$17,499.....	7
\$17,500 to \$19,999.....	13
\$20,000 to \$22,499.....	31
\$22,500 to \$24,999.....	18
\$25,000 to \$27,499.....	69
\$27,500 to \$29,999.....	45
\$30,000 to \$32,499.....	64
\$32,500 to \$34,999.....	31
\$35,000 to \$37,499.....	8
\$37,500 to \$39,999.....	25
\$40,000 to \$42,499.....	46
\$42,500 to \$44,999.....	21
\$45,000 to \$47,499.....	36
\$47,500 to \$49,999.....	7
\$50,000 to \$54,999.....	33
\$55,000 to \$59,999.....	82
\$60,000 to \$74,999.....	70
\$75,000 to \$99,999.....	64
\$100,000 to \$124,999.....	34
\$125,000 to \$149,999.....	8
\$150,000 or more.....	0

POVERTY STATUS IN 1989 BY AGE OF HOUSEHOLDER BY HOUSEHOLD TYPE (P127)

Universe: Households

Income in 1989 above poverty level:

Householder 15 to 64 years:

Married-couple family.....212

Other family:

Male householder, no wife present.....11

Female householder, no husband present.....40

Nonfamily households:

Householder living alone.....455

Householder not living alone.....159

Householder 65 to 74 years:

Married-couple family.....19

Other family:

Male householder, no wife present.....0

Female householder, no husband present.....5

Nonfamily households:

Householder living alone.....81

Householder not living alone.....0

Householder 75 years and over:

Married-couple family.....28

Other family:

Male householder, no wife present.....0

Female householder, no husband present.....0

Nonfamily households:

Householder living alone.....44

Householder not living alone.....	8
Income in 1989 below poverty level:	
Householder 15 to 64 years:	
Married-couple family.....	8
Other family:	
Male householder, no wife present.....	0
Female householder, no husband present.....	2
Nonfamily households:	
Householder living alone.....	16
Householder not living alone.....	21
Householder 65 to 74 years:	
Married-couple family.....	0
Other family:	
Male householder, no wife present.....	0
Female householder, no husband present.....	0
Nonfamily households:	
Householder living alone.....	0
Householder not living alone.....	7
Householder 75 years and over:	
Married-couple family.....	0
Other family:	
Male householder, no wife present.....	0
Female householder, no husband present.....	0
Nonfamily households:	
Householder living alone.....	0
Householder not living alone.....	0

For Group Quarters:

GROUP QUARTERS (P40)

Universe: Persons in group quarters

Institutionalized persons (00I-99I):

Correctional institutions (20I-24I, 27I, 28I, 95I).....	0
Nursing homes (60I-67I).....	0
Mental (Psychiatric) hospitals (45I-48I).....	0
Juvenile institutions (01I-05I, 10I-12I, 15I).....	0
Other institutions (00I, 06I-09I, 13I, 14I, 16I-19I, 25I, 26I, 29I-)	0

Other persons in group quarters (00N-99N):

College dormitories (87N).....	0
Military quarters (96N-98N).....	0
Emergency shelters for homeless persons (82N, 83N).....	0
Visible in street locations (84N, 85N).....	0
Other noninstitutional group quarters (00N-81N, 86N, 88N-95N, 99N).....	11

GROUP QUARTERS BY AGE (P41)

Universe: Persons in group quarters

Institutionalized persons:

Under 18 years.....	0
18 to 64 years.....	0
65 years and over.....	0

Other persons in group quarters:

Under 18 years.....4
 18 to 64 years.....7
 65 years and over.....0

Appendix B

Demographic Button Name Definitions

Following is a list of the demographic options that are available for households and the allowed values for each demographic item.

DATA	SIZE	BEGIN
D RECTYPE	1	1
	Record Type	
V H	.Housing Record	
D SERIALNO	7	2
V 0000000..		
V 9999999	.Housing unit/GQ person serial number unique identifier assigned within state or state group	
D SAMPLE	1	9
	Sample Identifier	
V 1	.5% sample	
V 2	.1% sample	
V 3	.Elderly	
D DIVISION	1	10
	Division code	
V 0	.Region/division not identifiable	
V	.(Selected MSA/PMSAs on 1% sample)	
V 1	.New England (Northeast region)	
V 2	.Middle Atlantic (Northeast region)	
V 3	.East North Central (Midwest region)	
V 4	.West North Central (Midwest region)	
V 5	.South Atlantic (South region)	
V 6	.East South Central (South region)	
V 7	.West South Central (South Region)	
V 8	.Mountain (West region)	
V 9	.Pacific (West region)	
D STATE	2	11
	State Code	
V 01..56	.FIPS state code (See appendix I-59)	
V 99	.PUMA boundaries cross state lines - 1% file	
D PUMA	5	13
	Public use microdata area (state dependent)	
V 00100..		
V 99999	.PUMA code (Includes tract groups) 1st 3	
V	.Digits = main PUMA - generally county place	
V	.Last 2 digits = groups of tracts, BNA, etc.	
D AREATYPE	2	18
	Area type revised for PUMS equivalency file	
	(See Appendix C-1)	
V 10	.Central city	
V 11	.Central city part	

V	20	.MSA/PMSA - Outside central city
V	21	.MSA/PMSA - Outside central city (part)
V	22	.Central City (part) & outside central city
V		.(part)
V	30	.Entire MSA
V	31	.2 or more MSAs/PMSAs
V	40	.Mixed MSA/PMSA/NON-MSA/PMSA area
V	50	.Outside MSA/PMSA
V	60	.Place
V	61	.Place - part
V	70	.MCDs/Towns (New England only)
V	80	.Counties/independent Cities (2 or more)
V	81	.County/independent city - part
V	82	.County/independent city
D	MSAPMSA	4 20
	MSA/PMSA	
V	0040..	
V	9360	.FIPS/MSA/PMSA code, selected MSA/PMSA
V		.(See appendix G)
V	9997	.Mixed MSA/PMSA NONMSA/PMSA area
V	9998	.2 or more MSAs
V	9999	.Not in MA
D	PSA	3 24
	Planning service area (elderly sample only -	
	state dependent)	
V	000	.N/A (Elderly PUMS only)
V	1..18B	.Planning service area codes (See appendix G)
D	SUBSAMPL	2 27
	Subsample number (Use to pull extracts - 1/1000/etc.)	
V	00..99	.See text. pp 4-45.
D	HOUSWGT	4 29
	Housing Weight	
V	0000..	
V	1152	.Integer weight of housing unit
D	PERSONS	2 33
	Number of person records following this housing	
	record	
V	00	.Vacant unit
V	01	.One person record (one person in household
V		.or any person in group quarters)
V	02..29	.Number of person records (number of persons
V		.in household)
D	GQINST	1 35
	Group quarters institution	
V	0	.N/A (housing unit)
V	1	.Institutionalized
V	2	.Not institutionalized
D	HFILLER	3 36
	Filler	
D	UNITS1	2 39
	Units in structure	
V	00	.N/A (GQ)
V	01	.Mobile home or trailer
V	02	.One-family house detached
V	03	.One-family house attached

V	04	.2 Apartments		
V	05	.3-4 Apartments		
V	06	.5-9 Apartments		
V	07	.10-19 Apartments		
V	08	.20-49 Apartments		
V	09	.50 or more apartments		
V	10	.Other		
D	HUSFLAG	1	41	
		All 100% housing unit data substituted		
V	0	.No		
V	1	.Yes		
D	PDSFLAG	1	42	
		All 100% person data substituted		
V	0	.No		
V	1	.Yes		
D	ROOMS	1	43	
		Rooms		
V	0	.N/A (GQ)		
V	1	.1 Room		
V	2	.2 Rooms		
V	3	.3 Rooms		
V	4	.4 Rooms		
V	5	.5 Rooms		
V	6	.6 Rooms		
V	7	.7 Rooms		
V	8	.8 Rooms		
V	9	.9 or more rooms		
D	TENURE	1	44	
		Tenure		
V	0	.N/A (GQ/vacant)		
V	1	.Owned with mortgage or loan		
V	2	.Owned free and clear		
V	3	.Rented for cash rent		
V	4	.No cash rent		
D	ACRE10	1	45	
		On ten acres or more		
V	0	.N/A (GQ/not a one-family house or mobile home)		
V	1	.House on ten or more acres		
V	2	.House on less than ten acres		
D	COMMUSE	1	46	
		Business or medical office on property		
V	0	.N/A (GQ/not a one-family house or mobile home)		
	1	.Yes		
V	2	.No		
D	VALUE	2	47	
		Property value		
V	00	.N/A (GQ/rental unit/vacant, not for sale only)		
V	01	.Less than \$ 10000		
V	02	.\$ 10000 - \$ 14999		
V	03	.\$ 15000 - \$ 19999		
V	04	.\$ 20000 - \$ 24999		
V	05	.\$ 25000 - \$ 29999		
V	06	.\$ 30000 - \$ 34999		
V	07	.\$ 35000 - \$ 39999		
V	08	.\$ 40000 - \$ 44999		
V	09	.\$ 45000 - \$ 49999		

V	0	.N/A (occupied/GQ)	
V	1	.For rent	
V	2	.For sale only	
V	3	.Rented or sold, not occupied	
V	4	.For seasonal/recreational/occasional use	
V	5	.For migratory workers	
V	6	.Other vacant	
D	VACANCY3	1	54
	Boarded up status		
V	0	.N/A (occupied/GQ)	
V	1	.Yes	
V	2	.No	
D	VACANCY4	1	55
	Months vacant		
V	0	.N/A (occupied/GQ)	
V	1	.Less than 1 month	
V	2	.1 up to 2 months	
V	3	.2 up to 6 months	
V	4	.6 up to 12 months	
V	5	.12 up to 24 months	
V	6	.24 or more months	
D	YRMOVED	1	56
	When moved into this house or apartment		
V	0	.N/A (GQ/vacant)	
V	1	.1989 or 1990	
V	2	.1985 to 1988	
V	3	.1980 to 1984	
V	4	.1970 to 1979	
V	5	.1960 to 1969	
V	6	.1959 or earlier	
D	BEDROOMS	1	57
	Bedrooms		
V	0	.N/A (GQ)	
V	1	.No bedrooms	
V	2	.1 Bedroom	
V	3	.2 Bedrooms	
V	4	.3 Bedrooms	
V	5	.4 Bedrooms	
V	6	.5 or more bedrooms	
D	PLUMBING	1	58
	Complete plumbing facilities		
V	0	.N/A (GQ)	
V	1	.Yes, all three facilities	
V	2	.No	
D	KITCHEN	1	59
	Complete kitchen facilities		
V	0	.N/A (GQ)	
V	1	.Yes	
V	2	.No	
D	TELEPHON	1	60
	Telephone in Unit		
V	0	.N/A (GQ/vacant)	
V	1	.Yes	
V	2	.No	

D AUTOS 1 61
 Vehicles (1 ton or less) available
 V 0 .N/A (GQ/vacant)
 V 1 .No vehicles
 V 2 .1 vehicle
 V 3 .2 vehicles
 V 4 .3 vehicles
 V 5 .4 Vehicles
 V 6 .5 Vehicles
 V 7 .6 Vehicles
 V 8 .7 or more vehicles

D FUELHEAT 1 62
 House heating fuel
 V 0 .N/A (GQ/vacant)
 V 1 .Gas: Underground pipes
 V 2 .Gas: Bottled, tank, or LP
 V 3 .Electricity
 V 4 .Fuel oil, kerosene, etc.
 V 5 .Coal or coke
 V 6 .Wood
 V 7 .Solar energy
 V 8 .Other fuel
 V 9 .No fuel used

D WATER 1 63
 Source of water
 V 0 .N/A (GQ)
 V 1 .Public system or private company
 V 2 .Individual drilled well
 V 3 .Individual dug well
 V 4 .Other source such as a spring, creek, etc.

D SEWAGE 1 64
 Sewage disposal
 V 0 .N/A (GQ)
 V 1 .Public sewer
 V 2 .Septic tank or cesspool
 V 3 .Other means

D YRBUILT 1 65
 When structure first built
 V 0 .N/A (GQ)
 V 1 .1989 or 1990
 V 2 .1985 to 1988
 V 3 .1980 to 1984
 V 4 .1970 to 1979
 V 5 .1960 to 1969
 V 6 .1950 to 1959
 V 7 .1940 to 1949
 V 8 .1939 or earlier

D CONDO 1 66
 House or apartment part of condominium
 V 0 .N/A (GQ)
 V 1 .Yes
 V 2 .No

D ONEACRE 1 67
 House on less than 1 acre
 V 0 .N/A (GQ, two or more units in structure)
 V 1 .Yes
 V 2 .No

D	AGSALES	1	68
	1989	Sales of Agriculture Products	
V	0	.N/A (less than 1 acre/GQ/vacant/	
V	2	.2 or more units in structure)	
V	1	.None	
V	2	.\$1 to \$999	
V	3	.\$1,000 to \$2,499	
V	4	.\$2,500 to \$4,999	
V	5	.\$5,000 to \$9,999	
V	6	.\$10,000 or more	
D	ELECCOST	4	69
	Electricity (yearly cost)*		
V	0000	.N/A (GQ/vacant)	
V	0001	.Included in rent or in condo fee	
V	0002	.No charge or electricity not used	
V	0003..		
	3099	.\$3 to \$3,099	
V	3100	.Topcode	
V	3101+	.\$3101 or more = state median of topcoded .values	
D	GASCOST	4	73
	Gas (yearly cost)*		
V	0000	.N/A (GQ/vacant)	
V	0001	.Included in rent or in condo fee	
V	0002	.No charge or gas not used	
V	0003..		
	2099	.\$3 to \$2,099	
V	2100	.Topcode	
V	2101+	.\$2101 or more = state median of topcoded .values	
D	WATRCOST	4	77
	Water (yearly cost)		
V	000	.N/A (GQ/vacant)	
V	001	.Included in rent or in condo fee	
V	002	.No charge	
V	003..999	.\$3 to \$999	
V	1000	.Topcode	
V	1000+	.\$1001+ or more = state median of topcoded .values	
D	FUELCOST	4	81
	House heating fuel (yearly cost)		
V	0000	.N/A (GQ/vacant)	
V	0001	.Included in rent or in condo fee	
V	0002	.No charge or these fuels not used	
V	0003..		
	1899	.\$3 to \$1,899	
V	1900	.Topcode	
V	1,901+	.\$1,901 or more = state median of topcoded .value	
D	RTAXAMT	2	85
	Property taxes (yearly amount)		
V	00	.N/A (GQ/vacant/not owned or being bought/not a .one-family house, mobile home or trailer or .condo)	
V	01	.None	
V	02	.\$ 2 - \$ 49	
V	03	.\$ 50 - \$ 99	
V	04	.\$ 100 - \$ 149	

V	05	.\$ 150 - \$ 199	
V	06	.\$ 200 - \$ 249	
V	07	.\$ 250 - \$ 299	
V	08	.\$ 300 - \$ 349	
V	09	.\$ 350 - \$ 399	
V	10	.\$ 400 - \$ 449	
V	11	.\$ 450 - \$ 499	
V	12	.\$ 500 - \$ 549	
V	13	.\$ 550 - \$ 599	
V	14	.\$ 600 - \$ 649	
V	15	.\$ 650 - \$ 699	
V	16	.\$ 700 - \$ 749	
V	17	.\$ 750 - \$ 799	
V	18	.\$ 800 - \$ 849	
V	19	.\$ 850 - \$ 899	
V	20	.\$ 900 - \$ 949	
V	21	.\$ 950 - \$ 999	
V	22	.\$1000 - \$1099	
V	23	.\$1100 - \$1199	
V	24	.\$1200 - \$1299	
V	25	.\$1300 - \$1399	
V	26	.\$1400 - \$1499	
V	27	.\$1500 - \$1599	
V	28	.\$1600 - \$1699	
V	29	.\$1700 - \$1799	
V	30	.\$1800 - \$1899	
V	31	.\$1900 - \$1999	
V	32	.\$2000 - \$2099	
V	33	.\$2100 - \$2199	
V	34	.\$2200 - \$2299	
V	35	.\$2300 - \$2399	
V	36	.\$2400 - \$2499	
V	37	.\$2500 - \$2599	
V	38	.\$2600 - \$2699	
V	39	.\$2700 - \$2799	
V	40	.\$2800 - \$2899	
V	41	.\$2900 - \$2999	
V	42	.\$3000 - \$3099	
V	43	.\$3100 - \$3199	
V	44	.\$3200 - \$3299	
V	45	.\$3300 - \$3399	
V	46	.\$3400 - \$3499	
V	47	.\$3500 - \$3599	
V	48	.\$3600 - \$3699	
V	49	.\$3700 - \$3799	
V	50	.\$3800 - \$3899	
V	51	.\$3900 - \$3999	
V	52	.\$4000 - \$4099	
V	53	.\$4100 - \$4199	
V	54	.\$4200 - \$4299	
V	55	.\$4300 - \$4399	
V	56	.\$4400 - \$4499	
V	57	.\$4500 = Topcode	
V	58	.\$4501 - \$54992Ä¿RANGE	
V	59	.\$5500 - \$7499 ³ FOR	
V	60	.\$7500 or more2Ä¿ MEDIAN	
D	HFILLER2	3	87
D	INSAMT	4	90
			Fire/hazard/flood insurance (yearly amount)
V	0000		.N/A (not owned or being bought/not a one
V			family house, mobile home, or condo/GQ/vacant)

V 0001 .None
V 0002..
1299 .\$.2 to \$1,299
V 1300 .Topcode
V 1301+ .\$.1,301 or more=state median of topcoded values

D MORTGAG 1 94
Mortgage status
V 0 .N/A (not owned or being bought/not a one family
.house, mobile home, or condo/GQ/vacant)
V 1 .Mortgage deed of trust, or similar debt
V 2 .Contract to purchase
V 3 .None

D MORTGAG3 5 95
Mortgage payment (monthly amount)
V 00000 .N/A (not owned or being bought/not a one
.family house, mobile home, or condo/GQ/vacant)
V 00001 .No regular payment required
V 00002..
01999 .\$.2 to \$1,999
V 02000 .Topcode
V 02001+ . \$2,001 or more = state median of topcoded
.values

D TAXINCL 1 100
Payment include real estate taxes
V 0 .N/A (GQ/vacant/not owned or being bought/
.not a one family house or condo/not mortgaged/
.No regular mortgage payment)
V 1 .Yes, taxes included in payment
V 2 .No, taxes paid separately or taxes not required

D INSINCL 1 101
Payment include fire/hazard/flood insurance
V 0 .N/A (GQ/vacant/not owned or being bought/
.Not a one family house, MHT or condo/not
.mortgaged/no regular mortgage payment)
V 1 .Yes, insurance included in payment
V 2 .No, insurance paid separately or no insurance

D MORTGAG2 1 102
Second mortgage or home equity loan status
V 0 .N/A (GQ/vacant/not owned or being bought/
.not a one family house, mobile home, trailer or
.condo/not mortgaged/no second mortgage)
V 1 .Yes
V 2 .No

D MORTAMT2 5 103
Second mortgage payment (monthly amount)
V 00000 .N/A (GQ/vacant/condo/not owned or being
.bought/not a one family house/not mortgaged/
.no second mortgage)
V 00001 .No regular payment required
V 00002..
00999 .\$.2 to \$999
V 01000 .Topcode
V 01001+ .\$.1001 or more = state median of topcoded
.values

D CONDOFEE 4 108
Condo fee (monthly amount)

V 0000 .N/A (not owned or being bought/not
 V .condo/GQ/vacant/no condo fee)
 V 0001..
 V 0599 \$.1 - \$599

 V 0600 .Topcode
 V 0601+ \$.601 or more = state median of topcoded values

 D MOBLHOME 4 112
 Mobile home costs (yearly amount)
 V 0000 .N/A (GQ/vacant/not owned or being bought/
 V .not mobile home/no costs)
 V 0001..
 V 3399 \$.1 - \$3,399 (cost in dollars)
 V 3400 .Topcode
 V 3401+ \$.3401 or more = state median of topcoded
 V .values

 D RFARM 1 116
 Farm/nonfarm status
 V 0 .N/A (GQ/urban)
 V 1 .Rural farm
 V 2 .Rural nonfarm

 D RGRENT 4 117
 Gross rent
 V 0000 .N/A (GQ/vacant, not rented for cash rent)
 V 0001..
 V 1499 .Gross rent (dollars)
 V 1500 .Topcode
 V 1501+ .1501 or more = state median of topcoded values

 D RGRAPI 2 121
 Gross rent as a percentage of household income in
 1989
 V 00 .N/A (GQ/vacant/not rented for cash rent/owner
 .occupied/no household income)
 V 01 . 1% to 9%
 V 02 .10% to 14%
 V 03 .15% to 19%
 V 04 .20% to 24%
 V 05 .25% to 29%
 V 06 .30% to 34%
 V 07 .35% to 39%
 V 08 .40% to 49%
 V 09 .50% to 59%
 V 10 .60% to 69%
 V 11 .70% to 79%
 V 12 .80% to 89%
 V 13 .90% to 99%
 V 14 .100% or more

 D HFILLER3 1 123
 Filler

 D ROWNRCST 5 124
 Selected monthly owner costs
 V 00000 .N/A (not owned or being bought/not a one
 V .family house, mobile home, or
 V .condo/GQ/vacant/no costs)
 V 00001..
 V 20299 .Monthly owner costs in dollars
 V 20300 .Topcode

D RNSMOCPI 3 129
 Selected monthly owner costs as a percentage of
 household income in 1989

V 000 .N/A (not owned or being bought/not a one family
 .house, mobile home, or condo/GQ/vacant/no HH
 .income)

V 001..100 .1% to 100%

V 101 .101% or more

D RRENTUNT 1 132
 Specified rent unit

V 0 .Not specified rent unit

V 1 .Specified rent unit

D RVALUNT 1 133
 Specified value unit

V 0 .Not specified value unit

V 1 .Specified value unit

D RFAMINC 7 134
 Family income

V 0000000 .N/A(GQ/vacant/no income)

V -999999..

V 9999999 .Total family income in dollars (See user notes
 .for state maximum and minimum values
 .Includes single person households.)

D RHHINC 7 141
 Household income

V 0000000 .N/A(GQ/vacant/no income)

V -999999..

V 9999999 .Total household income in dollars (See user notes
 .for state maximum and minimum values)

D RWRKR89 1 148
 Workers in family in 1989

V 0 .N/A (GQ/vacant/non-family household)

V 1 .No workers

V 2 .1 worker

V 3 .2 workers

V 4 .3 or more workers in family

D RHHLANG 1 149
 Household language

V 0 .N/A (GQ/vacant)

V 1 .English only

V 2 .Spanish

V 3 .Other Indo-European language

V 4 .Asian or Pacific Island language

V 5 .Other language

D RLINGISO 1 150
 Linguistic isolation

V 0 .N/A (GQ/vacant)

V 1 .Not linguistically isolated

V 2 .Linguistically isolated

D RHHFAMTP 2 151
 Household/family type

V 00 .N/A (GQ/vacant)

V 01 .Married-couple family household

V Other family household:

V 02 .Male householder

V	03	.Female householder		
V		Nonfamily household:		
V		Male householder:		
V	11	.Living alone		
V	12	.Not living alone		
V		Female householder:		
V	21	.Living alone		
V	22	.Not living alone		
D	RNATADPT	2	153	
		Number of own natural born/adopted children in household (unweighted)		
V	00	.N/A(GQ/vacant/no own natural born/adopted children)		
V	01..28	.Number of own children natural born/adopted children in household		
D	RSTPCHLD	2	155	
		Number of own stepchildren in household (unweighted)		
V	00	.N/A(GQ/vacant/no own stepchildren)		
V	01..28	.Number of own stepchildren in household		
D	RFAMPERS	2	157	
		Number of persons in family (unweighted)		
V	00	.N/A (GQ/vacant/non-family household)		
V	01..29	.Number of persons in family		
D	RNRLCHLD*	2	159	
		Number of related children in household (unweighted)		
V	00	.N/A (GQ/vacant/no related children)		
V	01..28	.Number of related children in household		
D	RNONREL	1	161	
		Presence of nonrelatives in household		
V	0	.N/A (No nonrelatives in household/GQ/vacant)		
V	1	.1 or more nonrelatives in household		
D	R18UNDR	1	162	
		Presence of person under 18 years in household		
V	0	.N/A (No person under 18 in household/GQ/vacant)		
V	1	.1 or more person under 18 in household		
D	R60OVER	1	163	
		Presence of persons 60 years and over in household		
V	0	.N/A (No person 60 and over/GQ/vacant)		
V	1	.1 person 60 and over (unweighted)		
V	2	.2 or more person 60 and over (unweighted)		
D	R65OVER	1	164	
		Presence of person 65 years and over in household		
V	0	.N/A (No person 65 and over/GQ/vacant)		
V	1	.1 person 65 and over (unweighted)		
V	2	.2 or more person 65 and over (unweighted)		
D	RSUBFAM	1	165	
		Presence of subfamilies in Household		
V	0	.N/A (No subfamilies or not applicable/GQ/vacant)		
V	1	.1 or more subfamilies		
D	AUNITS1	1	166	
		Units in structure allocation		
V	0	.No		

V	1	.Yes	
D	AROOMS	1	167
	Rooms allocation		
V	0	.No	
V	1	.Yes	
D	ATENURE	1	168
	Tenure allocation		
V	0	.No	
V	1	.Yes	
D	AACRES10	1	169
	On ten acres or more allocation		
V	0	.No	
V	1	.Yes	
D	ACOMMUSE	1	170
	Business or medical office on property allocation		
V	0	.No	
V	1	.Yes	
D	AVALUE	1	171
	Value allocation		
V	0	.No	
V	1	.Yes	
D	ARENT1	1	172
	Monthly rent allocation		
V	0	.No	
V	1	.Yes	
D	AMEALS	1	173
	Meals included in rent allocation		
V	0	.No	
V	1	.Yes	
D	AVACNCY2	1	174
	Vacancy status allocation		
V	0	.No	
V	1	.Yes	
D	AVACNCY3	1	175
	Boarded up status allocation		
V	0	.No	
V	1	.Yes	
D	AVACNCY4	1	176
	Months vacant allocation		
V	0	.No	
V	1	.Yes	
D	AYRMOVED	1	177
	When moved into this house or apartment allocation		
V	0	.No	
V	1	.Yes	
D	ABEDROOM	1	178
	Number of bedrooms allocation		
V	0	.No	
V	1	.Yes	
D	APLUMBNG	1	179

			Complete plumbing facilities allocation
V	0	.No	
V	1	.Yes	
D	AKITCHEN	1	180
			Complete kitchen facilities allocation
V	0	.No	
V	1	.Yes	
D	APHONE	1	181
			Telephones in house allocation
V	0	.No	
V	1	.Yes	
D	AVEHICLE	1	182
			Vehicles available by household allocation
V	0	.No	
V	1	.Yes	
D	AFUEL	1	183
			House heating fuel allocation
V	0	.No	
V	1	.Yes	
D	AWATER	1	184
			Source of water allocation
V	0	.No	
V	1	.Yes	
D	ASEWER	1	185
			Sewage disposal allocation
V	0	.No	
V	1	.Yes	
D	AYRBUILT	1	186
			When structure first built allocation
V	0	.No	
V	1	.Yes from not answered	
V	2	.Yes "don't know"	
D	ACONDO	1	187
			House or apartment part of condominium allocation
V	0	.No	
V	1	.Yes	
D	AONEACRE	1	188
			House on less than 1 acre allocation
V	0	.No	
V	1	.Yes	
D	AAGSALES	1	189
			1989 Sales of Agricultural Products allocation
V	0	.No	
V	1	.Yes	
D	AELECCST	1	190
			Electricity (yearly cost) allocation
V	0	.No	
V	1	.Yes	
D	AGASCST	1	191

Gas (yearly cost) allocation
V 0 .No
V 1 .Yes

D AWATRCST 1 192
Water (yearly cost) allocation
V 0 .No
V 1 .Yes

D AFUELCST* 1 193
House heating fuel (yearly cost) allocation
V 0 .No
V 1 .Yes

D ATAXAMT 1 194
Taxes on property allocation
V 0 .No
V 1 .Yes

D AINSAMT 1 195
Fire, hazard, flood insurance allocation
V 0 .No
V 1 .Yes

D AMORTG 1 196
Mortgage status allocation
V 0 .No
V 1 .Yes no answer
V 2 .Yes from junior mortgage

D AMORTG3 1 197
Regular mortgage payment allocation
V 0 .No
V 1 .Yes

D ATAXINCL 1 198
Payment include real estate taxes allocation
V 0 .No
V 1 .Yes

D AINSINCL 1 199
Payment include fire, hazard, flood insurance
allocation
V 0 .No
V 1 .Yes

D AMORTG2 1 200
Second mortgage status allocation
V 0 .No
V 1 .Yes

D AMRTAMT2 1 201
Second mortgage payment allocation
V 0 .NO
V 1 .Yes

D ACNDOFEE 1 202
Condominium fee allocation
V 0 .No
V 1 .Yes

D AMOBLHME 1 203
Mobile home costs allocation
V 0 .No

V 1 .Yes

D FILLER 28 204

Following is a list of the demographic options that are available for persons and the allowed values for each demographic item.

DATA	SIZE	BEGIN
D RECTYPE	1	1
Record Type		
V P		.Person Record
D SERIALNO	7	2
V 0000000..		
V 9999999		.Housing unit/GQ person serial number unique
V		.identifier assigned within state or state group
D RELAT1	2	9
Relationship		
V 00		.Householder
V 01		.Husband/wife
V 02		.Son/daughter
V 03		.Stepson/stepdaughter
V 04		.Brother/sister
V 05		.Father/mother
V 06		.Grandchild
V 07		.Other relative
		Not related
V 08		.Roomer/boarder/foster child
V 09		.Housemate/roommate
V 10		.Unmarried partner
V 11		.Other nonrelative
		Group quarters
V 12		.Institutionalized person
V 13		.Other persons in group quarters
D SEX	1	11
Sex		
V 0		.Male
V 1		.Female
D RACE	3	12
Recorded detailed race code (Appendix C)		
V 001-037		.(See appendix C)
V 301-327		.American Indian Tribes
D AGE	2	15
Age		
V 00		.Less than 1 year
V 01..89		.Age in years
V 90		.90 or more years old
D MARITAL	1	17
Marital status		
V 0		.Now married, except separated
V 1		.Widowed
V 2		.Divorced
V 3		.Separated
V 4		.Never married or under 15 years old
D PWGT1	4	18
Person's weight		
V 0001..		

V 1152 .Person's weight

D PFILLER1 4 22
Filler

D REMPLPAR 3 26
Employment status of parents

V 000 .N/A (not own child of householder, and not
.child in subfamily)

V Living with two parents:

V Both parents in labor force:

V 111 .Both parents at work 35 or more hours

V 112 .Father only at work 35 or more hours

V 113 .Mother only at work 35 or more hours

V 114 .Neither parent at work 35 or more hours

V Father only in labor force:

V 121 .Father at work 35 or more hours

V 122 .Father not at work 35 or more hours

V Mother only in labor force:

V 133 .Mother at work 35 or more hours

V 134 .Mother not at work 35 or more hours

V 141 Neither parent in labor force

V Living with one parent:

V Living with father:

V 211 .Father at work 35 or more hours

V 212 .Father not at work 35 or more hours

V 213 .Father not in labor force

V Living with mother:

V 221 .Mother at work 35 or more hours

V 222 .Mother not at work 35 or more hours

V 223 .Mother not in labor force

D RPOB 2 29
Place of birth (Recode)

V 10 .Born in State of residence

V Born in other State in the U.S.:

V 21 .Northeast

V 22 .Midwest

V 23 .South

V 24 .West

V U.S. outlying areas:

V 31 .Puerto Rico

V 32 .American Samoa

V 33 .Guam

V 34 .Northern Marianas

V 35 .US Virgin Islands

V 36 .Elsewhere

V 40 .Born abroad of American parents

V Foreign born:

V 51 .Naturalized citizen

V 52 .Not a citizen

D RSPOUSE 1 31
Married, spouse present/spouse absent

V 0 .N/A (less than 15 years old)

V 1 .Now married, spouse present

V 2 .Now married, spouse absent

V 3 .Widowed

V 4 .Divorced

V 5 .Separated

V 6 .Never married

D ROWNCHLD 1 32


```

                *Own child (see Appendix B, page 14)
V      * 1      .Own child
V      * 0      .Not own child

D  RAGECHLD          1          33
      Presence and age of own children
V      * 0      .N/A (male)
V      1      .With own children under 6 years only
V      2      .With own children 6 to 17 years only
V      3      .With own children under 6 years and 6 to 17
              .years
V      * 4      .No own children (.incl. females under 16 years)
D  RRELCHLD          1          34

                *Related child (see Appendix B, Page 14)
V      * 1      .Related child
V      * 0      .Not related child

D  RELAT2            1          35
      Detailed relationship (other relative)
V      0      .N/A (GQ/not other relative)
V      1      .Son-in-law/daughter-in-law
V      2      .Father-in-law/mother-in-law
V      3      .Brother-in-law/sister-in-law
V      4      .Nephew/niece
V      5      .Grandparent
V      6      .Uncle/aunt
V      7      .Cousin
V      8      .Other related by blood or marriage
V      9      .Other relative

D  SUBFAM2            1          36
      Subfamily number
V      0      .N/A (GQ/not in a subfamily)
V      1      .In subfamily 1
V      2      .In subfamily 2
V      3      .In subfamily 3

D  SUBFAM1            1          37
      Subfamily relationship
V      0      .N/A (GQ/not in a subfamily)
V      1      .Husband/wife
V      2      .Parent in a parent/child subfamily
V      3      .Child in subfamily

D  HISPANIC           3          38
      Detailed Hispanic origin code (See appendix I)
V  000,006.. .
      199 . Not hispanic
V  001,210..
      220 .Mexican, mex-am
V  002,261..
      270 .Puerto Rican
V  003,271..
      274 .Cuban
V  221..230 .Central American
V  231..249 .South American
V  275..289 .Dominican
V  004,200..
      209,250..
      260
V  290..401 .Other Hispanic

D  POVERTY           3          41

```

Person poverty status recode (See appendix B)

V 000 .N/A

V 001..500 .% Below or above poverty status value
V 501 .501% or more of poverty value

D POB 3 44
Place of birth (Appendix I)

V 001..056 .FIPS State code (U.S. States and D.C.)
V 060..099 .Puerto Rico (072) or U.S. outlying area
V 100..553 .Foreign country
V 554 .At sea
V 555 .Abroad, not specified

D CITIZEN 1 47
Citizenship

V 0 .Born in the U.S.
V 1 .Born in Puerto Rico, Guam, and outlying areas
V 2 .Born abroad of American parents
V 3 .U.S. citizen by naturalization
V 4 .Not a citizen of the U.S.

D IMMIGR 2 48
Year of entry

V 00 .Born in the U.S.
V 01 .1987 to 1990
V 02 .1985 to 1986
V 03 .1982 to 1984
V 04 .1980 or 1981
V 05 .1975 to 1979
V 06 .1970 to 1974
V 07 .1965 to 1969
V 08 .1960 to 1964
V 09 .1950 to 1959
V 10 .Before 1950

D SCHOOL 1 50
School enrollment

V 0 .N/A (less than 3 years old)
V 1 .Not attending school
V 2 .Yes, public school, public college
V 3 .Yes, private school, private college

D YEARSCH 2 51
Educational attainment

V 00 .N/A (less than 3 years old)
V 01 .No school completed
V 02 .Nursery school
V 03 .Kindergarten
V 04 .1st, 2nd, 3rd, or 4th grade
V 05 .5th, 6th, 7th, or 8th grade
V 06 .9th grade
V 07 .10th grade
V 08 .11th grade
V 09 .12th grade, no diploma
V 10 .High school graduate, diploma or GED
V 11 .Some college, but no degree
V 12 .Associate degree in college, occupational program
V 13 .Associate degree in college, academic program
V 14 .Bachelor's degree
V 15 .Master's degree
V 16 .Professional degree

V 17 .Doctorate degree

D ANCSTRY1 3 53
 Ancestry - first entry (See appendix I)

V 001..998 .Ancestry codes - first entry

V 999 .Not reported

D ANCSTRY2 3 56
 Ancestry - second entry (See appendix I)

V 000 .No secondary ancestry

V 001..998 .Ancestry codes

V 999 .Not reported

D MOBILITY 1 59
 Mobility status (lived here on April 1, 1985)

V 0 .N/A(less than 5 years old)

V 1 .Yes same house (nonmovers)

V 2 .No, different house (movers)

D MIGSTATE 2 60
 Migration - State or foreign country code
 (Appendix I)

V 00 .N/A (person less than 5 years old/lived
 .in same house in 1985)

V 01..56 .FIPS state code (U.S. States and D.C.)

V 72 .Puerto Rico

V 98 .Other abroad in 1985

V 99 .State not identified (B sample)

D MIGPUMA 5 62
 Migration PUMA (state dependent)

V 00000 .N/A (person less than 5 years old/lived in
 .same house in 1985)

V 00100..
 99800 .Migration PUMA (Not coded to tract level)

V 99900 .Abroad

D LANG1 1 67
 Language other than English at home

V 0 .N/A (less than 5 years old)

V 1 .Yes, speaks another language

V 2 .No, speaks only English

D LANG2 3 68
 Language spoken at home (See appendix I)

V 000..600 .N/A (less than 5 years old/speaks only
 .English)

V 601..999 .Specific language codes

D ENGLISH 1 71
 Ability to speak English

V 0 .N/A (less than 5 years old/speaks only English)

V 1 .Very well

V 2 .Well

V 3 .Not well

V 4 .Not at all

D MILITARY 1 72
 Military service

V 0 .N/A (less than 16 years old)

V 1 .Yes, now on active duty

V 2 .Yes, on active duty in past, but not now

V 3 .Yes, service in reserves or national guard only

V 4 .No service

D RVETSERV 2 73
Veteran period of service

V 00 .N/A (less than 16 years old, no active duty)
V 01 .September 1980 or later only
V 02 .May 1975 to August 1980 only
V 03 .May 1975 to August 1980 and September 1980
V .or later only
V 04 .Vietnam era, no Korean conflict, no WWII
V 05 .Vietnam era and Korean conflict, no WWII
V 06 .Vietnam era and Korean conflict and WWII
V 07 .February 1955 to July 1964 only
V 08 .Korean conflict, no Vietnam era, no WWII
V 09 .Korean conflict and WWII, no Vietnam era
V 10 .WWII, no Korean conflict, no Vietnam era
V 11 .Other service

D SEPT80 1 75
Served September 1980 or later

V 0 .(Did not serve this period/less than 16 years
V .old)
V 1 .Served this period

D MAY75880 1 76
Served May 1975 to August 1980

V 0 .(Did not serve this period/less than 16 years
V .old)
V 1 .Served this period

D VIETNAM 1 77
Served Vietnam era (August 1964 - April 1975)

V 0 .(Did not serve this period/less than 16 years
V .old)
V 1 .Served this period

D FEB55 1 78
Served February 1955 - July 1964

V 0 .(Did not serve this period/less than 16 years
V .old)
V 1 .Served this period

D KOREAN 1 79
Served Korean conflict (June 1950 - January 1955)

V 0 .(Did not serve this period/less than 16 years
V .old)
V 1 .Served this period

D WWII 1 80
Served World War II (September 1940 - July 1947)

V 0 .(Did not serve this period/less than 16 years
V .old)
V 1 .Served this period

D PFILLER2 1 81
Filler

D OTHRSERV 1 82
Served any other time

V 0 .(Did not serve this period/less than 16 years
V .old)
V 1 .Served this period

D YRSSERV 2 83

Years of active duty military service

V 00 .N/A (less than 16 years/no active duty military
V .service)

V 01 .1 Year or less of service

V 02..49 .2 to 49 years of service

V 50 .50 or more years of service

D DISABL1 1 85
Work limitation status

V 0 .N/A (less than 16 years, and selected persons in
V .GQs - See appendix C)

V 1 .Yes, limited in kind or amount of work

V 2 .No, not limited

D DISABL2 1 86
Work prevented status

V 0 .N/A (less than 16 years, and selected persons in
V .GQs - See appendix C)

V 1 .Yes, prevented from working

V 2 .No, not prevented from working

D MOBILIM 1 87
Mobility limitation

V 0 .N/A (less than 15 years/institutionalized
V .person, and selected persons in GQs -
V .See appendix C)

V 1 .Yes, has a mobility limitation

V 2 .No, does not have a mobility limitation

D PERSCARE 1 88
Personal care limitation

V 0 .N/A (less than 15 years/institutionalized
V .person, and selected persons in GQs -
V .See appendix C)

V 1 .Yes, has a personal care limitation

V 2 .No, does not have a personal care limitation

D FERTIL 2 89
Number of children ever born

V 00 .N/A (less than 15 years/male)

V 01 .No children

V 02 .1 Child

V 03 .2 Children

V 04 .3 Children

V 05 .4 Children

V 06 .5 Children

V 07 .6 Children

V 08 .7 Children

V 09 .8 Children

V 10 .9 Children

V 11 .10 Children

V 12 .11 Children

V 13 .12 or more children

D RLABOR 1 91
Employment status recode

V 0 .N/A (less than 16 years old)

V 1 .Civilian employed, at work

V 2 .Civilian employed, with a job but not at work

V 3 .Unemployed

V 4 .Armed forces, at work

V 5 .Armed forces, with a job but not at work

V 6 .Not in labor force

D WORKLWK 1 92
 Worked last week

V 0 .N/A (less than 16 years old/not at work/
 .unemployed/NILF/Q21A not reported)

V 1 .Worked

V 2 .Did not work

D HOURS 2 93
 Hours worked last week

V 00 .N/A (less than 16 years old/not at
 .work/unemployed/NILF)

V 01..98 .1 to 98 hours worked last week

V 99 .99 or more hours worked last week

D POWSTATE 2 95
 Place of work - state - (Appendix I)

V 00 .N/A (not a worker--not in the labor force,
 .including persons under 16 years; unemployed;
 .employed, with a job not at work; Armed Forces,
 .With a job but not at work)

V 01-56 .FIPS state code (U.S. States and D.C.)

V 98 .Abroad

V 99 .State not identified

D POWPUMA 5 97
 Place of work PUMA (State dependent)

V 00000 .N/A (not a worker--not in the labor force,
 .including persons under 16 years;
 .unemployed; employed, with a job but not at
 .work; Armed Forces, with a job but not at
 .work)

V 00100..
 99800 .PUMA of work (Not coded to tract level)

V 99900 .Abroad

D MEANS 2 102
 Means of transportation to work

V 00 .N/A (not a worker--not in the labor force,
 .including persons under 16 years; unemployed;
 .employed, with a job but not at work; Armed
 .Forces, with a job but not at work)

V 01 .Car, truck, or van

V 02 .Bus or trolley bus

V 03 .Streetcar or trolley car

V 04 .Subway or elevated

V 05 .Railroad

V 06 .Ferryboat

V 07 .Taxicab

V 08 .Motorcycle

V 09 .Bicycle

V 10 .Walked

V 11 .Worked at home

V 12 .Other method

D RIDERS 1 104
 Vehicle occupancy

V 0 .N/A (not a worker or worker whose means of
 .transportation to work was not car, truck,
 .or van)

V 1 .Drove alone

V 2 .2 People

V 3 .3 People

V 4 .4 People
 V 5 .5 People
 V 6 .6 People

 V 7 .7 to 9 people
 V 8 .10 or more people

 D DEPART 4 105
 Time of departure for work - hour and minute
 V 0000 .N/A (not a worker or worker who worked at
 .home)
 V 0001..
 2400 .Time (hour and minute of departure for
 .work) (2400 midnight)

 D TRAVTIME 2 109
 Travel time to work
 V 00 .N/A (not a worker or worker who worked at home)
 V 01..98 .1 to 98 minutes to get to work
 V 99 .99 Minutes or more to get to work

 D TMPABSNT 1 111
 Temporary absence from work
 V 0 .N/A (less than 16 years old/at work/did not
 .report Q25)
 V 1 .Yes, on layoff
 V 2 .Yes, on vacation, temporary illness, labor
 .dispute
 V 3 .No

 D LOOKING 1 112
 Looking for work
 V 0 .N/A (less than 16 years old/at work/did not
 .report Q26A)
 V 1 .Yes
 V 2 .No

 D AVAIL 1 113
 Available for work
 V 0 .N/A (less than 16 years/at work/not looking/
 .Q26A = 0/did not report Q26B)
 V 1 .No, already has a job
 V 2 .No, temporarily ill
 V 3 .No, other reasons (in school, etc.)
 V 4 .Yes, could have taken a job

 D YEARWRK 1 114
 Year last worked
 V 0 .N/A (less than 16 years old)
 V 1 .1990
 V 2 .1989
 V 3 .1988
 V 4 .1985 to 1987
 V 5 .1980 to 1984
 V 6 .1979 or earlier
 V 7 .Never worked

 D INDUSTRY 3 115
 Industry
 V 000 .N/A (less than 16 years old/unemployed who
 .never worked/nilf who last worked prior to
 .1985)
 V 010..992 .specific industry codes (see appendix I)

D	OCCUP	3	118
	Occupation		
V	000	.N/A (less than 16 years old/unemployed who never worked/nilf who last worked prior to .1985)	
V	003..909	.specific occupation codes (see appendix I)	
D	CLASS	1	121
	Class of worker		
V	0	.N/A (less than 16 years old/unemployed who never worked/NILF who last worked prior to .1985)	
V	1	.employee of a private for profit company or business or of an individual, for wages, salary, or commissions	
V	2	.Employee of a private not-for-profit, tax-exempt, or charitable organization	
V	3	.Local government employee (city, county, etc.)	
V	4	.State government employee	
V	5	.Federal government employee	
V	6	.Self-employed in own not incorporated business, professional practice, or farm	
V	7	.Self-employed in own incorporated business, professional practice or farm	
V	8	.Working without pay in family business or farm	
V	9	.Unemployed, last worked in 1984 or earlier	
D	WORK89	1	122
	Worked last year (1989)		
V	0	.N/A (less than 16 years old)	
V	1	.Worked last year	
V	2	.Did not work last year	
D	WEEK89	2	123
	Weeks worked last year (1989)		
V	00	.N/A (less than 16 years old/did not work in .1989)	
V	01..52	.1 to 52 weeks worked last year	
D	HOUR89	2	125
	Usual hours worked per week last year (1989)		
V	00	.N/A (less than 16 years old/did not work in .1989)	
V	01..98	.1 To 98 usual hours	
V	99	.99 Or more usual hours	
D	REARNING	6	127
	Total person's earnings		
V	000000	.N/A (no earnings)	
V	-19996	.Loss of \$19996 or more	
V	-19995..	.Total person's earnings in dollars	
V	283999	.Total person's earnings in dollars	
V	284000	.\$284000 = Topcode	
V	284001+	.State medians included	
D	RPINCOME	6	133
	Total person's income (signed)		
V	000000	.N/A (no income)	
V	-29997	.Loss of \$29997 or more	
V	-29996..	.Total person's income in dollars	
V	400999	.Total person's income in dollars	
V	401000	.Topcode of total person's income	
V	401001+	.State medians included	

D INCOME1 6 139
Wages or salary income in 1989
V 000000 .N/A (less than 16 years old/none)
V 000001..
V 139999 .\$.1 - 139,999
V 140000 .Topcode
V 140001+ .140001 or more = state median of topcoded
V .values

D INCOME2 6 145
Nonfarm self-employment income in 1989 (signed)
V 000000 .N/A (less than 16 years/none)
V -09999 .Loss of \$9,999 or more
V -00001..
V -09998 .Loss \$1 to \$9,998
V 000001 .Break even or \$1
V 000002..
089999 .\$.2 To \$89999
V 090000 .Topcode
V 090001+ .\$.90,001 or more = state median of topcoded
V .values

D INCOME3 6 151
Farm self-employment income in 1989 (signed)
V 000000 .N/A (less than 16 years/none)
V -09999 .Loss of \$9,999 or more
V -00001 to
-09998 .Loss \$1 to \$9,998
V 1 .Break even or \$1
V 000002..
053999 .\$.2 To \$53999
V 054000 .Topcode
V 054001+ .\$.54001 or more = state median of
V .topcoded values

D INCOME4 6 157
Interest, dividends, and net rental income in 1989 (signed)
V 000000 .N/A (less than 15 years/none)
V -09999 .Loss of \$9,999 or more
V -00001 to
-09998 .Loss \$1 to \$9,998
V 1 .Break even or \$1
V 000002..
039999 .\$.2 To \$39999
V 040000 .Topcode
V 040001+ .\$.40001 or more = state median of
V .topcoded values

D INCOME5 5 163
Social security income in 1989
V 00000 .N/A (less than 15 years/none)
V 00001..
16999 .\$.1 to \$16999
V 17000 .Topcode
V 17001+ .17001 or more = state median of topcoded
V .values

D INCOME6 5 168
Public assistance income in 1989
V 00000 .N/A (less than 15 years/none)
V 00001..
9999 .\$.1 To \$9999
V 10000 .Topcode

V 10001+ .\$.10001 or more = state median

D INCOME7 5 173
Retirement income in 1989

V 00000 .N/A (less than 15 years/none)

V 00001..
29999 .\$.1 to \$29999

V 30000 .Topcode

V 30001+ .\$.30001 or more = state median of topcoded
.values

D INCOME8 5 178
All other income in 1989

V 00000 .N/A (less than 15 years/none)

V 00001..
19999 .\$.1 to \$19999

V 20000 .Topcode

V 20001+ .\$.20,001 or more = state median of topcoded
.values

D AAUGMENT 1 183
Augmented person (see text pp. C-5)

V 0 .No

V 1 .Yes

D ARELAT1 1 184
Relationship allocation flag

V 0 .No

V 1 .Yes

D ASEX 1 185
Sex allocation flag

V 0 .No

V 1 .Yes

D ARACE 1 186
Detailed race allocation flag

V 0 .No

V 1 .Yes

D AAGE 1 187
Age allocation flag

V 0 .No

V 1 .Yes

D AMARITAL 1 188
Marital status allocation flag

V 0 .No

V 1 .Yes

D AHISPAN 1 189
Detailed Hispanic origin allocation flag

V 0 .No

V 1 .Yes

D ABIRTHPL 1 190
Place of birth

V 0 .No

V 1 .Yes

D ACITIZEN 1 191
Citizenship allocation flag

V 0 .No

V 1 .Yes

D AIMMIGR 1 192
 Year of entry allocation flag
 V 0 .No
 V 1 .Yes

D ASCHOOL 1 193
 School enrollment allocation flag
 V 0 .No
 V 1 .Yes

D AYEARSCH 1 194
 Highest education allocation flag
 V 0 .No
 V 1 .Yes

D AANCSTR1 1 195
 First ancestry allocation flag
 V 0 .No
 V 1 .Yes

D AANCSTR2 1 196
 Second ancestry allocation flag
 V 0 .No
 V 1 .Yes

D AMOBLTY 1 197
 Mobility status allocation flag
 V 0 .No
 V 1 .Yes

D AMIGSTATE 1 198
 Migration state allocation flag
 V 0 .No
 V 1 .Yes

D ALANG1 1 199
 Language other than English allocation flag
 V 0 .No
 V 1 .Yes

D ALANG2 1 200
 Language spoken at home allocation flag
 V 0 .No
 V 1 .Yes

D AENGLISH 1 201
 Ability to speak English allocation flag
 V 0 .No
 V 1 .Yes

D AVETS1 1 202
 Military service allocation flag
 V 0 .No
 V 1 .Yes

D ASERVPER 1 203
 Military periods of service allocation flag
 V 0 .No
 V 1 .Yes

D AYRSSERV 1 204
 Years of military service allocation flag
 V 0 .No

V 1 .Yes

D ADISABL1 1 205
 Work limitation status allocation flag
 V 0 .No
 V 1 .Yes

D ADISABL2 1 206
 Work prevention status allocation flag
 V 0 .No
 V 1 .Yes

D AMOELLIM 1 207
 Mobility limitation status allocation flag
 V 0 .No
 V 1 .Yes

D APERCARE 1 208
 Personal care limitation status allocation flag
 V 0 .No
 V 1 .Yes

D AFERTIL 1 209
 Children ever born allocation flag
 V 0 .No
 V 1 .Yes

D ALABOR 1 210
 Employment status recode allocation flag
 V 0 .No
 V 1 .Yes

D AHOURS 1 211
 Hours worked last week allocation flag
 V 0 .No
 V 1 .Yes

D APOWST 1 212
 Place of work state allocation flag
 V 0 .No
 V 1 .Yes

D AMEANS 1 213
 Means of transportation to work allocation flag
 V 0 .No
 V 1 .Yes

D ARIDERS 1 214
 Vehicle occupancy allocation flag
 V 0 .No
 V 1 .Yes

D ADEPART 1 215
 Time of departure to work allocation flag
 V 0 .No
 V 1 .Yes

D ATRAVTIME 1 216
 Travel time to work allocation flag
 V 0 .No
 V 1 .Yes

D ALSTWRK 1 217

```

Year last worked allocation flag
V      0  .No

V      1  .Yes

D  AINDUSTR      1      218
    Industry allocation flag
V      0  .No
V      1  .Yes

D  AOCCUP      1      219
    Occupation allocation flag
V      0  .No
V      1  .Yes

D  ACLASS      1      220
    Class of worker allocation flag
V      0  .No
V      1  .Yes

D  AWORK89      1      221
    Worked last year allocation flag
V      0  .No
V      1  .Yes

D  AWKS89      1      222
    Weeks worked in 1989 allocation flag
V      0  .No
V      1  .Yes

D  AHOUR89      1      223
    Usual hours worked per week in 1989 allocation flag
V      0  .No
V      1  .Yes

D  AINCOME1      1      224
    Wages and salary income allocation flag
V      0  .No
V      1  .No (derived)
V      2  .Yes

D  AINCOME2      1      225
    Nonfarm self-employment income allocation flag
V      0  .No
V      1  .No (derived)
V      2  .Yes

D  AINCOME3      1      226
    Farm self-employment income allocation flag
V      0  .No
V      1  .No (derived)
V      2  .Yes

D  AINCOME4      1      227
    Interest, dividend, and net rental income allocation flag
V      0  .No
V      1  .No (derived)
V      2  .Yes

D  AINCOME5      1      228
    Social security income allocation flag
V      0  .No
V      1  .No (derived)

```

```
V          2 .Yes
D AINCOME6      1          229
    Public assistance allocation flag
V          0 .No
V          1 .No (derived)
V          2 .Yes

D AINCOME7      1          230
    Retirement income allocation flag
V          0 .No
V          1 .No (derived)
V          2 .Yes

D AINCOME8      1          231
    All other income allocation flag
V          0 .No
V          1 .No (derived)
V          2 .Yes
```

References

1. Beckman, R. J., Keith A. B., and Michael D. M. Creating Synthetic Baseline Populations. *Transpn Res.-A*. Vol. 30. No. 6 (1996), 415-429.
2. Deming, W. E., and Frederick F. S. On a Least Squares Adjustment of a Sampled Frequency Table When the Expected Marginal Totals Are Known. *Annals of Math Stat.* Vol. 11 (1940), 427-444.
3. Los Alamos National Laboratory. *TRANSPORTATION ANALYSIS SIMULATION SYSTEM (TRANSIMS) Version: TRANSIMS-LANL-1.0*. NM (1999).
4. Los Alamos National Laboratory. *TRANSPORTATION ANALYSIS SIMULATION SYSTEM (TRANSIMS) Version: TRANSIMS-LANL-1.1*. NM (2000).
5. Sherali, H. D. Estimation of Cross-Classified Tables for the Population Synthesizer. *Notes*. Virginia Polytechnic Institute and State University, Blacksburg, VA (2000).